Generation of diffusion tensor imaging data of the human heart using deep convolutional adversarial networks

Yunlong He

N°d'ordre NNT : 2022ISAL0087

THESE de DOCTORAT DE L'INSA LYON, membre de l'Université de Lyon

Ecole Doctorale N° ED 160 EEA
(Electronique, électrotechnique, automatique)

Spécialité/ discipline de doctorat :
Traitement du Signal et de l'Image

Soutenue publiquement le 19/10/ 2022, par:
Yunlong HE

Generation of diffusion tensor imaging data of the human heart using deep convolutional adversarial networks

---

Devant le jury composé de :

MME. RUAN Su, Professeure à Université de Rouen	Rapporteuse
M. YANG jie, Professeur à Shanghai Jiao Tong University	Rapporteur
M. JOUK Pierre-Simon, Professeur des Universités au CHU de Grenoble	Examinateur
M. USSON Yves, Chargé de Recherche CNRS au laboratoire TIMC-IMAG	Examinateur
M. CLARYSSE Patrick, Directeur de recherche CNRS à INSA de Lyon	Examinateur
M. ZHU Yue-Min, Directeur de recherche CNRS à INSA de Lyon	Directeur de thèse
MME. WANG Lihui, Professeure à Guizhou University	Co-directrice de thèse

Département FEDORA – INSA Lyon - Ecoles Doctorales

SIGLE	ECOLE DOCTORALE	NOM ET COORDONNEES DU RESPONSABLE
CHIMIE	CHIMIE DE LYON https://www.edchimie-lyon.fr Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage secretariat@edchimie-lyon.fr	M. Stéphane DANIELE C2P2-CPE LYON-UMR 5265 Bâtiment F308, BP 2077 43 Boulevard du 11 novembre 1918 69616 Villeurbanne directeur@edchimie-lyon.fr
E.E.A.	ÉLECTRONIQUE, ÉLECTROTECHNIQUE, AUTOMATIQUE https://edeea.universite-lyon.fr Sec. : Stéphanie CAUVIN Bâtiment Direction INSA Lyon Tél : 04.72.43.71.70 secretariat.edeea@insa-lyon.fr	M. Philippe DELACHARTRE INSA LYON Laboratoire CREATIS Bâtiment Blaise Pascal, 7 avenue Jean Capelle 69621 Villeurbanne CEDEX Tél : 04.72.43.88.63 philippe.delachartre@insa-lyon.fr
E2M2	ÉVOLUTION, ÉCOSYSTÈME, MICROBIOLOGIE, MODÉLISATION http://e2m2.universite-lyon.fr Sec. : Bénédicte LANZA Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 secretariat.e2m2@univ-lyon1.fr	Mme Sandrine CHARLES Université Claude Bernard Lyon 1 UFR Bioscienc es Bâtiment Mendel 43, boulevard du 11 Novembre 1918 69622 Villeurbanne CEDEX sandrine.charles@univ-lyon1.fr
EDISS	INTERDISCIPLINAIRE SCIENCES-SANTÉ http://ediss.universite-lyon.fr Sec. : Bénédicte LANZA Bât. Atrium, UCB Lyon 1 Tél : 04.72.44.83.62 secretariat.ediss@univ-lyon1.fr	Mme Sylvie RICARD-BLUM Institut de Chimie et Biochimie Moléculaires et Supramoléculaires (ICBMS) - UMR 5246 CNRS - Université Lyon 1 Bâtiment Raulin - 2ème étage Nord 43 Boulevard du 11 novembre 1918 69622 Villeurbanne Cedex Tél : +33(0)4 72 44 82 32 sylvie.ricard-blum@univ-lyon1.fr
INFOMATHS	INFORMATIQUE ET MATHÉMATIQUES http://edinfomaths.universite-lyon.fr Sec. : Renée EL MELHEM Bât. Blaise PASCAL, 3e étage Tél : 04.72.43.80.46 infomaths@univ-lyon1.fr	M. Hamamache KHEDDOUCI Université Claude Bernard Lyon 1 Bât. Nautibus 43, Boulevard du 11 novembre 1918 69 622 Villeurbanne Cedex France Tél : 04.72.44.83.69 hamamache.kheddouci@univ-lyon1.fr
Matériaux	MATÉRIAUX DE LYON http://ed34.universite-lyon.fr Sec. : Yann DE ORDENANA Tél : 04.72.18.62.44 yann.de-ordenana@ec-lyon.fr	M. Stéphane BENAYOUN Ecole Centrale de Lyon Laboratoire LTDS 36 avenue Guy de Collongue 69134 Ecully CEDEX Tél : 04.72.18.64.37 stephane.benayoun@ec-lyon.fr

MEGA	**MÉCANIQUE, ÉNERGÉTIQUE, GÉNIE CIVIL, ACOUSTIQUE** **http://edmega.universite-lyon.fr** Sec. : Stéphanie CAUVIN Tél : 04.72.43.71.70 Bâtiment Direction INSA Lyon mega@insa-lyon.fr	**M. Jocelyn BONJOUR** INSA Lyon Laboratoire CETHIL Bâtiment Sadi-Carnot 9, rue de la Physique 69621 Villeurbanne CEDEX **jocelyn.bonjour@insa-lyon.fr**
ScSo	**ScSo*** **https://edsciencessociales.universite-lyon.fr** Sec. : Mélina FAVETON INSA : J.Y. TOUSSAINT Tél : 04.78.69.77.79 melina.faveton@univ-lyon2.fr	**M. Bruno MILLY** Université Lumière Lyon 2 86 Rue Pasteur 69365 Lyon CEDEX 07 **bruno.milly@univ-lyon2.fr**

*ScSo : Histoire, Géographie, Aménagement, Urbanisme, Archéologie, Science politique, Sociologie, Anthropologie

Acknowledgements

Time flies, and suddenly four years of doctoral studies at INSA-Lyon have passed. I remember when I first came to France four years ago, everything was new to me and I was fascinated and excited. Unfortunately, a year later a serious epidemic suddenly swept the world and changed the way we studied and lived during our PhD. Today, the epidemic seems to have passed and with it comes the imminent graduation of my PhD. During this time, I must thank many people who have helped me tremendously in my four-year PhD career.

First of all, I would like to express my heartfelt gratitude to my supervisor, Professor Yuemin Zhu. During these four years, he guided me day by day, helped me solve complex scientific problems in detail, gave me valuable advice on all aspects of this work, encouraged me when I lacked motivation and confidence. From him I learned the spirit of rigorous, responsible and diligent research, which will always inspire me to face and solve any difficulty in my future life. Besides being a supervisor, he was more like a friend, sharing with us his stories and enthusiasm for research and life, which inspired me during my thesis. Thanks to him, I had a great time as a PhD student in France, and I successfully defended my graduation.

Second, I would like to express my sincere gratitude to my co-supervisor, Professor Lihui Wang who is working in Guizhou University, China. Although we don't have many opportunities to meet face to face, she always provided me with important guidance whenever I encountered complex research problems. I would also like to thank many of her master students with whom I solved one algorithmic problem after another through discussions and exchanges. Without their guidance and detailed help, I would not have been able to carry out my thesis research successfully.

Finally, I am very grateful for the tremendous support of my colleagues and students, including but not limited to Dr. Bingqing Xie, Dr. Pei Niu, Dr. Maxime Di Folco, Dr. Nina Ghigo, Dr. Anne-Lise Duroy, and Dr. Jingfeng Lu, Dr.Shunli Wang, Dr. Zexian Wang. In addition, I would like to thank everyone for the friendship we had in Lyon, and also for their constant care, support, and encouragement. I will remember everything here for the rest of my life.

Generation of diffusion tensor imaging data of the human heart using deep convolutional adversarial networks

Abstract

Diffusion tensor imaging (DTI) has recently emerged as unique medical imaging modality to noninvasively investigate the fiber architecture of the human heart, providing a novel tool for the diagnosis of cardiac diseases. However, clinical applications of cardiac DTI are often penalized by the limited quantity and/or quality of acquired diffusion weighted (DW) images due to limited acquisition times in practice. Recent advances in deep learning, in particular in generative adversarial network (GAN) hints at a new way to cope with this problem. The goal of this thesis is to assist in improving the acquisition efficiency and clinical applicability of cardiac DTI by investigating GAN-based image generation methods. To this end, super angular resolution (SAR) method for cardiac DTI data was first proposed for improving the estimation accuracy of diffusion tensors. Then, the developed SAR method was applied to compensate for motion-induced signal loss in *in vivo* cardiac DTI. Our work has the following four main contributions:

1) We systematically studied the joint influence of angular resolution and noise on cardiac DTI, demonstrating the validity of SAR for improving the accuracy of diffusion tensor estimation.
2) We proposed a GAN-based SAR framework, which can be run on both simulated and real cardiac DTI data to improve the estimation accuracy of diffusion tensor.
3) We investigated the motion-induced signal loss in vivo cardiac DTI and provided a novel data fitting method capable of generating realistic motion-induced DW images directly from existing cardiac DTI datasets.
4) A novel GAN-based SAR has been proposed that can produce high angular-resolution DW images from low angular-resolution motion-induced DW images to compensate for motion-induced signal loss in in vivo cardiac DTI.

Génération de données d'imagerie du tenseur de diffusion du cœur humain à l'aide de réseaux convolutifs adversariaux profonds

Résumé

L'imagerie du tenseur de diffusion (ITD) est récemment apparue comme une modalité d'imagerie médicale unique permettant d'étudier de manière non invasive l'architecture des fibres du cœur humain, offrant ainsi un nouvel outil pour le diagnostic des maladies cardiaques. Cependant, les applications cliniques de l'ITD cardiaque sont souvent pénalisées par la quantité et/ou la qualité limitées des images pondérées en diffusion (diffusion weighted—DW) acquises en raison des temps d'acquisition limités en pratique.

Les récentes avancées dans l'apprentissage profond, en particulier dans le réseau adversarial génératif (RAG), laissent entrevoir une nouvelle façon de faire face à ce problème. L'objectif de cette thèse est d'aider à améliorer l'efficacité de l'acquisition et l'applicabilité clinique de l'ITD cardiaque en étudiant les méthodes de génération d'images basées sur le RAG. À cette fin, une méthode de super résolution angulaire (SRA) pour les données ITD cardiaques a d'abord été proposée pour améliorer la précision de l'estimation des tenseurs de diffusion. Ensuite, la méthode SAR développée a été appliquée pour compenser la perte de signal induite par le mouvement dans l'ITD cardiaque *in vivo*. Notre travail comporte les quatre contributions principales suivantes :

1) Nous avons étudié systématiquement l'influence conjointe de la résolution angulaire et du bruit sur l'ITD cardiaque, démontrant ainsi la validité de la méthode SAR pour améliorer la précision de l'estimation des tenseurs de diffusion.
2) Nous avons proposé un cadre SRA basé sur le RAG, qui peut être appliqué sur des données ITD cardiaques aussi bien simulées que réelles pour améliorer la précision de l'estimation du tenseur de diffusion.
3) Nous avons étudié la perte de signal induite par le mouvement dans l'ITD cardiaque *in vivo* et fourni une nouvelle méthode d'ajustement des données capable de générer des images DW réalistes induites par le mouvement directement à partir des ensembles de données ITD cardiaques existants.
4) Nous avons proposé un nouveau SRA basé sur le RAG qui peut produire des images DW à haute résolution angulaire à partir d'images DW à faible résolution angulaire et corrompues par le mouvement afin de compenser la perte de signal induite par le mouvement dans l'ITD cardiaque *in vivo*.

Contents

List of Figures

List of Tables

Synthèse en Français de la thèse

Introduction Générale

L'imagerie du tenseur de diffusion (en anglais diffusion tensor imaging ou DTI) est une technique avancée d'imagerie par résonance magnétique (IRM) qui mesure la diffusion des molécules d'eau dans les tissus biologiques pour étudier de manière non invasive la microstructure de ces tissus. En particulier, dans le cœur humain, la diffusion des molécules d'eau est beaucoup moins limitée dans la direction des faisceaux de fibres musculaires que dans la direction qui les traverse et reflète donc l'orientation principale de ces fibres. La DTI permet de décrire la diffusion des molécules d'eau à l'intérieur de chaque voxel avec un tenseur de diffusion estimé à partir de signaux pondérés par diffusion (en anglais diffusion weighted ou DW) acquis dans différentes directions de gradient de diffusion. Les valeurs propres et les vecteurs propres du tenseur de diffusion caractérisent, respectivement, les amplitudes et les directions de la diffusion de l'eau dans l'espace tridimensionnel (3D) (Kingsley, 2006a ; Kingsley, 2006b ; Kingsley, 2006c). Il est ainsi possible de représenter de manière non invasive les architectures de fibres sous-jacentes du cœur humain. La fonction du cœur étant fondamentalement liée à l'architecture spécifique de ses fibres, la DTI cardiaque peut être utilisée comme un moyen puissant pour déduire les changements dans le myocarde en relation avec les maladies cardiovasculaires, par exemple l'infarctus du myocarde (Wu, et al., 2006), la cardiomyopathie hypertrophique (Tseng, et al., 2006) et les cardiopathies ischémiques (Mekkaoui, et al., 2017).

Malgré la grande utilité de la DTI cardiaque, un problème courant dans la pratique est que la quantité et/ou la qualité des images pondérées en diffusion acquises sont limitées en raison de la contrainte de temps, ce qui entraîne des estimations inexactes des tenseurs de diffusion. Tout d'abord, la quantité d'images acquises, en particulier dans plusieurs directions de gradient de diffusion, est souvent limitée et inadéquate. Bien que le tenseur de diffusion puisse être calculé mathématiquement à partir de six signaux DW associés à différentes directions de gradient de diffusion (Kingsley, 2006a), davantage de directions sont généralement nécessaires pour obtenir des estimations plus précises du tenseur de diffusion (Jones, 2004). Le nombre de directions de gradient de diffusion utilisées détermine la résolution angulaire des données DTI. Cependant, lorsque le temps d'acquisition est limité, l'augmentation du nombre de directions de gradient de diffusion (résolution angulaire) se fait au prix d'un rapport signal/bruit (SNR) plus faible et/ou d'un nombre limité d'acquisitions pour chaque direction. Deuxièmement, la qualité des images DW est généralement faible, ce qui est particulièrement évident lors de l'application de la DTI au cœur in vivo. Ceci est principalement dû au fait que la DTI mesure essentiellement le mouvement microscopique des molécules d'eau et est donc très sensible aux mouvements du sujet, y compris les battements cardiaques et les mouvements respiratoires. Les mouvements qui se produisent pendant l'imagerie induisent une perte de signal importante dans les images DW et dégradent ainsi considérablement la précision des tenseurs de diffusion estimés. De plus, afin de

minimiser les artefacts causés par le mouvement et d'améliorer le confort du patient, les images DW doivent être acquises rapidement pendant la DTI cardiaque in vivo, ce qui réduit le SNR.

Pour surmonter ces problèmes et améliorer ainsi l'applicabilité de la DTI cardiaque, un large éventail de nouvelles séquences d'ITD ont été proposées. Certaines de ces séquences ont permis d'améliorer la résolution angulaire de la DTI et d'obtenir une plus grande précision dans l'estimation des tenseurs de diffusion, comme l'imagerie de diffusion à haute résolution angulaire (HARDI) et ses variantes (Descoteaux, 1999 ; Descoteaux, et al., 2007 ; Descoteaux, et al., 2009 ; Descoteaux, 2008). Ces séquences sont particulièrement efficaces pour décrire des configurations plus complexes (comme le croisement et la courbure des fibres) qui ne peuvent pas être bien mesurées par un simple modèle DTI. D'autres séquences se sont concentrées sur la réduction des effets causés par le mouvement pendant la DTI cardiaque in vivo. Par exemple, Edelman et al. (Edelman, et al., 1994) ont proposé une séquence en mode d'acquisition par écho stimulé (STEAM) dans laquelle les signaux de diffusion étaient acquis en fonction de la position spatiale et de la forme du myocarde sur l'ensemble du cycle des battements cardiaques de manière à minimiser les effets du mouvement des battements cardiaques. Gamper et al. (Gamper, et al., 2007) ont conçu une séquence basée sur l'écho de spin (SE) pour diminuer la sensibilité du codage de diffusion au mouvement cardiaque tout en maintenant un SNR élevé. En outre, il existe également une variété de séquences qui peuvent être considérées comme des extensions de STEAM ou de SE et qui ont amélioré les performances de la DTI in vivo du cœur humain (Reese, et al., 1995 ; Tseng, et al., 1999 ; Nielles-Vallespin, et al., 2013 ; Stoeck, et al., 2016 ; Nguyen, et al., 2014). Malgré la puissance de ces séquences d'imagerie, l'amélioration de la résolution angulaire et la réduction des effets de mouvement se font au prix de temps d'acquisition de données relativement longs pour ces séquences. Ces temps d'acquisition prolongés sont souvent prohibitifs en routine clinique car ils compromettent le confort du patient et augmentent le risque d'induire des artefacts de mouvement. Cela devient la principale difficulté d'application de ces séquences en pratique.

Ces dernières années, les chercheurs ont commencé à développer des techniques de post-traitement qui peuvent être utilisées comme un outil auxiliaire pour compenser les inconvénients des séquences DTI actuelles et améliorer leur applicabilité. Un grand nombre de méthodes de post-traitement ont été développées pour la super résolution angulaire en ITD cardiaque (Tuch, 2004 ; Descoteaux, et al., 2007 ; Baraniuk, 2007 ; Cheng, et al., 2015 ; Mani, et al., 2015). Bien que ces méthodes soient efficaces, elles nécessitent souvent des protocoles d'imagerie spécialisés et sont donc difficiles à appliquer en pratique. Plusieurs méthodes post-acquisition pouvant être utilisées pour compenser la perte de signal induite par le mouvement dans la DTI cardiaque in vivo ont également été développées (Pai, et al., 2011 ; Wei, et al., 2014 ; Deng, et al., 2020). L'idée commune de toutes ces méthodes de post-traitement est de récupérer les informations souhaitées à partir des données DTI acquises cliniquement. En plus des méthodes ci-dessus, il existe très peu de méthodes pour traiter le problème de l'acquisition des données DTI.

Les récentes avancées dans le domaine de l'apprentissage profond, en particulier dans le réseau adversarial génératif (en anglais generative adversarial network ou GAN), ont montré

les grandes capacités de ce réseau pour la génération d'images médicales. Nous nous intéressons aux applications du GAN pour la génération d'images par RM puisque les données DTI sont essentiellement des images DW-MR. Les applications actuelles de la génération d'images RM basées sur le GAN peuvent être divisées en deux catégories principales. La première catégorie consiste à améliorer la qualité de l'image. Les problèmes de cette catégorie comprennent le débruitage, la super-résolution, etc. La deuxième catégorie est appelée synthèse d'image à image, qui permet de générer des images RM souhaitées à partir d'images RM cliniquement disponibles, ou d'estimer des images RM à partir d'images d'une autre modalité (par exemple, le CT).

L'objectif de cette thèse est d'aider à améliorer l'efficacité d'acquisition et l'applicabilité clinique des données DTI cardiaques en étudiant les méthodes de génération d'images basées sur les GAN. À cette fin, une méthode de super résolution angulaire (SAR) pour les données DTI cardiaques a été proposée pour améliorer la précision de l'estimation des tenseurs de diffusion. Avant de commencer ce travail, nous avons systématiquement étudié l'influence du bruit et de la résolution angulaire sur la précision de l'estimation des tenseurs de diffusion. Nous avons démontré que lorsque le temps d'acquisition est suffisant, la précision des tenseurs de diffusion augmente avec la résolution angulaire. Nous avons construit et entraîné le SAR proposé pour produire des images DW à plus haute résolution angulaire à partir d'images DW à plus basse résolution angulaire. De plus, nous avons modifié le réseau SAR proposé et l'avons appliqué pour compenser la perte de signal induite par le mouvement dans le DTI cardiaque in vivo. Le principal défi de ce travail est le manque de données DTI cardiaques in vivo avec des pertes de signal induites par le mouvement pour permettre au réseau de compenser la perte de signal dans le DTI in vivo. Pour combler cette lacune, nous avons étudié les pertes de signal dans les données de l'ITD cardiaque in vivo, nous avons fourni une méthode simple mais efficace pour générer des pertes de signal induites par le mouvement, puis nous les avons ajoutées aux images DW ex vivo pour imiter les données de l'ITD in vivo et nous avons ensuite entraîné le réseau proposé. Notre travail comporte les quatre contributions principales suivantes :

5) Nous avons étudié systématiquement l'influence conjointe de la résolution angulaire et du bruit sur l'ITD cardiaque, démontrant ainsi la validité de la méthode SAR pour améliorer la précision de l'estimation des tenseurs de diffusion.
6) Nous avons proposé un cadre SRA basé sur le GAN, qui peut être appliqué sur des données ITD cardiaques aussi bien simulées que réelles pour améliorer la précision de l'estimation du tenseur de diffusion.
7) Nous avons étudié la perte de signal induite par le mouvement dans l'ITD cardiaque *in vivo* et fourni une nouvelle méthode d'ajustement des données capable de générer des images DW réalistes induites par le mouvement directement à partir des ensembles de données ITD cardiaques existants.
8) Nous avons proposé un nouveau SRA basé sur le GAN qui peut produire des images DW à haute résolution angulaire à partir d'images DW à faible résolution angulaire et corrompues par le mouvement afin de compenser la perte de signal induite par le mouvement dans l'ITD cardiaque *in vivo*.

Chapter 1 — Imagerie par résonance magnétique de diffusion et imagerie tensorielle de diffusion cardiaque

L'imagerie par résonance magnétique de diffusion (IRMd) est une technique d'imagerie prometteuse qui permet d'obtenir des images non invasives du corps humain en mesurant la diffusion des molécules d'eau dans les tissus. Il s'agit essentiellement d'un type particulier de technique d'imagerie par résonance magnétique (IRM). L'IRM est une technique d'imagerie non invasive capable de fournir des images très détaillées des structures du corps humain (telles que les nerfs, les muscles et les vaisseaux sanguins) en utilisant des champs magnétiques et des impulsions de radiofréquence (RF). L'eau est le principal composant du corps humain. Trois phénomènes se produisent dans le corps humain lorsqu'un champ magnétique externe est appliqué aux molécules d'eau, à savoir la précession, la résonance et la relaxation. Le signal IRM est obtenu à partir de ces trois phénomènes physiques. Le signal acquis reflète les propriétés des tissus contenant des molécules d'eau dans le corps humain.

L'imagerie du tenseur de diffusion (ITD) est l'une des techniques d'IRMd les plus courantes et est de plus en plus utilisée pour caractériser les structures des fibres cardiaques et déduire les anomalies microstructurelles liées aux maladies cardiovasculaires. Elle a été largement utilisée dans l'imagerie cérébrale, et des solutions à la plupart des défis associés à la neuroimagerie basée sur l'ITD ont déjà été développées (Neil, et al., 2002). Jusqu'en 1994, l'utilisation du ITD pour l'imagerie cardiaque a été décrite pour la première fois (Edelman, et al., 1994). Depuis, la technique ITD a été de plus en plus appliquée à la caractérisation non invasive des structures myocardiques du cœur humain sain (Edelman, et al., 1994 ; Tseng, et al., 2000 ; Dou, et al., 2002 ; Gamper, et al., 2007 ; Nielles-Vallespin, et al., 2013) et malade (Wu, et al., 2006 ; Wu, et al., 2009). Pour le cœur humain sain, l'ITD a été appliquée avec succès pour l'évaluation de l'agencement des fibres myocardiques. Bien que plusieurs études histologiques précoces aient montré que les fibres du myocarde normal sont disposées de manière ordonnée selon un schéma hélicoïdal croisé, leurs expériences étaient toujours réalisées sur des cœurs d'animaux (Streeter Jr & Hanna, 1973 ; Stejskal, 1965). Dans ce contexte, les études d'ITD cardiaque sur des cœurs humains ex vivo (Geerts, et al., 2002 ; Hahn, 1950 ; Lombaert, et al., 2012) et in vivo (Reese, et al., 1995 ; Tseng, et al., 2000 ; Tseng, et al., 1999 ; Dou, et al., 2002 ; Gamper, et al., 2007 ; Toussaint, et al., 2013) ont obtenu des résultats cohérents avec les études histologiques, ce qui a permis une validation plus solide de cette structure. Cela offre également des connaissances intéressantes pour la modélisation de l'ITD cardiaque. Dans les cœurs humains malades, l'ITD cardiaque a également été utilisée comme un outil efficace pour décrire le désarroi des fibres induit par la maladie (Wu, et al., 2006 ; Wu, et al., 2009). L'intégrité des fibres après un infarctus du myocarde a souvent été quantifiée en utilisant l'anisotropie fractionnelle (FA) et la diffusivité moyenne (MD) (Wu & Alexander, 2007 ; Sosnovik, et al., 2009). Tous ces résultats suggèrent le potentiel de l'ITD pour le diagnostic précoce et le traitement des maladies cardiovasculaires.

Bien que la plupart des techniques d'imagerie cérébrale basées sur l'ITD puissent être appliquées à l'imagerie cardiaque, de nombreux défis se posent dans la pratique. Le principal défi réside dans le manque de qualité et/ou de quantité. Comme nous le savions, un tenseur de diffusion peut être mathématiquement calculé à partir de six signaux pondérés par diffusion

(en anglais diffusion weighted ou DW) acquis dans six directions de gradient de diffusion (Kingsley, 2006a). En pratique, cependant, le nombre de directions de gradient de diffusion, qui définit la résolution angulaire des données d'ITD, a une influence importante sur l'estimation des tenseurs de diffusion (Tournier, et al., 2013 ; Jones, 2004 ; Zhan, et al., 2013). De nombreux chercheurs ont conclu qu'une résolution angulaire plus élevée (par exemple, plus de 30 directions de gradient de diffusion) est nécessaire pour obtenir une estimation plus précise des tenseurs de diffusion (Jones, 2004). Cependant, l'acquisition d'un grand nombre d'images associées à différentes directions de gradient est souvent impossible en raison des contraintes de temps. Un autre défi réside dans le mouvement global causé par le déplacement pendant l'imagerie cardiaque in vivo. Contrairement au cerveau humain, le mouvement global du cœur est un mélange complexe de mouvement cardiaque combiné à un mouvement cyclique de pompage et de respiration. Comme l'ITD mesure essentiellement le mouvement microscopique des molécules d'eau, elle est très sensible au mouvement du sujet. L'ampleur du mouvement global du cœur qui se produit pendant l'imagerie induit une perte de signal importante dans les images DW, comme illustré à la Figure 1, ce qui dégrade ensuite la précision des tenseurs de diffusion estimés.

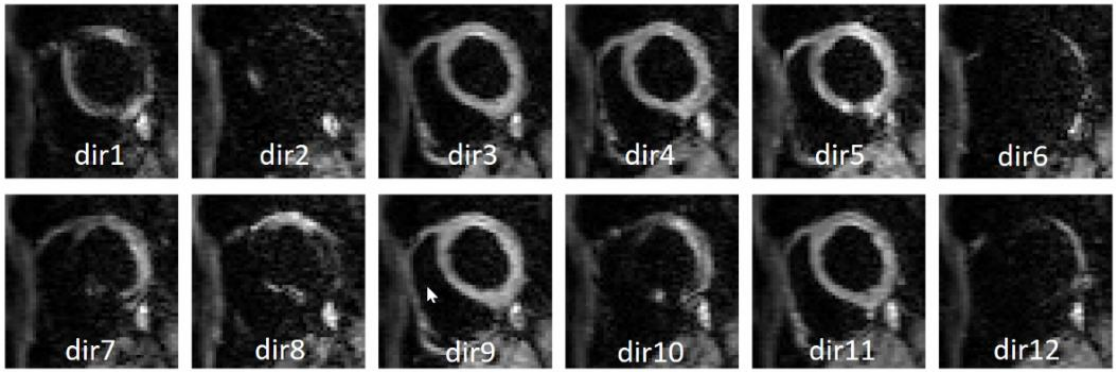

Figure 1 *Images DW cardiaques in vivo acquises dans 12 directions de gradient de diffusion et corrompues par la perte de signal induite par le mouvement.*

Chapter 2 — Réseaux convolutifs profonds adverses pour la génération de données DTI

Le réseau adversarial génératif (abréviations en anglais: GAN) est un cadre d'apprentissage profond prometteur qui suscite un intérêt croissant dans le monde universitaire et l'industrie. Ceci est principalement dû à ses performances remarquables dans la génération d'images souhaitées mais infaisables à acquérir en pratique. Le GAN a récemment été appliqué avec succès à un large éventail de tâches de génération d'images médicales. Cette thèse se concentre sur les applications basées sur le GAN spécifiquement pour la génération de données DTI. Dans ce chapitre, nous présentons une introduction au GAN, qui constitue la méthodologie de cette thèse.

Le GAN est construit à partir d'un réseau de neurones conventionnel (abréviations en anglais: CNN ou ConvNet). Le CNN (ou ConvNet) est une technique d'apprentissage profond bien connue qui a été largement utilisée pour résoudre une variété de problèmes de référence en vision par ordinateur (Alzubaidi, et al., 2021). D'un point de vue computationnel, CNN est

essentiellement une série de réseaux neuronaux profonds qui peuvent prendre une ou plusieurs images en entrée et ensuite produire une autre représentation de l'entrée. La relation entre l'entrée et la sortie est définie par les fonctions mathématiques apprises qui peuvent calculer la somme pondérée pour les entrées multiples et sortir les valeurs d'activation. Les variables de cette fonction sont principalement déterminées par l'architecture du réseau utilisée pour la formation. Bien qu'il existe de nombreuses variantes d'architectures CNN dans la littérature, la plupart d'entre elles sont constituées de composants de base similaires qui peuvent être divisés en trois types de couches de réseau : couche convolutionnelle, couche de mise en commun et couche entièrement connectée. La couche convolutionnelle est le composant principal d'un CNN qui effectue les opérations de convolution du réseau. En général, les couches convolutionnelles initiales capturent les caractéristiques de bas niveau des images, comme les lignes ou les blobs, tandis que les couches ultérieures extraient les caractéristiques de haut niveau, comme les formes ou les objets (Ribeiro, et al., 2018). En outre, une fonction d'activation (par exemple, sigmoïde ou RELU) est souvent utilisée après la convolution pour forcer la relation entre l'entrée et la sortie dans une cartographie non linéaire. La couche de mise en commun est un autre composant important du CNN qui est ajouté après la couche de convolution pour garantir que la sortie produite est indépendante de l'emplacement des caractéristiques dans l'entrée. Elle fonctionne dans les cas où les grandes structures ou les éléments de forme sont plus importants que les détails fins. Les couches entièrement connectées (FC) sont généralement les dernières couches des réseaux CNN. Dans cette couche, chaque élément de la sortie est connecté à tous les éléments de l'entrée.

GAN est maintenant un cadre d'apprentissage profond populaire qui sont devenus un sujet de recherche chaud ces dernières années. Il a été proposé pour la première fois en 2014 par Goodfellow et al. (Goodfellow, et al., 2014), ce qui implique l'utilisation d'une stratégie contradictoire pour la formation de deux réseaux neuronaux profonds. Il se compose de deux sous-réseaux qui sont formés simultanément : un générateur G et un discriminateur D, où G se concentre sur la génération d'images et D sur la discrimination. L'entrée de G, z, est un vecteur de bruit aléatoire échantillonné à partir d'une distribution antérieure $p_z(z)$. La sortie de G, G(z), est censée être aussi similaire que possible à l'échantillon réel (ou image) r qui provient de la distribution de données réelles $p_r(r)$. L'entrée du discriminateur D est soit un échantillon réel r, soit un échantillon généré G(z). Ensuite, D essaie de sortir une valeur unique P_{out} qui représente la probabilité que l'entrée soit un échantillon réel ou faux.

Le GAN fournit un idéal ingénieux pour construire des réseaux neuronaux profonds pour le problème de la génération de données. Avec l'augmentation des applications du GAN dans de nombreux domaines, les limites du modèle original ont également été constatées. Afin de surmonter ces limites et de permettre au réseau d'être appliqué efficacement à divers problèmes, des variantes du GAN ont été proposées. Selon la manière dont elles modifient le modèle GAN original, ces variantes peuvent être divisées en deux catégories : 1) la variation de la fonction objective du GAN, 2) la variation de l'architecture du réseau. Ces dernières années, un grand nombre d'applications du GAN et de ses variantes ont vu le jour dans un large éventail de domaines. Dans le domaine de l'imagerie médicale, il existe généralement deux façons d'utiliser le réseau GAN. La première se concentre sur la génération d'images

médicales, qui peut aider à décrire la distribution sous-jacente des images et à produire des images souhaitées mais difficiles à obtenir en pratique. La deuxième façon fait référence à l'utilisation du discriminateur appris pour traiter certaines tâches de classification. Un grand nombre d'applications connexes ont montré que le réseau neuronal convolutif profond basé sur le GAN peut atteindre des performances de pointe dans la génération d'images RM, ce qui a démontré un grand potentiel pour l'application du GAN aux données DTI.

Chapter 3 — Étude systématique de l'influence conjointe de la résolution angulaire et du bruit dans l'imagerie du tenseur de diffusion cardiaque

L'imagerie du tenseur de diffusion (ITD) est une technique d'imagerie prometteuse qui permet d'étudier de manière non invasive les propriétés de diffusion et les structures des fibres des tissus myocardiques en mesurant la diffusion des molécules d'eau en leur sein. Une estimation précise des tenseurs de diffusion est importante pour mesurer la diffusion de l'eau et ainsi caractériser la structure des fibres du myocarde. Le bruit et la résolution angulaire sont deux facteurs courants qui affectent la précision de l'estimation du tenseur de diffusion dans la pratique. Des études antérieures ont examiné l'influence du bruit ou de la résolution angulaire indépendamment sur l'estimation des tenseurs de diffusion en ITD. Cependant, l'influence conjointe de ces deux facteurs dans l'ITD reste peu claire. Dans cet article, nous proposons d'étudier systématiquement l'influence conjointe des résolutions angulaires et des niveaux de bruit sur l'estimation des tenseurs de diffusion et l'anisotropie fractionnelle (FA) et la diffusivité moyenne (MD) dérivées des tenseurs. À cette fin, des ensembles d'images DW avec différentes résolutions angulaires ont été échantillonnés à partir d'ensembles de données d'ITD cardiaques simulés et réels, puis ont été ajoutés avec différents niveaux de bruit de Rician pour comparaison et analyse. Pour chaque direction, des images DW à plusieurs points temporels ont également été collectées pour étudier la précision de l'estimation du tenseur de diffusion dans le cas de temps de balayage fixes. Les résultats expérimentaux ont montré que, comme prévu, pour un certain niveau de bruit et un temps d'acquisition suffisant, la précision du tenseur de diffusion, de l'AF et du MD augmente avec la résolution angulaire. La Figure 2 montre les courbes de distance de Frobenius en fonction des résolutions angulaires pour différents niveaux de bruit. Ici, le nombre de directions de gradient de diffusion utilisées était Nd = {6, 12, 24, 36, 48, ..., 240}. Sur cette figure, nous pouvons clairement voir que pour chaque SNR, la distance de Frobenius diminue lorsque la résolution angulaire augmente, et que pour chaque résolution angulaire, la distance diminue lorsque le SNR augmente. Les valeurs des distances de Frobenius pour la résolution angulaire maximale (240) sont inférieures à 0.2 même pour un SNR plus faible. En outre, les valeurs de la distance de Frobenius pour l'ensemble de données sans bruit (ligne bleue) sont toujours inférieures à $1.2 \times 10^{-15} mm^2/s$, ce qui signifie que la variation de la résolution angulaire de l'ensemble de données sans bruit n'a aucun impact sur les tenseurs de diffusion estimés.

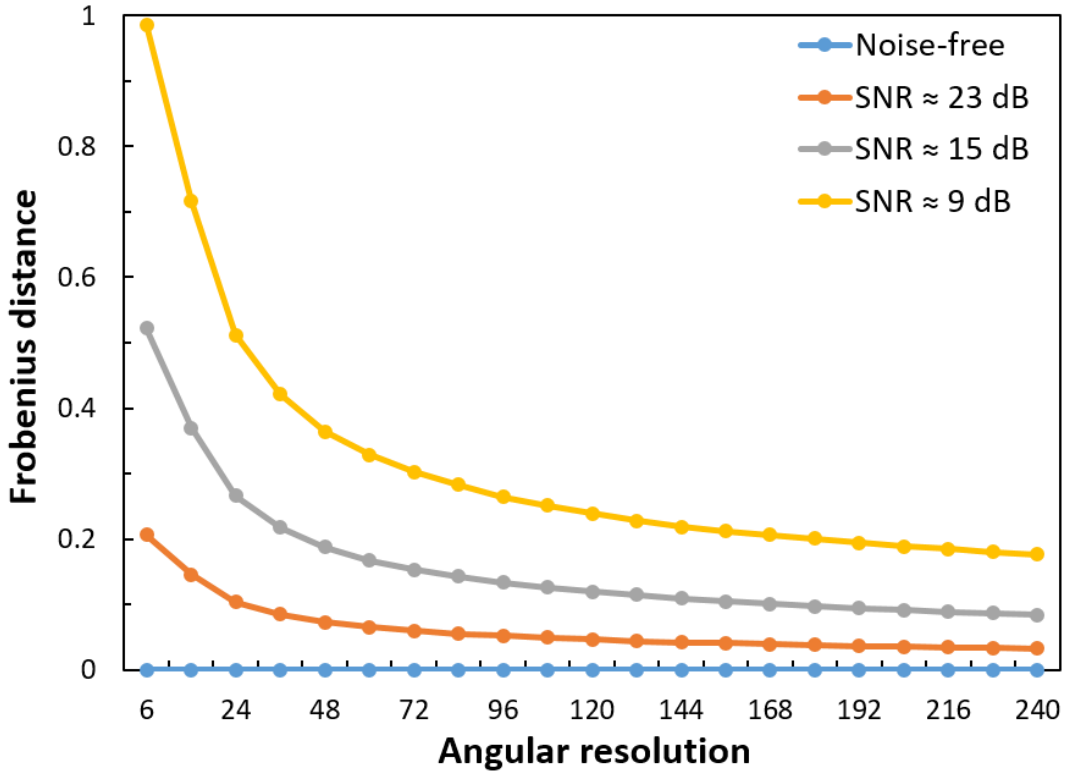

Figure 2 Différences (distance de Frobenius) entre les champs tensoriels estimés et ceux du sol en fonction de la résolution angulaire et du niveau de bruit.

De plus, lorsque la résolution angulaire atteint une certaine valeur, l'augmentation du nombre de résolutions angulaires a peu d'effet sur l'estimation du tenseur de diffusion, de la FA et de la MD. En outre, la moyenne et la variance de FA ou MD diminuent lorsque la résolution angulaire augmente. Pour un temps d'acquisition imposé, l'augmentation de la résolution angulaire réduit le SNR des images DW. En fixant le SNR, une résolution angulaire plus élevée peut être obtenue au prix d'un temps d'acquisition plus long. Ces résultats suggèrent la nécessité d'un compromis optimisé lors de la conception de protocoles d'ITD.

Chapter 4 — Super résolution angulaire basée sur le GAN dans la DTI cardiaque

L'imagerie du tenseur de diffusion (en anglais diffusiion tensor imaging ou DTI) permet l'étude non invasive des structures des fibres cardiaques en utilisant le tenseur de diffusion pour caractériser la diffusion des molécules d'eau en leur sein. Notre étude précédente a démontré que l'augmentation appropriée de la résolution angulaire est une alternative intéressante pour améliorer la précision de l'estimation des tenseurs de diffusion. En pratique, cependant, la DTI à plus haute résolution angulaire est généralement limitée par des temps d'acquisition de données plus longs qui réduisent le confort du patient et introduisent des

problèmes de mouvement. Le réseau adversarial génératif (generative adversarial network ou GAN) offre une solution efficace à ce problème en générant des images pondérées par la diffusion (en anglais diffusion weighted ou DW) dans un plus grand nombre de directions de gradient de diffusion. Malheureusement, la plupart des méthodes existantes basées sur les GAN et développées pour la traduction d'images uniques ne sont pas applicables aux données DTI en raison des informations spatio-angulaires continues des séquences d'images DW plutôt que des informations spatiales indépendantes dans chaque image.

Dans cette étude, nous avons proposé une nouvelle méthode de génération d'images basée sur un réseau GAN pour améliorer la résolution angulaire des données DTI cardiaques et donc la précision de l'estimation des tenseurs de diffusion. Cette méthode consiste à utiliser un réseau GAN pour estimer une séquence d'images à partir d'images DW données, acquises dans un nombre limité de directions de gradient de diffusion. La figure 4.1 illustre un aperçu du réseau GAN proposé. Le générateur G applique la transformation à n images DW observées X pour prédire un ensemble de n images DW G(X) associées à n nouvelles directions. Les entrées du discriminateur sont à la fois les images DW observées X et les n images DW supplémentaires associées à n nouvelles directions, où les n images DW supplémentaires peuvent être soit les images de sortie G(X), soit les images réelles R qui proviennent de l'ensemble de données original. Le discriminateur tente ensuite de décider si l'ensemble d'images d'entrée est la vraie paire {X, G(X)} ou la fausse paire {X, R}. Dans la Figure 3, les flèches en pointillés rouges indiquent la rétropropagation du réseau qui ajuste les paramètres du réseau pour mettre à jour le générateur. Ce processus se poursuit de manière itérative jusqu'à ce que le générateur puisse produire des séquences d'images DW très réalistes.

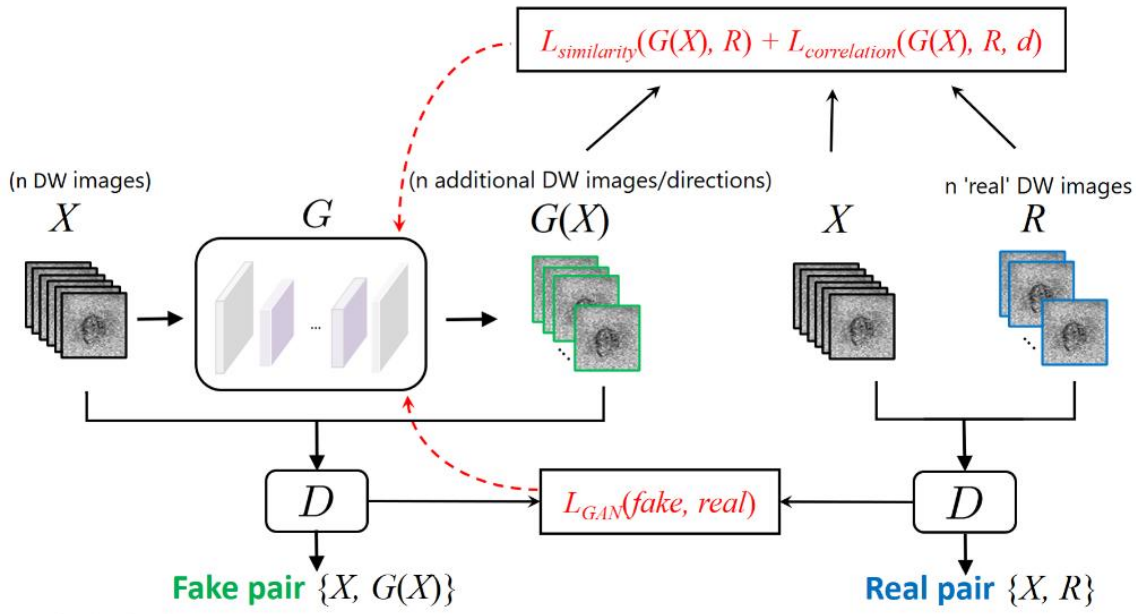

Figure 3 Illustration schématique du GAN régularisé par le manifold proposé. Le réseau mappe les images DW acquises dans n directions de gradient dans n nouvelles directions.

De plus, la corrélation entre plusieurs images DW associées à différentes directions de gradient est exploitée en introduisant un terme régularisé de collecteur supervisé dans la fonction objectif du réseau proposé. Les résultats expérimentaux sur des données DTI

cardiaques ont démontré que la méthode proposée améliore significativement la qualité de la reconstruction DTI. Cela suggère que la méthode proposée est un outil efficace pour la résolution des super-angles en DTI cardiaque.

Chapter 5 — Étude de la compensation de la perte de signal dans la DTI cardiaque in vivo

L'imagerie du tenseur de diffusion (en anglais diffusion tensor imaging ou DTI) est un moyen puissant d'explorer de manière non invasive les propriétés de diffusion et les structures des fibres des tissus biologiques en mesurant la diffusion de l'eau dans les tissus. Elle a été largement utilisée pour étudier les structures des fibres du myocarde et accéder aux anomalies liées à de nombreuses maladies cardiaques. Cependant, la DTI du cœur humain in vivo reste un défi ouvert car les mouvements cardiaques et respiratoires peuvent induire une perte de signal importante dans les images pondérées par diffusion (diffusion weighted—DW). Les séquences DTI avancées peuvent réduire l'effet du mouvement mais généralement au prix de temps d'acquisition plus longs, ce qui limite leur application en pratique. Les techniques actuelles de post-traitement, en particulier les réseaux de neurones convolutionnels (convolutional neural networks—CNNs), offrent une solution pour compenser la perte de signal induite par le mouvement dans les images DW acquises. Malheureusement, le peu de données DTI cardiaques in vivo disponibles en pratique rend difficile l'entraînement de CNNs efficaces. Cette étude vise à étudier les méthodes de compensation de la perte de signal pour la DTI cardiaque in vivo. À cette fin, la perte de signal causée par le mouvement dans les ensembles de données DTI cardiaques in vivo a été étudiée. Ensuite, une nouvelle méthode d'ajustement des données a été proposée pour générer des images DW induites par le mouvement directement à partir d'ensembles de données DTI cardiaques existants afin d'imiter les données DTI cardiaques in vivo. Enfin, la méthode proposée a été validée par l'entraînement d'un nouveau réseau CNNs à l'aide des images DW induites par le mouvement. Les résultats de cette étude ont montré que la méthode proposée peut être utilisée en combinaison avec les CNNs pour compenser la perte de signal induite par le mouvement dans la DTI cardiaque in vivo.

Chapter 6 — Compensation de la perte de signal induite par le mouvement dans la DTI in vivo du cœur humain à l'aide d'une super résolution angulaire basée sur le GAN

Ces dernières années, l'imagerie du tenseur de diffusion (diffusion tensor imaging—DTI) est devenue une modalité indispensable pour caractériser de manière non invasive l'architecture des fibres du cœur humain. La DTI du cœur humain a été largement utilisée pour caractériser l'architecture des fibres myocardiques du cœur et accéder à la disposition histologique des fibres myocardiques dans les tissus. Cependant, les applications cliniques de La DTI cardiaque in vivo restent extrêmement limitées en raison de l'importante perte de signal causée par les mouvements pendant l'imagerie. Dans ce chapitre, nous proposons une

nouvelle méthode qui permet de compenser la perte de signal induite par les mouvements dans la DTI cardiaque in vivo. Cette méthode consiste à entraîner un réseau d'apprentissage profond pour produire des images pondérées par diffusion (diffusion weighted—DW) de plus haute résolution angulaire à partir d'images DW de plus basse résolution angulaire acquises en pratique. Le réseau proposé est construit sur la base d'un réseau adversatif génératif (GAN) combiné à une norme L2 pour produire des images DW cardiaques réalistes. En outre, la cohérence angulaire entre les images DW de même résolution angulaire est utilisée comme terme de régularisation supplémentaire pour améliorer les performances de la compensation de la perte de signal.

La Figure 4 illustre une vue d'ensemble du réseau de super-résolution angulaire proposé, basé sur un GAN. Il contient deux sous-réseaux entraînés à se concurrencer l'un l'autre, un générateur G et un discriminateur D. Le générateur G est construit pour prendre en entrée des images DW existantes de résolution angulaire inférieure L et produire des images DW de résolution angulaire supérieure G(L). D est essentiellement un modèle discriminant qui prend en entrée L et un ensemble d'images DW à plus haute résolution angulaire, où les images à plus haute résolution angulaire peuvent être soit G(L), soit des images de base H. Il essaie de déterminer si l'entrée est une vraie paire {L, H} ou une fausse paire {L, G(L)}. Comme le montre la Figure 4, le générateur G apprend la correspondance entre les images DW de résolution angulaire inférieure L et les images DW de résolution angulaire supérieure H en se basant sur trois contraintes principales : la similarité des images, la cohérence angulaire et la stratégie d'apprentissage contradictoire.

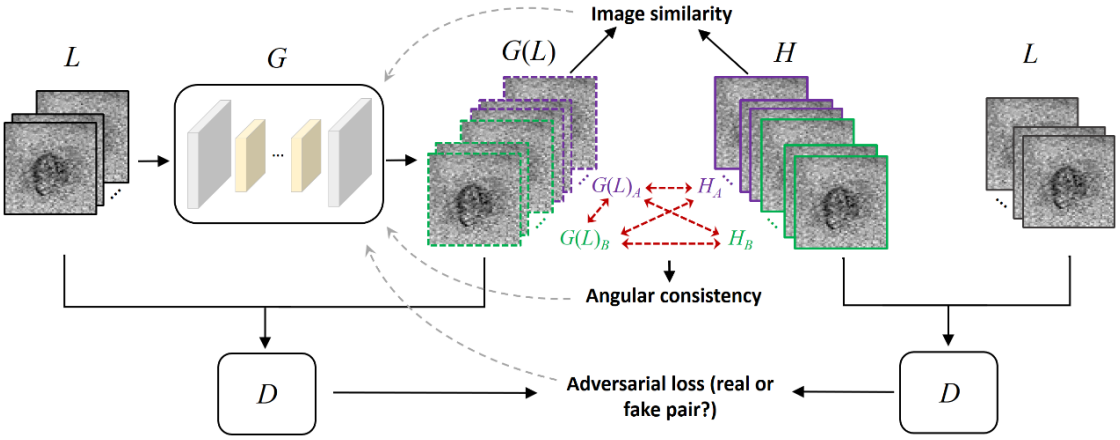

Figure 4 Illustration schématique du réseau de super-résolution angulaire proposé basé sur un GAN. Les flèches pointillées grises indiquent la rétropropagation pour ajuster les paramètres du réseau afin de mettre à jour G, et les flèches pleines bidirectionnelles rouges représentent la cohérence angulaire entre les ensembles d'images DW.

Afin d'entraîner et de tester le réseau proposé, des ensembles de données DTI cardiaques in vivo sont utilisés en combinaison avec des ensembles de données DTI cardiaques ex vivo pour imiter de nombreuses images DW induites par le mouvement avec une résolution angulaire inférieure et leurs images DW de base avec une résolution angulaire supérieure.

La Figure 5 montre les cartes d'orientation des fibres (cartes d'angle d'élévation et d'azimut) obtenues en exécutant le réseau proposé et les réseaux de pointe directement sur les images DW cardiaques d'un cœur humain in vivo. On observe que les angles d'élévation originaux (calculés à partir des images DW in vivo originales) dans le ventricule gauche varient de

manière irrégulière de l'épiocarde à l'endocarde, et que les angles d'azimut originaux dans le ventricule gauche ne présentent pas de modèle de variation circulairement périodique. Après compensation de la perte de signal à l'aide de PCATMIP (Pai, et al., 2011), WIFPCA (Wei, et al., 2014), DIF-net (Jung, et al., 2020), U2Fusion-net (Xu, et al., 2020) et de la méthode proposée, les angles d'élévation dans le ventricule gauche ont montré une variation grossièrement linéaire de l'épiocarde à l'endocarde, et la variation circulaire périodique correspondante des angles d'azimut a également été récupérée. En particulier, les variations régulières des angles d'élévation et d'azimut pour la WIFPCA et la méthode proposée ont été observées plus clairement par rapport aux autres méthodes. Mais les angles pour la méthode proposée ont montré un niveau de bruit relativement plus faible par rapport à celui de la WIFPCA. Le Table 1 résume la moyenne et l'écart type (Moyenne ± SD) des valeurs FA et MD obtenues après avoir utilisé la méthode proposée et celles de l'état de l'art sur les ensembles de données DTI cardiaques des six cœurs humains in vivo. Nous observons que la moyenne et les variations de FA et de MD pour WIFPCA étaient inférieures à celles de PCATMIP, ce qui est cohérent avec les résultats de (Wei, et al., 2014). En revanche, la méthode proposée a donné la moyenne et les variations les plus faibles de FA (0.42 ± 0.08) et de MD (0.65 ± 0.09 × 10^{-3} mm^2/s) par rapport à toutes les autres méthodes.

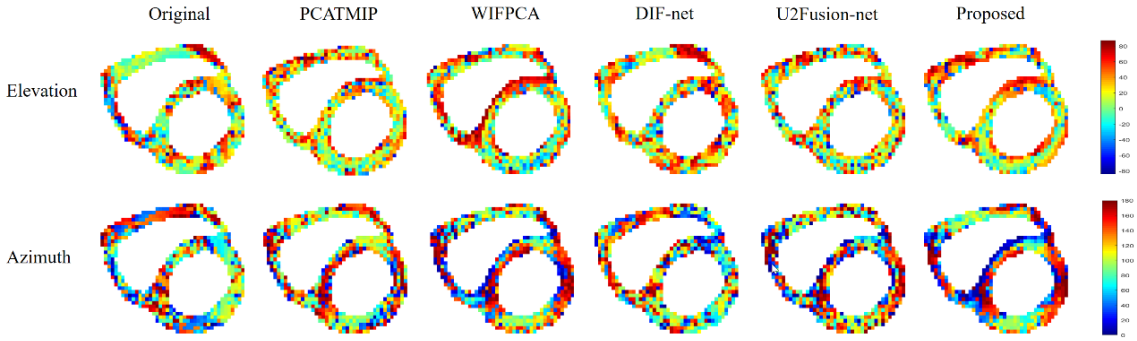

Figure 5 Cartes d'angle d'élévation (rangée supérieure) et cartes d'angle d'azimut (rangée inférieure) calculées à partir des images DW cardiaques in vivo originales (première colonne à gauche) et des images DW après compensation de la perte de signal à l'aide du réseau proposé et des réseaux de pointe (de la deuxième à la dernière colonne).

Les résultats expérimentaux, dont ceux de la Figure 5 et du Tableau 1, ont tous montré que les images produites à plus haute résolution angulaire permettent une amélioration significative de la précision de l'estimation des tenseurs de diffusion, de l'anisotropie fractionnelle (FA) et de la diffusivité moyenne (MD). De plus, les orientations des fibres sont mieux caractérisées grâce à la compensation des pertes de signal. La méthode proposée est donc une méthode de post-traitement intéressante pour compenser la perte de signal induite par le mouvement dans la DTI cardiaque in vivo.

Table 1 Mean ± SD de FA et MD dans le LV de six coeurs humains in vivo. L'unité de MD est 10^{-3} mm^2/s.

	FA ± SD	**MD ± SD**
Original	0.63 ± 0.14	0.63 ± 0.14
PCATMIP (Pai, et al., 2011)	0.51 ± 0.12	0.85 ± 0.11
WIFPCA (Wei, et al., 2014)	0.46 ± 0.10	0.73 ± 0.09
DIF-net (Jung, et al., 2020)	1.52 ± 0.67	2.40 ± 0.18
U2fusion-net (Xu, et al., 2020)	0.61 ± 0.16	1.42 ± 0.12
Proposed	0.42 ± 0.08	0.65 ± 0.09

Conclusions générales

L'ITD est une modalité d'imagerie médicale unique permettant d'étudier de manière non invasive l'architecture des fibres du cœur humain, offrant ainsi un nouvel outil pour le diagnostic des maladies cardiaques. Cependant, les applications cliniques de la DTI cardiaque sont souvent pénalisées par la quantité et/ou la qualité limitées des images DW acquises en raison des temps d'acquisition limités en pratique. Les récentes avancées dans l'apprentissage profond, en particulier dans le GAN, laissent entrevoir une nouvelle façon de faire face à ce problème. Dans cette thèse, nous présentons le travail original d'investigation des réseaux basés sur GAN pour la super résolution angulaire en DTI cardiaque et la compensation de la perte de signal induite par le mouvement en DTI cardiaque in vivo. Avant de présenter notre travail, l'introduction à l'IRMd incluant l'ITD et les applications de l'ITD cardiaque sont présentées dans le chapitre 1, et les bases et concepts du CNN et du GAN sont présentés dans le chapitre 2. Les conclusions de notre travail du chapitre 3 au chapitre 6 sont résumées comme suit :

Dans le chapitre 3, nous avons étudié systématiquement l'influence conjointe des résolutions angulaires et des niveaux de bruit dans la DTI cardiaque. Les résultats obtenus sur des images DW synthétiques et réelles ont montré que, compte tenu d'un temps d'acquisition suffisant et d'un certain niveau de bruit, la précision des tenseurs de diffusion, des mesures FA et MD augmente avec l'augmentation des résolutions angulaires. En outre, continuer à augmenter la résolution angulaire au-delà d'une certaine valeur a peu d'effet sur la précision des tenseurs de diffusion, de l'AF et de la MD. Pour un SNR plus élevé (23 dB), les tenseurs de diffusion, FA et MD quasi-optimaux ont été obtenus avec 48 directions de gradient. Avec un SNR plus faible, le tenseur de diffusion, l'AF et le MD quasi-optimaux ont été obtenus avec, par exemple, 120 directions si le SNR = 15 dB et avec 156 directions si le SNR = 9 dB. En outre, les résultats sur des données réelles montrent que la moyenne et la variance de FA ou MD diminuent lorsque la résolution angulaire augmente. Lorsque le temps d'acquisition est imposé, l'augmentation de la résolution angulaire se fait souvent au prix d'une réduction du nombre d'excitations utilisées pour le moyennage du signal, ce qui réduit le SNR des images DW. Des temps d'acquisition plus longs permettent d'utiliser une résolution angulaire plus élevée tout en conservant un certain niveau de SNR. Ces résultats démontrent la validité d'une résolution angulaire plus élevée pour améliorer la précision de l'estimation du tenseur de diffusion.

Sur la base de ces résultats présentés au chapitre 3, nous avons proposé un réseau basé sur le GAN pour la super résolution angulaire dans l'ITD cardiaque, qui sont détaillés au chapitre 4. Avec le réseau proposé, les images DW dans des directions de gradient supplémentaires sont estimées à partir d'une séquence d'images DW donnée acquise dans un nombre limité de directions de gradient. Les expériences sur les ensembles de données de l'ITD cardiaque ont montré que le réseau proposé est capable de générer des images de plus haute résolution à partir d'images DW existantes avec une résolution angulaire plus faible et donc d'améliorer la précision de l'estimation des tenseurs de diffusion.

Afin d'appliquer le réseau proposé aux données de l'ITD cardiaque in vivo, nous avons mené une étude sur la compensation de la perte de signal dans l'ITD cardiaque in vivo, décrite

au chapitre 5. Dans cette étude, le nombre d'images et de pixels pour différents degrés de pertes de signal dans les ensembles de données DTI cardiaques in vivo ont été utilisés pour générer des pertes de signal réalistes induites par le mouvement. Après avoir ajouté les pertes de signal générées aux images DW simulées et ex vivo, on a obtenu une grande quantité d'images DW induites par le mouvement, qui peuvent imiter visuellement et de manière distributive les images DW in vivo corrompues par le mouvement. Les expériences sur les images DW induites par le mouvement ont démontré que les méthodes proposées constituent un moyen efficace de générer des images DW réalistes induites par le mouvement qui peuvent être utilisées pour développer des méthodes avancées de post-traitement basées sur les CNN dans le DTI cardiaque in vivo.

Enfin, dans le chapitre 6, un nouveau réseau GAN à super-résolution angulaire a été proposé pour compenser la perte de signal induite par le mouvement dans l'ITD in vivo du cœur humain. Le réseau GAN proposé a été entraîné en utilisant une fonction objective basée sur l'apprentissage contradictoire, la similarité des images et la cohérence angulaire. Plusieurs stratégies de prétraitement ont été utilisées pour générer suffisamment d'images DW in vivo à basse et haute résolution angulaire pour entraîner et tester le réseau proposé. Les résultats expérimentaux ont montré que le réseau proposé est capable de compenser la perte de signal induite par le mouvement en termes d'estimation précise des tenseurs de diffusion et des paramètres dérivés de la DTI tels que FA et MD. En outre, les résultats ont également démontré que l'utilisation de la cohérence angulaire a amélioré le réseau en produisant des images DW à partir desquelles les orientations des fibres ont été dérivées avec plus de précision. La méthode proposée fournit donc une méthode intéressante pour la compensation des mouvements dans la DTI cardiaque in vivo.

General Introduction

Diffusion tensor imaging (DTI) is an advanced magnetic resonance imaging (MRI) technique that measures the diffusion of water molecules in biological tissues to noninvasively study the the microstructure of these tissues. In particular, in the human heart, the diffusion of water molecules is much less restricted in the direction along muscle fiber bundles than in the direction through them and thus reflects the principal orientation of those fibers. DTI allows to describe the diffusion of water molecules inside each voxel with a diffusion tensor estimated from diffusion weighted (DW) signals acquired in different diffusion gradient directions. The eigenvalues and eigenvectors of the diffusion tensor characterize, respectively, the magnitudes and directions of the water diffusion in three-dimentional (3D) space (Kingsley, 2006a; Kingsley, 2006b; Kingsley, 2006c). This makes it possible to noninvasively depict the underlying fiber architectures of the human heart. Because the function of the heart is fundamentally linked to its specific fiber architecture, cardiac DTI can be used as a powerful means to infer the changes in the myocardium in relation to cardiovascular diseases, e.g., myocardial infarction (Wu, et al., 2006), hypertrophic cardiomyopathy (Tseng, et al., 2006), and ischemic heart diseases (Mekkaoui, et al., 2017).

Despite the great utility of cardiac DTI, a common problem in practice is that the quantity and/or quality of acquired diffusion weighted (DW) images are limited due to the time constraint, resulting in inaccurate estimates of the diffusion tensors. First, the quantity of image acquisitions, especially in multiple diffusion gradient directions, is often restricted and inadequate. Although the diffusion tensor can be calculated mathematically from six DW signals associated with different diffusion gradient directions (Kingsley, 2006a), more directions are generally required to achieve more accurate diffusion tensor estimates (Jones, 2004). The number of used diffusion gradient directions determines the angular resolution of DTI data. However, when the acquisition time is limited, increasing the number of diffusion gradient directions (angular resolution) comes at the cost of lower signal-to-noise ratio (SNR) and/or limited number of acquisitions for each direction. Second, the quality of DW images is typically low, which is particularly evident when applying DTI to the *in vivo* heart. This is mainly due to the fact that DTI essentially measures the microscopic motion of water molecules and thus is very sensitive to subject motion including heartbeat and respiratory motions. Motions occur during imaging induce large signal loss in DW images and thereby greatly degrade the accuracy of estimated diffusion tensors. Besides, in order to minimize artifacts caused by motion and improve patient comfort, the DW images have to be acquired rapidly during *in vivo* cardiac DTI, thus reducing the SNR.

To overcome these problems and thus improve the applicability of cardiac DTI, a wide range of new DTI sequences have been proposed. Some of these sequences have improved the angular resolution of DTI and achieved higher accuracy in estimating the diffusion tensors, such as high-angular resolution diffusion imaging (HARDI) and its variants (Descoteaux, 1999; Descoteaux, et al., 2007; Descoteaux, et al., 2009; Descoteaux, 2008). These sequences are particularly effective in describing more complex configurations (such as fiber crossing

and bending) that can not be measured well by a simple DTI model. Some other sequences focused on reducing the effects caused by motion during *in vivo* cardiac DTI. For example, Edelman *et al*. (Edelman, et al., 1994) proposed a stimulated echo acquisition mode (STEAM) sequence where diffusion signals were acquired based on the spatial position and shape of the myocardium over the entire heartbeat cycle so that the effects of heartbeat motion were minimized. Gamper *et al*. (Gamper, et al., 2007) designed a spin echo (SE) based sequence to decrease the sensitivity of diffusion-encoding to cardiac motion while maintaining a high SNR. Besides, there also exist a variety of sequences that can be seen as extensions of STEAM or SE and have improved performance of *in vivo* DTI of the human heart (Reese, et al., 1995; Tseng, et al., 1999; Nielles-Vallespin, et al., 2013; Stoeck, et al., 2016; Nguyen, et al., 2014). Despite the power of these imaging sequences, both the improvement of angular resolution and the reduction of motion effects come at the expense of relatively long data acquisition times for these sequences. Such prolonged acquisition times are often prohibitive in clinical routine because it compromises patient comfort and increases the risk of inducing motion artifacts. This becomes the main difficulty in applying these sequences in practice.

In recent years, researchers began to develop post-processing techniques that can be used as an auxiliary tool to compensate for the drawbacks of current DTI sequences and improve their applicability. A large number of post-processing methods have been developed for the super angular resolution in cardiac DTI (Tuch, 2004; Descoteaux, et al., 2007; Baraniuk, 2007; Cheng, et al., 2015; Mani, et al., 2015). Although these methods are effective, they often require specialized imaging protocols and are therefore difficult to apply in practice. Several post-acquisition methods that can be used for compensation of motion-induced signal loss in *in vivo* cardiac DTI have also been developed (Pai, et al., 2011; Wei, et al., 2014; Deng, et al., 2020). A common idea of all these post-processing methods is to recover the desired information from the clinically acquired DTI data. In addition to the above methods, there are very few methods on dealing with the DTI data acquisition problem.

Recent advances in deep learning, in particular in generative adversarial network (GAN) have shown great capabilities of this network for medical image generation. We are interested in the applications of GAN for MR image generation since the DTI data is essentially the DW-MR image. Current applications of GAN-based MR image generation can be devided into two main categories. The first category involves improving the quality of the image. Issues in this category include denoising, super-resolution, etc. The second category is called image-to-image synthesis, which can generate desired MR images from the clinically available MR images, or estimating MR images from images of another modality (e.g., CT).

The goal of this thesis is to assist in improving the acquisition efficiency and clinical applicability of cardiac DTI data by investigating GAN-based image generation methods. To this end, a super angular resolution (SAR) method for cardiac DTI data was proposed for improving the estimation accuracy of diffusion tensors. Before we started this work, we systematically investigated the influence of noise and angular resolution on the estimation accuracy of diffusion tensors. We demonstrated that when the acquisition time is sufficient, the accuracy of diffusion tensors increases as the angular resolution. We constructed and trained the proposed SAR to produce higher angular resolution DW images from lower angular resolution DW images. Further, we modified the proposed SAR network and applied

it to compensate for motion-induced signal loss in *in vivo* cardiac DTI. The main challenges of this work is lacking of *in vivo* cardiac DTI data with motion-induced signal losses for enabling the network to achieve signal loss compensation in *in vivo* DTI. To fill with this gap, we studied the signal losses in *in vivo* cardiac DTI data, provides a simple but effective method to generate motion-induced signal losses and then added to *ex vivo* DW images to imitate *in vivo* DTI data and subsequently trained the proposed network.

The presentation of this thesis is organized as below.

The first part of this thesis is devoted to the presentation of medical foundations of cardiac DTI, and the methodology of our work:

In Chapter 1, we describes the principles and concepts of diffusion MRI and cardiac DTI in the human heart. In Chapter 2, we introduces basics of convolutional neural network (CNN) and the concept of GAN model and its variants. Then we discuss current applications of GAN in MR image generation.

The second part is dedicated to the simulation study and *in vivo* cardiac DTI contributions:

The work presented in Chapter 3 systematically studied the joint influence of angular resolution and noise on cardiac DTI. This demonstrated the validity of SAR for improving the accuracy of diffusion tensor estimation. The work provided in Chapter 4 proposed a GAN-based SAR framework, which can be run on both simulated and real cardiac DTI data to improve the estimation accuracy of diffusion tensor. In Chapter 5, we investigated the motion-induced signal loss in vivo cardiac DTI and provided a novel data fitting method capable of generating realistic motion-induced DW images directly from existing cardiac DTI datasets. The work in Chapter 6 proposed a novel GAN-based SAR that can produce high angular-resolution DW images from low angular-resolution motion-induced DW images to compensate for motion-induced signal loss in in vivo cardiac DTI.

In Chapter 7, we give the conclusions and the perspectives.

PART I

Medical and methodological background

Chapter 1

Diffusion Magnetic Resonance Imaging and Cardiac Diffusion Tensor Imaging

Contents

1.1 Introduction

Diffusion magnetic resonance imaging (dMRI) refers to a promising magnetic resonance imaging (MRI) modility that allows non-invasive views of the human body by measuring the diffusion of water molecules in tissues. Diffusion tensor imaging (DTI) — one of the most common dMRI techniques — is increasingly used for characterizing cardiac fiber structures and infering microstructural abnormalities related to cardiovascular diseases. The present thesis focuses on the DTI of the human heart. In order to better understand the medical background of this thesis, in this chapter, an introduction to dMRI techniques especially the DTI is provided, and then the applications and challenges of cardiac DTI are presented.

dMRI, which includes DTI, is fundamentally a technique that acquires images based on MRI. In order to understand how to obtain an image from magnetic resonance (MR) signals, the principle of MRI is first presented, including the basic concepts of MRI (e.g., precession, resonance and relaxation), k-space encoding technique, and two common MRI sequences (spin echo and gradient echo). Then, the concepts of diffusion-weighted (DW) MRI is introducing by explaining phenomenon of diffusion and diffusion effect on the acquired MR signals. Meanwhile, the diffusion-weighted sequence and the apparent diffusion coefficient (ADC) contribute to the acquisition of diffusion signals to form a DW image. Finally, the DTI model as well as the commonly used diffusion characteristics are presented in detail.

DTI is currently the unique potential method that can non-invasively characterize the myocardial structures in the human heart. In order to understand cardiac DTI, the structures and functions of the human heart are presented. The basic anatomy of the heart includes several parts with different physical properties and functions. One of the most important parts is the myocardium, as changes in its fiber arrangement are highly correlated with many cardiovascular diseases. Three of most common cardiovascular diseases are briefly described, including myocardial infarction, hypertrophic cardiomyopathy, and dilated congestive cardiomyopathy. Subsequently, a large number of studies of DTI techniques applied to the human heart are listed, which also demonstrated the great potential of DTI for the diagnosis of cardiac disease. Moreover, current challenges of applying cardiac DTI to the *in vivo* human heart are reported, highlight that the major challenge is caused by the motion of the heart. Current advanced and novel MR sequences that can reduce the effects of cardiac motion during cardiac DTI are listed. Unfortunately, most of them come at the cost of longer acquisition times, which remain difficult to apply in practice. Therefore, there is a great need for new solutions to deal with the motion problem in *in vivo* cardiac DTI.

The following sections are originized as follow. The dMRI and DTI are introduced in section 1.2, where the principle of MRI, concepts of diffusion-weighted MRI and DTI presented in section 1.2.1, 1.2.2 and 1.2.3, respectively. In section 1.3, the human heart structions and functions (section 1.3.1) and the DTI-based cardiac imaging (section 1.3.2) including the applications and challenges of cardiac DTI are presented, followed by the conclusion in section 1.4.

1.2 Diffusion Magnetic Resonance Imaging

Diffusion magnetic resonance imaging (dMRI) is a magnetic resonance imaging (MRI)-based technique that originated in the mid-18th century (Merboldt et al., 1985; Le Bihan, 1985; Taylor et al., 1985). It enables non-invisively probe the human body's tissues by measuring the diffusion of water molecules in tissues. In dMRI, the description of water diffusion are usually obtained by integrating the acquired diffusion weighted (DW) signals into some mathematical models. Diffusion tensor imaging (DTI) is one of the most commonly used dMRI techniques that adopts the simplest mathematical model to describe the properties of water molecules within each voxel. To deal with the limitations of DTI model, other advanced dMRI techniques have emerged, where the DW signals are usually acquired in a relatively high angular resolution. This section firstly explains the principles of MRI, then the theory of DTI and other advanced dMRI technques are provided after introducing the physical phenomenon of water diffusion.

1.2.1 Magnetic Resonance Imaging principles

MRI refers to a medical imaging technique that is capable of providing highly detailed pictures of the human body's structures (such as nerves, muscles, and blood vessels) using magnet fields and radio-frequency (RF) pulses. Compared with an x-ray or CT scan, MRI allows imaging without exposure to radiation. Besides, MRI takes advantage of the body's most abundant molecule – water, and thus is far more flexible. Some detailed reviews concerning the MRI development history, physical principles and applications have been presented in previous works (Goldman, 2004; Westbrook & Talbot, 2018). Here we briefly present the basic concepts and principles of MRI in order to understand how dMRI works.

1.2.1.1 Basic concepts of MRI

It is well known that the human tissue contains a lot of water. Especially in the brain and heart, more than 70% of the composition is water (Mitchell, et al., 1945). For the purpose of MRI, we are interested in hydrogen nuclei (or protons ^{1}H) of water molecules. The little positive electrical charged proton constantly rotates on its axis in the nucleus, behaving similar to planets and stars, which is called spin, as shown as in Fig. 1.1(a). The spinning of each proton generates a tiny bar magnet. Under normal condition, the protons in the body spin with their axes randomly and their tiny magnets cancel each other out so that the magnetic forces equalize, as shown in Fig. 1.1(b). However, when these protons are placed in a stronger magnetic field B_0, all of these spins or proton magnets line up with or against (parallel or anti-parallel to) the main magnetic field, precessing like a spinning top on the solid surface, as illustrated by Fig. 1.1(c). The rate of precession is determined by the Larmor frequency equation (Levitt, 2013) expressed as:

$$\omega_0 = \gamma B_0 \tag{1.1}$$

where B_0 and γ denote the strength of magnetic field and the gyromagnetic ratio of proton ^{1}H, respectively.

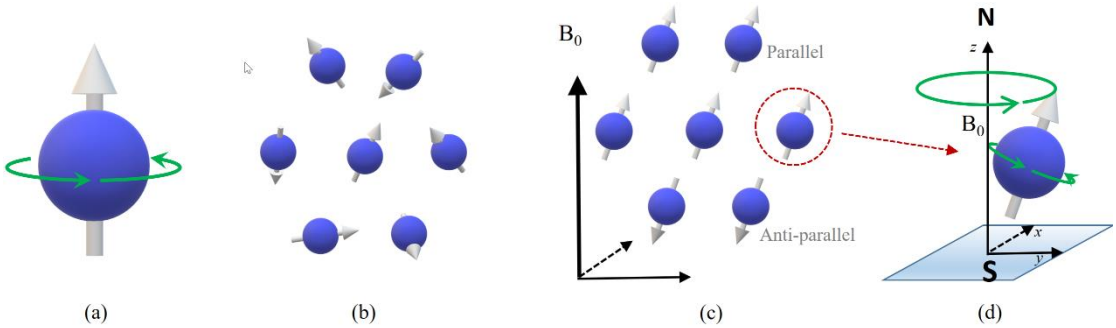

Fig. 1.1 Spin and precession of protons. (a) A proton spinning around the nucleus of an atom. (b) Protons in normal condition without external magnetic field. (c) Protons placed in a magnetic field B_0. (d) Precession of a proton within the magnetic field B_0.

By adding up all magnetic fields of protons, a net magnetization M_0 aligned in the same direction as the large external magnetic field B_0 can be ontained, and such M_0 is called longitudinal magnetization. In MRI, we mainly focus on studying what happenes with M_0 since it reflects the density of water protons. However, M_0 is difficult to be detected directly when it is in the same direction with B_0. Therefore, a critical step in MRI is to manipulate the direction of the net magnetization by the application of RF pulses. Specifically, if we transmit a Radio Frequency (RF) pulse with exactly Larmor frequency of hydrogen (42,58MHz for 1 Tesla) in the vicinity of the protons, the protons will absorb that energy to move out of alignment away from B_0, thus changing the direction of net magnetization. This phenomenon is known as resonance. With a 90° RF pulse, the direction of net magnetization can be moved from the z plane into the transverse (xy) plane, that is, the longitudinal magnetization M_0 are transferred into the transverse magnetization M_{xy}, as shown in Fig. 1.2. When the net magnetization vector lies in the transverse plane, it is precessing (rotating) around the *z* axis.

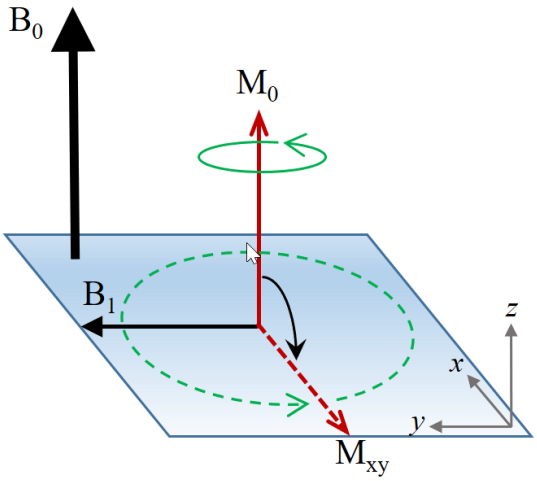

Fig. 1.2 Net magnetization before (solid red arrow) and after (dashed red arrow) application of a 90° RF pulse named by B_1. The solid and dashed green arrow designate the pression directions for the longitudinal magnetization M_0 and transverse magnetization M_{xy}, respectively.

Once the RF pulse is removed, the net magnetization is again influenced by the magnetic field B_0 and returns to the original direction of M_0 parallel to B_0. This phenomenon is called

relaxation. During the relaxation, two effects happen over time. One is that the amount of magnetization in the longitudinal plane gradually increases, as the enegy that was previously obsorbed by the protons will be released into the surrounding external environment. This is called longitudinal relaxation or T1 relaxation. The other occurs at the same time and involves that the amount of magnetization in the transverse plane gradually decreases due to the protons exchanging energy with each other. This is called transverse relaxation or T2 relaxation. The the amount of magnetization for both T1 and T2 relaxations can be modeled as exponential functions of relaxation time *t*, as shown in Fig. 1.3. The magnetizations as a function of time during T1 and T2 relaxations are given by Bloch equation (Bloch, 1946):

$$\frac{d\vec{M}}{dt} = \gamma(\vec{M} \times \vec{B}_0) - \begin{pmatrix} \frac{M_x}{T_2} \\ \frac{M_y}{T_2} \\ \frac{M_z - M_0}{T_1} \end{pmatrix} \tag{1.2}$$

where M_x, M_y, M_z are the component of $\vec{M}$ along the directions x, y, z, T_1 and T_2 denotes the time for T1 and T2 relaxations. If we place a receiver coil in the transverse plane at this moment, the variations of magnetization will create magnetic field fluctuations within the coil and therefore induce a voltage. Such voltage causes a small electrical current that constitutes the magnetic resonance (MR) signal, resulting in a two-dimensional (2D) MR image.

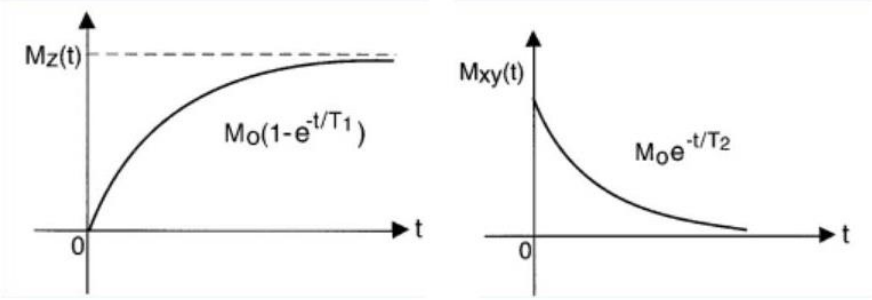

Fig. 1.3 Schema for T1 relaxation (left) and T2 relaxation (right).
https://warwick.ac.uk/fac/sci/physics/research/condensedmatt/imr_cdt/students/stephen_day/relaxation

1.2.1.2 K-space encoding

In dMRI, the MR signal needs to be located in three-dimensional (3D) space for imaging 3D tissues. This involves the k-space encoding technique which was introduced by (Likes, 1981) and in1983 by (Twieg, 1983; Ljunggren, 1983). The process of k-space encoding consists of two main steps: slice selection and phase-frequency encoding.

Slice selection

Slice selection is the first step of k-space encoding. As stated earlier, the MR signal is essentially an electrical current produced in the receiver coil by the precession of the magnetization during resonance. The resonance occurs only when the RF pulse is applied at

the Larmor frequency of the spins (or protons). If the magnetic field is homogenous through the patient, all the spins are in resonance and constitute the MR signal. When a gradient appears (i.e., the magnetic field varies across space), the Larmor frequency of the spins changes along the direction of gradient. For example, if there is a gradient being applied along the *z* aixs, the total magnetic field $B(z)$ at a location (*z*) will become:

$$B(z) = B_0 + z \cdot G_z \tag{1.3}$$

where G_z is the strength of the gradient. Then the precession frequency $\omega(z)$ of spins at the location (z) is:

$$\omega(z) = \gamma(B_0 + z \cdot G_z) \tag{1.4}$$

If we simultaneously apply a RF pulse along the z axis, only the spins located at a slice plane perpendicular to z aixs will be in resonance and thus produce signals. The location of slice along the perpendicular direction can be selected by changing the frequency of the RF pulse. Conversely, the spins located outside the selected slice are not in resonance and thus will not produce a signal. Fig. 1.4 gives the illustratation of slice selection when the gradient G_z being applied along the field direction z, where Z1, Z2, and Z3 represent the spins rotate at three different frequencies along z, respectively. In fact, gradients can also be applied in x and y directions to locate the spins for resonance in the transverse plane.

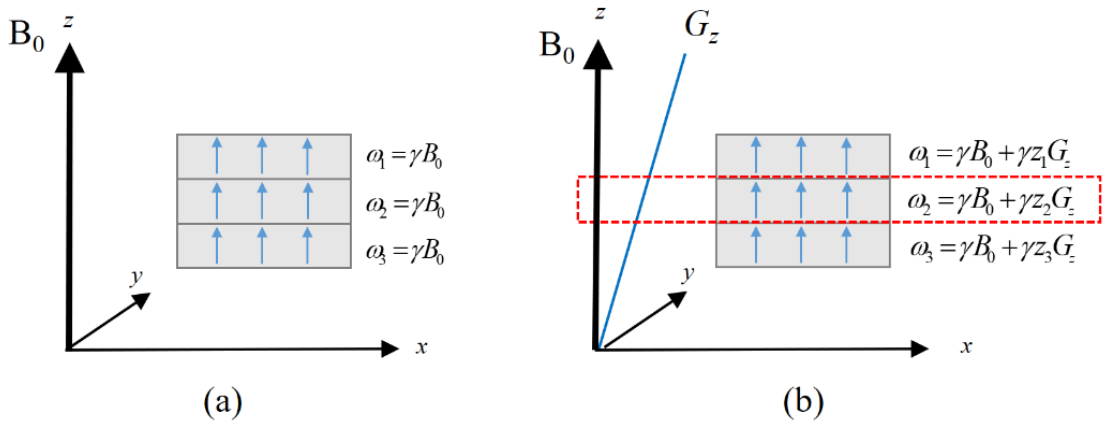

Fig. 1.4 Slice selection with application of a gradient. (a) Precession frequency of spins in a homogenous magnetic field. (b) Precession frequency of spins when applying a gradient G_z along the z direction.

Each selected slice has a thickness of tn determined by the frequency bandwidth $\Delta\omega$ of the RF pulse and the G_z:

$$tn = \frac{\Delta\omega}{\gamma G_z} \tag{1.5}$$

Phase and frequency encoding

After selecting the slice along the direction of gradient, the second step of k-space encoding is to locate the MR signal along both the x axis and y axis of this slice plane. This refers to the phase and frequency encoding technique. Phase encoding is achaieved by applying a phase encoding gradient in the direction along one axis of the selected slice plane. When such

gradient is applied, the magnetic field is changed so that the precession frequency of spins along the axis of the phase gradient is changed. The result is that all the spins along that axis are precessing in the same frequency but in different phases. Similary, by applying a frequency encoding gradient, magnetic field strength is changed and so does the precession frequency of spins along the remaining axis of the slice. Therefore, the signal can be produced according the different frequencies along such axis. Fig. 1.5 illustrates the process of k-space encoding from "slice selection" to MR signal acquisition. The MR signal $S(t)$ that multiplied by G_y and G_x with the time and t_x and t_y, is:

$$S(t_x,t_y)=\int \rho(x,y)exp[i\gamma(\int_0^{t_x} G_x(t)xdt+\int_0^{t_y} G_y(t)ydt)]dxdy \tag{1.6}$$

where $\rho(x,y)$ is the spatial distribution of spins within a sample. The MR signal is often expressed in terms of spatial frequencies k (Twieg, 1983; Ljunggren, 1983), where the two frequency coordinates axes k_x and k_y (the right side of Fig. 1.5) corresponding to the read out and phase encoding directions in k-space, respectively, are defined by:

$$\begin{aligned} k_x &= \int_0^{t_x} G_x(t)xdt \\ k_y &= \int_0^{t_y} G_y(t)ydt \end{aligned} \tag{1.7}$$

The final MR image is reconstructed using the 2D inverse Fourier transform.

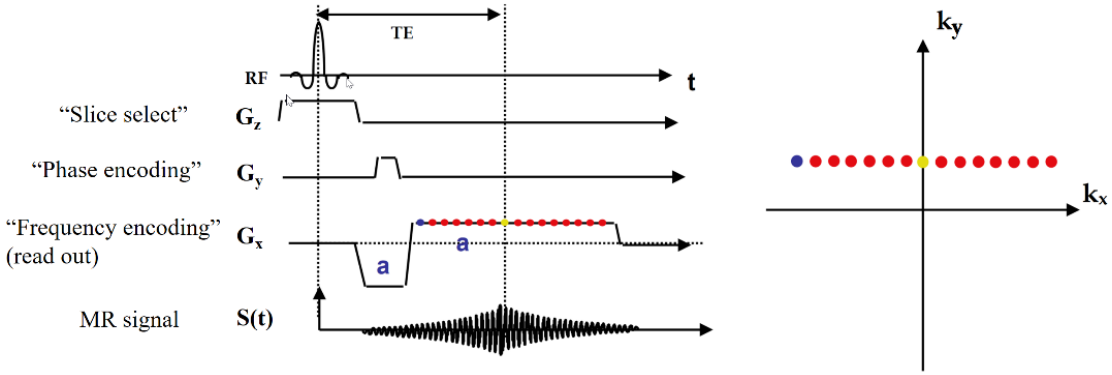

Fig. 1.5 Diagram of k-space encoding for the acquisition of MR signal. (Rosen & Wald, 2006)

1.2.1.3 MRI sequences

In MRI, the strategy of applying pulses and gradients to form MR images depends on the MRI sequence used. There exists a wide range of MRI sequences available. Here we briefly discuss two commonly used sequences: spin echo and gradient echo sequence.

Spin-echo sequence

The spin-echo (SE) sequence is frequently used in MRI studies and experiments due to its good imaging effects. In this sequence, the SE is produced with a 90º RF pulse followed by 180º rephasing RF pulses. Fig. 1.6 gives an illustration of SE sequence. A 90º pulse is first used to flip the magnetization into the transverse plane. Due to the T2 relaxation, some of the spins slow down. A 180º pulse is then applied so that previously slower spins are effectively

precessing ahead of the faster ones. After repetition time TR, the echo is reproduced by using this method again. In general, a gradient is applied along the reading out direction before 180º RF pulse. This is in order to compensate the dephasing effects induced by the inhomogeneity of the gradient.

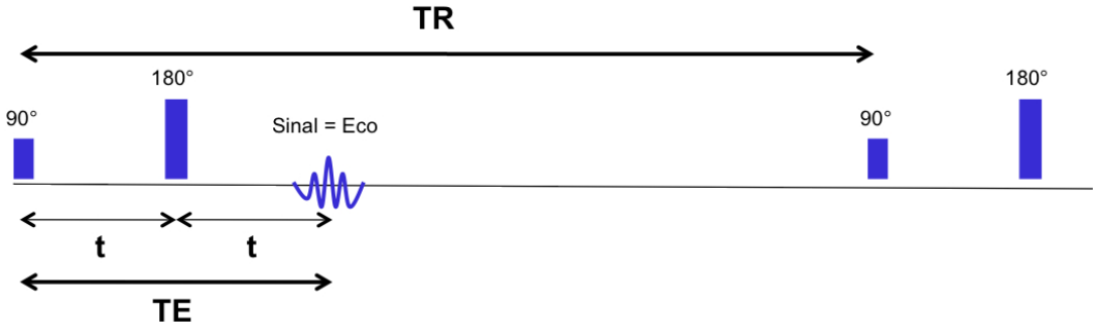

Fig. 1.6 Illustration of Spin-echo sequence. [Jung et al., 2013]

In most MRI sequences, followed by the slice selection gradient there is often a gradient with half amplitude strength of G_{sl}, which is used to compensate the dephasing effects caused by the inhomogeneity of the gradient. With the same reason, before 180º RF pulse, a gradient is applied along the reading out direction.

After spatial encoding with the gradients, the intensity of each voxel in the acquired MR image now relates to three properties: the density of protons, T1 and T2 relaxations. The contrast of image highly influenced by the used parameters of the SE sequence. One drawback of the SE sequence is that its acquisition time is relatively long.

Gradient echo sequence

The gradient echo sequence (GE) is an alternative technique to SE sequences. The gradient echo is generated using reversed polarity gradients, which means that it is used twice in in opposite directions. GE starts with the applications of an RF pulse with flip angle of αº to enforce transverse dephasing of spinning protons (due to the T2 relaxation). The dephasing process continues until the polarity of G_{ro} is changed), as shown in Fig. 1.7.

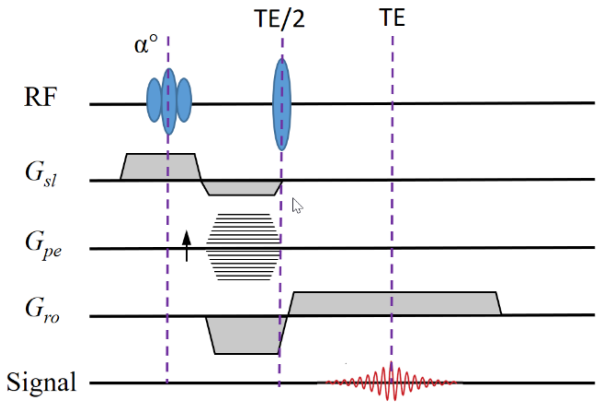

Fig. 1.7 Gradient echo sequence and its principle. G_{sl}, G_{pe}, and G_{ro} denote the gradients used for slice selection, phase encoding, and frequency encoding (read out), respectively.
https://www.radiologycafe.com/frcr-physics-notes/mr-imaging/gradient-recalled-echo-sequence

After that, the spins begin to rephrase (like in spin-echo MRI) and hence acquire signal. Compared with SE, the T2* decay caused by the lack of a 180º refocusing pulse makes GE sequence more repid, but this also makes it more susceptible to static field inhomogeneities.

1.2.2 Diffusion-weighted MRI concepts

Diffusion weighted magnetic resonance imaging (DWI, or DW-MRI) is a specific MR imaging that uses the diffusion of water molecules in biological tissues (Mori & Van Zijl, 2002) to generate images. This makes it possible to probe the microscopic structures of biological tissues. This section present a brief introduction to the DW-MRI techniques.

1.2.2.1 Diffusion phenomena and effect

Diffusion a mass transport of substances in nature from regions of higher concentration to regions of lower concentration due to concentration gradients. “Diffusion” in biological tissues usually refers to the random motion of water molecules results from the thermal energy carried by these molecules, which is also called Brownian motion (Brown, 1828). The Brownian motion is usually described by the random walk model (Knight, 1962). This model provides the trajectory and the diffusion displacement of water molecules during a given time interval, as illustrated in Fig. 1.8. The dotted white line (vector $\vec{r}$) in Fig. 1.8 represents the displacement of a single water molecule during the diffusion time between t1 and t2. Such displacement is determined by the time interval $\Delta = t2 - t1$, the diffusion coefficient *D* which depends on the viscosity of the medium, and the types of diffusion types (e.g., free, hindered, or restricted).

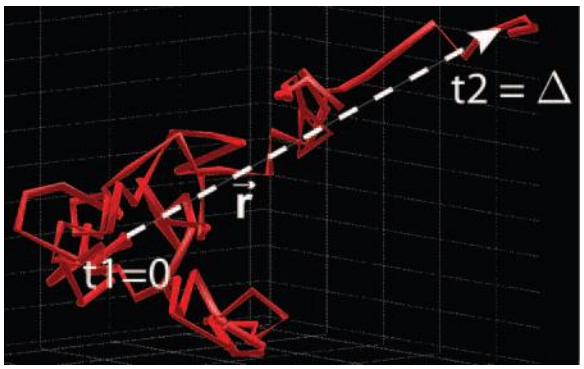

Fig. 1.8 Diagram of the molecular diffusion trajectory of a single water molecule. (Hagmann, et al., 2006)

As water diffusion is a random process, for the same time interval Δ, a molecule diffuses from the same position and may end at the different positions. The author in (Callaghan, 1993) introduced used displacement *probability density function* (PDF) for better describing the phenomenon of water diffusion. After the time interval Δ, the probability of a water molecule located at position $\vec{p}_0$ moving to the position $\vec{p}_\Delta$ can be described as a PDF which can be written as $P(\vec{p}_\Delta \left| \vec{p}_0, \Delta \right.)$. The isotropic diffusion (in a homogeneous medium) for a 3D system is expressed as:

$$P(\vec{p}_{\Delta}\left|\vec{p}_{0},\Delta\right)=\frac{1}{\sqrt{\left|D\right|(4\pi\Delta)^{3}}}\exp(-\frac{(\vec{p}^{\mathrm{T}}D^{-1}\vec{p})2}{4\Delta}) \tag{1.8}$$

with $\vec{p}=\vec{p}_{\Delta}-\vec{p}_{0}$. The variance of $\vec{p}$ is proportional to the diffusion time Δ:

$$<\left|\vec{p}\right|^{2}>=6D\Delta \tag{1.9}$$

The Eq. (1.9) is the famous Einstein equation (Einstein, 1956).

In general, the diffusion of water molecules in the biological tissues is not isotropic but often restricted or hindered by many mechanisms, such as cell membrane, fiber boundaries of tissues, etc. Hindered diffusion occurs in heterogeneous media, where the diffusion is usually anisotropic, while restricted diffusion refers to intracellular diffusion, in which water is not allowed to diffuse into the extracellular space.

The effects of diffusion of water molecules on the MRI signal was first noticed by Hahn in 1950 (Hahn, 1950). Later, Torrey (Torrey, 1956) introduced a generalized Bloch equation which mathematically defined the variation of nuclear magnetization when diffusion of water molecules is present. As described in section 1.2.1.1, the original Bloch equation Eq. (1.2) describes the variation of nmagnetization with time in a magnetic field. In the presence of molecular diffusion, the Eq. (1.2) becomes:

$$\frac{d\vec{M}}{dt}=\gamma\vec{M}\times\vec{B}_{0}+\begin{pmatrix}\frac{M_{x}}{T_{2}}\\ \frac{M_{y}}{T_{2}}\\ \frac{M_{0}-M_{z}}{T_{1}}\end{pmatrix}+D\nabla^{2}\vec{M} \tag{1.10}$$

where ∇^2 denotes the Laplace operator, the first two terms are exactly those of Eq. (1.2), the third term is represents the additional variation caused by molecular diffusion. The Eq. (1.10) is the famous Bloch-Torrey equation.

Now we know what diffusion is and how it affects the MRI signal. In order to acquire the signals induced by molecular diffusion, the original MRI pulse sequence was altered and produced another imaging modality. In order to extract the information of water diffusion, Stejskal and Tanner developed a specific diffusion measurement sequence based on SE (Stejskal, 1965). In this sequence, one symmetric pair of diffusion-weighted gradients is inserted to either side of the refocusing 180º RF pulse in each of the three directions of k-space encoding (slice selection, phase encoding, and frequency encoding), as shown in Fig. 1.9.

1.2.2.2 *Diffusion-weighted sequence*

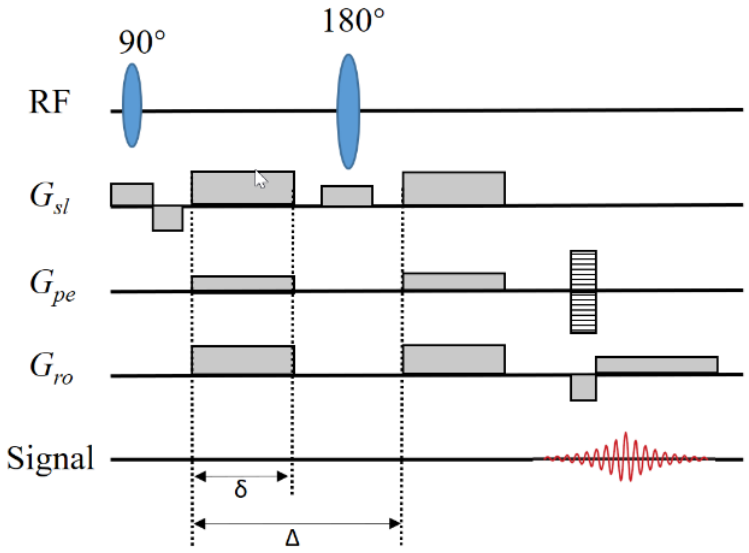

Fig. 1.9 Stejskal-Tanner diffusion MRI sequence

In the three encoding directions, diffusion-weighted gradients with different magnitudes provide a specific diffusion-weighted direction corresponding to the coordinate system of encoding. The first of these gradients offsets the phase of the spins, and the amount of the shift depends on their position. The second gradient provides equal and opposite (due to 180º pulse) rephasing if the spins have not moved during the diffusion time Δ (the time between the application of the two gradients). The rephasing effects of 180º will disappear since the spins move randomly due to diffusion. During the diffusion time, the further the spins have diffused, the less perfect the rephasing will be, resulting in a smaller amplitude of the final signal. The diffusion signal detected by this sequence is:

$$S(b) = M_0 \exp(-\frac{t}{T_2}) e^{(-bD)} \tag{1.11}$$

where b is the diffusion-weighting factor called b-value expressed as:

$$b = \gamma^2 \delta^2 (\Delta - \delta / 3) G_d^2 \tag{1.12}$$

where G_d is the strength of diffusion gradient, δ is the pulse duration, γ is the gyromagnetic ratio, and Δ is the time between the application of the two gradients pulses. In the absence of diffusion gradients, the b-value is 0 and then the Eq. (1.11) becomes:

$$S(b) = S_0 e^{(-bD)} \tag{1.13}$$

where $S_0 = M_0 \exp(-t/T_2)$.

1.2.2.3 *Apparent diffusion coefficient (ADC)*

As mentioned in section 1.2.2.1, the Brownian motion of water molecules in biological tissues are usually restricted by cell membrane or fiber boundaries, which means that the diffusion will be restricted to a limited volume. If the diffusion time is sufficient, the diffusion

distance may be smaller than the expected range of free diffusion, resulting the reduced diffusion coefficient D. In 1986, LeBihan refers to this diffusion coefficient as the apparent diffusion coefficient (ADC), which is a function of the geometry of the confining volume and the diffusion time (Le Bihan, et al., 1986).

The ADC is a global parameter reflecting the diffusion properties of a tissue. It depends on the direction, diffusion time and strength of diffusion gradient. When water molecules diffuse without any preferential direction within an image voxel, the ADC will be the same in every direction (or isotropic) and can be expressed as Eq.(1.13), where the *D* represents ADC. In the presence of restricted diffusion, that is, the amplitude of diffusion is different in different directions (or anisotropic), then the ADC is changed depending on the gradient direction and can be expressed as

$$S_i = S_0 e^{(-b \times ADC_i)} \tag{1.14}$$

where S_i is the diffusion-weighted signal at a given voxel with the diffusion-gradients applied along direction *i*, S_0 is the signal at the same voxel without the additional diffusion gradients, ADC_i is the ADC of water molecules along the i^{th} direction of the diffusion gradients, indicating that the diffusion process is not free in tissues.

Note that the ADC is essentially a single (scalar) parameter. When the diffusion is isotropic, the use of a single ADC is enough since the ADC is all the same value in each direction of diffusion. However, in the presence of anisotropy, it requires other parameters to describe the molecular diffusion along each spatial direction. Such parameters are involved in the diffusion tensor imaging technique, as detailed in the subsequent section.

1.2.3 Diffusion Tensor Imaging (DTI)

In order to fully describe the molecular diffusion in anisotropic systems, in 1994, Basser et al. (Basser, et al., 1994) introduced the concept of diffusion tensors into dMRI, which leads to diffusion tensor imaging (DTI). DTI has now become one of the most popular MRI techniques capable of describing diffusion anisotropy.

1.2.3.1 DTI model

In DTI, the diffusion properties of water molecules within each voxel are described by a diffusion tensor estimated from a series of diffusion-weighted (DW) images associated with noncollinear difusion gradient directions. The diffusion magnetic resonance signal as a function of the diffusion tensor is expressed by (Le Bihan, et al., 2001):

$$S_i = S_0 \exp(-b\vec{g}_i^T \underline{D}\vec{g}_i) \tag{1.15}$$

where S_i is the diffusion weighted signal acquired at one of the diffusion direction, S_0 is the signal at the same voxel without the additional diffusion gradients, $\vec{g}_i = (x_i, y_i, z_i)$ is the unit vector that designates the diffusion gradient direction, $\vec{g}_i^T \underline{D}\vec{g}_i$ denotes measured ADC_i along the i^{th} diffusion gradient direction, and $\underline{D}$ denotes the diffusion tensor which is mathematically expressed as:

$$\underline{D} = \begin{pmatrix} D_{xx} & D_{xy} & D_{xz} \\ D_{yx} & D_{yy} & D_{yz} \\ D_{zx} & D_{zy} & D_{zz} \end{pmatrix} \tag{1.16}$$

Note that $\underline{D}$ is a symmetric matrix with six unique elements, D_{xx}, D_{yy}, D_{zz}, D_{xy}, D_{xz} and D_{yz}. If S_0, $\vec{g}_i$ and b are given, and the signals S_i for all diffusion gradient directions are acquired, then the diffusion tensors can be calculated by solving the system of linear equation as:

$$\mathbf{T} = \mathbf{B}^{-1}\mathbf{E} \tag{1.17}$$

where $\mathbf{B} = \begin{bmatrix} x_1^2 & y_1^2 & z_1^2 & 2x_1y_1 & 2y_1z_1 & 2z_1x_1 \\ x_2^2 & y_2^2 & z_2^2 & 2x_2y_2 & 2y_2z_2 & 2z_2x_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_i^2 & y_i^2 & z_i^2 & 2x_iy_i & 2y_iz_i & 2z_ix_i \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_n^2 & y_n^2 & z_n^2 & 2x_ny_n & 2y_nz_n & 2z_nx_n \end{bmatrix}$, $\mathbf{T} = \begin{bmatrix} D_{xx} \\ D_{yy} \\ D_{zz} \\ D_{xy} \\ D_{yz} \\ D_{zx} \end{bmatrix}$, and $\mathbf{E} = -\frac{1}{b}\begin{bmatrix} \ln(S_1/S_0) \\ \ln(S_2/S_0) \\ \vdots \\ \ln(S_i/S_0) \\ \vdots \\ \ln(S_n/S_0) \end{bmatrix}$,

here the number of diffusion gradients $n \geq 6$. When $n = 6$, the diffusion tensor can be directly calculated by Eq. (1.17). If $n > 6$, there is not true inverse $\mathbf{B}^{-1}$, but $\mathbf{B}^{-1}$ can be replaced by a pseudo inverse matrix $\hat{\mathbf{B}}$ (Lenglet, 2006), where $\hat{\mathbf{B}} = (\mathbf{B}^T\mathbf{B})^{-1}\mathbf{B}^T$. Then the diffusion tensor $\underline{D}$ can be estimated using a reconstruction algorithm, such as least squares estimation (Barrett, et al., 1994).

1.2.3.2 Diffusion characteristics extraction

After calculating the diffusion tensor, the structures of the diffusion environment are characterized by extracting some diffusion characteristics from the tensor matrix. Here we introduce three commonly used characteristics: the main direction of diffusion, the local mean diffusivity and the diffusion anisotropy.

Main direction of the diffusion

In DTI, the diffusion tensor $\underline{D}$ indicates the local structures of the tissue. Specifically, by diagonaliizing the diffusion tensor matrix in Eq. (1.16), the diffusion tensor $\underline{D}$ can be present by

$$\underline{D} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}\begin{pmatrix} V_1 \\ V_2 \\ V_3 \end{pmatrix} \tag{1.18}$$

where λ_1, λ_2, and λ_3 are three eigenvalues of $\underline{D}$, corresponding to the there eigenvectors V_1, V_2, and V_3, respectively. The diffusion in a local region (e.g., within a voxel) can be represented a ellipsoid, where the three eigenvalues and eigenvectors defined the lengths and

directions of three radii of the ellipsoid (Basser, 1995). When the diffusion is isotropic, the three eigenvalues λ_1, λ_2, and λ_3 are equal. If the water diffuses in biological tissue such as muscle fibers, the three eigenvalues is possibly different, for example, $\lambda_3 \le \lambda_2 \le \lambda_1$. Then V_1 indicates the main direction of water diffusion that is least restricted in the muscle fiber tracts and thus reflects the principal orientation of the fibers inside that voxel. Consequently, the fiber architecture can be depicted.

In practice, the diffusion orientation extracted from DTI is commonly used for doing fiber tracking in the brain (Mori & Van Zijl, 2002; Hüppi & Dubois, 2006; Assaf, et al., 2008), sometimes in the heart muscle (Zhukov & Barr, 2003; Zhang, et al., 2010) and some other tissues. Besides, it is also the unique way to study non-invasively the "*in vivo*" cerebral connectivity.

Mean Diffusivity (MD)

The MD is a frequently used parameter in DTI. It represents the overall mean-squared displacement of the water molecules along the three diffusion directions. It can also help us determine if there is restrictive structures within the local region. MD can be defined by the ADC along the three eigenvector directions of the diffusion tensor. For example, if the tensor axis are aligned along x, y and z axis, then the MD is:

$$MD = (ADC_x + ADC_y + ADC_z)/3 \tag{1.19}$$

Alternatively, the MD can be mathematically computed from the three eigenvalues of the tensor:

$$MD = (\lambda_1 + \lambda_2 + \lambda_3)/3 \tag{1.20}$$

Fractional Anisotropy (FA)

The FA reflects the spatial heterogeneity of displacements of water molecules in relation to the presence of orientated structures (Pierpaoli & Basser, 1996), it is defined as

$$FA = \sqrt{\frac{3}{2}}\sqrt{\frac{(\lambda_1 - MD)^2 + (\lambda_2 - MD)^2 + (\lambda_3 - MD)^2}{(\lambda_1^2 + \lambda_2^2 + \lambda_3^2)}} \tag{1.21}$$

Principal orientations

In many studies on the human brain or heart, the principal direction of the fibers, which is represented by the major eigenvector of the diffusion tensor, V_1, is also a common metric for evaluating the performance of diffusion tensor estimation. Two commonly used orientation metrics are the azimuth angle and elevation angle. The azimuth angle indicates the angle between the projection of the fiber (or V_1) on the xy-plane and the east-west axis of the plane (shown as α in Fig. 1.10). The elevation angle refers to the angle between the fiber and the selection xy-plane (shown as ε in Fig. 1.10). The azimuth angle should have a value between 0 and π, while the elevation angle has a value between -π/2 and π/2. The two metrics completely determines the orientation of the muscle fibers within the brain or heart.

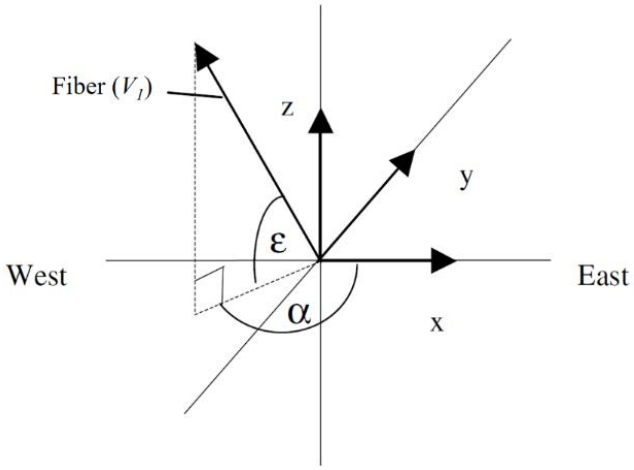

Fig. 1.10 Azimuth angle α and elevation angle ε of a fiber. [Yang et al. 2011]

1.3 Cardiac Diffusion Tensor Imaging

The human heart is the main organ of the cardiovascular system. It pumps pump blood throughout the body using the blood vessels of the circulatory system to provide oxygen and nutrients to the tissues and remove carbon dioxide and other wastes. DTI is currently the unique potential method that can non-invasively study the myocardial structures and diagnoze the abnormality of cardiac muscles (Frindel, et al., 2010; Yang, et al., 2012). In the following, we will first introduce structures and functions of the huaman heart, and then present the applications and challenges of DTI in the human heart.

1.3.1 Human heart structures and functions

1.3.1.1 Basic anatomy of the heart

The heart is located in the center of the chest - slightly to the left of the posterior aspect of the sternum. It lies between the right and left lungs, being protected by the pericardium and the ribcage. The size of the heart varies from person to person. Generally, the adult heart has the size of two clenched fists, and the child's heart has the size of one clenched fist.

Fig. 1.11 illustrates the anatomical structure of the heart. In addition to the blood vessels and cardiac conduction system, the heart can be divided into three parts based on their physical properties and functions: chambers, walls, and valves, which are detailed in the following paragraphs.

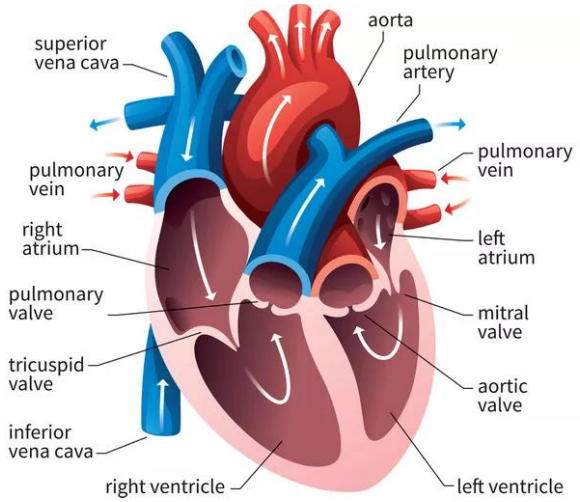

Fig. 1.11 Heart anatomical structure at macroscopic scale.
https://www.thoughtco.com/evolution-of-the-human-heart

Heart chambers

As shown in Fig. 1.11, the whole human heart has four chambers, including two chambers on the top called left and right atrium and two on the bottom named by left ventricle (LV) and right ventricle (RV). The atrium and ventricle in the right side pump the deoxygenated blood and those in the left side pump the oxygenated blood.

- **Right atrium**: It is located in the upper right side of the heart. The right atrium receives the de-oxygenated blood though the superior and the inferior vena cava, and then pumps the blood into the right ventricle (located below) via the tricuspid valve. The received de-oxygenated blood are came from other parts of the body such as the head, neck, arm, chest, legs, back, abdomen, etc.
- **Right ventricle**: This chamber lies below the right atrium. It receives de-oxygenated blood from the right atrium and pumps it to lungs via the pulmonary artery so that the blood can be reloaded with oxygen.
- **Left atrium**: It is located beside the right atrium. After reloading the blood with oxygen in the lungs, the de-oxygenated blood becames oxygenated and then returns into the left atrium through the pulmonary veins. The left atrium subsequently pumps the oxygenated blood to the left ventricle (located below) via the bicuspid valve or mitral valve.
- **Left ventricle**: It is the lower left part of the heart. The left ventricle is much longer than the right ventricle. It receives oxygenated blood from the left atrium via the mitral valve, and then sends it back to other parts of the body.

Compared with the right atrium and ventricle, the left atrium and ventricle are thicker. This is bacause the left part of the heart requires to pump blood throughout the body, while the right part pump blood only through the lungs.

Heart valves

The valves in the heart can be thought of as the doors between the chambers. The presence of the valves ensures that the blood circulating in the heart flows in the right direction so that

the chambers can perform their functions effectively. There are four main valves in the heart, which can be devided into two categories: atrioventricular (AV) valves and semilunar (SL) valves.

The AV valves open between the top and and bottom heart chambers. They guarantee the flow of blood from the atria to the ventricles. The AV valves in the heart include:

- **Tricuspid valve**: Located between your right atrium and right ventricle. It opens when the right atrium is in systole phase to allow the de-oxygenated blood to flow into the right ventricle.
- **Mitral valve**: Located between the left atrium and left ventricle. It opens during the atrium's systole phase, forcing the oxygenated blood to flow from the lungs to the left ventricle.

The SL valves open when blood flows out of your ventricles. The SL valves in the heart include:

- **Aortic valve**: Located at the exit of the aorta and the left ventricle. It opens when blood flows out of the left ventricle in order to deliver oxygenated blood to the body.
- **Pulmonary valve**: Located between right ventricle and the pulmonary artery. It opens as blood flows from the right ventricle to the pulmonary artery in order to carry oxygen-poor blood to the lungs.

Heart walls

The heart walls consists of three tissues layers: the endocardium, myocardium and epicardium, shown as in Fig. 1.12. The presence of these three layers ensures that the heart is able to pump blood properly. In the following, we describe in detail the structure and functions of these three layers.

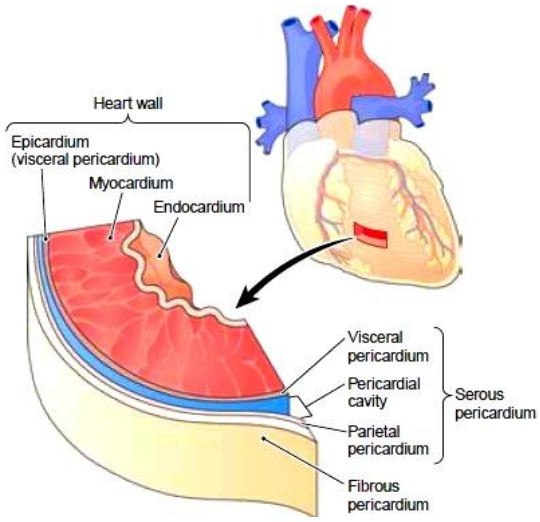

Fig. 1.12 Layers of the heart wall.
http://encyclopedia.lubopitko-bg.com/Structure_of_the_Heart.html

- **Epicardium:** The epicardium refers to the outer layer of heart tissue. It is also one layer of the pericardium (visceral pericardium, see Fig. 1.12). It is composed of

epithelial cells, which cover the fragile connective tissue for protecting the heart. Besides, the epicardium also protects the heart by producing fluid to keep the heart from rubbing against other organs.

- **Myocardium:** The myocardium is the middle layer of the heart wall between the epicardium and the endocardium, which is the thickest of all the layers. It is the basic muscle structure of the heart. The cardiac muscle consists of cardiac muscle cells (or myocyte) arranged in overlapping spiral patterns. The cardiac muscle provides power for the contraction and relaxation of the chambers, so that blood flows in and out of the heart, supplying oxygen to all parts of the body. The structure and function of the cardiac muscle in the myocardium is quite important for human health. For example, as we know the the heart needs to beat every minute of every day, so the heart muscle must constantly provide the power needed for the heartbeat. Any abnormality in the function or structure of the heart muscle can cause cardiac arrest and death.
- **Endocardium** The endocardium is the innermost layer of the heart wall. It is a thin, smooth layer of epithelial tissue, which is located on the inner surface of all cardiac chambers and valves. It helps to prolong the contraction time of the cardiac muscle.

1.3.1.2 Cardiovascular diseases

Cardiovascular diseases (CVDs) are the leading cause of death worldwide. According to World Health Statistics report, each year about 17.9 million people die from CVDs, accounting for approximately 32% of all deaths worldwide. In general, many CADs are caused by structural or functional damage to the heart muscle. In the following we briefly include three types of cardiovascular diseases: myocardial infarction, hypertrophic cardiomyopathy, and dilated congestive cardiomyopathy.

Detailed cardiomyopathy

Detailed cardiomyopathy is the most common type of heart muscle disease. It can cause the heart chambers to thin and stretch, becoming larger (Jefferies & Towbin, 2010), as shown in Fig. 1.13. In the early stages, it occurs in the left ventricle. As the disease worsens, it may spread to the right ventricle and atria. The disease causes abnormal blood supply to the heart and blood builds up in the lungs, eventually leading to heart failure. Patients may not have any obvious symptoms in the early stages of the disease. Common symptoms include fatigue and shortness of breath, or other symptoms of heart failure. As the disease progresses it can become life-threatening. The complication rate of this disease is higher in men than in women. Early detection and treatment may help improve prognosis and delay progression.

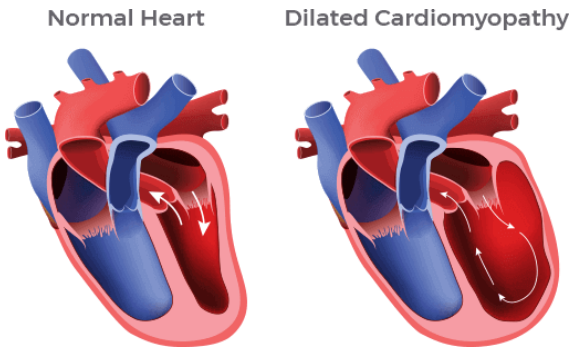

Fig. 1.13 Dilated Cardiomyopathy
https://dcmfoundation.org/dilated-cardiomyopathy-causes/

Myocardial infarction

Myocardial infarction (MI) is another Cardiac Condition and also known as "heart attack". MI occurs when the blood flow to the heart's muscle is blocked from flowing properly, parts of the myocardium become necrotic due to constant lack of oxygen, as shown in Fig. 1.14. During 1994-2017, about more than 500,000 people have died from acute myocardial infarction (AMI) (Ioacara, et al., 2020). According to National Board of Health and Welfare (Socialstyrelsen) statistics from 2018-2020, more than 20,000 people suffered from acute myocardial infarction (AMI) and about 5000 people died with this disease. The main symptoms may not be obvious, but may also be present with chest discomfort or pressure that can radiate to the neck, jaw, or arm. Early diagnosis of this disease followed by rapid restoration of blood flow to the ischemic myocardium can minimize cardiac injury and improve early and long-term morbidity and mortality.

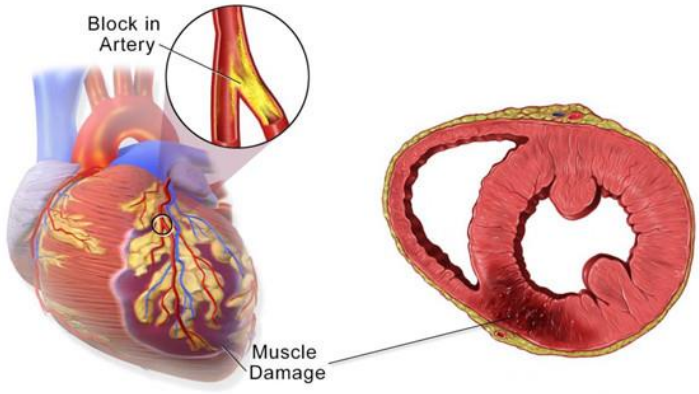

Fig. 1.14 Myocardial Infarction
http://www.miracormedical.com/picso-therapy/picso-in-ami/clinical-background/

Hypertrophic cardiomyopathy

Hypertrophic cardiomyopathy (HC) is the second most common type of cardiomyopathy, which is also called idiopathic hypertrophic subaortic stenosis (IHSS) or asymmetric septal hypertrophy (ASH) (Maron & Maron, 2013). It is due to the uncontrolled growth of myocardial fibers, causing the chambers to become thick-walled and bulky, as shown in Fig. 1.15.

This disease may affect all chambers, most notably the thickening of the left ventricular wall. The most common case is that the wall between the left and right ventricles enlarges, causing the left ventricle to twist and distort, ultimately affecting the systolic function of the left ventricle. In severe cases, this can lead to disruption of mitral valve function, with blood flowing backward. Common symptoms of HC include chest pain, shortness of breath, fatigue, lightheadedness, and swelling in the ankles, feet, legs, abdomen, etc.

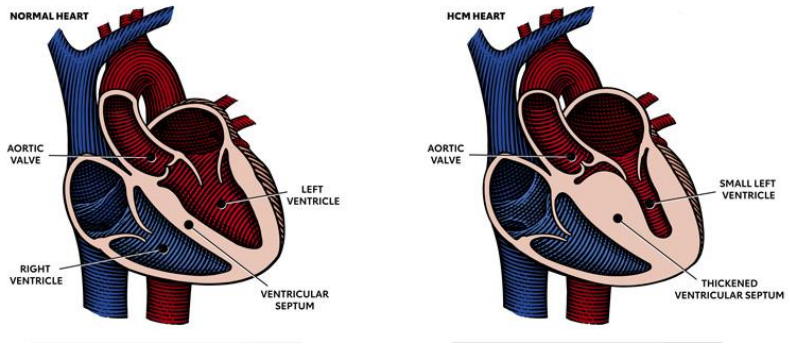

Fig. 1.15 Hypertrophic cardiomyopathy
https://www.heart.org/en/health-topics/cardiomyopathy/what-is-cardiomyopathy-in-adults/hypertrophic-cardiomyopathy

1.3.2 DTI in the human heart

1.3.2.1 Cardiac DTI techniques

Over the last decades, DTI has been extensively used in brain imaging, and solutions to most of the challenges associated with DTI-based neuroimaging have already been developed (Neil, et al., 2002). Until 1994, the use of DTI for cardiac imaging was first described (Edelman, et al., 1994). Since then, the DTI technique has been increasingly applied to the non-invasive characterization of the myocardial structures in the human heart (Mekkaoui, et al., 2017).

DTI have been widely used to depict the fiber architecture of the human heart in healthy (Edelman, et al., 1994; Tseng, et al., 2000; Dou, et al., 2002; Gamper, et al., 2007; Nielles-Vallespin, et al., 2013) and diseased (Wu, et al., 2006; Wu, et al., 2009) human heart. For the healthy human heart, DTI have been successfully applied for the assessment of myocardial fiber arragement. Although several early histological studies have shown that the fibers in the normal myocardium are orderly arranged in a crossing helical pattern, their experiments were still performed on conducted on animal hearts (Streeter Jr & Hanna, 1973; Stejskal, 1965). In this point, cardiac DTI studies on both *ex vivo* (Geerts, et al., 2002; Hahn, 1950; Lombaert, et al., 2012) and *in vivo* (Reese, et al., 1995; Tseng, et al., 2000; Tseng, et al., 1999; Dou, et al., 2002; Gamper, et al., 2007; Toussaint, et al., 2013)human heart have obtained results consistent with histological studies, which provided more robust validation of this structure. This also offers interesting knowledge for cardiac DTI modeling. In diseased human heart, cardiac DTI has also been used been used as an effective tool to describe the disease-induced fiber disarray (Wu, et al., 2006; Wu, et al., 2009). The fiber integrity after myocardial

infarction has often been quantified using fractional anisotropy (FA) and mean diffusivity (MD) (Wu & Alexander, 2007; Sosnovik, et al., 2009). All these findings suggest the potential of DTI for early diagnosis and treatment of cardiovascular diseases.

There is now a wide range of advanced DTI-based techniques available for the human heart. In fact, many of these techniques can be regarded as some extensions of DTI model, and most of them were initially developed for brain imaging but can also be applied to cardiac imaging since both of them are measuring the diffusion of water melcules within fibers. The advanced DTI-based techniques commonly used in human heart can be devided into two categories accurding to the angular resolution: low-angular resolution diffusion imaging and High-Angular Resolution Diffusion Imaging (HARDI). The angular resolution of DTI data is usually determined by the number of diffusion gradient directions used in the imaging process. Low-angular resolution diffusion imaging includes standard DW-MRI, apparent diffusion coefficient weighting imaging (Moseley, et al., 1990), and DTI (Kingsley, 2006a). HARDI includes a large number of models, such as mixture models (Descoteaux, 1999), QBI (Tuch, 2004), DSI (Wedeen, et al., 2005), Diffusion Orientation Transformation (DOT) (Özarslan, et al., 2006), and Diffusion Propagator Imaging (DPI) (Descoteaux, et al., 2009), etc.

1.3.2.2 Current challenges of cardiac DTI

Although most of DTI-based brain imaging techniques can be applied to cardiac imaging, many challenges are present in practice. The primary challenge lies in the lack of quality and/or quantity. As we knew, a diffusion tensor can be mathematically calculated from six DW signals acquired in six diffusion gradient directions (Kingsley, 2006a). In practice, however, the number of diffusion gradient directions, which defines the angular resolution of DTI data, has an important influence on the estimation of diffusion tensors (Tournier, et al., 2013; Jones, 2004; Zhan, et al., 2013). Many researchers have concluded that a higher angular resolution (e.g., more than 30 diffusion gradient directions) is required to achieve more accurate diffusion tensor estimation (Jones, 2004). However, acquiring a large number of images associated with different gradient directions is often not possible due to time constraints. Another challenge lies in the bulk motion caused by motion during *in vivo* cardiac imaging. Different from the human brain, the overall motion of the heart is a complex mixture of cardiac motion combined with both cyclic pumping and respiratory motion. Since the DTI essentially measures the microscopic motion of water molecules, it is very sensitive to subject motion. The magnitude of overall motion of the heart that occurs during imaging will induce large signal loss in DW images, as illustrated in Fig. 1.16, which subsequently degrade the accuracy of estimated diffusion tensors.

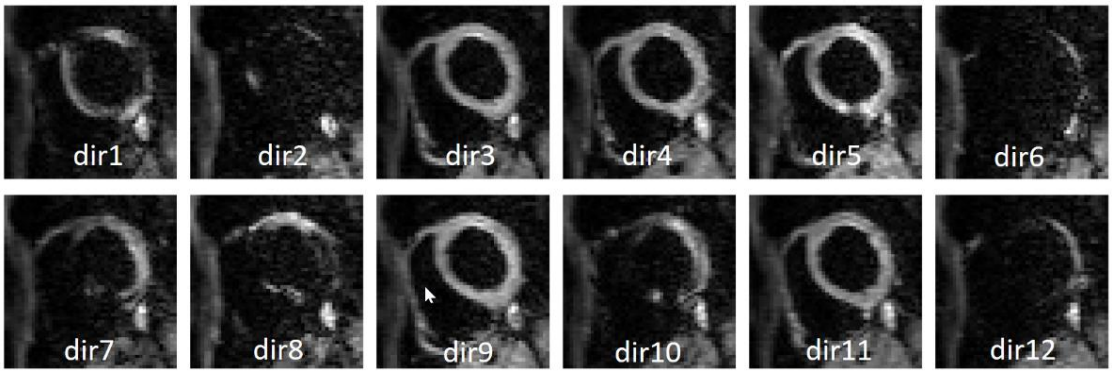

Fig. 1.16 *In vivo* cardiac DW images acquired in 12 diffusion gradient directions and corrupted with motion-induced signal loss.

In order to overcome the motion problem during imaging, several new MR sequences have been developed in the past few decades. Most of them focus on reducing the effect of cardiac motion, and combining some strategies at the same time to deal with respiratory motion. For example, (Edelman, et al., 1994) proposed a stimulated echo acquisition mode (STEAM) sequence where diffusion signals were acquired based on the spatial position and shape of the myocardium over the entire heartbeat cycle so that the effects of heartbeat motion were minimized. Tseng et al. (Tseng, et al., 1999) exploited an additional strain measurement to counteract the effects of myocardial strain produced during STEAM imaging. Gamper et al. (Gamper, et al., 2007) designed a spin echo (SE) based sequence to decrease the sensitivity of diffusion-encoding to cardiac motion while maintaining a high signal-to-noise ratio (SNR). Most other sequences can be regarded as as extensions of STEAM or SE, which have achieved improved performance of motion-effect reduction in *in vivo* DTI of the human heart (Reese, et al., 1995; Tseng, et al., 1999; Tseng, et al., 2000; Dou, et al., 2002; Nielles-Vallespin, et al., 2013; Toussaint, et al., 2013; Nguyen, et al., 2014; Stoeck, et al., 2016). However, most of them are time consuming due to the use of dedicated strategies for respiratory motion, resulting in a relatively long data acquisition time. The long scan time will compromise patient comfort and increases the risk of inducing motion artifacts. Thus, the reduction of motion effects for most sequences will come at the cost of limited amount of data or else a relatively poor SNR, making the use of these sequences in routine clinical settings still challenging. But these challenges have also contributed to the increasing interest in developing post-processing methods capable of recovering motion-induced signal loss in clinically acquired DTI data.

1.4 Conclusion

In this chapter, the diffusion magnetic rersonance imaging (dMRI), especially, diffusion tensor imaging (DTI) technique was introduced. The applications and challenges of DTI in the human heart were also presented. DTI has recently been used in cardiac imaging studies and can be a potential tool for early diagnosis and treatment of cardiovascular diseases. However, in vivo cardiac DTI remains challenging due to the motion problem of the heart. Current new MR sequences designed for cardiac DTI are difficult to apply in practice due to

the long acquisition time. Therefore, developing new methods to deal with the motion problem *in vivo* cardiac DTI becomes so significant for the clinical diagnosis and treatment.

Chapter 2

Deep convolutional adverserial networks for the generation of DTI data

Contents

2.1 Introduction

Generative adversarial network (GAN) is a promising deep learning framework that has attracted increasing research interests in both academia and industry. This is mainly due to its remarkable performance in generating images that are desired but infeasible to acquire in practice. GAN has recently been successfully applied to a wide range of medical image generation tasks. This thesis focuses on the GAN-based applications specifically for DTI data generation. In this chapter, we present an introduction to GAN, which constitutes the methodology of this thesis.

In order to better understand the GAN arthitecture, the basic components of convolutional neural network (CNN) are first presented in section 2.2. CNN generally consists of three parts: convolutional layer, pooling layer, and fully-connected layer. Convolutional layer is the core part of CNN that mainly performs the convolution operations of the networks. Pooling layer makes the output of the network partially invariant to small translations. Fully connected layer is usually used in the last few layers of the network, calculating images without convolution operations.

The concepts of GAN are introduced in section 2.3. GAN model refers to the use of an adversarial strategy to train two deep neural networks: a generator and a discriminator. The discriminator is trained to classify between real and fake images, while the generator is trained to generate images that are difficult to be classified by the discriminator. In section 2.3.1, the models of the two networks are presented. Then, numerous variants of GAN are described in section 2.3.2. These variants were developed based on eighter varying the objective of GAN model or varying the GAN's architecture.

In section 2.4, a large number of GAN-based applications specifically for MR image generation are reviewed. Some of these applications involves image reconstruction that focuses on improving the quality of MR images, and image-to-image synthesis that tries to generate a new image from the same or different modality image. Some of these applications involve image reconstruction that focused on improving the quality of MR images, some others attempt to generate desired images from images of the same or different modalities. These applications have shown that the GAN-based deep convolutional neural network can achieve state-of-the-art performance in generating MR images. As we know that DTI acquisitions are essentially DW-MR images, these applications therefore demonstrate the great potential of GAN for the generation of DTI data.

2.2 Basics of convolutional neural network

Convolutional neural network (CNN, or ConvNet) is a well-known deep learning technique which has been extensively used for solving a variety of benchmark problems in computer vision (Alzubaidi, et al., 2021). The modern framework of CNN was first established in (LeCun, et al., 1989) and later was improved in (LeCun, et al., 1998). From a computational point of view, CNN is essentially a series of deep neural networks which can take one or multiple images as input and then output another representation of the input. The relationship between the input and output is defined by the learned mathematical functions that can calculate the weighted sum for multiple inputs and output the activation values. The variables

of this function are mainly determined by the network architecture used for training. Although numerous variants of CNN architectures exist in the literature, most of them consist of similar basic components which can be devided into three types of network layers: convolutional layer, pooling layer, and fully-connected layer. In the following section, we will briefly introduce the CNN architecture by describing these three types of layers.

2.2.1 Convolutional layer

Convolutional layer is the main component of a CNN that performs the convolution operations of the network. A convolution is a linear operation that refers to the multiplication of a set of weights with the input using a kernel (or filter). Fig. 2.1 shows an example of the convolution process, where a kernel of size 3×3 is used to produce a map (output) from a 5×5 image. For each pixel in the input image, a square window with the same size as the kernel is used to determine a neighbourhood (or patch) centered on that pixel (marked blue border in Fig. 2.1). For each center pixel, a dot product is calculated between elements of the kernel and the pixels within the neighbourhood. The window slides over the height and width of the input image so that each pixel and its neighbours can be located to compute the dot product, resulting in each element of the output map. From the brain's point of view, each element in the output can be imagined as an output of a neuron that is connected to a small local region in the input image. Each neuron is obtained by calculating the weighted sum of the connected region. All the neurons in the output map that resulted from the input images share the weights (elements of the kernel) with each other.

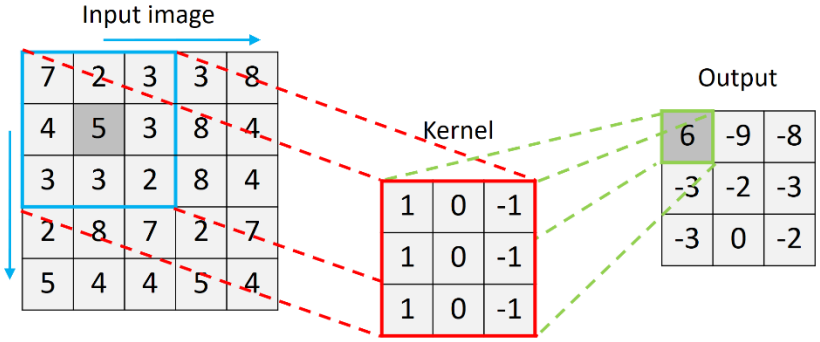

Fig. 2.1 An example of convolution. The element of the output (marked with green border) is calculated by convolving the filter/kernel with the first patch of the input image (marked with blue border).

The above example explains the computation of one convolutional layer to produce one output map/image. However, in reality, a CNN generally contains multiple convolutional layers, and the output from each layer is usually not a single image but a series of images (also called feature maps). Fig. 2.2 illustrates the famous LeNet-5 network (LeCun, et al., 1998), where 3 layers are used for convolution. The first convolutional layer (represented by C1 in Fig. 2.2) produces 6 feature maps by convoluting the original image using 6 kernels of size 5×5. The second convolutional layer (represented by C3 in Fig. 2.2) takes six 14×14 feature maps as input and produces sixteen 10×10 feature maps using 16 kernels of size 5×5. The third convolutional layer (represented by C5 in Fig. 2.2) is also a fully connected layer, which produces 120 feature maps with size 1×1 from sixteen 5×5 feature maps using 120 kernels of size 5 × 5. In each convolutional layer, each point in the output is guaranteed to

extract features from its corresponding local region in the input image. This property is usually called local connectivity. In general, the initial layers capture the low-level features of images, like lines or blobs, while the later layers extract the high-level features, like shapes or objects (Ribeiro, et al., 2018). Besides, an activation function (e.g., sigmoid or RELU) is often used after the convolution to force the relationship between input and output into a nonlinear mapping.

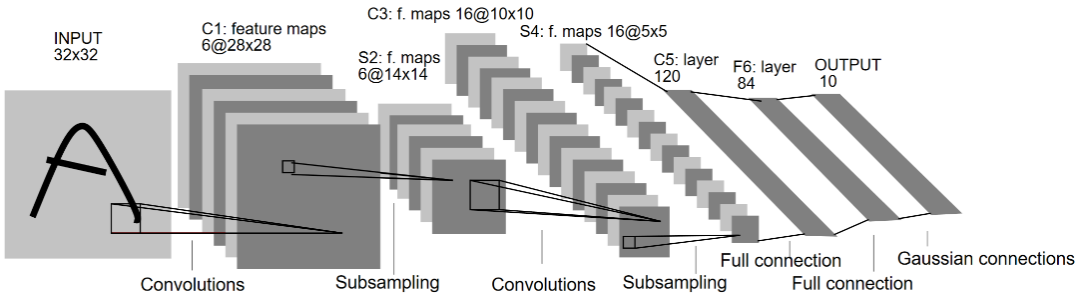

Fig. 2.2 Architecture of the LeNet-5 network. “6@28×28” means there are 6 feature maps of size 28x28. (LeCun, et al., 1998)

Note that after convolution, the dimention of the output feature is always smaller than the input. This phenomenon can also be found in the example of Fig. 2.1. This is because pixels at the edges of the input image do not have full neighbors for convolution like the internal pixels. The size of the output depends on the size of both the input image and the used convolutional filter. For an input image of size n×n and filter/kernel of size f×f, the dimension of the output feature resulted from convolution is $(n - f + 1) \times (n - f + 1)$. However, in many tasks of computer vision and medical image analysis, it is necessary to preserve the original size of the input image when applying a convolutional filter.

A commonly used method to overcome this problem is called padding, which involves adding enough zeros (or non-zeros) to the edge regions of the input image to allow the filter to perform a full convolution of the edge pixels. For example, padding can be applied to the input image in Fig. 2.1 to make the outer border of the input image zero so that the dimension of the output is equal to the input, as shown in Fig. 2.3. Besides, in some cases, the stride of the sliding window during the convolution process needs to be specified. If we denote the dimension of the zero-padding on the image boundary by P (e.g., P = 1 in Fig. 2.3), and the sliding stride S is 1, then we can set the zero padding to P=(f-1)/2 to ensure that the input and output have the same dimensions, where f is the size of filter. Zero-padding is a common way to perform convolution in CNN, and is often necessary when dealing with some problems such as image segmentation or generation where the original input is required to be the same size as the final output.

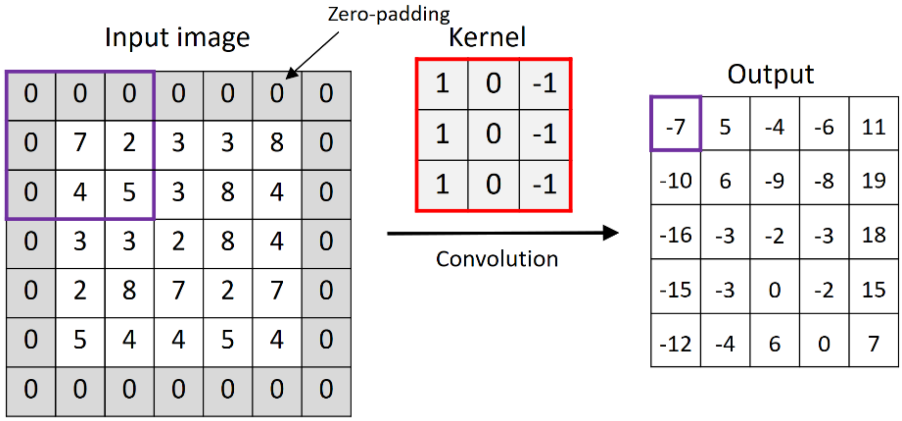

Fig. 2.3 Zero-padding. Before convolution, the output map is obtained by adding border of the input image.

2.2.2 Pooling layer

Pooling layer is another important component of CNN that is added after the convolutional layer. With the convolutional layer, each point in the output maps can correspond exactly to one region in the input. However, this also makes the output very sensitive to the location of the input, which means that small changes in the position of the feature in the input image will cause a different feature map. These changes include rotation, shifting, and other minor movements to the input image.

Pooling layers is a common down-sampling method (shown as subsampling in Fig. 2.2) to ensure that the produced output is independent of the location of the features in the input. It works in the cases where large structures or shape elements are more important than fine details. There are two types of pooling operations that are often used in CNN: max pooling and average pooling, as shown in Fig. 2.4. Max pooling returns the maximum value from the region of the image covered by the kernel, while average pooling calculate the average of all values from that region.

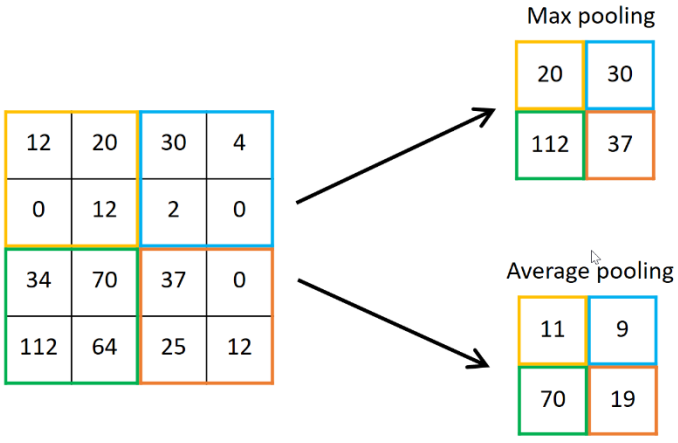

Fig. 2.4 Max pooling and average pooling. The kernel of size 4×4 is used for computation.

The convolution followed by a pooling layers is useful for extracting the characteristics which are positional invariant, making the output unchanged by the change of location of the input contents. Besides, it also reduce the computational power required to process the data.

2.2.3 Fully-connected Layer

Fully-connected (FC) layer are usually the last few layers in CNN networks. For example, the fifth layer of the famous LeNet-5 (represented by C5 in Fig. 2.2) is also a FC layer. In this layer, every element of the output is connected to all elements of the input. Taking the C5 layer in the LeNet-5 as an example, it uses 120 kernels of size 5×5 to produce 120×1×1 feature maps from the 16×5×5 feature maps of the previous layer S4. In fact, each of the produced 120 pixels is connected to the 5×5 neighborhood on all 16 feature maps of S4. Therefore, C5 is completely connected to S4.

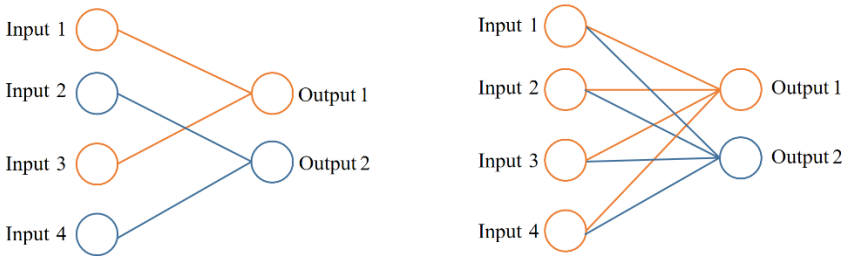

Fig. 2.5 Diagram for convolutional layer (left) and fully-connected layer (right).

Fig. 2.5 clarifies the difference between the convolutional layer and FC layer. As we can see, each of the outputs in the convolutional layer is only connected to two inputs, while each of the outputs in the FC layer is connected to all the inputs. In general, Maps requires one step before entering the fully connected layer. Compared to the use of convolutional layer, applying a FC Layer is simpler way to learn non-linear combinations of the high-level features, but usually requires larger computation cost to implement.

2.3 Generative adversarial network

Generative Adversarial Network (GAN) is a special deep learning framework that has become a hot research topic in recent years. It was first proposed in 2014 by Goodfellow *et al.*, (Goodfellow, et al., 2014), which involves the use of an adversarial strategy for training two deep neural networks. To date, numerous variants of GAN have been successfully applied to the generation of realistic data in a range of fields, including both computer vision and medical imaging (Gui, et al., 2021; Yi, et al., 2019; Qin, et al., 2019). In the following sections, we begin with an introduction of the principles of GAN model in section 2.3.1, followed by a description of some advanced GAN variants in section 2.3.2.

2.3.1 GAN model

The model of GAN was initially introduced as an interesting alternative for the generation of realistic image using deep neural networks (Goodfellow, et al., 2014). It consists of two sub-networks that are trained simultaneously: a generator G and a discriminator D, where G focuses on image generation and D on discrimination. Fig. 2.6 illustrates the configuration of a GAN network. The input to G, z is a random noise vector sampled from a prior distribution

$p_z(z)$. The output of G, $G(z)$ is expected to be as similar as possible to the real sample (or image) r that comes from the real data distribution $p_r(r)$. The input to the discriminator D is either a real sample r or a generated sample $G(z)$. Then, D tries to output a single value P_{out} that represents the probability of the input being a real or fake sample.

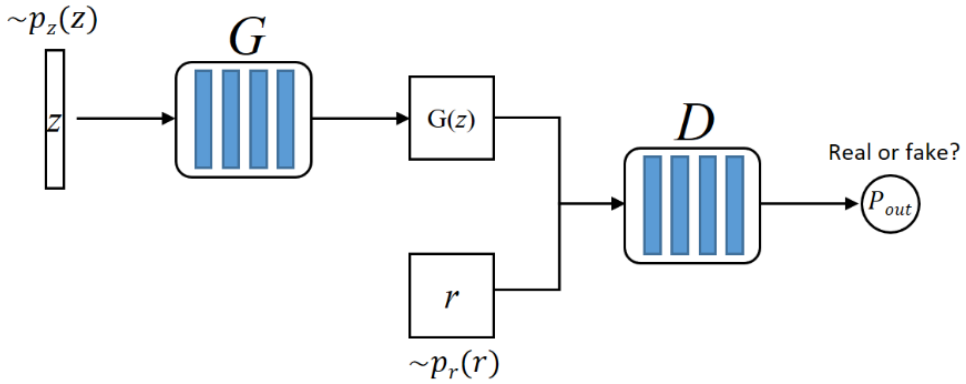

Fig. 2.6 Configuration of GAN network

The objective function of D is expressed mathematically as:

$$L_D^{GAN} = \max_D E_{r\sim p_r(r)}[\log D(r)] + E_{z\sim p_z(z)}[\log(1-D(G(z)))] \tag{1.22}$$

With D, the objective function G can be expressed as:

$$L_G^{GAN} = \min_G \mathrm{E}_{z\sim p_z(z)}[\log(1-D(G(z)))] \tag{1.23}$$

In general, Eqs. (1.22) and (1.23) are generally integrated into one min-max function:

$$L_{(G,D)}^{GAN} = \max_D \min_G E_{r\sim p_r(r)}[\log D(r)] + E_{z\sim p_z(z)}[\log(1-D(G(z)))] \tag{1.24}$$

Eq. (1.24) indicates an adversarial strategy for training both D and G. During training D, the network parameters are updated by minimizing $D(G(z))$ while maximizing $\log D(r)$ to improve its ability to classify real sample r and fake sample $G(z)$. Then, the trained D can be used as an additional network to calculate the $D(G(z))$ so that G can be trained by minimizing $\log(1-D(G(z)))$ to generate sample $G(z)$ that are difficult to be classified by D. After that, the generated $G(z)$ can in turn be used as new fake sample to minimize $D(G(z))$ for updating the parameters of D. This process continues until D can no longer to distinguish between $G(z)$ and r. The authors in (Goodfellow, et al., 2014) have demonstrated that if the D is trained to optimality before the next G updates, then minimizing L_G^{GAN} is essentially minimizing the Jensen-Shannon divergence between the distribution of generated data and the real data distribution $p_r(r)$, making the generated samples very close to the real samples in terms of data distribution.

2.3.2 Variants of GAN

GAN provides an ingenious ideal to build deep neural networks for the data generation problem. With the increasing applications of GAN in many fields, the limitations of the original model have also been found. In order to overcome these limitations and allows the network to be effectively applied to various problems, some GAN variants have been proposed. According to the way they modify the original GAN model, these variants can be devided into two categories: 1) varying the objective function of GAN, 2) varying architecture of the network. To learn more details, we can refer to the recent review articles (Yi, et al., 2019; Wang, et al., 2021). The following is a brief introduction of some representative variants of GAN.

2.3.2.1 Varying the objective of GAN

As described in section 2.3.1, the original GAN model does not have any restrictions on the patterns of the data it generates. A natural idea to change the network is to provide extra information during the generation process in order to generate samples with properties we desired. This is refered as a conditional GAN (cGAN) proposed by Mirza and Osindero (Mirza & Osindero, 2014). The objective function of cGAN is obtained by simply modifying the min-max function of Eq. (1.24), which is expressed as:

$$L_{(G,D)}^{GAN} = \max_{D}\min_{G} E_{r\sim p_r(r)}[\log D(r,c)] + E_{z\sim p_z(z)}[\log(1-D(G(z,c)))] \tag{1.25}$$

where c denotes a kind of auxiliary information such as an images. As shown in Eq. (1.24), both the generator G and the discriminator D are conditioned on c. In other wards, G is now forced to learn a mapping from a fake joint representation $\{z,c\}$ to a real joint representation $\{r,c\}$.

Previous works have further varied cGAN model by using different kinds of auxiliary information as conditional input c (Mirza & Osindero, 2014; Denton, et al., 2015; Reed, et al., 2016). One of the most common conditions is an image, which has been applied to handle various image generation problems (Mathieu, et al., 2015; Wang & Gupta, 2016; Karacan, et al., 2016). In 2016, Isola *et al*. proposed a general-purpose cGAN model for image-to-image translation tasks. They have shown that even without noisy z, the cGAN network still effectively learns the mapping from the input image to the target image. Similar to the Eq. (1.25), their objective function is expressed as:

$$L_{(G,D)}^{GAN} = \max_{D}\min_{G} E_{r\sim p_r(r)}[\log D(r,img)] + E_{z\sim p_z(z)}[\log(1-D(G(img)))] \tag{1.26}$$

where *img* can denote a label map used for synthesizing the photo, or an edge map that need to be colorized. Despite the use of image or labels, an alternative way to add the extra information to the network is to combine two generators together from head to toe so that the cycle consistancy between two networks can be used to transform images between two unpaired sets of images. A representative network of this ideal is CycleGAN (Zhu, et al., 2017). Besides, some other papers have not used cGAN but applied the original GAN with a common loss terms (such as L2 regression) to force the output to be conditioned on the input

and achieved impressive results in image-to-image translation (Pathak, et al., 2016; Zhou & Berg, 2016; Zhu, et al., 2016).

In original GAN model, G and D should be trained alternately from scratch so that the generation ability of G and the discrimination ability of D can increase together. However, in practice, there is no guarantee that the training of G and D is stable. In other wards, the ability of one may inevitably be better than the other, and in most cases D beomes stronger. This is also called mode collapse. When this situation occurs, the gradient from D will be close to zero, which provides no guidance for training G. To balance the training of G and D and avoid the mode collapse, new losses for D were proposed, including least-square loss (Mao, et al., 2017), hinge loss (Miyato & Koyama, 2018), f-divergence loss (f-GAN) (Nowozin, et al., 2016), and Wasserstein distance loss (Arjovsky, et al., 2017). In addition, similar to the work in (Mirza & Osindero, 2014), the (Donahue, et al., 2016) also proposed to use joint representation as D's objective and thus D was centered on separating real and fake joint samples.

2.3.2.2 Varying network architecture

The original GAN model was built based on fully-connected layers (Goodfellow, et al., 2014). This building block was later replaced by Radford et al. with a fully convolutional layer, resulting in the deep conventional GAN (DCGAN) (Radford, et al., 2015). In their study, the conventional network demonstrated better training stability compared to the fully-connected layers for constructing the GAN networks. Besides, the BatchNorm (Ioffe & Szegedy, 2015) and LeakyRelu (Maas, et al., 2013) are two important strategies of this network. Beside, in the pix2pix network (Isola, et al., 2017), the "U-Net"-based architecture was used instead of using traditional deep neural network as the generator (Goodfellow, et al., 2014), where a "skip connection" strategy was used to to improve the performance of the image-to-image transform. A diagram of the pix2pix model is shown in Fig. 2.7. As shown, this network contains two main parts: the encoding and decoding units. Each unit is essentially a convolutional layer of CNN, where the convolution operations are performed followed by ReLU activation. The encoder downsamples the image to extract multi-scale level features. The decoder use upsampling operators to progressively recover the multi-scale level features of the image from low to high resolution. The "skip connections" combines the downsampled features with the upsampled features so that it can let the information to bypass a large number of unrelated layers.

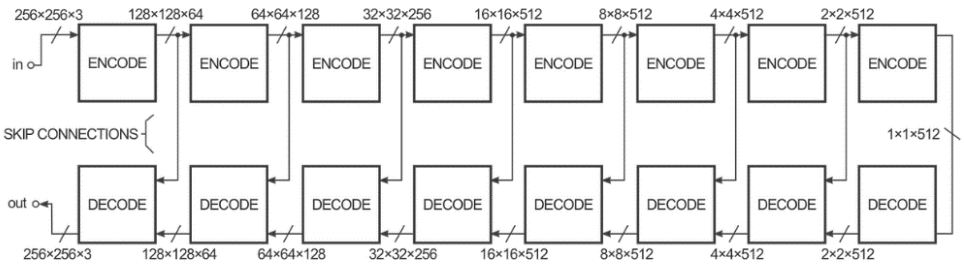

Fig. 2.7 Artchitecture of pix2pix GAN.
https://affinelayer.com/pix2pix/

2.4 Current applications of GAN in MR image generation

In recent years, a large number of applications of GAN have emerged in a wide range of fields. In the field of medical imaging, there are generally two ways to use GAN network. The first way focuses on medical image generation, which can help in depicting the underlying distribution of the images and producing images that are desired but difficult to obtain in practice. The second way refers to the use of the learned discriminator to deal with some classification tasks. In the following paragraphs we center on the first way, briefly reviewing current applications of GAN specifically for MR image generation.

A common problem in MRI is the low quality of acquired images caused by the limitations of the hardware and physics and also due to the limited scan time in practice. GAN-based post-preprocessing methods for solving this problem include denoising, super-resolution, etc. This type of problem is also called MR reconstruction in the literature. For example, Chen *et al.* (Chen, et al., 2018) proposed a multi-level densely connected network for MRI super-resolution, where a GAN was trained to output more realistic-looking images. Kim et al. (Kim, et al., 2018) also used an adversarial network to generate high-resolution MR images from one contrast to another. In (Dar, et al., 2018), the cGAM was successfully applied for synergistic reconstruction-synthesis of multi-contrast MRI. Besides, the adversarial neural network was also used to estimate the missing samples during partial MRI sampling and finally accelerated the time for MRI scans (Shitrit & Riklin Raviv, 2017). With the cyclic consistency loss, a deep learning-based generative adversarial model was proposed for accurate compressed sensing (CS) MRI reconstruction (Quan, et al., 2018). These studies demonstrated that it is feasible to use clinically available acquisitions combined with dedicated GAN to cope with MR reconstruction. Among these applications, pix2pix (Isola, et al., 2017) is one of the most commonly used building blocks of them. More publications about GAN-based MR image reconstruction are listed in Table 2.1.

Table 2.1 MR image generation (reconstruction) publications.

Publications	Network architecture	Tasks
(Quan, et al., 2018)	pix2pix	CS-MRI reconstruction
(Mardani, et al., 2017)	pix2pix	CS-MRI reconstruction
(Ran, et al., 2019)	pix2pix	Denoising
(Chen, et al., 2018)	pix2pix	Superresolution
(Kim, et al., 2018)	pix2pix	Superresolution
(Armanious, et al., 2019b)	pix2pix	Inpainting
(Abramian & Eklund, 2019)	CycleGAN	Facial anonymization
(Sánchez & Vilaplana, 2018)	pix2pix	Superresolution
(Armanious, et al., 2019a)	pix2pix	Motion correction
(Oksuz, et al., 2018)	pix2pix	Motion correction

Another common application of GAN in MR images is called image-to-image synthesis, which involves generating desired MR images from the clinically available MR images, or estimating MR images from images of another modality (e.g., CT). For example, Calimeri *et al.* employed a Laplacian pyramid GAN framework to generate MRI slices of the human brain, which can be seen as a biomedical data augmentation method (Calimeri, et al., 2017). With the same purpose, Han *et al.* (Han, et al., 2018) demonstrated that the Wasserstein GAN (WGAN) (Arjovsky, et al., 2017) can generate realistic muti-sequence brain MR images

better than the DCGAN (Radford, et al., 2015). Zhang *et al*. (Zhang, et al., 2017) presented a semi-coupled-GAN to automatically generate adversarial examples which can help to detect missing basal and apical slices in Cardiac magnetic resonance (CMR) images. Besides, Bermudez *et al*. developed a novel DCGAN to learn implicit brain MRI manifolds from MRI data and further applied their method to MR image denoising (see Fig. 2.8 for their denoising results).

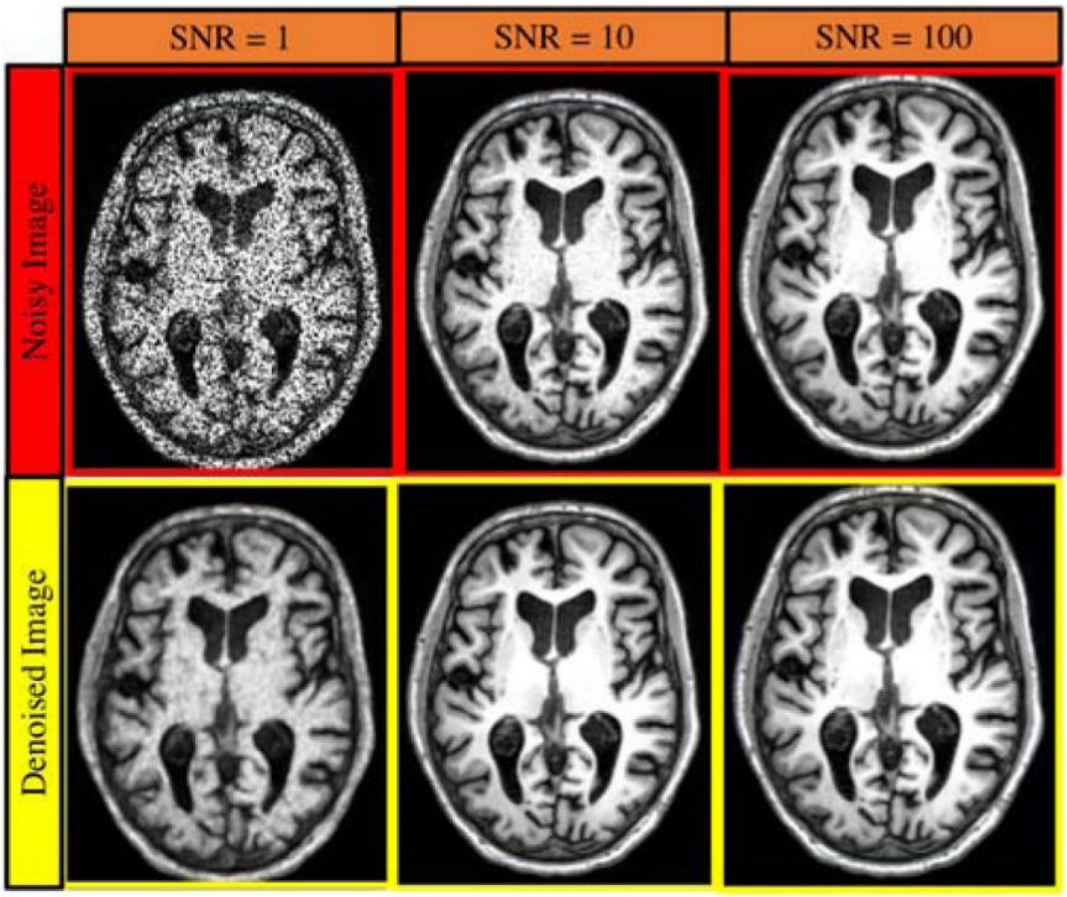

Fig. 2.8 MR image denoising results presented in (Bermudez *et al*., 2015).

In addition to the above applications of GAN, some other studies focused on the transformation between different modal images. For example, numerous GAN-based methods have been developed for estimating MR images from CT images, or estimating CT images from MR images (Chartsias, et al., 2017; Zhang, et al., 2018; Huo, et al., 2018; Hiasa, et al., 2018; Wolterink, et al., 2017; Maspero, et al., 2018). Besides, GAN-based methods for estimating PET images from MR modalities have also been proposed (Wei, et al., 2018; Pan, et al., 2018). Moreover, in (Nie, 2018), generating 7T MR images from 3T MR images has also been achieved effectively combining fully convolutional network with an adversarial learning strategy. An example of this network's results is shown in Fig. 2.9.

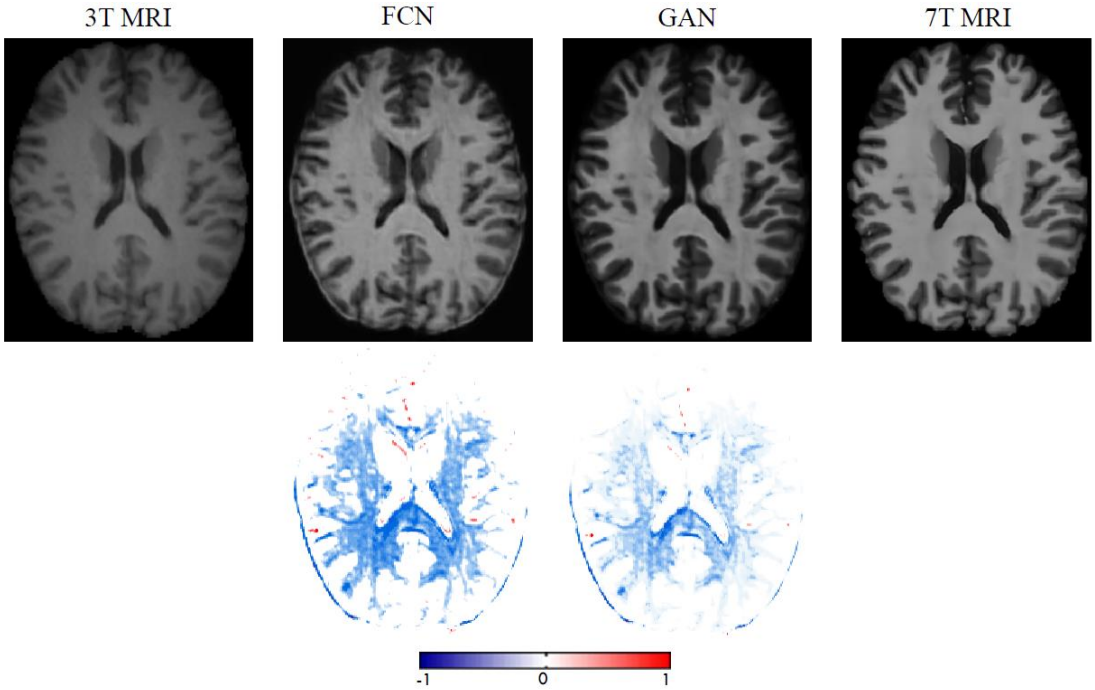

Fig. 2.9 Image generation from 3T to 7T MRI using GAN-fully convolutional network. The 1st row shows 3T MRI, two synthetic 7T MRI by two methods, and ground-truth 7T MRI. The 2nd row shows difference maps between each synthetic 7T MRI and ground-truth 7T MRI. Note that FCN means the case without adversarial learning, and GAN means the case with adversarial learning. (Nie et al., 2018)

2.5 Conclusion

In this chapter, we mainly introduced the deep convolutional GAN, which constitutes the methodology of this thesis. For a better understanding of GAN arthitecture, the basic components and general architecture of CNN were first explained. Then, the GAN model and its variants were described. Finally, current GAN arthitectures that have been applied for MR image generation were reviewed. A large number of related applications have shown that the GAN-based deep convolutional neural network can achieve state-of-the-art performance in generating MR images, which demonstrated great potential towards the application of GAN in DTI data.

PART II

Contributions

Chapter 3

Systematic study of joint influence of angular resolution and noise in cardiac diffusion tensor imaging

Contents

Abstract

Diffusion tensor imaging (DTI) is a promising imaging technique that allows non-invasively study of diffusion properties and fiber structures of myocardial tissues by measuring the diffusion of water molecules within them. Accurate estimation of diffusion tensors is important for the measurement of water diffusion and thus characterizing the myocardial fiber structure. Noise and angular resolution are two common factors that affect the estimation accuracy of the diffusion tensor in practice. Previous studies have investigated the influence of noise or angular resolution independently on the estimation of diffusion tensors in DTI. However, the joint influence of these two factors in DTI remains unclear. In this chapter, we propose to systematically study the joint influence of angular resolutions and noise levels on the estimation of diffusion tensors and tensor-derived fractional anisotropy (FA) and mean diffusivity (MD). For this purpose, DW image sets with different angular resolutions were sampled from simulated and real cardiac DTI datasets and then were added with different levels of Rician noise for comparison and analysis. For each direction gradient direction, DW images at multiple time points were also collected for studying the accuracy of diffusion tensor estimation in the case of fixed scan times. Experimental results showed that, as expected, given a certain noise level and sufficient acquisition time, the estimation accuracy of diffusion tensor, FA and MD all increase as the angular resolution. Moreover, when the angular resolution reached a certain value, further increasing the number of diffusion gradient directions has little effect on the estimation of diffusion tensor, FA and MD. Also, both the mean and variance of FA or MD decrease as the angular resolution increases. For an imposed acquisition time, increasing the angular resolution reduces SNR of DW images. When fixing SNR, higher angular resolution can be obtained at the expense of longer acquisition time. These findings suggest the necessity of an optimized trade-off when designing DTI protocols.

3.1 Introduction

Diffusion tensor imaging (DTI) refers to a magnetic resonance imaging (dMRI) technique that measures diffusion of water molecules within biological tissues using diffusion-weighted pulse sequences. DTI makes it possible to explore diffusion properties and fiber structures of tissues non-invasively compared to conventional imaging modalities (Tuch, et al., 2003). In cardiac imaging, for example, DTI has been increasingly used to investigate myocardial microstructure changes related to many cardiac disorders such as hypertrophic cardiomyopathy (Mekkaoui, et al., 2013), accute ischemia (Sosnovik, et al., 2014), myocardial infarction (Wu, et al., 2006), etc.

As mentioned in section 1.2.3, in DTI, water diffusion in tissues is described by diffusion tensors estimated from a set of diffusion-weighted (DW) images associated with multiple noncollinear diffusion gradient directions. The number of gradient directions defines the angular resolution of DTI data acquired. Specifically, it is possible to estimate the diffusion tensor $\underline{D}$, a 3×3 symmetric positive definite matrix, according to the Stejskal-Tanner equation (Stejskal & Tanner, 1965): $S_i = S_0 \exp(-b\vec{g}_i^T \underline{D}\vec{g}_i)$, for i = 1,2, …, n (n≥6), where n is the number of diffusion gradient directions, S_i is the DW signal intensity acquired in the i^{th} gradient direction, $\vec{g}_i = (x_i, y_i, z_i)$, and b is the diffusion weighting factor (Kingsley, 2006b). At a given voxel, the water diffusion properties are quantified by calculating the eigenvalues of the diffusion tensor and tensor-derived measures (Taylor, et al., 2004) such as fractional anisotropy (FA) and mean diffusivity (MD).

Accurate estimation of diffusion tensors is very important for characterizing water diffusion and thus assessing the myocardial fiber structure. In practice, however, the estimation accuracy of diffusion tensors can be influenced by several factors, e.g., image quality, estimation algorithm, tissue complexity, etc. Noise and angular resolution are two common factors that have attracted much attention in both research and clinical practice. The noise sources in DW images refer to random signals arising from the hardware (e.g., gradient-coil noise (Hurwitz, et al., 1989), field inhomogeneity (Vovk, et al., 2007), etc.) and inherent motions of the subject (e.g., cardiac and respiratory motion). These noises give rise to perturbed DW signals, and thus produce errors in the calculated diffusion tensor as well as its derived measures. The angular resolution in DTI determined by the number of diffusion gradient directions is used in acquiring DW images. It also influences the estimation of diffusion tensors.

The influence of noise or angular resolution in DTI has been independently investigated by researchers. For example, Pierpaoli et al. (Pierpaoli & Basser, 1996) showed that the MD derived from diffusion tensors decreases as the noise level increases. Anderson in (Anderson, 2001) demonstrated that higher noise levels increase the maximum eigenvalue of tensors, and thus result in a larger FA value. For the angular resolution, Papadakis et al. (Papadakis, et al., 2000) compared various DTI angular sampling schemes, and showed that the minimum number of gradient directions for accurate estimation of FA is around 18-21. Jones (Jones & Basser, 2004) found that at least 20 gradient directions are required for robust estimation of FA, and at least 30 directions are required for robust estimation of tensor orientation and MD.

Despite these useful findings, it remains unclear how the angular resolution and noise jointly affect the diffusion tensor estimation in DTI.

We propose in this chapter to systematically study the joint influence of angular resolution and noise in cardiac DTI. Specifically, a set of DW images associated with different diffusion gradient directions are generated using a simulated diffusion tensor field. Besides, real datasets of four human hearts are acquired in a large number of diffusion gradient directions. Then, different levels of Rician noise are added to both the simulated and real DW images to produce datasets with different signal-to-noise ratio (SNR). Different angular resolutions of data are achieved by varying the number of considered diffusion gradient directions for each SNR. Finally, the quality of diffusion tensors and tensor-derived FA, MD for different angular resolutions and different noise levels is quantitatively assessed. In addition, the relationship between angular resolution and SNR for a fixed acquisition time is also studied.

This study mainly has two contributions. First, this is, to our knowledge, the first work to systematically investigate the joint influence of different angular resolutions and different levels of noise in cardiac DTI. Second, we quantitatively defined the minimum angular resolution required to obtain near-optimal diffusion tensors, FA and MD at a certain noise level, and also investigated the maximum angular resolution for keeping certain SNRs when imposing acquisition time.

3.2 Materials and methods

3.2.1 Simulated datasets

Simulated DW data were generated to analyze the influence of angular resolution and noise level on the estimation accuracy of diffusion tensors. The simulation strategy is similar to the method proposed in (Frindel, et al., 2009). Specifically, a 20 × 20 × 20 diffusion tensor field containing 9 homogeneous tensor regions with discontinuities of different amplitudes was created to represent 9 different structures of a human heart. Each z-slice of the tensor field is defined by

$$\underline{D}_{fd} = \begin{pmatrix} R_0 & R_1 & R_0 & R_2 \\ R_0 & R_3 & R_0 & R_4 \\ R_0 & R_5 & R_0 & R_6 \\ R_0 & R_7 & R_0 & R_8 \end{pmatrix} \tag{2.1}$$

where R_i represents a 5 × 5 homogeneous region containing 25 of the same diffusion tensors. Here we set these tensors by using the same coefficients as in (Frindel, et al., 2009), where the eigenvalues of each tensor were (2, 1, 1) × 10^{-3}mm^2/s. The b-value was 700 s/mm^2, which is the same value that is used in the real cardiac datasets. The FA and MD values for the simulated noisy DW images associated with 12 directions were FA = 0.33 ± 0.06 and MD = 1.3 ± 0.11 × 10^{-3}mm^2/s, which are typical for the left ventricular (LV) myocardium of *ex vivo* human hearts according to the statistical results in (Zhang, et al., 2013).

The simulated DW images were computed from the above diffusion tensor field using Stejskal-Tanner equation (Stejskal & Tanner, 1965) with associated diffusion gradient

directions. In order to analyze the effects of noise, different levels of Rician noise (Gudbjartsson & Patz, 1995) were added to the ideal DW images with different standard deviation values. The noise levels of the simulated DW images were measured via SNR (as shown in Fig. 3.1). Different angular resolutions of DW images were achieved by sampling different number of associated diffusion gradient directions. The sampling method for diffusion gradient directions is detailed in section 3.2.3. In this chapter, the number of excitations used for signal averaging for each direction was set only when acquisition time is fixed; otherwise, the number was set to 1.

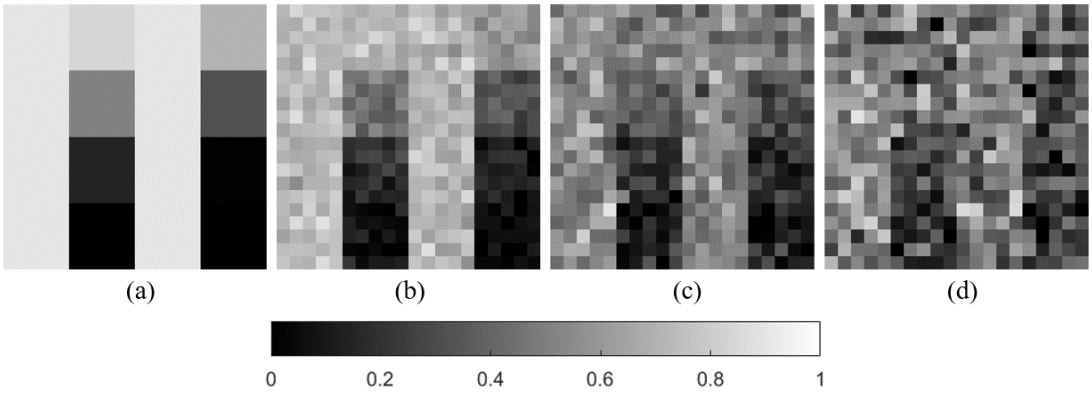

Fig. 3.1 An example of simulated DW images with: (a) no noise (ground-truth), noise levels with standard deviation (b) σ = 0.02 (SNR = 23 dB), (c) σ = 0.05 (SNR = 15 dB), and (d) σ = 0.1 (SNR = 15 dB). The encoding gradient direction g_i related to the DW image is (0.894, 0, 0.447).

3.2.2 Real datasets

DTI of *ex vivo* human hearts including two infants and two adults were performed. The infant datasets were acquired in clinical conditions with a Siemens 3T MRI Magnetom Verio. The imaging parameters are the following: TE = 74 ms, TR = 7900 ms, FOV = 144 × 144 mm^2, slice thickness=1.4 mm, in-plane resolution = 2 mm, slice spacing = 1.4 mm, slice duration = 123.2 ms, number of slices = 35, slice size: 104 × 104 pixels, diffusion sensitivity b = 700 s/mm^2, and number of gradient directions = 192, 64 or 12. In each direction, MRI scans were acquired 6 to 10 times for noise reduction.

The adult datasets were acquired using Siemens 3T MRI Magnetom Prisma with following parameters: TE = 71 ms, TR = 9600 ms, FOV = 177 × 177 mm^2, slice thickness=1.5 mm, slice spacing = 1.5mm, slice duration = 123.1 ms, number of slices = 70, slice size: 122 × 122 pixels, diffusion sensitivity b = 700 s/mm^2, and number of gradient directions = 192. In each direction, MRI scans were performed 3 times for noise reduction. An example of real DW images from an infant and an adult is given in Fig. 3.2.

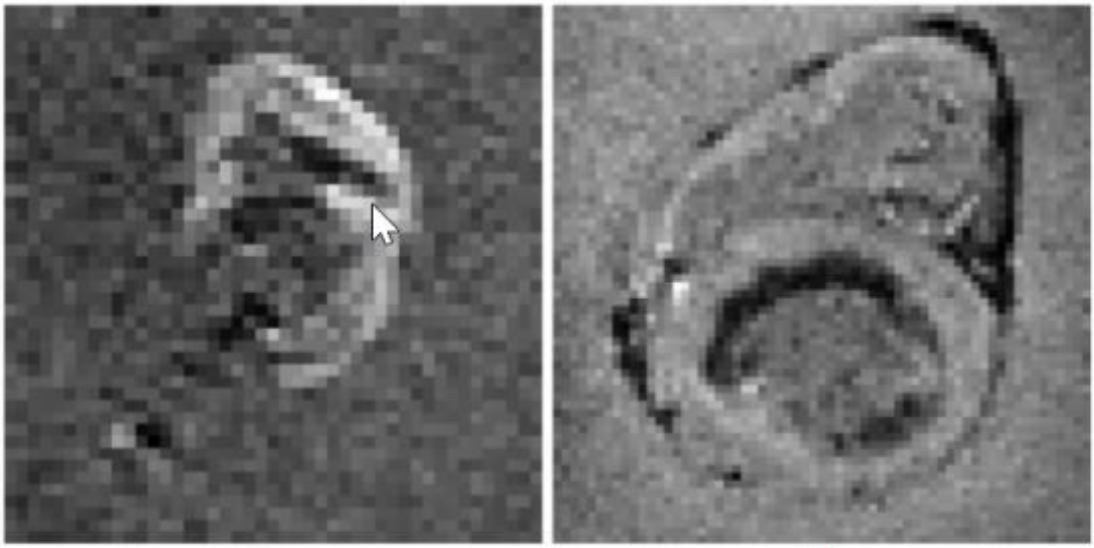

Fig. 3.2 An example of DW images from real datasets. (a) and (b) show the DW image of and infant and an adult, respectively.

3.2.3 Gradient direction sampling

We considered two commonly used angular sampling schemes to generate spherically uniform distributed gradient directions for different angular resolutions of simulated datasets: Spherical Tessellation (ST) (Teanby, 2006) and Electrostatic Energy Minimization (EEM) (Jones, et al., 1999). The ST employs spherical polyhedrons to generate specified numbers of directions geometrically uniformly distributed on a sphere. The EEM is capable of generating an arbitrary number of uniformly distributed directions by minimizing the electrostatic energy based on Coulomb's law. In order to systematically analyze the effects of various angular resolutions, we used EEM to generate different angular resolutions with uniformly-spaced numbers of diffusion gradient directions and one resolution with 6 directions (minimum number of gradient directions required for tensor estimation). For the real datasets, the popular spherical code sampling method (Cheng, et al., 2018) was used to sampling different angular resolutions from the real 192 directions. The sampling results are illustrated in Fig. 4.3. The uniformity distribution of the directions on the sphere is measured by the covering radius (CR) defined in (Cheng, et al., 2018).

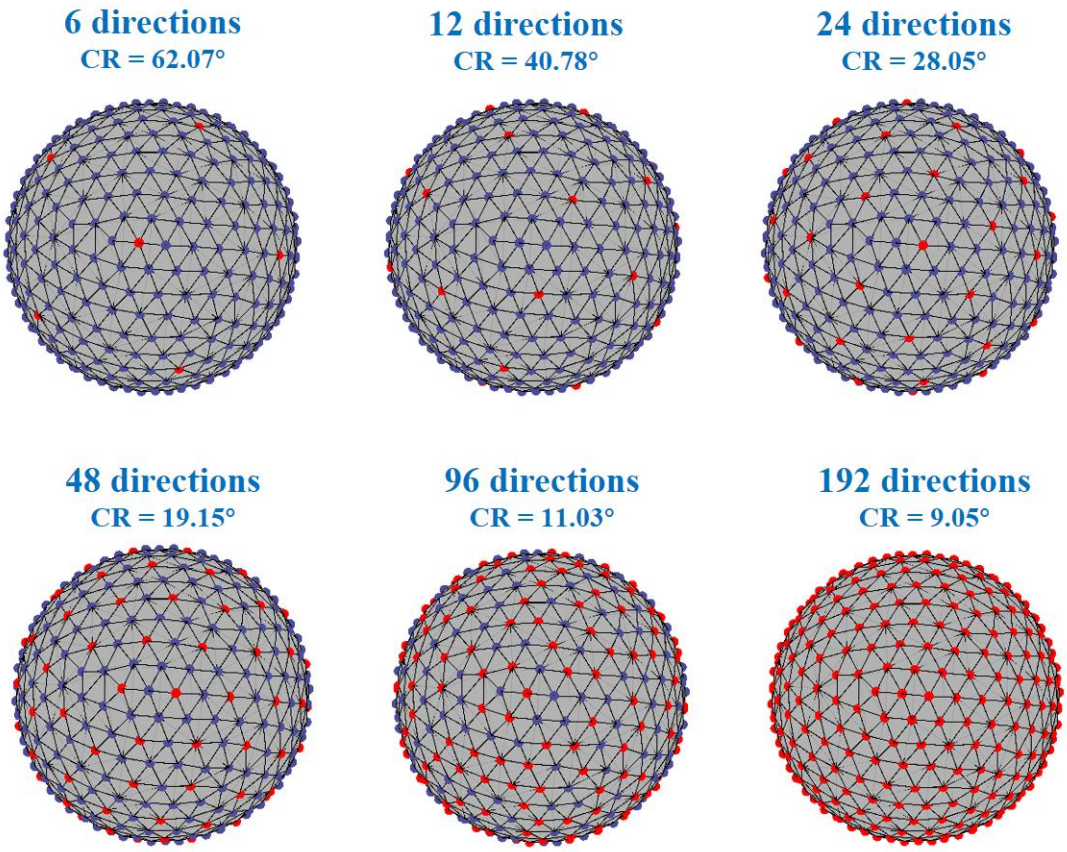

Fig. 3.3 Sampling directions of diffusion gradients for different angular resolution on a sphere. The blue and red points represent the set of 192 directions and its subset directions sampled from it, respectively. CR designates the covering radius.

3.3 Evaluation

In the case of the simulated data, the SNR (dB) is defined by:

$$SNR = 20 \cdot \log_{10}\left(\frac{V_n}{\frac{1}{N_v}\left\|V_n - V_f\right\|_2}\right) \tag{2.2}$$

where V_n denotes a discrete volume corrupted by Rician noise and V_f its noise-free representation. $\|\bullet\|_2$ is the standard Euclidean norm, and N_v is the number of voxels. In the case of real cardiac data, the SNR (dB) is defined by

$$SNR = 20 \cdot \log_{10}\left(\frac{\mu_{rm}}{\sigma_{rm}}\right) \tag{2.3}$$

where μ_{rm} and σ_{rm} are the mean and standard deviation (SD) of the voxel intensity values over the region of the myocardium.

We also evaluated the simulated and real data with different angular resolutions and different levels of noise by estimating diffusion tensors using standard least-squares estimation. The FA and MD were then derived from the diffusion tensors. For the simulated data, the mean Frobenius distance, root mean square errors (RMSE) of FA and MD (In the following sections, we call them FA or MD errors.) between ground-truth (the simulated tensor field) and the estimated diffusion tensors were computed. For the real data, the mean and standard deviation of FA and MD were computed.

3.4 Results

3.4.1 Results on simulated data

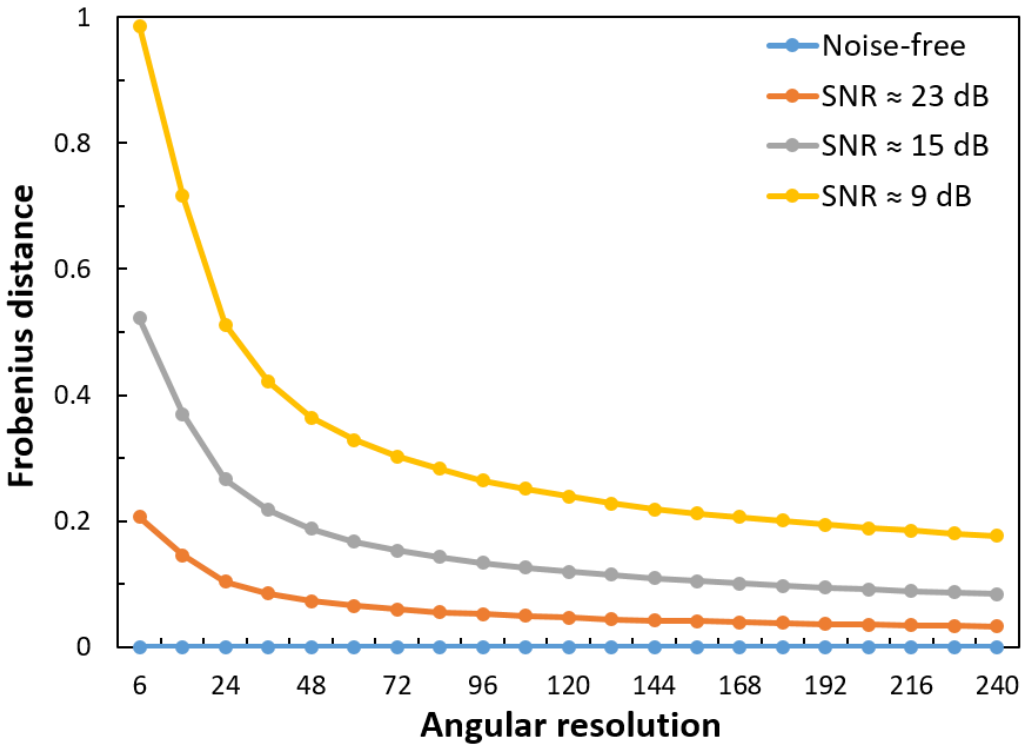

Fig. 3.4 Differences (Frobenius distance) between estimated and ground-truth tensor fields as a function of angular resolution and level of noise

In Fig. 3.4, Frobenius distance curves are illustrated for different angular resolutions and different levels of noise. Here the used numbers of diffusion gradient directions were N_d = {6,12,24,36,48,…,240}. From this figure, we can clearly see that for each SNR, Frobenius distance decreases as the angular resolution increases, and that for each angular resolution, the distance decreases as the SNR increases. The values of Frobenius distances for the maximum angular resolution (240) are smaller than 0.2 even for smaller SNR. Besides, the values of Frobenius distance for the dataset without noise (blue line) are always less than $1.2 \times 10^{-}$

^{15}mm^2/s, which means that varying the angular resolution of noise-free dataset has no impact on the estimated diffusion tensors.

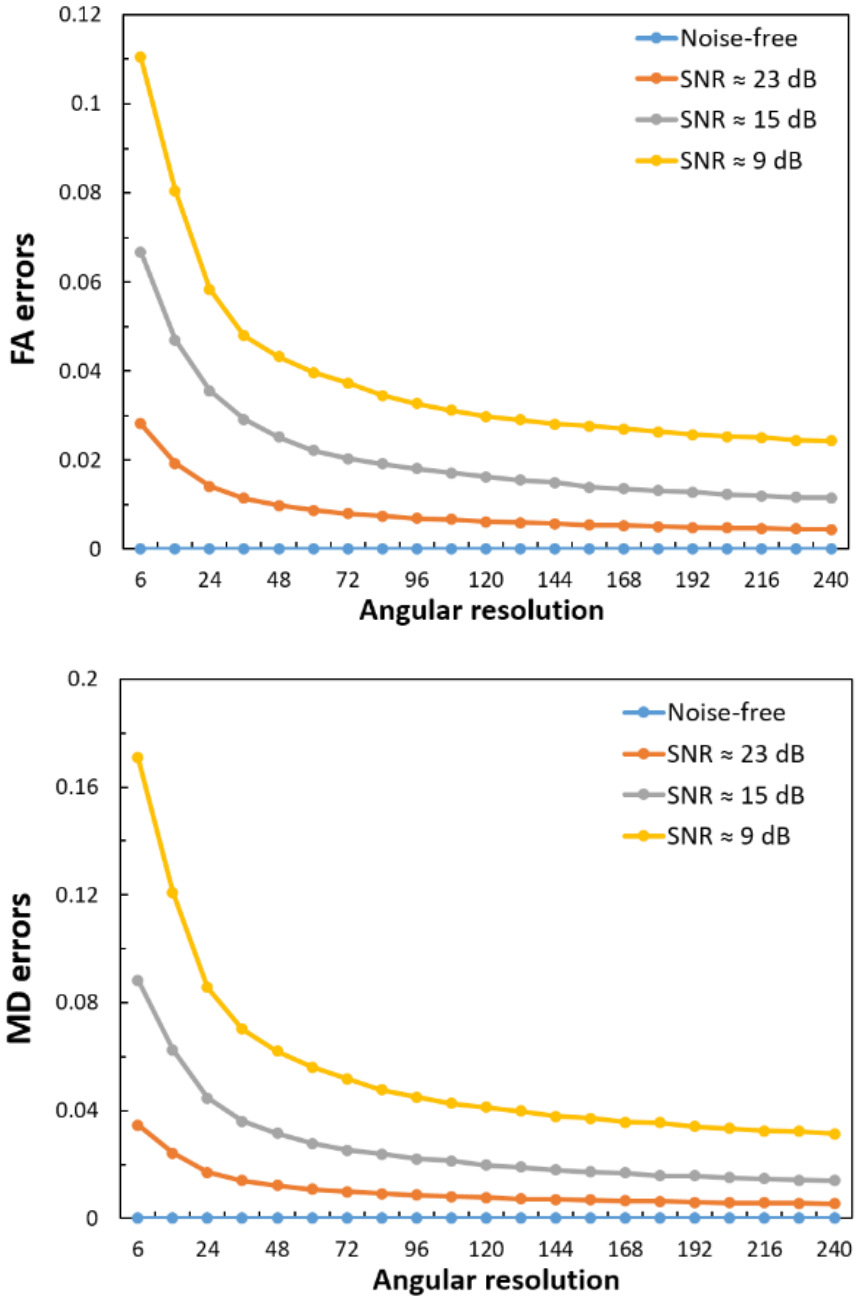

Fig. 3.5 Errors on tensor-derived measures for different angular resolution and level of noise. Top: FA errors. Bottom: MD errors. The unit of MD is 10^{-3}mm^2/s.

Fig. 3.5 shows the curves of FA and MD errors as a function of the angular resolution and noise. The values of FA and MD were computed from the diffusion tensors corresponding to the results in Fig. 3.4). As observed, for all SNRs, the errors of FA or MD decrease as the angular resolution increases. Moreover, for each SNR, there is a tendency for both FA and MD errors to stabilize when the number of diffusion gradients (i.e. angular resolution)

reached a certain number. This tendency can also be found in Fig. 3.4. To describe it clearly, we defined this number as the minimum angular resolution required for which the Frobenius distance, FA or MD errors come within 1% of their range (subtract the minimum value from the maximum value). For example, with this rule, the near-optimal diffusion tensors (the range of Frobenius distances smaller than 0.02) for SNR = 23 dB, 15 dB and 9 dB were achieved with 48, 120 and 156 diffusion gradient directions, respectively. The near-optimal FAs (the range of FA errors smaller than 0.006) for SNR = 23 dB, 15 dB and 9 dB were achieved with 48, 120 and 132 diffusion gradient directions, respectively, while for the near-optimal MDs (the range of MD errors smaller than 0.007), these numbers were 48, 108 and 124, respectively.

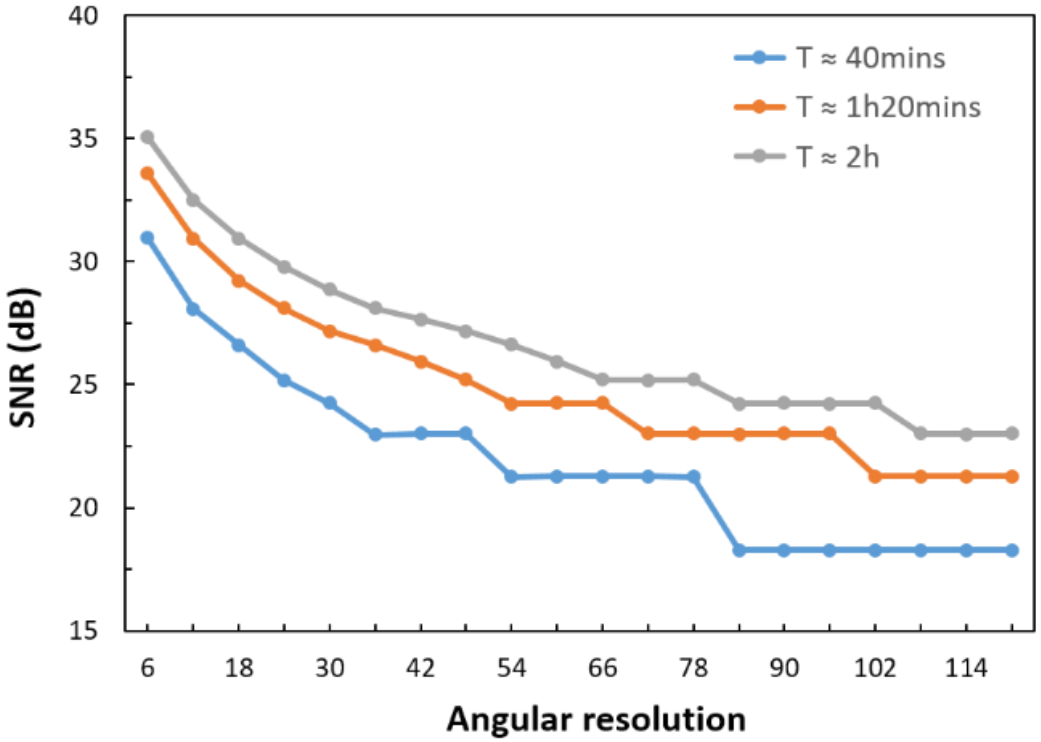

Fig. 3.6 SNR as a function of angular resolution for three fixed acquisition times.

Fig. 3.6 gives the SNR curves as a function of the angular resolution for three acquisition times. To facilitate comparison, we assumed that the scan time for each simulated DW image was approximately 1s, and that the total acquisition time T is defined by the number N_d of diffusion gradient directions multiplied by the number N_e of excitations for each direction (used for signal averages): $T = N_d \times N_e$. Here $N_d = \{6,12,18,24,30, \ldots ,120\}$ were sampled for achieving different angular resolutions. N_e was set such that for a given angular resolution, we kept roughly the same acquisition time, e.g., $N_e = \{20,10,7,5,4,\ldots,1\}$ was chosen for a acquisition time of approximately 40 minutes. From this figure, we can see that for each acquisition time, the SNR value decreases as the angular resolution increases. Moreover, longer acquisition time allows the use of higher angular resolution while keeping a certain

value of SNR. For example, when SNR ≥ 25 dB, the maximum available angular resolutions were 25, 48, and 78 for T ≈ 40 mins, 1h20min, and 2h, respectively.

3.4.2 Results on real data

Fig. 3.7 shows how the mean and standard deviation (SD) of FA and MD vary with the angular resolution and noise on real cardiac data. To evaluate the effect of noise, three additional datasets (SNR= 17 dB, 11 dB and 4 dB) were generated by adding different levels of Rician noise to the original DW images. The SNR was computed using Eq. (2.2), where V_n and V_f are the discrete volume corrupted with Rician noise and without noise, respectively. It can be seen that for all the noisy datasets with different levels of noise, the mean FA and MD value decreases as the angular resolution increases. Besides, the results also show that the MD value decreases as the SNR increases. This is consistent with the results in Pierpaoli et al's study (Pierpaoli & Basser, 1996).

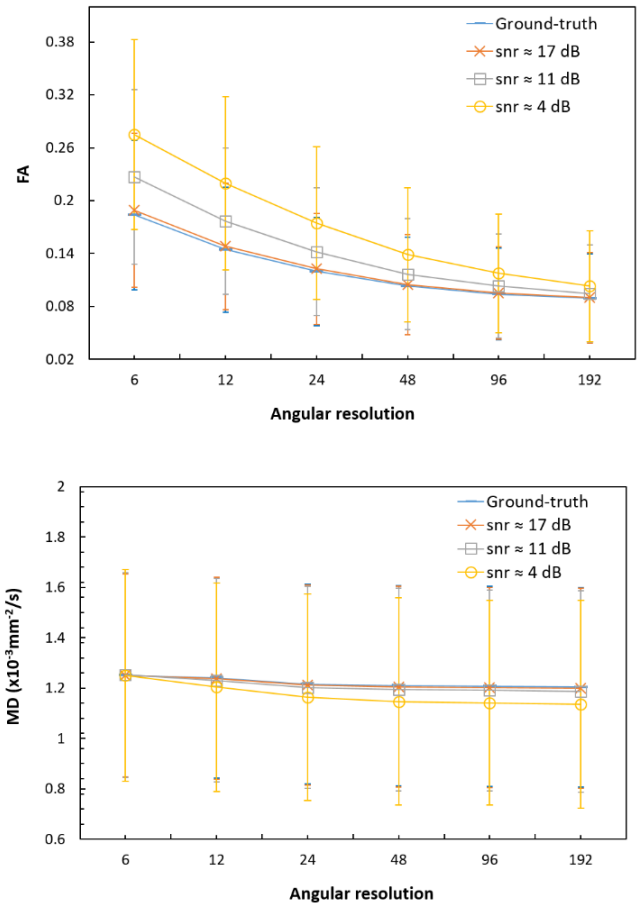

Fig. 3.7 Mean ± SD of FA (top) and MD (bottom) values for different angular resolutions and different levels of noise on a real dataset of an infant heart.

Fig. 3.8 illustrates the variation of SNR as a function of the angular resolution on a real cardiac dataset (an infant heart) when the acquisition time is fixed to approximately 1h40min. Different angular resolutions were achieved by sampling Nd={32,38,44,50, …, 140} from the

original 192 diffusion gradient directions. Ne ={6,5,4,3, …, 1} so that we kept a roughly equivalent acquisition time. The step shape occurs in Fig. 3.7 because for some continuous N_d, a same Ne was set for them in order to make sure that their total acquisition time were closer to 1h40min. For example, for Nd from 62 to 92, the Ne was set to 2; but for Nd ⩾ 98, Ne was set to 1. It can be seen that for a fixed acquisition time, SNR decreases as angular resolution increases. Besides, the maximum angular resolution for SNR ⩾ 12 dB is 62.

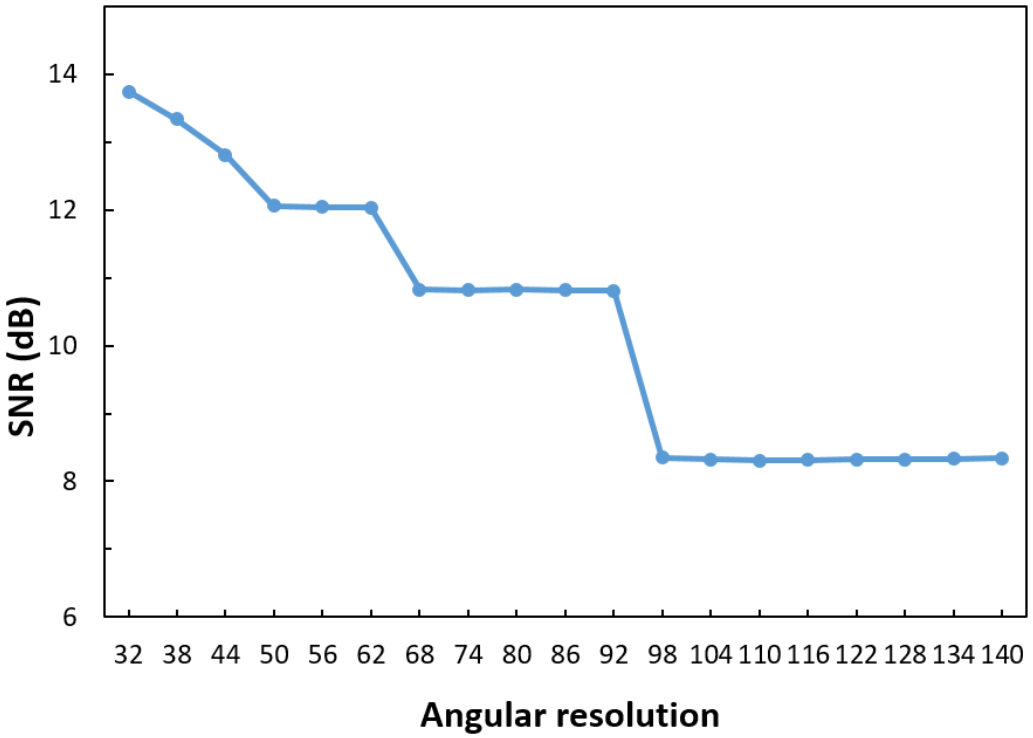

Fig. 3.8 SNR as a function of the angular resolution for fixed acquisition time on a real cardiac dataset.

Fig. 3.9 shows the main fiber orientations for different angular resolutions with 6 directions, 12 directions, 24 directions, 48 directions, 96 directions, and 192 directions with FA image as backgrounds. It can be seen that the arrangement of fiber orientation becomes more regular with the increase of angular resolution. When the number of directions used reaches 96 directions, the fiber directions show a periodic circular arrangement in the left ventricle. This indicates a cross-helix anatomical structure of myocardial fibers, as described in 1.3.1.1.

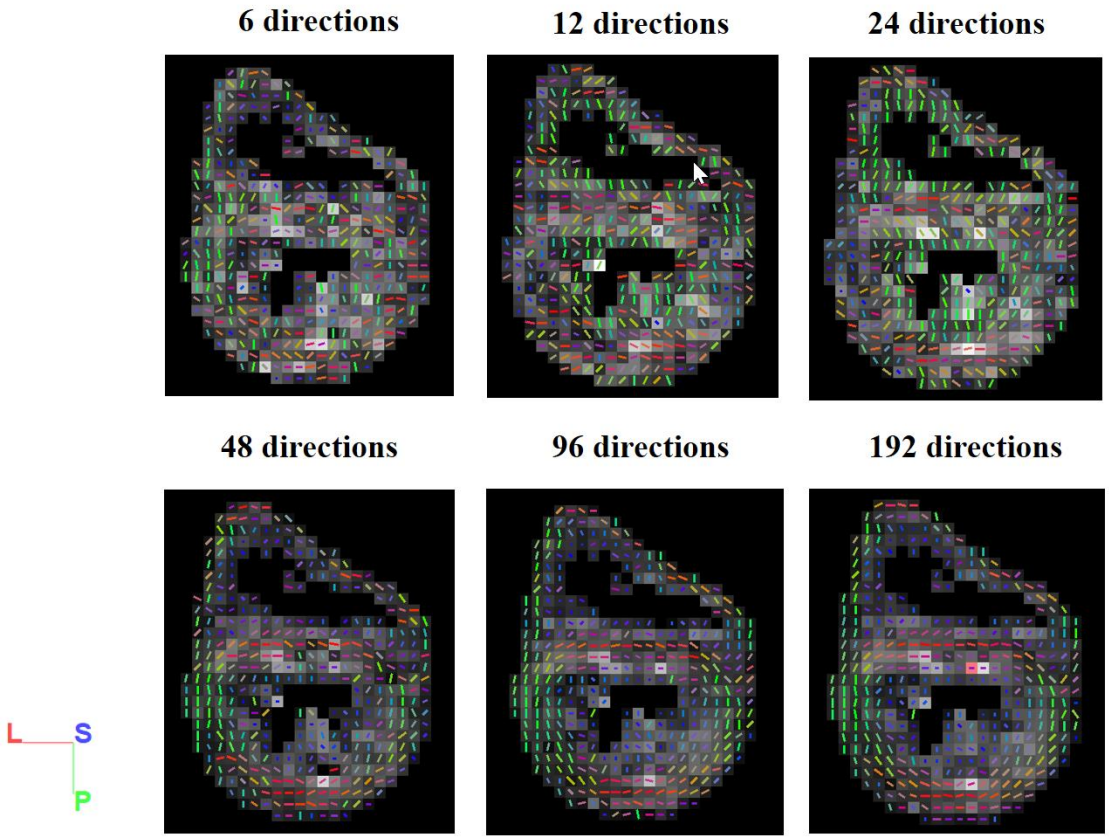

Fig. 3.9 Main fiber orientation for differention angular resolutions. The P, S, and L designate the axes of anterior-posterior, inferior-superior, and left-right in anatomical coordinate system, respectively.

3.5 Conclusion

In this study we investigated the joint influence of angular resolutions and noise levels in cardiac DTI. The results on both synthetic and actual DW images showed that, as expected, given sufficient acquisition time and a certain noise level, the accuracy of diffusion tensors, FA and MD measurements increase as angular resolutions increase. Besides, continuing to increase the angular resolution beyond a certain value has little effect on the accuracy of diffusion tensor, FA and MD. For higher SNR (23 dB), the near-optimal diffusion tensor, FA and MD were obtained with 48 gradient directions. With lower SNR, the near-optimal diffusion tensor, FA and MD were obtained with, for example, 120 directions if SNR = 15 dB and with 156 directions if SNR = 9 dB. In addition, the results on real data showed that both the mean and variance of FA or MD decrease as the angular resolution increases. When acquisition time is imposed, increasing angular resolution often comes at the cost of a reduced number of excitations used for signal averaging, thus reducing the SNR of DW images. Longer acquisition times allow the use of higher angular resolution while keeping a certain level of SNR. These findings suggest the necessity of an optimized trade-off when designing DTI protocols. In the future, we would like to conduct a similar study on *in vivo* DTI of the human heart.

Chapter 4

GAN-based super angular resolution in cardiac DTI

Contents

Abstract

Diffusion tensor imaging (DTI) allows for the non-invasive study of cardiac fiber structures by using the diffusion tensor to characterize the diffusion of water molecules within them. Our previous study has demonstrated that increasing the angular resolution appropriately is an interesting alternative to improve the estimation accuracy of diffusion tensors. In practice, however, higher angular-resolution DTI is usually limited by longer data acquisition times that reduces patient comfort and introduces motion problems. Generative adversarial network (GAN) provides an effective avenue to cope with this problem by generating diffusion-weighted (DW) images in more diffusion gradient directions. Unfortunately, most existing GAN-based methods developed for single image translation are not applicable to the DTI data due to the continuous spatial-angular information of DW image sequences rather than independent spatial information in each image. In this study, we proposed a novel GAN-based image generation method to improve the angular resolution of cardiac DTI data and thus the estimation accuracy of diffusion tensors. It consists in using a GAN network to estimate a sequence of images from given DW images acquired in a limited number of diffusion gradient directions. Moreover, the correlation between multiple DW images associated with different diffusion gradient directions is exploited by introducing a supervised manifold regularized term in the objective function of the proposed network. Experimental results on cardiac DTI data demonstrated that the proposed method significantly improved the quality of DTI reconstruction. This suggests that the proposed method is an effective tool for super angler resolution in cardiac DTI.

4.1 Introduction

Diffusion tensor imaging (DTI) is a non-invasive imaging modality that probes the displacement of the water molecules using diffusion-weighted pulse sequences (Basser, et al., 1994). This makes it possible to study cardiac fiber structures at a microscopic scale and thus used in various applications and diseases (Peyrat, et al., 2007). In DTI, the diffusion of water molecules is characterized by the diffusion tensor estimated from a sequence of diffusion-weighted (DW) images acquired using diffusion gradients along multple directions. The angular resolution of the imaging is determined by the number of diffusion gradient directions used, each corresponding to one sample point in q-space. A simple DTI model requires DW images associated with six diffusion gradient directions to be able to mathematically calculate a diffusion tensor matrix that describes the orientation of a single fiber for each voxel. More complex configurations (such as fiber crossing and bending) usually require a more sophisticated model, e.g., High Angular Resolution Diffusion Imaging (HARDI) (Tuch, et al., 2002) acquires DW images in a higher angular resolution. Besides, our previous study has also demonstrated that an appropriate increase in angular resolution is an alternative way of improving the estimation accuracy of diffusion tensors. However, the acquisition of a larger number of DW images also lengthens the acquisition time. Such longer acquisition time is often undesirable in clinical practices as it reduces the patient confort and induces more motion artifacts. This becomes a common difficulty in increasing the angular resolution of cardiac DTI.

To deal with this problem, many post-acquisition methods were proposed. For example, Tuch et al. (Tuch, 2004) interpolated signals by weighted averaging of angular neighboring measurements in q-space. Descoteaux et al. (Descoteaux, et al., 2007) proposed to employ spherical harmonic coefficient vectors for the measurement interpolation. Moreover, several compressive sensing (CS) methods were developed, in which joint x-q space information from DTI data was also exploited to generate HARDI signals with a higher quality (Baraniuk, 2007; Cheng, et al., 2015; Mani, et al., 2015). However, these post-acquisition methods usually need dedicated imaging protocols and thus are difficult to apply in practice. In recent years, deep learning methods have attracted great research attention in the field of medical image analysis due to their ability to automatically and accurately predict the desired but difficult-to-obtain medical images (Shen, et al., 2017). Deep learning networks for super resolution of images already exist. For instance, Dong et al. (Dong, et al., 2015) addressed the single image super resolution using convolutional neural network (CNN) and gained impressive results. Kim et al. (Kim, et al., 2016) improved the performance of the super-resolution scheme using a recursive CNN and a simple parametric learning model. More recently, generative adversarial network (GAN) (Goodfellow, et al., 2014) has been applied to the super resolution of MR images (Kim, et al., 2018). However, these studies are centered on super spatial resolution, with very few networks focused on improving the angular resolution of DTI data.

Disregarding the super-resolution problem, current deep learning methods, in particular, GAN-based methods have achieved remarkable results in many medical image generation tasks, as detailed in section 2.4 of this thesis. From this point on, GAN offers an interesting

alternative to improve the angular resolution of DTI by generating DW images in more diffusion gradient directions. Unfortunately, existing GAN-based methods were mainly focused on other modility images, such as computerized tomography (CT), positron emission tomography (PET), or standard MR images. Moreover, most of these methods were designed for single medical image transformation, while it is nontrivial to directly apply them to DW image generation due to the spatial-angular information within multiple images rather than independent spatial information in one image.

In this study, we proposed a novel GAN-based method for super angular resolution of cardiac DTI. Unlike previous work that focused on single image translation (as detailed in section 2.4), our approach is a group translation strategy that consists of training a generative model to simultaneously estimate sequences of new DW images from a low-resolution DW images acquired in a limited number of gradient directions. Instead of directly using traditional GAN networks, we used the U-net architecture (Radford, et al., 2015) to build the generator and then take the low-resolution DW images as a constraint to learn the networks so that the realistic DW images can be generated by learning multi-level features. Moreover, we modified the traditional GAN loss function using a supervised manifold regularized term to exploit the correlation between DW images in different gradient directions.

4.2 Methods

4.2.1 Problem Formulation

The goal of our method is to train a deep generative network that can predict DW images with higher angular resolution from DW images with low angular resolution. Given a set of low-angular resolution DW images *X*: $\{x_1, x_2, x_3 \dots, x_n\}$ that have been acquired in *n* diffusion gradient directions, assuming that there is a set of higher-angular resolution DW images associated with *n* new diffusion gradient directions: *H*: $\{r_1, r_2, r_3, \dots, r_{2n}\}$. With the DL network, we try to learn the mapping from *X* to *H*. For clarity but without loss of generality, we set where *X* is a subset of *H*, for example, *H*: $\{x_1, x_2, x_3 \dots, x_n, r_1, r_2, r_3 \dots, r_n\}$. Then our problem can be changed into leaning a map *G* from existing *X* to the remaining images *R*: $\{r_1, r_2, r_3 \dots, r_n\}$. The diffusion gradient directions for both *X* and *R* are distributed uniformly on a sphere. By the training of the network, we wish that the generated data *G*(*X*) $\{g_1, g_2, g_3 \dots, g_n\}$ is as similar as possible to thee ground-truth image set *R*.

4.2.2 GAN-based super angular resolution network

We propose to use conditional GANs (Mirza & Osindero, 2014) due to their ability to generate highly realistic images with a limited number of observed DW images *X*. More specifically, we trained two networks to compete with each other, a generator *G* and a discriminator *D*. Fig. 4.1 illustrates an overview of proposed GAN network, the generator *G* applies the transformation to *n* observed DW images *X* to predict a set of *n* DW images *G*(*X*) associated with *n* new directions. The input to the discriminator are both the observed DW images *X* and the *n* additional DW images associated with *n* new directions, where the n additional DW images can be either the output images *G*(*X*) or the real images *R* that comes

from the original dataset. Then the discriminator tries to decide if the input image set is the real pair $\{X, G(X)\}$ or fake pair $\{X, R\}$. In Fig. 4.1, the red dashed arrows denote the back propagation of the network that adjust the network parameters to update the generator. This process continues iteratively until the generator can produce highly realistic DW image sequences. Fig. 4.2 gives the detailed network architecture of the generator and discriminator. These architectures are built based on the idea in (Isola, et al., 2017).

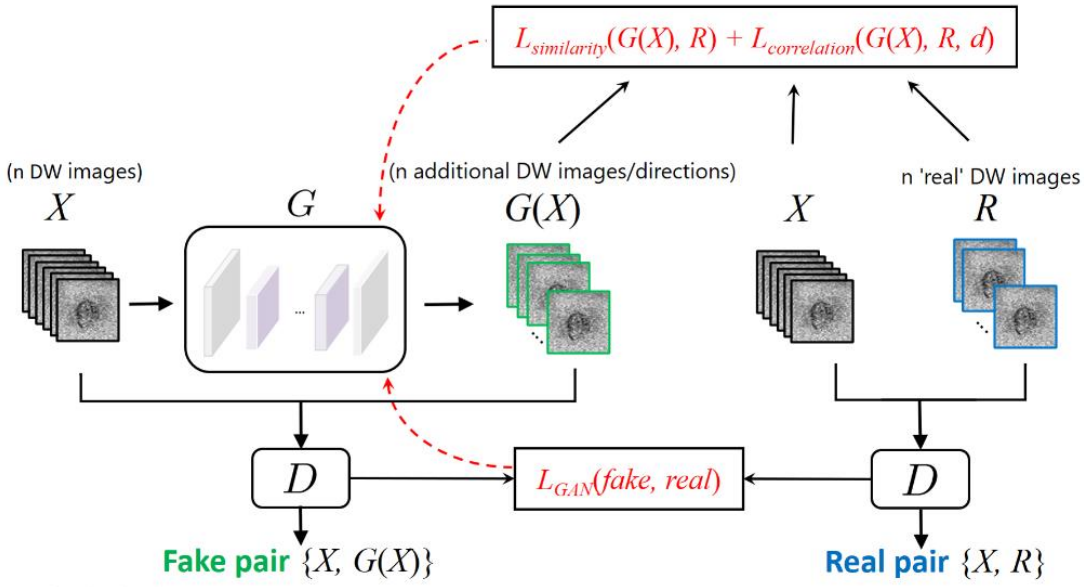

Fig. 4.1 Schematic illustration for the proposed manifold regularized GAN. The network maps DW images acquired in *n* gradient directions into new *n* directions.

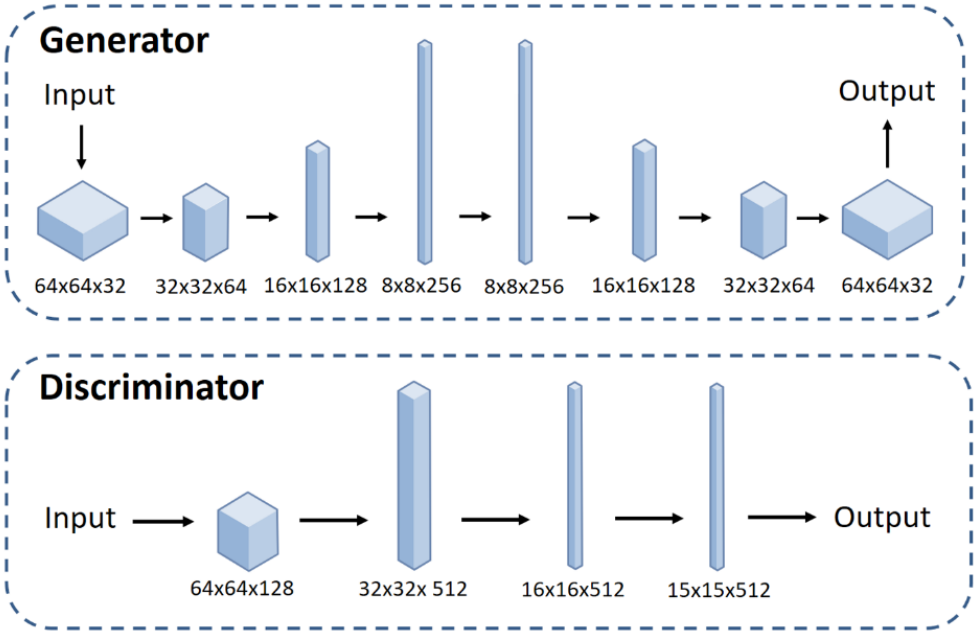

Fig. 4.2 The detailed architecture of proposed network. The solid arrows indicate the forward propagation of the network. The number below the box indicates the size of the feature map in its corresponding layer.

4.2.3 Objective function

The main objective function of the proposed GAN network is designed based on the conditional network in (Mirza & Osindero, 2014) and can be expressed as:

$$L_{(G,D)}^{GAN} = \max_{D} \min_{G} E_{X,R}[\log D(X,R)] + E_{X}[\log(1 - D(X, G(X)))] \tag{3.1}$$

where $G(X)$: $\{g_1, g_2, g_3 \ldots, g_n\}$ represents the generated DW images for n new gradient directions compared to the ground-truth of DW images R from real dataset. The discriminator D tries to maximize the probability of assigning the correct label to both real R and generated $G(X)$ groups. Simultaneously, G tries to minimize the second term against D. By minimizing this function, the capability of generator G and discriminator D will gradually increase from the competition process until $D(X,R)$ and $D(X,G(X))$ have the same distributions.

Previous studies have found that it is beneficial to combine the traditional GAN loss with L1 or L2 norm (Pathak, et al., 2016). Therefore, in our method, the task of the generator is not only to produce realistic images but also to output difference between real and fake DW images in L2 sense. This L2 loss function characterizes the similarity between generated and ground-truth images and can be expressed as:

$$L_{similarity} = E_{X,Y} \left\| R - G(X) \right\|^2 \tag{3.2}$$

where $\|\bullet\|^2$ represents some distance function, such as Frobenius norm. For a sequence of DW images from different gradient directions, this term is expressed as $\frac{1}{n}\sum_{i=1}^{n}\|r_i - g_i\|^2$.

4.2.4 Manifold regularized loss function

GAN provides an ingenious competition scheme with the ability of guiding the generative model to produce data very similar to the distribution of real data. However, our experiments showed that the traditional GAN even with L2 loss does not perform well in practice, as it neglects the continuous spatial-angular information (e.g., the correlation of pixels in multiple directions within a 3D neighborhood) in DW images. To deal with this problem, we additionally minimize a manifold regularized function (Li, et al., 2018) to increase the angular correlation and thus improve performance of the proposed GAN network. Such regularized function can be expressed as:

$$L_{correlation} = E_{X,R} \left\| f(R) - f(G(X)) \right\|^2 W \tag{3.3}$$

where $f(\bullet)$ is an embedding function to extract the useful information from the raw data. W is a weight characterizes the relationships between different sample points in real data R. $\left\| f(R) - f(G(X)) \right\|^2$ measures the difference between the real and generated data. If $f(R) = x$, then minimizing this term is equivalent to minimize Eq. (3.2). In our cases, both the generated and real data are a sequence of DW images for a number of directions. Then Eq. (3.3) can be can be approximated as:

$$L_{correlation} = \frac{1}{n^2} \sum_{i,j}^{n} \left\| r_i - g_i - r_j + g_j \right\|^2 W_{i,j} \tag{3.4}$$

for $i, j \in \{1,2,\ldots,n\}$. $W_{i,j}$ is the correlation weight defined by considering the angular similarity between *i*th and *j*th directions of DW images. It can be set according to the Locality Preserving Projections strategy (He & Niyogi, 2004) as follows:

$$W_{i,j} = e^{-\frac{\|d_{i,j}\|^2}{\varphi}} \tag{3.5}$$

where $d_{i,j}$ denotes the angular difference between *i*th and *j*th diffusion gradient directions, φ controls the scale of the correlation weight $W_{i,j}$. In DTI, the diffusion gradient directions are often represented by vectors in 3D q-space, e.g., $\vec{v}_i$ and $\vec{v}_j$ corresponding to direction *i* and *j*. Here we compute the angular difference $d_{i,j}$ by using the angle of two vectors:

$$d_{i,j} = \frac{\vec{v}_i \cdot \vec{v}_j}{\left|\vec{v}_i\right| \cdot \left|\vec{v}_j\right|} \tag{3.6}$$

It should be noted that this choice of Eq. (3.4) with $W_{i,j}$ incurs a heavy constraint to control the generated results. Here, if the data samples r_i and r_j are from different submanifolds in 3D q-space (in terms of angle location), it encourages the generated g_i and g_j to lie in different manifolds. Finally, our loss function is:

$$G = \arg\min\max L_{GAN} + \lambda L_{similarity} + \beta L_{correlation} \tag{3.7}$$

where λ and β are regularization parameters which control the importance of the spatial similarity and angular correlation, respectively.

4.3 Experiments and results

4.3.1 Datasets

Three *ex vivo* human hearts were obtained and processed in compliance with French legal and ethical guidelines (Yang, 2018). The investigations conformed to the principles outlined in the Declaration of Helsinki (Carlson, et al., 2004). They relate to infants who died after birth up to 14 months of life. After being medically declared dead, the body was kept in a morgue at a temperature of 4°C for more than 24 hours. This was to give time for the administration to confirm that there were no contraindications to autopsy and for the heirs of the deceased to authorize an autopsy. After completion of these legal formalities, an autopsy was performed on the corpse. The extracted hearts were fixed in formalin solution to ensure good preservation of cells and tissues and thus resemble as closely as possible the living structures. The average weight of the ventricular masses of these hearts ranged from 12 to 32 g, and the average thickness of interventricular septum at the equatorial level ranged from 3 to 8 mm.

The DTI datasets of these hearts were acquired in clinical conditions with Siemens 3T MRI Magnetom Verio. The imaging parameters are the following: TE = 74 ms, TR = 7900 ms,

FOV = 144 × 144 mm^2, slice thickness = 1.4 mm, in-plane resolution = 2 mm, slice spacing = 1.4 mm, slice duration = 123.2 ms, number of slices = 35, slice size = 104 × 104, diffusion sensitivity b = 700 s/mm^2, number of diffusion gradient directions = 192, 64 and 12. In each direction, MR scans were acquired 6 times for noise reduction.

4.3.2 Data pre-processing

To conduct the experiments, DTI datasets of the *ex vivo* human hearts were employed to imitate lower and higher angular-resolution DW images. Specificcally, from the DTI datasets, we used two data pre-processing steps to sample sets of diffusion gradient directions and generate the associated DW images for training the network. Given below are details about the two data pre-processing steps.

4.3.2.1 Gradient directions sampling

A necessary property for an optimal sampling scheme in DTI is that the directions should be uniformly distributed on a sphere and have no directional preference. To this end, from a set of the 192 directions that have been used for acquiring DTI datasets, we sampled a number of sets of directions using an direction sampling algorithm called IMOC (Cheng, et al., 2018). Each set consists of 2*n* directions that distributed uniformly on a sphere. Afterwards, from each set of 2n diirections, *n* directions were sampled using IMOC. Finally, a pair of direction sets were obtained, including *n* directions and *n* new directions, and both of them are distributed uniformly on a sphere. Finally, the DW images (from the original cardiac DTI datasets) corresponding to each pair of direction sets were regarded as a "sample" containing *n* input and *n* output images to train and test the proposed network.

4.3.2.2 Directional normalization

By definition, our network processes DW image set without using any directional information. Therefore, it is necessary to design a strategy that matches the diffusion gradient directions of input and output DW images. Previous studies have shown that several independent sets of directions uniformly spread on the sphere can be regarded as rigid rotations of the others (Caruyer, et al., 2013). Based on this, we first determine a fixed set of 2*n* directions as the reference for the directions of input and output DW images. For each set with different directions, we find the optimal rigid rotations with respect to the reference using a point matching algorithm. Then, the most uniform coverage possible is achieved by projecting the rotated sets of points onto the same reference sphere. In other words, each new set of directions can be aligned to a same set of directions. Then, each voxel in DW images associated with new 2*n* diffusion gradient directions was normalized into the fixed set of 2*n* directions using the equation below:

$$S_i^{new} = S_0 \exp(-b\vec{g}_i^T \underline{D} \vec{g}_i) \tag{3.8}$$

where $\underline{D}$ denotes the diffusion tensor computed from the un-normalized DW images for that voxel, $\vec{g}_i = (x_i, y_i, z_i)$ denotes the unit vector that designates the *i*th diffusion gradient direction, S_i^{new} is its normalized value and *b* is the diffusion weighting factor (Kingsley,

2006b). The use of this equation ensures that the diffusion tensor, b-value and S_0 are kept constant during the normalization.

4.3.3 Results

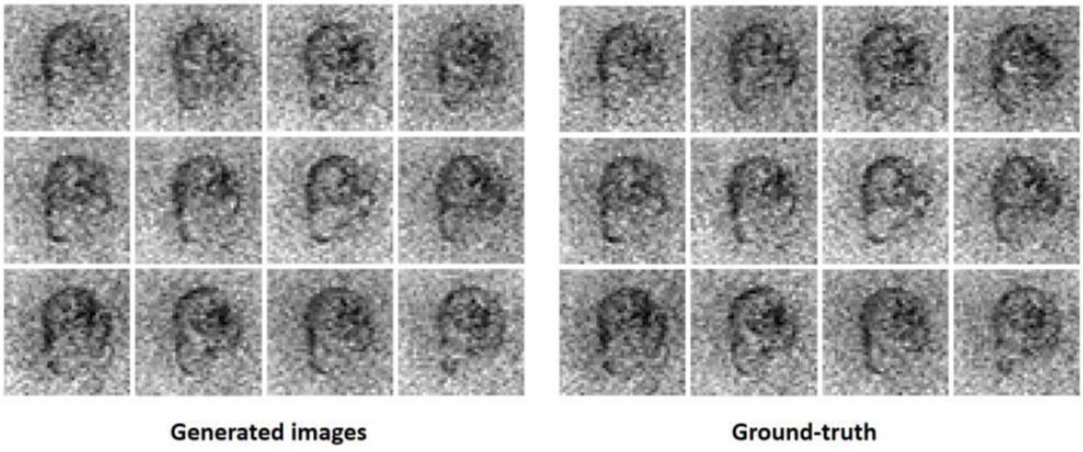

Fig. 4.3 An example of generated DW images using our method. Here the 12 DW images were generated from 6 given directions for 12 gradient directions.

Fig. 4.3 gives 12 DW images generated from 6 acquired DW images by the proposed network. The network was trained and tested by 600 and 400 samples, respectively. In the generated images, both the structural and detail information are preserved. This indicates that the generative model has well captured the underlying distribution of the input data, and thus generated realistic images. Furthermore, we also computed DT field and mapped it with three commonly used metrics computed from the three eigenvalues of the tensors: mean diffusivity (MD), axial diffusion (AD), and fractional anisotropy (FA) (Kumar, et al., 2009)., The quantitatively differences between the MD, AD, FA from the generated DT and ground-truth are reported in Table 4.1. The results were tested in terms of mean squared error (MSE), root Mean squared error (RMSE), peak signal to noise ratio (PNSR), and structural similarity index (SSIM) (Sara, et al., 2019). We found that the MD and AD are more similar to their corresponding ground-truth compared with FA. This may be because FA are more sensitive to anisotropic diffusion within a voxel. In other words, the generated data can describe the isotropic properties of water molecules more accurately than the anisotropic properties of water molecules.

Table 4.1 Differences between generated and ground-truth DW images in terms of FA, MD, and MD.

	MSE($\times 10^{-3}$)	RMSE	PSNR	SSIM
MD	0.2	0.01	36.11	0.97
AD	2.6	0.05	25.81	0.87
FA	3.8	0.06	24.21	0.76

Table 4.2 gives a quantitative comparison of diffusion properties and fiber orientations for original DW images in 24 directions. The network was trained and tested by 1000 and 500 samples, respectively. The final resulting DW images consists of the original 24 DW images and new 24 DW images (in 24 directions) generated by the proposed method. The images in 192 direction were used as ground-truth for comparison. The values in the first three columns

are the RMSE differences between the estimated DW images and ground-truth in terms of MD, AD, and FA, respectively. The values in the last column are the angle differences between the underlying critical fiber orientations of the estimated images and ground truth.

Table 4.2 RMSE of diffusion properties and fiber orientations for original DW images in 24 directions (top row) and the generated DW images in 48 directions (bottom row). FD represents the frobenius distances.

	MD (10^{-3})	AD (10^{-3})	FA	Tensor (FD)
Original	0.02	0.05	0.04	6.05
Obtained	0.01	0.04	0.03	4.58

An example of qualitative results in terms of main fiber orientations on one human heart is given in Fig. 4.4 with FA image as backgrounds. It can be seen in that the main fiber orientation from the our resulted DW images are much closer to the ground-truth, compared with the fiber orientation from original images. Note that the fiber orientations from the original DW images are very noisy, while the fiber orientations from the resulted DW images are locally well-aligned and regularized.

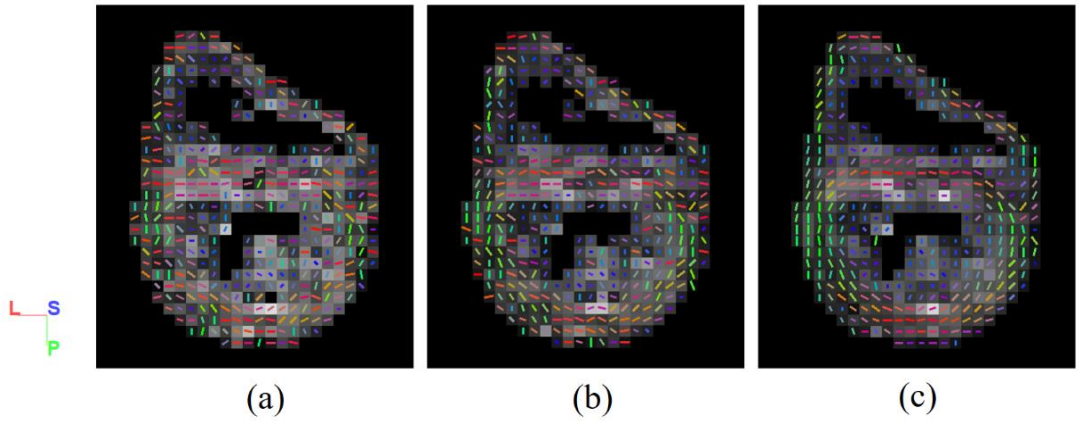

Fig. 4.4 Results of main fiber orientations. (a), (b) and (c) are the fiber orientations estimated from original DW images in 24 directions, higher angular resolution DW images in 48 directions generated by the proposed method, and the ground-truth DW images in 192 directions, respectively.

4.4 Conclusion

In this chapter, we have presented a GAN-based translation method for super angular resolution in cardiac dMRI data. DW images in additional gradient directions are estimated from given DW image sequence acquired in a limited number of gradient directions. The correlation between multiple DW images acquired in different gradient directions are exploited by introducing a supervised manifold regularized term which takes advantage of the information of diffusion gradient directions in 3D q-space. Experiments on cardiac DTI data showed that our method is capable of producing highly realistic DTI data and improving the accuracy of DTI measurements.

Chapter 5

Study of motion-induced signal loss in *in vivo* cardiac DTI

Contents

Abstract

Diffusion tensor imaging (DTI) provides a powerful means to non-invasively explore diffusion properties and fiber structures of biological tissues by measuring water diffusion in tissues. It has been widely used in studying the myocardial fiber structures and access abnormalities related to many cardiac diseases. However, DTI of the *in vivo* human heart remains an open challenge since cardiac and respiratory motions can induce large signal loss in diffusion weighted (DW) images. Advanced DTI sequences can reduce the effect of motion but usually at the cost of longer acquisition times and thus limits their application in practice. Current post-processing techniques, in particular the convolutional neural networks (CNNs), offer a solution to compensate for motion-induced signal loss in the acquired DW images. Unfortunately, the limited *in vivo* cardiac DTI data available in practice makes it difficult to train effective CNNs. This study aims to investigate motion-induced signal loss in *in vivo* cardiac DTI. Subsequently, a novel method was proposed to generate motion-induced DW images directly from existing cardiac DTI datasets to imitate *in vivo* cardiac DTI data. Finally, the proposed method was validated by training a new CNNs-based network using the generated motion-induced DW images. The results of this study showed that the proposed method can be used to generate realistic motion-induced signal loss and then add to the *ex vivo* DW images to imitate *in vivo* cardiac DTI data.

5.1 Introduction

Diffusion tensor imaging (DTI) provides an avenue to non-invasively study diffusion properties and fiber structures of biological tissues by measuring the water diffusion in tissues. In cardiac imaging, for example, the diffusion of water molecules is described by diffusion tensors estimated from a set of diffusion-weighted (DW) images associated with noncollinear diffusion gradient directions. The eigenvalues and eigenvectors of the diffusion tensor express, respectively, the magnitude and direction of water diffusion in a voxel, and thus indicate local architecture of fiber tracts passing through this voxel. The depiction of fiber architecture in the myocardium makes it able to investigate the microstructural changes related to many cardiac disorders (Mekkaoui, et al., 2013; Sosnovik, et al., 2014;Wu, et al., 2006).

DTI techniques are now widely used in *ex vivo* cardiac studies of both healthy (Frindel, et al., 2009) and diseased human heart (Le Bihan, et al., 2001). However, translating DTI to *in vivo* human heart remains an open challenge. This is because the cardiac or respiratory motions during imaging induce large signal loss in DW images, making it difficult to correctly measure the water diffusion and myocardial fiber structures. The conventional solution to this problem is to reduce the motion effects by designing advanced MR sequences for *in vivo* DTI (Edelman, et al., 1994; Stoeck, et al., 2014; Gamper, et al., 2007; Nguyen, et al., 2014). Despite the power of these imaging sequences, most of them require a long acquisition time to achieve a significant performance. Prolonging the acquisition time is often prohibitive in clinical routine because it compromises patient comfort and induces motion artifacts. This makes these sequences difficult to apply in practice.

An interesting alternative is to use post-preprocessing methods as an aid to compensate for signal loss caused by motions. As a first attempt, Wei et al. (Wei, et al., 2013) has investigated the effects of cardiac motion on the fiber architecture in beating human hearts, where the motion-induced signal loss was simulated by estimating the relationship between cardiac displacement and DW signals. Unfortunately, such estimation requires a special acquisition sequence, which can not be applied to other acquired datasets. In recent years, deep learning methods, especially, convolutional neural networks (CNNs) have become a method of choice for processing or analyzing medical images. To date, however, there are few CNNs applied to *in vivo* cardiac DTI. The main difficulties lie in the fact that CNNs have to be trained by a huge amount of images (Shen, et al., 2017), but in reality available *in vivo* cardiac DTI data is extremely scarce. This becomes a bottleneck for developing new CNNs to solve this problem.

The goal of this study is to to investigate motion-induced signal loss in *in vivo* cardiac DTI. Subsequently, a novel data fitting method was proposed to generate motion-induced signal loss directly from existing *in vivo* cardiac DTI datasets. The generated signal loss was then added to the *ex vivo* DW images to imitate *in vivo* cardiac DTI data currpted by motion artifacts. Finally, the generated images were validated by training a GAN-based network for signal loss compensation in *in vivo* cardiac DTI.

5.2 Materials and methods

5.2.1 Data acquisition

***Ex vivo* datasets**

One *ex vivo* human heart was obtained and processed in compliance with French legal and ethical guidelines. The investigations conformed to the principles outlined in the Declaration of Helsinki (Carlson, *et al*., 2004). More details about the acquisition process of these hearts were reported in (Yang et al., 2018). The DTI datasets were acquired in clinical conditions with Siemens 3T MRI Magnetom Verio. The imaging parameters are the following: TE = 74 ms, TR = 7900 ms, FOV = 144 × 144 mm^2, slice thickness = 1.4 mm, in-plane resolution = 2 mm, slice spacing = 1.4 mm, slice duration = 123.2 ms, number of slices = 35, slice size = 104 × 104, diffusion sensitivity b = 700 s/mm^2, number of diffusion gradient directions = 192, 64 and 12. In each direction, MRI scans were acquired 6 times for noise reduction.

***In vivo* datasets**

The *in vivo* human subjects used in this method are from six healthy volunteers. The cardiac DTI of these volunteers were performed with a 1.5 T clinical scanner (MAGNETOM Avanto, Siemens AG) at the Neuro-Cardiology Hospital of Lyon in France. All subjects gave informed consent to the institutional review board-approved study protocol prior to participation. The acquisition parameters are: TE/TR = 51/100 ms, b-value = 200 s/mm^2, spatial resolution = 2.63 × 2.63 × 6 mm^3, matrix size = 90 × 160, FOV = 420 × 236 mm^3, bandwidth = 1302 Hz, acceleration rate = 2, partial Fourier = 6/8. For each heart, 10 slices (slice thickness = 6 mm) were acquired in the short-axis view without any interslice gap. An individual mask was obtained to detect the myocardial region. The DW images were acquired in 10 different time points. For each time point, 13 DW images including 12 images associated with 12 diffusion gradient directions and one b_0 image (without diffusion weighting) were acquired. More information on data acquisition was detailed in (Wei, *et al*., 2013).

5.2.2 Observation of signal loss in *in vivo* DW images

As shown in Fig. 5.1, the signal losses caused by spatially variable in-scan motion are visible in DW images acquired in multiple diffusion gradient directions. According to our observations on several hearts, we found that the degree of signal attenuation due to cardiac motion is different in different gradient directions. For example, in Fig. 5.1, huge signal loss occurred in the dir7, dir8, and dir11, but minor signal loss occurred in the dir1, dir3, dir5, dir10, and dir12. Besides, we found that the similar degree of signal loss (e.g., huge singal loss) usually occurs randomly in different directions. For example, for volunteer 1, larger signal loss occurred in the 7th, 8th, and 11th directions, while for volunteer 2, it occurred in the 1st, 3rd, and 11th directions.

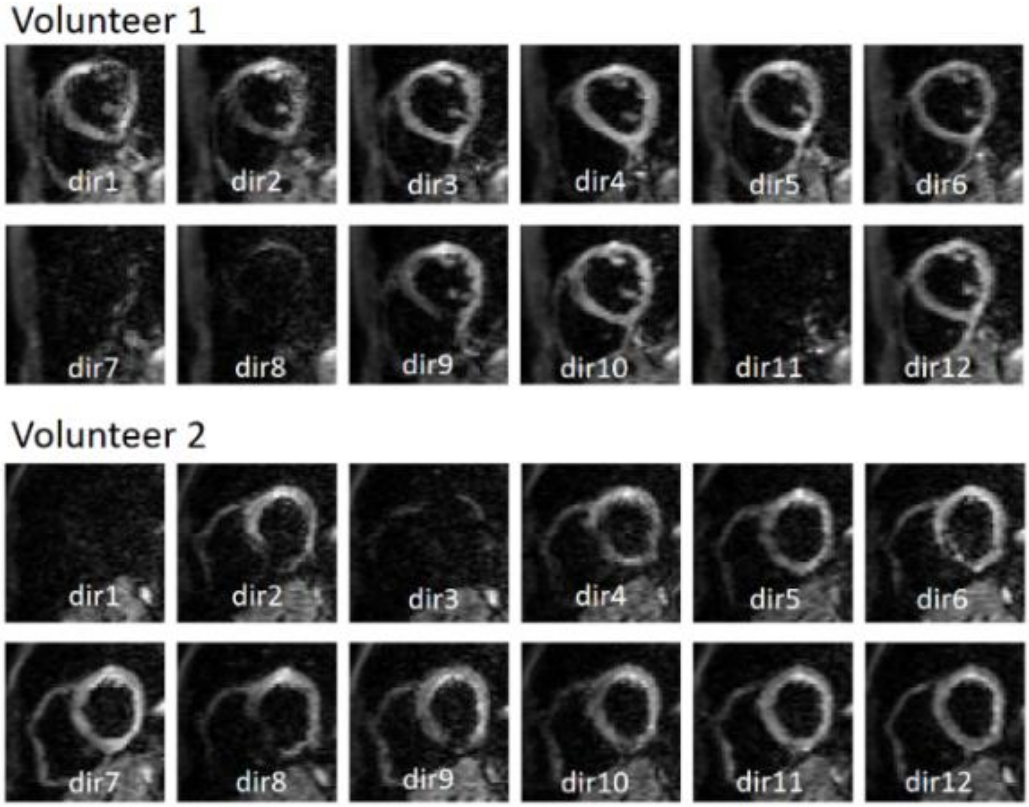

Fig. 5.1 In vivo DW images acquired with 12 diffusion gradient directions at single time point during diastole on two volunteer. "dir*N*" designates the image associated with *N*th diffusion gradient direction.

Based on the above observations, we defined three different categories of signal losses, named as minor, moderate and huge signal losses, as shown in Fig. 5.2. More specifically, given a set of *in vivo* DW images acquired in 12 diffusion gradient directions, the motion estimation algorithm (Pai *et al*., 2011) was used to generate an image without motion artifacts for each *in vivo* DW image. Then the mean difference between each *in vivo* DW image and its generated motion-free image is calculated. Afterwards, the mean difference values for the 12 DW images are normalized to 0-1 for measuring the degree of signal loss. Three different categories of signal losses were classified according to the normalized signal loss values. For example, the range of minor, moderate and huge signal losses are [0, 0.2], [0.2, 0.45], and [0.45, 1], respectively. As observed in Fig. 5.2, different degrees of signal loss we observe in the DW images can also be reflected by the data points in the chart below.

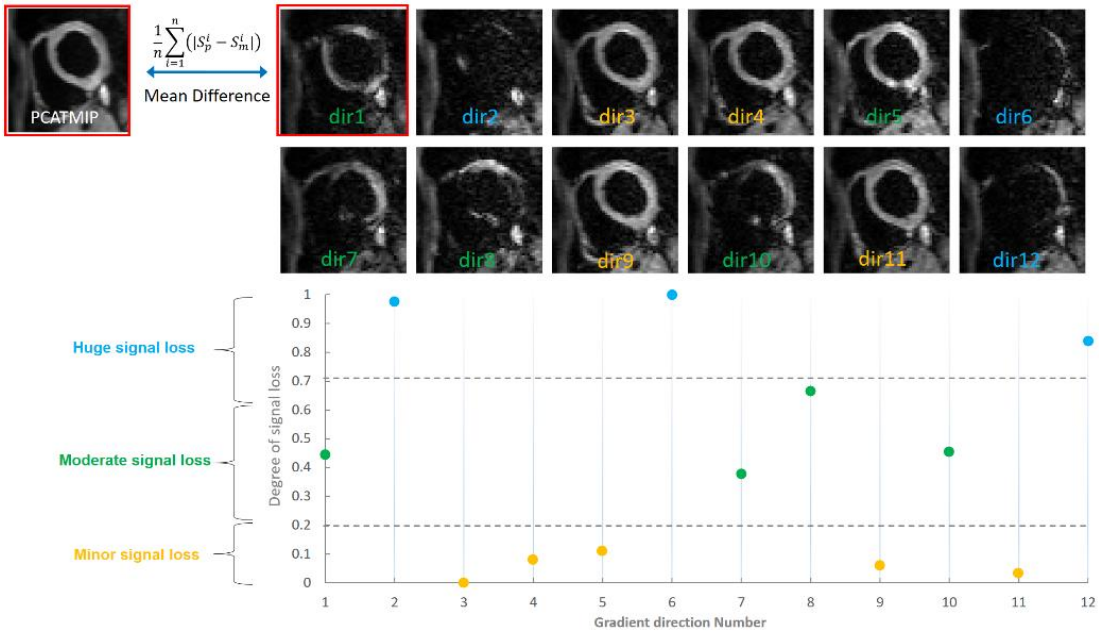

Fig. 5.2 Three categories of motion-induced signal-loss in *in vivo* cardiac DW images. Mean difference between each of the DW images (from dir1 to dir12) and its corresponding PCATMIP image is reported in the bottom chart, representing signal loss occurring in this DW image.

Table 5.1 reported the number of images for the three categories of signal loss-degrees in 12 diffusion gradients directions for 10 time points. An interesting phenomenon is that the number of images with both huge and minor signal losses is usually larger than the number of images with moderate signal losses. This phenomenon can also be found in Table 5.2, in which the numbers for all the six *in vivo* human heart are reported. These results provide us with a mean number of DW images associated with 12 diffusion gradient directions for simulating different degree of signal losses.

Table 5.1 Number of DW images for three categories of signal-loss on one volunteer. The numbers in each row were counted in DW images acquired in 12 directions at the same time point.

TD(ms)	Number of DW images		
	Minor	Moderate	Huge
850	7	1	4
860	9	0	3
870	9	0	3
880	8	1	3
890	5	3	4
900	6	0	6
910	5	3	4
920	9	0	3
930	4	4	4
940	4	1	7
mean (≈)	7	1	4

Table 5.2 Number of DW images for three degrees of signal-loss over six human hearts/subjects.

	Minor	Moderate	Huge
Subject 1	7	1	4
Subject 2	6	2	4
Subject 3	4	3	4
Subject 4	4	4	4
Subject 5	3	4	5
Subject 6	5	3	4
mean	5	3	4

5.2.3 I*n vivo* signal loss generation

Based on above observations and statistics, we proposed a new method to generate the realistic signal loss from *in vivo* cardiac DTI datasets. As we already knew, DTI signal attenuation caused by the diffusion of water molecules is expressed by:

$$S_i = S_0 \exp(-b\vec{g}_i^T \underline{D}\vec{g}_i) \tag{5.1}$$

where S_i is the diffusion weighted signal acquired at the *i*th diffusion direction $\vec{g}_i = (x_i, y_i, z_i)$, S_0 is the signal at the same voxel without the additional diffusion gradients, $\vec{g}_i^T \underline{D}\vec{g}_i$ denotes the diffusion coefficient related to molecular mobility, and $\underline{D}$ denotes the 3×3 diffusion tensor matrix. In the presence of motion, the model of DTI can be considered according to the principle of *in vivo* DTI sequence such as DENSE (Aletras et al., 1999), that is, the diffusion gradients not only encode molecular motion (diffusion) but also generate an additional signal attenuation term:

$$S_i^m = S_0 \exp(-bADC_i)\exp(-i\varphi_i) \tag{5.2}$$

where φ_i represents the phase shift and depends on the proton gyromagnetic ratio, the displacement component of the spins, and the gradient pulse in the ith diffusion gradient direction. Unlike the diffusion attenuation term, $S_0 \exp(-bADC_i)$, the attenuation, $\exp(-i\varphi_i)$, resulting from spatially variable tissue motion, is unpredictable and has several orders of magnitude larger than diffusion. The precession of spins in an external magnetic field as well as the simultaneous relaxation and diffusion effects are a very complex phenomenon. Wei *et al*. (Wei, et al., 2013) simplified the description of this phenomenon by modeling the relationship between the relative longitudinal displacement amplitude of cardiac motion and the acquired motion-induced signal as an exponential function:

$$y = A\exp(Bx) + c \tag{5.3}$$

where y represents the normalized motion-induced singal S_i^m inside a local region, x represents the motion of this region, A and B represent the function parameter that can be estimated with data of x and y via regression estimation, and c represents the remaining

difference between the data and the model. One limitation of this method is that the corresponding motion displacement variables *x* needs to be obtained before estimation.

Our model for estimating the motion-induced signal loss is also based on Eq. (5.3), but we redefined the meaning of variables in this equation. We assumed that the additional signal attenuation term $\exp(-i\varphi_i)$ caused by motion in Eq. (5.2) can be modeled as *y* in Eq. (5.3), and the *x* can be defined as a reflection of the degree of signal loss, such as the data points in the below chart of Fig. 5.2, in other words, we defined $S_i^m = S_0 \exp(-bADC_i)y$. For each degree value *x*, there is a signal loss term *y* corresponding to it. Then, Eq. (5.3) can be solved by fitting *y* as the exponential function of *x* using regression estimation. Fig. 5.3 plots the results of regression estimation, where the fitting parameters were RMAE = 0.038 and R-square = 0.983. We found that the fitted results roughly follow the behavior of monotonically decreasing exponential function. This fitted model provides a mean value to generate realistic motion-induced signal loss with an arbitrary value of *x*.

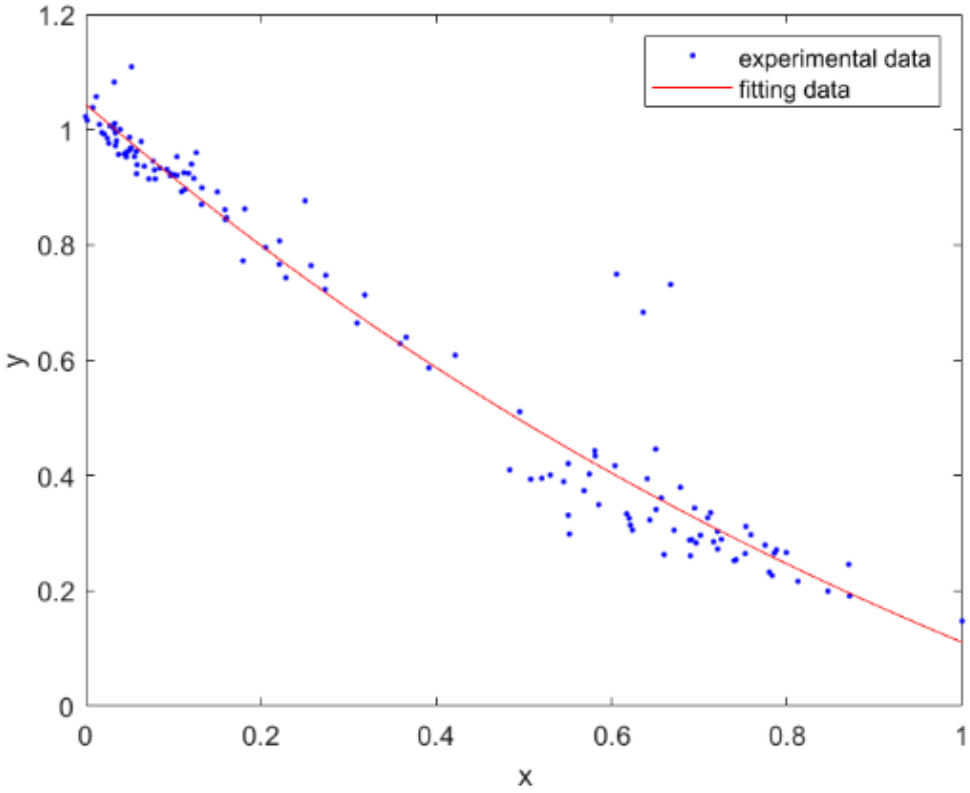

Fig. 5.3 Motion-induced attenuation term as a function of signal loss degree over the six human hearts. For each blue point (the experimental data), the *x* value denotes the degree of signal loss and the *y* value the mean signal loss computed from all pixels averaged inside each of image on the myocardium area.

In section 5.2.2, we observed that the DW images can be divided into three categories with minor, moderate, and huge signal losses. For each category, we counted the mean number of DW images. These findings guided us to be able to determine the value x and the number of images for the different degrees of signal losses. Besides, we also simulated the emergence pattern of signal loss for each DW image. Specifically, for each image, its motion effects were

first reduced using PCATMIP algorithm and the resulting image is denoted as PCATMIP image. Then the differences between in vivo DW images and their PCATMIP images can be computed, reflecting the signal losses caused by motion, as shown in Fig. 5.4, where the difference images lateled with {dir1, dir4, dir6, dir9, dir10}, {dir3, dir5, dir12}, and {dir2, dir7, dir8, dir11} designate minor, moderate, and huge signal losses, respectively. Afterwards, for each degree of signal losses, we specify an intensity range in the difference images so that the number of primary signal-loss pixels for each degree can be counted. Finally, for all DW images belonging to a certain degree, the number of primary pixels within their intensity range on each image was counted, and then a mean number from all images for this degree can be obtained.

Table 5.4 gives the specified range and the counted results for the DW images shown in Fig. 5.4. Finally, realistic signal losses due to motion can be generated by setting the values of x according to the results of the above estimations and statistics.

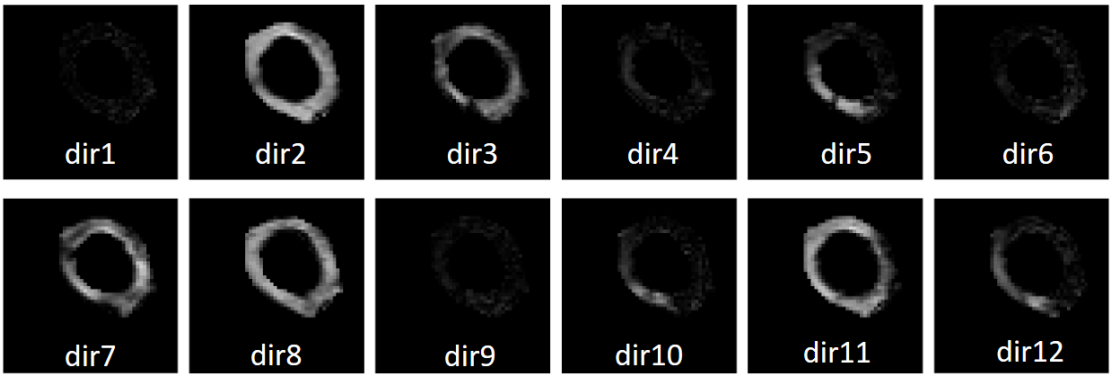

Fig. 5.4 Differences between *in vivo* DW images associated with 12 diffusion gradient directions and their PCATMIP images. The label superimposed on the images designate the number of associated diffusion gradient direction, e.g., dir1 designates the 1st diffusion gradient direction.

Table 5.3 Ranges and primary pixel numbers for the three degrees of signal losses. For each degree, the number of primary pixels designates the mean number of pixels within its intensity range for one DW image.

Signal-loss degree	Intensity range	Number of primary pixels
Minor	[0, 0.3]	359
Partial	[0.3, 0.6]	193
Huge	[0.6, 1]	326

5.2.4 *Generate realistic motion-induced DW images*

After considering the above statistics, the realistic motion-induced signal loss can be generated for each pixel of the DW images in 12 directions. Subsequently, we added these generated signal losses to both simulated and *ex vivo* DW images for imitating *in vivo* DTI data with motion effects. The simulated images were designed such that they are very similar to a slice of real human heart DTI volumes. Visually, each image mainly consists of two regions: the myocardium region (region of interest—ROI) and the background. In the background, isotropic diffusion tensors were designed with three equal eigenvalues 0.75, and

in the myocardium region, the diffusion tensors were generated by setting the eigenvalues of each tensor to (2, 1, 1) × 10^{-3}mm^2/s. The principal eigenvectors of these tensors make a fixed angle (45°) with the slice plane, and their projections on the plane take the direction of the tangents of the concentric circles centered on the center of the left ventricle (LV). Fig. 5.5 illustrates the process of generating a simulated cardiac image from a slice of a real human heart DTI volume.

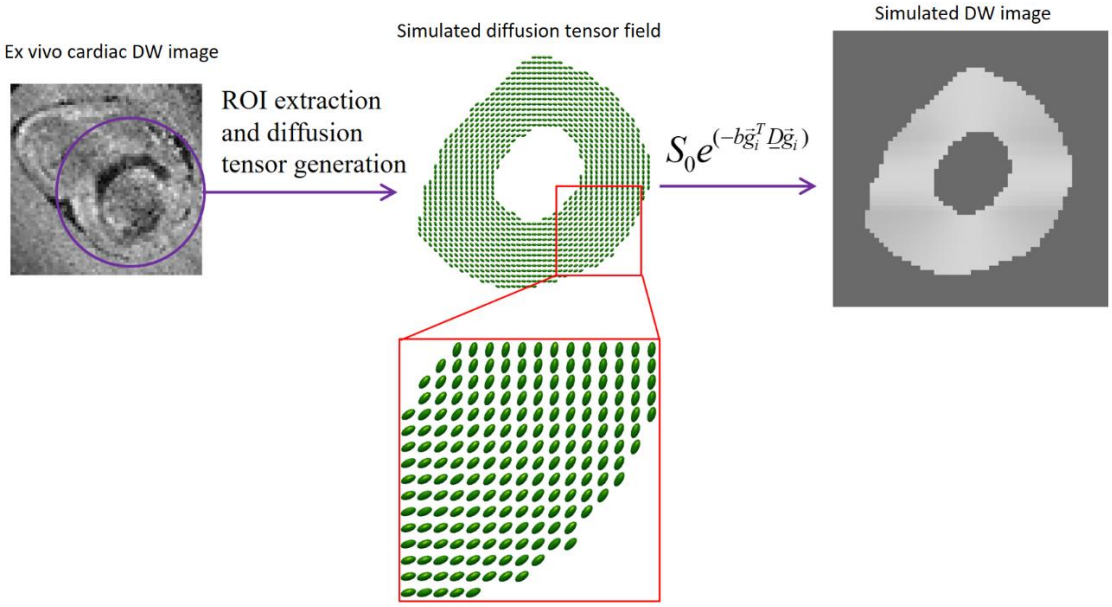

Fig. 5.5 Simulation of cardiac DW images from a slice of an *ex vivo* human heart DTI volumes.

We also added the generated signal losses directly to the *ex vivo* cardiac DW images to imitate the *in vivo* cardiac DTI data, since the ex vivo data is the closest to the in vivo data when disregarding the motion effects. For the *ex vivo* datasets, we sampled sever subsets of 12 diffusion gradient directions from the original 192 directions using an direction sampling algorithm called IMOC (Cheng, et al., 2018), each set of 12 can be used to sample a set of 12 DW images from the 192 *ex vivo* DW images. Subsequently, the generated signal losses can be added to each pixel of the sampled *ex vivo* DW image.

5.3 Results

Fig. 5.6 presented the results of adding the generated motion-induced signal losses and/or Rician noise to the simulated DW images associated with 12 diffusion gradient directions. It can be seen that the motion-induced signal losses are generated according to the previous statistical results, i.e., the mean numbers of images in the 12 diffusion gradient directions were 5, 3, and 4, respectively for minor, moderate and huge signal losses. Besides, the number of primary pixels for minor or huge signal losses is visually larger than the number of pixels with moderate losses. For minor signal-loss images, the mean proportion of primary pixels per image is 86%, and this proportion is 51% and 6% for moderate and huge signal losses, respectively. By adding signal loss to the simulated image in row (a), three degrees of signal loss can be seen clearly, which effectively imitates highly realistic *in vivo* motion-induced DW images shown as in Fig. 5.2.

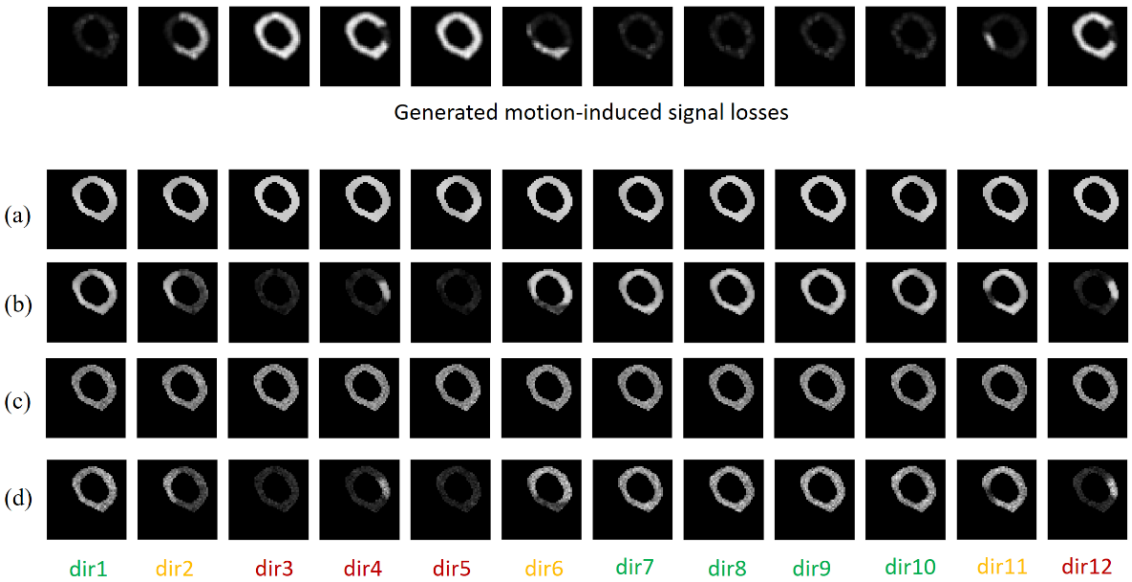

Fig. 5.6 Imitation of *in vivo* cardiac DW images for 12 diffusion gradient directions. Row (a) represents simulated DW images without noise and artifacts. Row (b), (c), and (d) are obtained by adding the motion-induced signal losses, Rician noise, and both signal losses and Rician noise to the images in row (a).

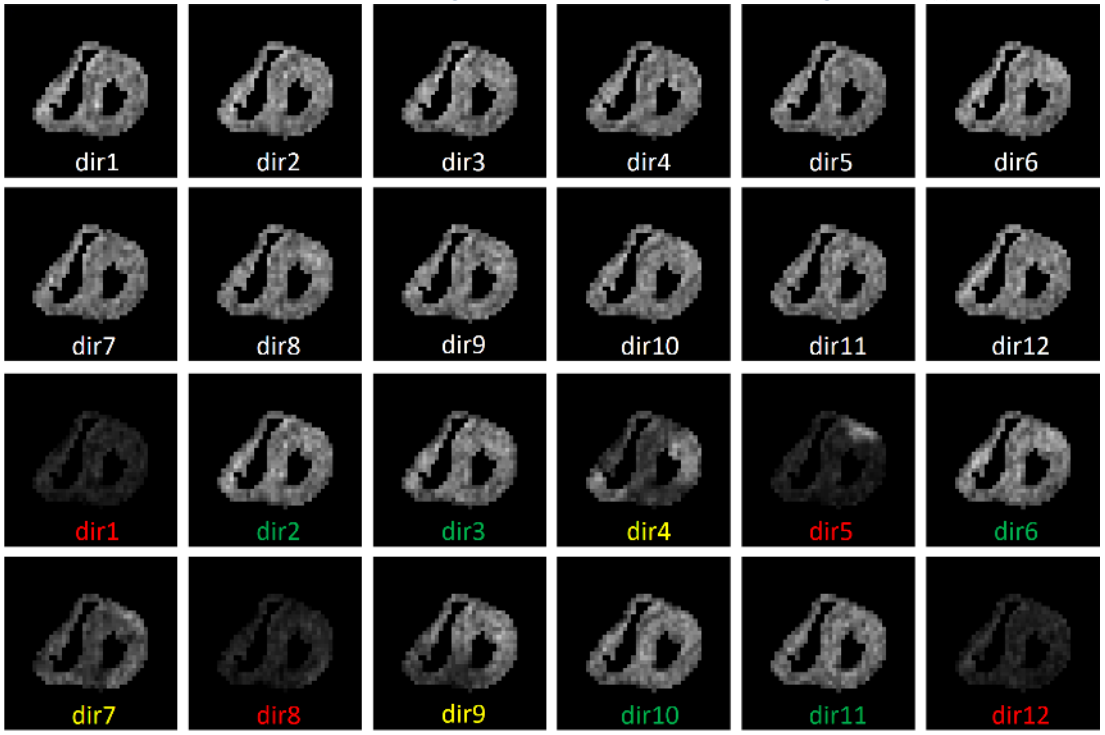

Fig. 5.7 E*x vivo* DW images added with motion-induced signal losses. Top two rows are the *ex vivo* DW images associated with 12 diffusion gradient directions, and the bottom two rows are the motion-induced DW images. The direction labels superimposed on the image with green, yellow and red colors designate minor, moderate and huge signal losses, respectively.

Fig. 5.7 gives the results of adding the motion-induced signal losses directly to the DW images of an *ex vivo* human heart. The DW images with minor, moderate and huge signal

losses designated by green, yellow and red labels, respectively, are clearly observed. These added inhomogeneous signal losses make the *ex vivo* DW images visually and distributively closer to realistic *in vivo* DW images. Besides, we also used the generated motion-induced DW images to train the GAN-based network proposed in the previous chapter to compensate for motion-induced signal losses. Table 5.4 reports mean squared error (MSE) of diffusion tensors (in terms of frobenius distance), mean diffusivity (MD), and fractional anisotropy (FA) for the motion-induced and signal loss-compensated DW images. The results show that the GAN-based signal loss compensation improved the estimation accuracy of diffusion tensors, FA, and MD.

Table 5.4 RMSE of diffusion properties for original motion-induced DW images (Top row) and the GAN-based motion-compensated DW images (Bottom row).

	Diffusion tensors	FA	MD
Motion-induced	0.07	0.22	0.25
Signal loss-compensated	0.02	0.03	0.07

5.4 Conclusion

In this chapter, we focused on investigating motion-induced signal loss in *in vivo* cardiac DTI. After observation and statistics, we presented a novel method capable of generating motion-induced signal losses directly from *in vivo* cardiac DTI datasets. This method was achieved by counting the number of images for different degree of signal losses in *in vivo* cardiac DTI data and the proportions of pixels in a DW image exhibiting certain degree of signal losses. After adding the generated signal losses to both simulated and *ex vivo* DW images, a large amount of motion-induced DW images were obtained, which can visually and distributively imitate the *in vivo* DW images corrupted by motion. The generated motion-induced DW images were further used to train a GAN-based super angular-resolution network proposed in the previous chapter for motion-compensation in *in vivo* cardiac DTI. The results demonstrated that the GAN-based network presents effective performance of motion compensation. This offers a novel way to generate realistic motion-induced DW images that can be used for developing advanced CNN-based post-processing methods for *in vivo* cardiac DTI.

Chapter 6

Compensation of motion-induced signal loss in *in vivo* DTI of the human heart using GAN-based super angular resolution

Contents

Abstract

Diffusion tensor imaging (DTI) has become an indispensable modality for non-invasively characterizing fiber architectures of the human heart in recent years. It has been widely used for characterizing the myocardial fiber architectures of the heart and accessing the histological arrangement of the myocardial fibers in tissues. However, clinical applications of *in vivo* cardiac DTI remain extremely limited due to large signal loss caused by motions during imaging. In this chapter, we propose a novel method that can compensate for motion-induced signal loss in *in vivo* cardiac DTI. This is achieved by training a deep learning network to produce higher angular-resolution diffusion weighted (DW) images from lower angular-resolution DW images acquired in practice. The proposed network is constructed based on a generative adversarial network (GAN) combined with L2 norm to produce realistic cardiac DW images. Besides, angular consistency between DW images with the same angular resolution is employed as an additional regularization term to improve the performance of signal loss compensation. In order to train and test the proposed network, *in vivo* cardiac DTI datasets are used in combination with *ex vivo* cardiac DTI datasets to imitate plenty of motion-induced DW images with lower angular resolution and their ground-truth DW images with higher angular resolution. Experimental results showed that the produced higher angular-resolution images enable significant improvement of the estimation accuracy of diffusion tensors, fractional anisotropy (FA) and mean diffusivity (MD). Moreover, fiber orientations are better characterized with the compensation of signal losses. The proposed method thus provides an interesting post-processing way to compensate for motion-induced signal loss in *in vivo* cardiac DTI.

6.1 Introduction

Diffusion tensor imaging (DTI) noninvasively probes diffusion properties and fiber structures of human tissues by measuring the diffusion of water molecules in them. In cardiac imaging, for example, water diffusion process in each voxel can be described by a diffusion tensor estimated from a set of diffusion-weighted (DW) signals acquired in various diffusion gradient directions (Kingsley, 2006a; Kingsley, 2006b; Kingsley, 2006c). The major eigenvector corresponding to the largest eigenvalue of the diffusion tensor defines the direction of water diffusion that is least restricted in the muscle fiber tracts and thus reflects the principal orientation of the fibers inside that voxel (Taylor, et al., 2004). This makes it possible to depict the fiber architecture of human myocardium and infer the microstructural changes related to cardiovascular diseases, e.g., myocardial infarction (Wu, et al., 2006), hypertrophic cardiomyopathy (Tseng, et al., 2006), and ischemic heart diseases (Mekkaoui, et al., 2017), etc.

Cardiac DTI has been widely used to characterize the myocardial fiber architectures of the heart and evaluate the histological arrangement of the myocardial fibers (Edelman, et al., 1994; Wu, et al., 2009; Dou, et al., 2002; Basser, 2011; Wu, et al., 2007; Sosnovik, et al., 2009; Sosnovik, et al., 2014). These studies have shown that fibers in the normal myocardium are orderly arranged in a crossing helical pattern. In patients, DTI has also been considered a useful tool to describe the disease-induced fiber disarray (Wu, et al., 2006; Wu, et al., 2009). The fiber integrity after myocardial infarction has often been quantified using fractional anisotropy (FA) and mean diffusivity (MD) (Wu, et al., 2007; Basser, 2011; Sosnovik, et al., 2014). All these findings suggest the potential of DTI for early diagnosis and treatment of cardiovascular diseases.

However, the applications of *in vivo* DTI of the human heart remain limited in clinical practice. The main reason lies in the fact that DTI essentially measures the microscopic motion of water molecules and thus is very sensitive to subject motion. Respiratory and heartbeat motions that occur during imaging induce large signal loss in DW images and thereby degrade substantially the accuracy of estimated diffusion tensors, which potentially leads to incorrect clinical diagnoses.

To cope with this problem, researchers have made efforts to design new DTI sequences that can reduce the effects caused by motion during imaging. For example, Edelman *et al.* (Edelman, et al., 1994) proposed a stimulated echo acquisition mode (STEAM) sequence where diffusion signals were acquired based on the spatial position and shape of the myocardium over the entire heartbeat cycle so that the effects of heartbeat motion were minimized. Tseng *et al.* (Tseng, et al., 1999) exploited an additional strain measurement to counteract the effects of myocardial strain produced during STEAM imaging. Gamper *et al.* (Gamper, et al., 2007) designed a spin echo (SE) based sequence to decrease the sensitivity of diffusion-encoding to cardiac motion while maintaining a high signal-to-noise ratio (SNR). Besides, there exist a variety of other sequences that can be seen as extensions of STEAM or SE and have improved performance of *in vivo* DTI of the human heart (Reese, et al., 1995; Tseng, et al., 1999; Nielles-Vallespin, et al., 2013; Stoeck, et al., 2016; Nguyen, et al., 2014). Despite the power of these imaging sequences, most of them are time consuming and have to

be used in combination with dedicated strategies (e.g., navigator gating and breath holding), resulting in a relatively long data acquisition time. Unfortunately, prolonging the acquisition time is often prohibitive in clinical routine because it compromises patient comfort and increases the risk of inducing motion artifacts. Therefore, for most sequences, the reduction of motion effects often comes at the expense of acquiring a limited amount of data or else the SNR is relatively poor. This makes these sequences remain challenging to use in practice.

In recent years, researchers began to develop post-processing methods capable of recovering motion-induced signal loss in clinically acquired DTI data. In 2011, Pai *et al*. (Pai, et al., 2011) introduced a combined principal component analysis and temporal maximum intensity projection (PCATMIP) method to restore motion-induced signal losses in porcine liver DTI scans. Later, Wei et al. (Wei, et al., 2014) proposed a wavelet-based image fusion and principal component analysis (WIFPCA) method to retrieve myocardial fiber structures in free-breathing DTI acquisitions of the human heart. Recently, deep learning (DL) especially the convolutional neural network (CNN), which has become the method of choice for processing and analyzing medical images (Shen, et al., 2017; Dong, et al., 2015), has been applied to compensation of signal loss in *in vivo* cardiac DW images (Deng, et al., 2020). These studies demonstrated that it is feasible to use clinically available acquisitions combined with dedicated post-processing methods to achieve signal loss compensation in *in vivo* cardiac DTI. A common idea used in most of the studies is to repeatedly acquire DW images at multiple time points, and then a post-processing algorithm is used to recover motion-induced signal loss by exploiting the diffusion information from multiple acquisitions over time. However, up to now, there exists little work on the use of information from images acquired in various diffusion gradient directions to deal with signal loss in *in vivo* DW images.

In fact, in addition to collecting images at multiple time points, it is also desirable in many cases to acquire DW images in more diffusion gradient directions. Theoretically, a diffusion tensor can be mathematically calculated from six DW signals acquired in six diffusion gradient directions (Kingsley, 2006a). In practice, however, the number of diffusion gradient directions, which defines the angular resolution of DTI data, has an important influence on the estimation of diffusion tensors (Tournier, et al., 2013; Jones, 2004; Zhan, et al., 2013). Many researchers have concluded that a higher angular resolution (e.g., more than 30 diffusion gradient directions) is required to achieve more accurate diffusion tensor estimation (Jones, 2004). Moreover, some DTI studies of both human brain (Zhan, et al., 2013) and heart (He, et al., 2021) have shown that when the acquisition time is sufficient, the accuracy of diffusion tensor increases as the angular resolution. Seen in this light, increasing the angular resolution in our case would be an interesting alternative to compensate for motion-induced signal loss in estimating diffusion tensors. For the time being, available angular resolution in *in vivo* cardiac DTI is quite limited (typically no more than 12 diffusion gradient directions) (McGill, et al., 2012; Scott, et al., 2015; Stoeck, et al., 2016). This motivates the requirement for developing post-processing methods to improve the angular resolution of DTI data.

In this chapter, we propose a novel super angular-resolution method to compensate for motion-induced signal loss in *in vivo* cardiac DTI. The proposed method consists of training a DL algorithm that can produce higher angular-resolution DW images from lower angular-resolution DW images corrupted by motion. Instead of using traditional CNN-based models

(Dong, et al., 2015) that only minimize Euclidean distance between produced and ground-truth images, we construct a generative adversarial network (GAN) (Goodfellow, et al., 2014) where two competing sub-networks can be trained iteratively to minimize errors at a higher level, e.g., urges the produced images to be perceptually similar to the ground-truth images with higher angular resolution. Meanwhile, the similarity between the produced image and its corresponding ground truth in the pixel space is guaranteed by combining adversarial learning with L2 norm. Moreover, in order to enhance the angular correlation of produced DW images, we incorporate an additional regularized term into the final objective function to narrow the angular inconsistency not only between produced images and ground-truth images, but also between image subsets with the same angular resolution.

The rest of this chapter is structured as follows. In section 6.2, we describe the details of the proposed method. The data acquisition and processing, experimental setup, evaluation metrics and results are presented in section 6.3. Section **Error! Reference source not found.** p rovides the conclusions and discussions.

6.2 Methods

6.2.1 Problem formulation

The goal of the proposed method is to train a DL network that can increase the angular resolution of DW images for signal loss compensation in *in vivo* cardiac DTI. More specifically, given a set of lower angular-resolution *in vivo* DW images L: $\{l_1, l_2, \dots, l_n\}$ that have been acquired in n diffusion gradient directions $\{d_1, d_2, \dots, d_n\}$ and a set of higher angular-resolution DW images H: $\{h_1, h_2, \dots, h_n, h_{n+1}, \dots, h_m\}$ associated with m diffusion gradient directions $\{d_1, d_2, \dots, d_n, d_{n+1}, \dots, d_m\}$, we assume that the directions for L is a subset of directions for H. For clarity but without loss of generality, we set m=2n in our study. With the DL network, we try to learn the mapping G from L to H. Our final goal is to recover from L a set of images $G(L)$: $\{g_1, g_2, \dots, g_m\}$ that are as similar as possible to the ground-truth higher-resolution images H.

6.2.2 GAN-based super angular resolution

In the field of deep learning, a natural idea to learn the mapping G is to use a CNN-based super-resolution network (Dong, et al., 2015). However, the traditional CNN-based model that only minimizes the Euclidean distance between output and ground truth tends to neglect the information between neighboring pixels and thus causes blurring in the output image. Therefore, we adopted a generative adversarial network (GAN) (Goodfellow, et al., 2014) that utilizes an adversarial learning strategy to overcome the blurring problem and produce DW images that are perceptually similar to real DW images.

Fig. 6.1 illustrates an overview of the proposed GAN-based super angular-resolution network. It contains two sub-networks that are trained to compete with each other, a generator G and a discriminator D. The generator G is constructed to take existing lower angular-resolution DW images L as input and output higher angular-resolution DW images $G(L)$. D is essentially a discriminative model that takes both L and a set of higher angular-resolution DW

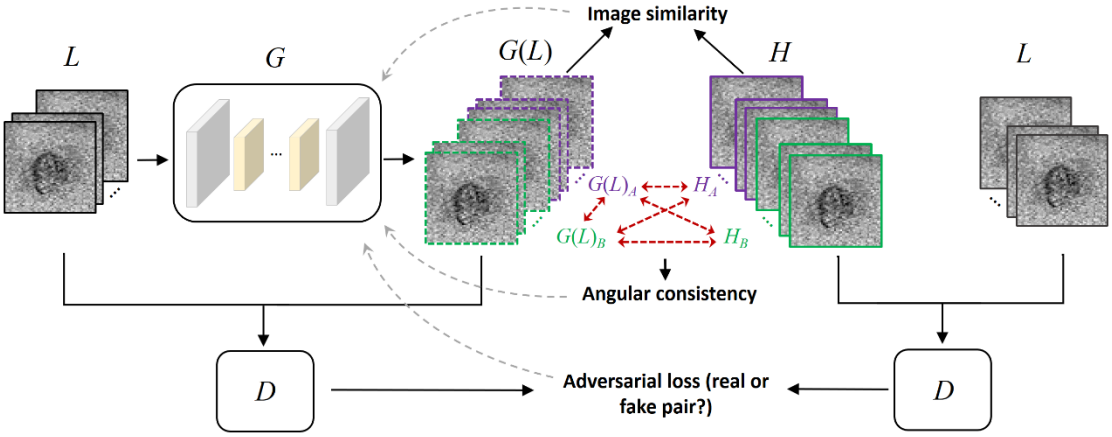

Fig. 6.1 Schematic illustration of the proposed GAN-based super angular-resolution network. The generator *G* learns the mapping from lower angular-resolution DW images *L* to higher angular-resolution DW images *H* based on three main constraints: image similarity, angular consistency, and adversarial learning strategy. The gray dashed arrows denote the back propagation to adjust the network parameters for updating *G*, and the red two-way solid arrows represent the angular consistency between DW image sets.

images as input, where the higher angular-resolution images can be either $G(L)$ or ground-truth images H. It tries to determine whether the input is real pair $\{L, H\}$ or fake pair $\{L, G(L)\}$. During training the D, the network parameters are updated by maximizing the following loss:

$$L_{GAN}(D) = E_{L,H}[\log D(L,H)] + E_L[\log(1 - D(L,G(L)))] \tag{6.1}$$

where the D minimizes $D(L,G(L))$ while maximizing $D(L,H)$ in order to improve its ability to classify real pair $\{L, H\}$ and fake pair $\{L, G(L)\}$. During training G, the network parameters are updated by minimizing:

$$L_{GAN}(G) = E_L[\log(1 - D(L,G(L)))] \tag{6.2}$$

where D is served as an additional network that has been updated using Eq.(6.1), then G minimizes log(1-$D(L,G(L))$) to generate images more difficult to be classified by D. In general, Eqs. (6.1) and (6.2) are integrated into a min-max function expressed as:

$$L_{GAN}(G,D) = \min_G \max_D E_{L,H}[\log D(L,H)] + E_L[\log(1 - D(L,G(L)))] \tag{6.3}$$

where G attempts to minimize this loss function in opposition to D that tries to maximize it.

The adversarial learning strategy in Eq. (6.3) guides the network to produce DW images $G(L)$ perceptually similar to real images H, but it cannot guarantee the similarity between produced and ground-truth images in pixel space. Many previous studies have found that combining the adversarial loss with a traditional loss (e.g, L1 or L2 norm) is beneficial for improving the performance of image generation/translation (Isola, et al., 2017; Pathak, et al., 2016). Therefore, in our method, we combine Eq. (6.3) with L2 norm to improve the similarity between the produced and ground-truth DW images. The L2 norm L_{sim} here can be expressed as:

$$L_{sim} = E_{L,H}\left\|H - G(L)\right\|^2 \tag{6.4}$$

where $\|\bullet\|^2$ denotes the Euclidean distance function. For the DW images associated with multiple diffusion gradient directions, Eq. (6.4) can be equivalent to minimizing $\frac{1}{2n}\sum_{i=1}^{2n}\|h_i - g_i\|^2$.

6.2.3 Regularized learning with angular consistency

As mentioned before, the ground-truth DW images H consist of 2*n* DW images associated with diffusion gradient directions $\{d_1, d_2, \dots, d_{2n}\}$. A direction sampling algorithm (Cheng, et al., 2017) can be used to split the 2*n* directions into two subsets and each subset contains *n* directions uniformly distributed on a sphere. Then we can obtain two subsets of images H_A: $\{h_1, h_2, \dots, h_n\}$ and H_B: $\{h_{n+1}, \dots, h_{2n}\}$ according to the two subsets of directions. Similarly, we specify that the network's outputs $G(L)$ also consist of two subsets $G(L)_A$ and $G(L)_B$ which have the same angular resolution as their targets H_A and H_B, respectively. To further improve the performance of the proposed network for producing realistic DW images, we propose to incorporate an additional regularization term into the final objective function to enforce the angular correlation between these image sets. The underlying assumption behind this idea is that diffusion tensors estimated from the image sets with the same angular resolution should be consistent with each other, although their diffusion gradient directions are different.

Specifically, we denote by $\underline{\mathrm{M}}(G(L))$ and $\underline{\mathrm{M}}(H)$ two diffusion tensor maps estimated from the produced images $G(L)$ and their ground-truth images H, respectively, then the angular correlation between the two image sets can be enhanced by minimizing a consistency loss defined as:

$$L_{\varphi} = \frac{1}{N}\sum_{s=1}^{N}\left\|\underline{\mathrm{M}}(H)_s - \underline{\mathrm{M}}(G(L))_s\right\|_F^2 \tag{6.5}$$

where N is the total number of diffusion tensors in each diffusion tensor map and s a spatial location of the tensor map. $\underline{\mathrm{M}}(H)_s$ and $\underline{\mathrm{M}}(G(L))_s$ denote sth 3×3 diffusion tensor matrices in $\underline{\mathrm{M}}(G(L))$ and $\underline{\mathrm{M}}(H)$, respectively, and $\|\bullet\|_F$ denotes the Frobenius norm. Besides, in order to enhance the consistency between image subsets with the same angular resolution, a similar idea is to minimize the following function:

$$L_{\eta} = \frac{1}{N}\left(\sum_{s=1}^{N}\left\|\underline{\mathrm{M}}(H_A)_s - \underline{\mathrm{M}}(G(L)_A)_s\right\|_F^2 + \left\|\underline{\mathrm{M}}(H_B)_s - \underline{\mathrm{M}}(G(L)_B)_s\right\|_F^2\right) \tag{6.6}$$

where $\underline{\mathrm{M}}(H_A)_s$, $\underline{\mathrm{M}}(G(L)_A)_s$, $\underline{\mathrm{M}}(H_B)_s$, and $\underline{\mathrm{M}}(G(L)_B)_s$ are sth 3×3 diffusion tensor matrices in the diffusion tensor maps estimated from the image subsets H_A, $G(L)_A$, H_B, and $G(L)_B$, respectively. Eq. (6.5) intends to reduce the angular difference between the higher angular-resolution DW images $G(L)$ and H, and Eq. (6.6) encourages the produced image subsets $G(L)_A$ and $G(L)_B$ are consistent with their targets H_A and H_B, respectively.

In theory, minimizing Eq. (6.5) combined with Eq. (6.6) is sufficient to ensure angular consistency between produced and ground-truth images. However, it is worth noting that the angular consistency also exists in DW images with different diffusion gradient directions.

There should be high angular consistency among at least three image sets like $G(L)_A$, $G(L)_B$ and H_A (see Fig. 6.1). Therefore, we further improve the first term in Eq. (6.6) as follows:

$$L_\eta = \frac{1}{N}\left(\sum_{s=1}^{N}\left\|\underline{M}(G(L)_A)_s + \underline{M}(G(L)_B)_s - 2\underline{M}(H_A)_s\right\|_F^2\right) \tag{6.7}$$

In fact, Eq. (6.7) establishes a 3-cycle-consistency (Zhou, et al., 2015) strategy for minimizing the differences between $\underline{M}(H_A)_s$, $\underline{M}(G(L)_A)_s$, and $\underline{M}(G(L)_B)_s$. Minimizing Eq. (6.7) ensures that the differences between any two of the three diffusion tensor maps would be minimized. Similarly, by changing the second term in Eq. (6.6), a new loss function can be written as:

$$L_\eta^{3c} = \frac{1}{N}\sum_{s=1}^{N}\left(\left\|\underline{M}(G(L)_A)_s + \underline{M}(G(L)_B)_s - 2\underline{M}(H_A)_s\right\|_F^2 + \left\|\underline{M}(G(L)_A)_s + \underline{M}(G(L)_B)_s - 2\underline{M}(H_B)_s\right\|_F^2\right) \tag{6.8}$$

The final regularized term based on angular consistency is:

$$L_{cons} = L_\varphi + L_\eta^{3c} \tag{6.9}$$

By combining (6.1), Eq. (6.4) and (6.9), we obtain our full objective function for super angular resolution:

$$L_{SR} = \min_G \max_D \alpha L_{GAN}(G,D) + \beta L_{sim} + \gamma L_{cons} \tag{6.10}$$

where α, β, and γ are parameters for controlling the trade-off between adversarial learning, image similarity and angular consistency, respectively.

6.2.4 Network architecture and training Process

Fig. 6.2 provides the detailed architecture of the proposed generator and discriminator used in the present network. We adjusted the architecture to learn the mapping from 12 diffusion gradient directions to 24 diffusion gradient directions for super angular-resolution in *in vivo* cardiac DTI. The generator was constructed based on the Batch normalization modules (Ioffe & Szegedy, 2015) and U-Net (Ronneberger, et al., 2015) as encoder-decoder that fuses the features from shallow and deep layers through multi-path information to optimize the generator quickly. The generator's inputs and outputs are DW images of size 128 × 128 ×12 and 128 × 128 × 24, respectively, where their last channels correspond to the angular resolutions. A skip connection strategy (Isola, et al., 2017) was used to connect the information between encoder and its corresponding decoder directly instead of letting the information flow pass through a mass of unrelated layers. The discriminator's inputs are 128 × 128 × 36 DW images containing both the generator's input and higher angular-resolution output (or ground-truth images). Its output is a 15 × 15 image map where each pixel value is a probability that represents how believable the corresponding patch of the input image is. All convolutions are c × 4 × 4 × f filters applied with stride 2, where c and f are the number of channels for the previous and next layers, respectively.

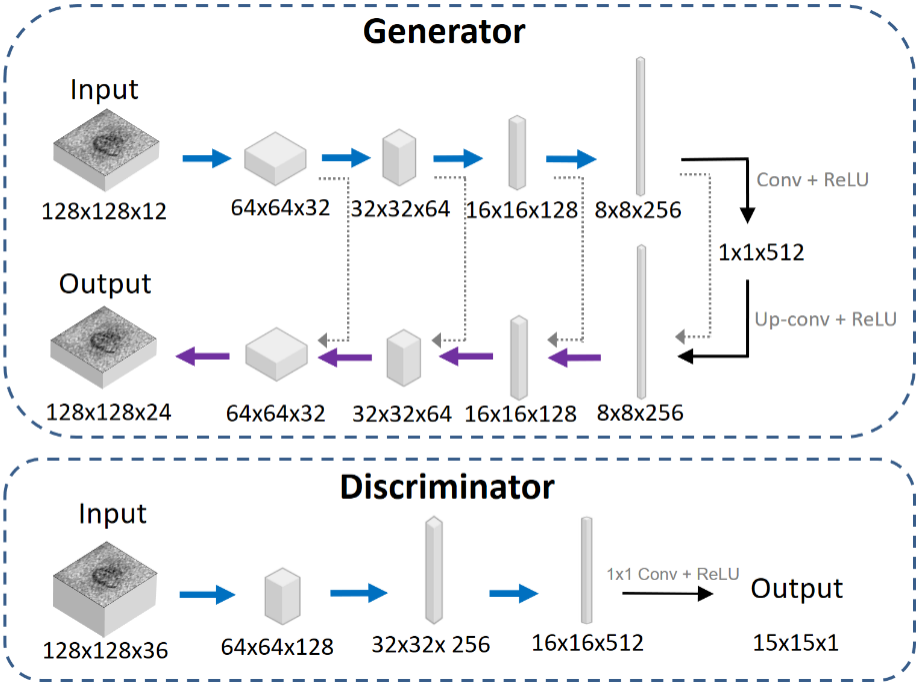

Fig. 6.2 The detailed network architecture of the proposed generator and discriminator. The blue arrow denotes repeated 4 blocks of 4×4 convolutional layers followed by a rectified linear unit (ReLU) and a 2×2 max pooling operation with stride 2 for downsampling. The purple arrow denotes a 2×2 deconvolution layer with two 4×4 convolution layers followed by a ReLU. The gray dotted line denotes skip connections between encoder and decoder.

At the beginning of the training stage, we initially updated parameters to train D using Eq. (6.1) and used the generator G to output $G(L)$. Then, the trained D was used to calculate an adversarial loss for the G's output based on Eq. (6.2). Meanwhile, the images $G(L)$ and its ground-truth images H were employed to compute the similarity loss and angular consistency loss according to Eqs. (6.4) and (6.9), respectively. After that, the parameters for G were updated by combining all the computed losses. The images produced from the trained G can, in turn, be regarded as "fake" images to update parameters for D. Finally, the capability of both G and D was increased gradually by repeating this process iteratively until all the real data $\{L, H\}$ and fake data $\{L, G(L)\}$ have no great influence on updating D.

6.3 Experiments and results

We conducted a number of experiments using human cardiac DTI datasets to evaluate the performance of the proposed method. In practice, DTI of the *in vivo* human heart is quite difficult to perform with up to 24 diffusion gradient directions. In order to train and test the proposed super angular-resolution network, in the experiments, *in vivo* cardiac DTI datasets were used in combination with *ex vivo* cardiac DTI datasets to imitate large amounts of lower angular-resolution motion-induced DW images (in 12 directions) and their corresponding ground-truth higher angular-resolution DW images (in 24 directions). In the following, we first describe the acquisition of original *in vivo* and *ex vivo* cardiac DTI datasets. Subsequently, the method including three pre-processing steps for imitating lower and higher

angular-resolution DW images are introduced. After presenting the experimental setting and evaluation metrics, the experimental results are reported.

6.3.1 Datasets

***Ex vivo* datasets**

Six *ex vivo* human hearts includes four infants and two adults were obtained and processed in compliance with French legal and ethical guidelines. The investigations conformed to the principles outlined in the Declaration of Helsinki (Carlson, et al., 2004). More details about the acquisition process of these hearts were reported in (Yang et al., 2018). The DTI datasets were acquired in clinical conditions with Siemens 3T MRI Magnetom Verio. The imaging parameters are the following: TE = 74 ms, TR = 7900 ms, FOV = 144 × 144 mm^2, slice thickness = 1.4 mm, in-plane resolution = 2 mm, slice spacing = 1.4 mm, slice duration = 123.2 ms, number of slices = 35, slice size = 104 × 104, diffusion sensitivity b = 700 s/mm^2, number of diffusion gradient directions = 192, 64 and 12. In each direction, MRI scans were acquired 6 times for noise reduction. The adult datasets were acquired using Siemens 3T MRI Magnetom Prisma with following parameters: TE = 71 ms, TR = 9600 ms, FOV = 177 × 177 mm^2, slice thickness=1.5 mm, slice spacing = 1.5mm, slice duration = 123.1 ms, number of slices = 70, slice size: 122 × 122 pixels, diffusion sensitivity b = 700 s/mm^2, and number of gradient directions = 192. In each direction, MRI scans were performed 3 times for noise reduction.

***In vivo* datasets**

The *in vivo* human subjects used in this method are from six healthy volunteers. The cardiac DTI of these volunteers were performed with a 1.5 T clinical scanner (MAGNETOM Avanto, Siemens AG) at the Neuro-Cardiology Hospital of Lyon in France. All subjects gave informed consent to the institutional review board-approved study protocol prior to participation. The acquisition parameters are: TE/TR = 51/100 ms, b-value = 200 s/mm^2, spatial resolution = 2.63 × 2.63 × 6 mm^3, matrix size = 90 × 160, FOV = 420 × 236 mm^3, bandwidth = 1302 Hz, acceleration rate = 2, partial Fourier = 6/8. For each heart, 10 slices (slice thickness = 6 mm) were acquired in the short-axis view without any interslice gap. An individual mask was obtained to detect the myocardial region. The DW images were acquired in 10 different time points. For each time point, 13 DW images including 12 images associated with 12 diffusion gradient directions and one b_0 image (without diffusion weighting) were acquired. More information on data acquisition was detailed in (Wei, et al., 2013).

6.3.2 Data pre-processing

1) Sampling of *ex vivo* cardiac DTI data

We used the ex vivo cardiac DW images acquired in 192 diffusion gradient directions to sample data for training the mapping from low to high angular-resolution DW images. Specifically, for the 192 diffusion gradient directions, we sub-sampled 24 diffusion gradient directions from them using a q-space sampling algorithm called iterative maximum overlap construction (IMOC) (Cheng, et al., 2017), such that 24 directions are uniformly distributed

on a sphere. By using the same algorithm, 12 directions uniformly distributed on a sphere were sub-sampled from the 24 directions. Then, two sets of DW images associated with the obtained 12 directions and 24 directions, respectively, were extracted from the 192 *ex vivo* DW images, which were served as the network's input and ground-truth, respectively.

However, training the network on a fixed set of 24 diffusion gradient directions (including the subset with 12 directions) may not be practical enough because the diffusion gradient directions used for a given angular resolution may be varied with different DTI machines that adopt different sampling schemes in practice. To design a network that can be used for different imaging machines or schemes, we further generated plenty of DW image sets as network's training data, where the associated 24 diffusion gradient directions are different from one set to another. It was shown in (Caruyer, et al., 2013) that any independent set of diffusion gradient directions (uniformly spread on the same unit sphere) can be regarded as a rigid rotation of another set of directions with the same angular resolution. Subsequently, we rotated the fixed set of 24 directions 192 times on the sphere with the rotation angles determined by the 192 directions. Then, for each rotated direction, we extracted its nearest neighbor in the 192 directions based on Euclidean distance between directions so that we can obtain 192 sets of varied 24 directions. After that, from each set of the 24 directions, 12 directions were sub-sampled using the IMOC algorithm. Finally, 192 pairs of DW image sets were extracted from the *ex vivo* DW images according to the obtained sets of 24 directions and their corresponding 12 directions.

2) Direction normalization

By definition, our network processes DW image set without using any directional information. Therefore, it is necessary to design a strategy that matches the diffusion gradient directions of input and output DW images. To achieve this, we first used the IMOC algorithm to generate a set of 24 diffusion gradient directions and its subset with 12 directions as two fixed sets of directions assigned to the output and input DW images, respectively. Then, each voxel in DW images associated with any new set of 12 or 24 diffusion gradient directions was normalized into the fixed set of directions using the equation below:

$$S_i^{norm} = S_0 \exp(-b\vec{\mu}_i^T \underline{D} \vec{\mu}_i) \tag{6.11}$$

where $\underline{D}$ denotes the diffusion tensor estimated from un-normalized signal S_i with *i*th original diffusion gradient direction $\vec{g}_i : S_i = S_0 \exp(-b\vec{g}_i^T \underline{D} \vec{g}_i)$ for $\vec{g}_i = (x_i, y_i, z_i)$ and *i*=1,2, …, n(n ≥ 6), $\vec{\mu}_i = (u_i, v_i, w_i)$ designates the *i*th new diffusion gradient direction in the fixed set, S_i^{norm} is its normalized value and *b* is the diffusion weighting factor (Kingsley, 2006). The use of this equation ensures that the diffusion tensor, b-value and S_0 are kept constant during the normalization.

3) Imitation of in vivo motion-induced DTI data

After normalizing the *ex vivo* DW images, a method was used to model motion-induced effects and imitate motion-induced DW images. More precisely, given a set of *in vivo* DW images acquired in 12 diffusion gradient directions, the motion estimation algorithm PCATMIP was used to calculate mean motion-induced signal loss for each DW image.

Then, a value x was computed by normalizing the mean signal loss of all images to 0-1 for measuring the degree of signal loss. We defined the signal decay in the presence of motion as $S_i^m = S_i \cdot y$, where y represents the signal attenuation term caused by cardiac and/or respiratory motion. For each motion-induced signal intensity S_i^m , its signal loss-free signal S_i can be estimated using PCATMIP algorithm so that the value of y can be obtained. After that, we modeled the relationship between the signal loss degree x and the motion-induced attenuation term y as an exponential function defined by $y=A\exp(Bx+c)$, where A, B and c are the estimation coefficients. This model was solved by the regression estimation using DW images over six *in vivo* human hearts. The obtained model provides a mean signal loss value given an arbitrary degree x. Besides, based on the value of x, we roughly divided DW images into three categories with minor, moderate, and huge signal losses. This allows us to determine the mean numbers of images for different categories of signal loss. Furthermore, for each category of signal loss, the proportion of smaller and larger pixel values for one image were also calculated. Based on these results, realistic motion-induced signal loss for each pixel was generated and then added to *ex vivo* DW images to imitate *in vivo* DTI data corrupted with signal loss.

6.3.3 Experimental Setting

To conduct the experiments, DTI datasets of six *ex vivo* human hearts were used to imitate lower and higher angular-resolution DW images. For each dataset, we selected 10 of the 35 slices (from the middle third of the heart) of the *ex vivo* DW images to imitate the number of slices often used in *in vivo* setting. A myocardium mask for each slice was obtained by thresholding the apparent diffusion coefficient (ADC) images calculated from the DW images using an upper threshold of 2×10^{-3} mm^2/s (following the method in (Yang et al., 2018)). With the myocardium mask, the myocardial area of each image was extracted as region of interest (ROI). For each slice, 192 pairs of DW image sets were generated using the three data pre-processing steps detailed in section 6.3.2. Each pair of sets was regarded as a training sample consisting of a set of motion-induced 12 DW images and a set of 24 ground-truth DW images without motion effects. That is, for each heart, a total of 192 × 10 = 1920 training samples for the 10 slices were generated.

We first trained and tested the proposed network on the imitated DW images of six human hearts. For each heart, we randomly selected 1400 samples from the generated 1920 training samples for training and used the remaining 520 samples for testing. Subsequently, we tested the proposed network directly on the original *in vivo* cardiac DTI datasets for further validation. To this end, 6000 samples were randomly selected from the training samples of all the six human hearts (1000 samples per heart) to train the proposed network.} According to our experimental experiences, we set parameters α = 1, β = 1 and γ = 50 to achieve an optimal balance for minimizing the loss function. All data pre-processing steps were performed using MATLAB (MathWorks, Inc., Natick, MA, USA). The network algorithm was run on a Linux cluster with a Matrox GPU and 32 GB of system memory (RAM).

6.3.4 Evaluation Metrics

Several metrics were employed to evaluate the performance of the proposed network in both qualitative and quantitative manners. First, diffusion tensors were estimated from the DW images to be evaluated using the least squares estimation. Then, for each estimated diffusion tensor, we calculated the FA and MD that describe the degree of anisotropy and the overall diffusion level of water molecules, respectively. Finally, main fiber orientations represented by azimuth angle and elevation angle were calculated using the primary eigenvector of the diffusion tensor.

For experiments performed on imitated *in vivo* DW images, we employed a similar method to that in (Yang, et al., 2011) to measure estimation errors in terms of the above metrics. Specifically, the *ex vivo* DW images acquired in 192 directions were averaged over all time points and used as reference images. The errors in estimating diffusion tensors were measured by computing the Frobenius distance (FD) between estimated tensors and reference tensors (from the reference images). The errors in estimating FA, MD, azimuth angle and elevation angle were calculated by computing the mean absolute error (MAE) between estimated values and reference values (computed from reference tensor). These estimation errors were first computed for each myocardial voxel and then the mean of all errors (ME) was computed. The standard deviation of estimation errors (ErrSD) was also computed using:

$$ErrSD = \left(\frac{1}{N_{mv}} \sum_{i=1}^{N_{mv}} (Err_i - ME)^2 \right)^{\frac{1}{2}} \tag{6.12}$$

where N_{mv} is the number of myocardial voxels, Err_i represents the estimation error (FD for the diffusion tensor and MAE for other metrics) at the ith voxel. For experiments directly performed on *in vivo* cardiac DTI datasets, the values of FA, MD, azimuth angle and elevation angle from the resulting DW images were used for comparison and analysis.

6.3.5 Results

In this section, we provide experimental results in three main parts. First, we show the results of imitating the motion-induced DTI data used for training the proposed network. Second, we give some results of analyzing the performance of the proposed network based on adverserval learning and angular consistency loss. Finally, we present the detailed evaluation results of the proposed network for the compensation of motion-induced signal loss in *in vivo* cardiac DTI.

Imitated motion-induced DW images

Fig. 6.3 plots the estimated relationship between the signal loss degree x and the mean signal loss value y over six human hearts. The results were obtained by fitting y as an exponential function of x using regression estimation (as mentioned in section 6.3.2, the 3rd step). The fitting parameters of this model were RMAE = 0.038 and R-square = 0.983. As we can see, the fitted results roughly follow the behavior of monotonically decreasing exponential function. This fitted model provides a mean value to generate realistic motion-induced signal loss.

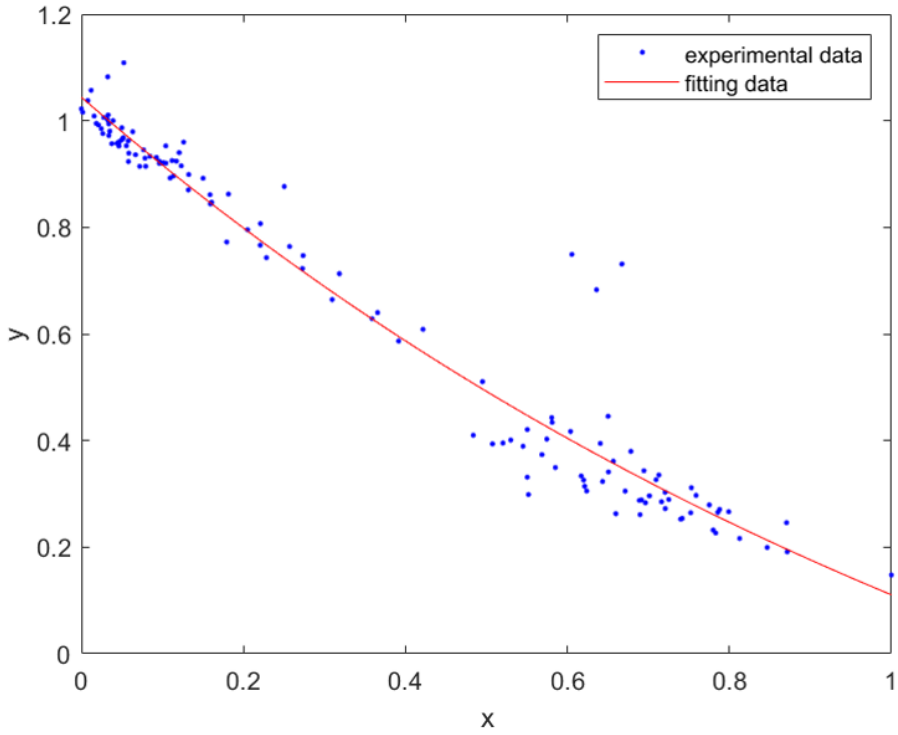

Fig. 6.3 Motion-induced attenuation term as a function of signal loss degree over the six human hearts. For each blue point (the experimental data), the *x* value denotes the degree of signal loss and the *y* value the mean signal loss computed from all pixels averaged inside each of image on the myocardium area.

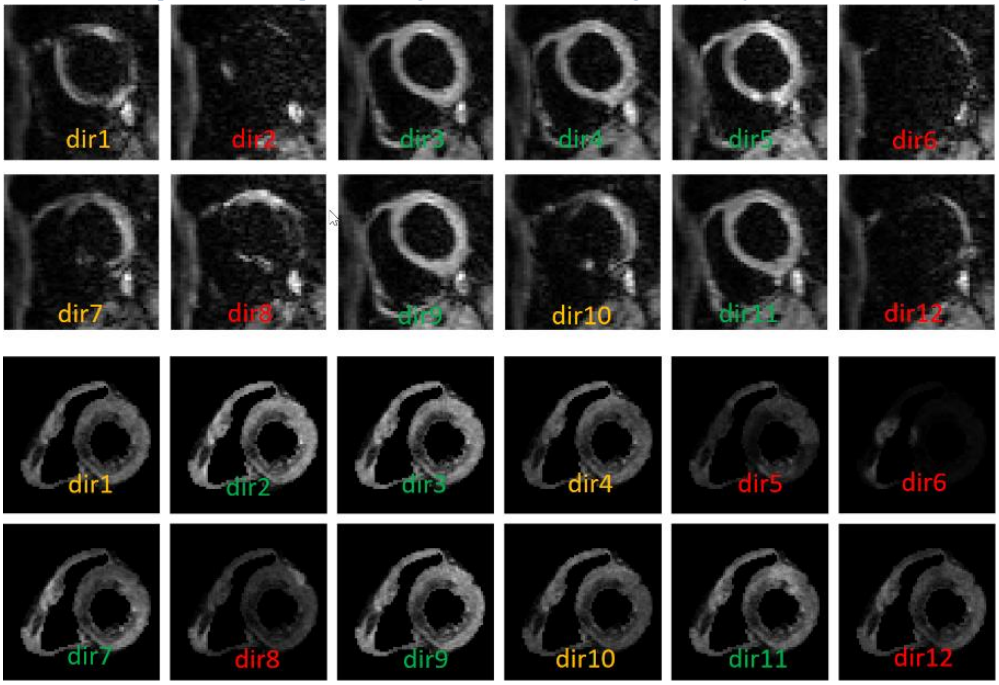

Fig. 6.4 An example of imitated motion-induced DW images for 12 directions. Top two rows: *In vivo* cardiac DW images. Bottom two rows: motion-induced DW images obtained by adding the estimated signal loss to *ex vivo* DW images. The direction labels superimposed on the image with green, yellow and red colors designate minor, moderate and huge signal losses, respectively.

Fig. 6.4 shows an example of motion-induced DW images imitated by adding the realistic signal loss to *ex vivo* DW images. Both the *in vivo* and original *ex vivo* DW images were acquired in 12 diffusion gradient directions. The mean numbers of images in the 12 directions over all *in vivo* human heart were 5, 3, and 4, respectively for minor, moderate and huge signal losses. For minor signal-loss images, the mean proportion of larger pixel values h_v ($h_v \geq 2\varepsilon/3$, where ε is the maximum intensity in the image) per image are 86%, and this proportion is 51% and 6% for moderate and huge signal losses, respectively. As observed, for a set of 12 diffusion gradient directions, the method imitates highly realistic motion-induced DW images.

Impact of adversarial learning and angular consistency

Fig. 6.5 gives the qualitative results of the network trained with and without the adversarial loss $L_{GAN}(G, D)$. The results of training without adversarial learning were achieved by setting $\alpha = 0$, $\beta = 1$ and $\gamma = 0$, which means the case that the network only minimizes the similarity between the produced and ground-truth DW images using L2 norm. The images produced by running the network with adversarial loss were obtained by changing α into 1 for training. We can observe that without using the adversarial loss (Fig. 6.5, top row), the intensities of the produced DW signals are blurred and visually unrealistic. In contrast, the produced images using the adversarial learning combined with the L2 norm (Fig. 6.5, middle row) present more variations and are perceptually closer to the ground truth.

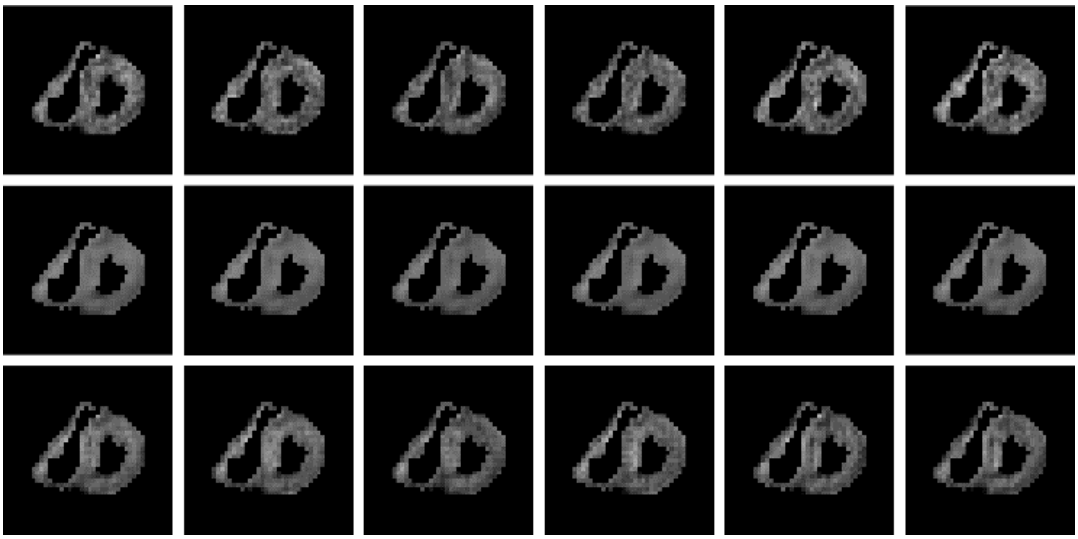

Fig. 6.5 Visual comparison of the cardiac DW images produced by our network without (middle row) and with (bottom row) adversarial learning, and ground-truth DW images (top row). Each column corresponds to one of six diffusion gradient directions.

Table 6.1 reports the ME ± ErrSD of the diffusion tensors, FA, MD, elevation and azimuth angles for the motion-induced DW images, motion-compensated DW images produced by the proposed network with and without adversarial learning on imitated DW images of an infant human heart. The unit of MD is 10^{-3} mm^2/s. As observed, the ME of all the metrics was decreased after performing signal loss compensation using the proposed network. When using the network without adversarial learning, the ME of these metrics was slightly but not significantly reduced. The mean errors of elevation angles and azimuth angles for the network

with adversarial learning were decreased by 0.57° and 0.62°, respectively, compared with the network without adversarial learning.

Table 6.1 Performance of signal loss compensation using the proposed network with and without adversarial learning (Adv.). The numbers designate the ME ± ErrSD values.

	Motion induced	**Motion-compensated**	
		Without Adv.	With Adv.
Diffusion tensor	0.07 ± 0.16	0.07 ± 0.14	0.05 ± 0.10
FA	0.01 ± 0.03	0.01 ± 0.03	0.01 ± 0.02
MD	0.03 ± 0.05	0.02 ± 0.05	0.01 ± 0.03
Elevation	8.91 ± 26.26	8.21 ± 25.67	7.84 ± 25.11
Azimuth	5.78 ± 18.13	5.61 ± 17.99	5.12 ± 17.10

Table 6.2 summarizes the ME ± ErrSD of diffusion tensors, azimuth and elevation angles for the proposed method with and without using the consistency loss (Eq. (6.9)). We performed the training without angular consistency by setting $\gamma = 0$ and kept the optimal parameters for the others. From Table 6.2, we can see that, compared with the network without angular consistency loss, the errors in estimating the elevation and azimuth angles are further decreased by 0.60° and 0.53°, respectively, and that the diffusion tensors are decreased by 0.01×10^{-3} mm^2/s. This indicates the interest of enhancing angular consistency for improving the effect of signal loss compensation.

Table 6.2 Comparison between with and without angular consistency (AC) when using the proposed network for the estimation errors (ME ± ErrSD) of diffusion tensor and fiber orientation.

	With AC	**Without AC**
Diffusion tensor	0.05 ± 0.10	0.04 ± 0.08
Elevation	7.84 ± 25.11	7.74 ± 24.60
Azimuth	5.12 ± 17.10	4.67 ± 16.20

Performance of Signal Loss Compensation

Fig. 6.6 gives the 3D visualization of diffusion tensor fields derived from imitated motion-induced DW images, signal loss-compensated DW images, and reference DW images for one slice of the left ventricle of an adult human heart. The three zoom views represent the same region of the left ventricle (LV) containing fiber structures from epicardium to endocardium. The motion-induced DW images used here correspond to the images shown in Fig. 6.4. From Fig. 6.6(b), it can be observed that the shapes of most tensor cuboids for the motion-induced images associated with 12 diffusion gradient directions are not very different from each other, and that their major axes (main orientations of tensor cuboids) are almost identical. After performing the signal loss compensation, these cuboids (Fig. 6.6(c)) became more anisotropic and closer to the reference tensors (Fig. 6.6(a)). By further observing the zoom views of Fig. 6.6(c), we can find that the main orientations of tensor cuboids linearly vary from epicardium to endocardium. This variation roughly reflects the helical structure of normal myocardial fibers of the human heart.

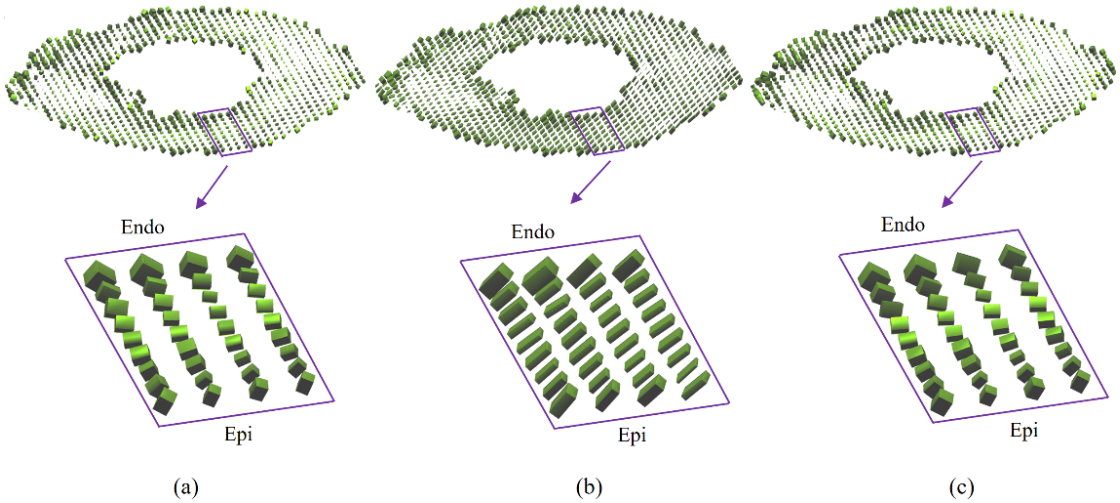

Fig. 6.6 3D visualization of diffusion tensor fields using superquadric glyphs derived from (a) reference DW images, (b) imitated motion-induced DW images, and (c) signal loss-compensated DW images obtained using the proposed network. (Epi: epicardium. Endo: endocardium.)

Fig. 6.7 shows the elevation and azimuth angle maps calculated from imitated motion-induced DW images with and without signal loss compensation. Here we take the LV myocardium as our region of interest. From Fig. 6.7(a), it is observed that most elevation angles in the LV myocardium vary from about 60° at the endocardium to about 10° at the mid-wall, and back to about 50° at the epicardium. Also, as observed in Fig. 6.7(d), there is an approximately continuous circular variation for the azimuth angles. When observing clockwisely around the LV, the azimuth angles vary periodically from 0° to 180°, and areas having nearly similar color appear twice. The variations of elevation and azimuth angles in the reference maps reflect the cross-helical structure of cardiac fibers. Such variations are difficult to be found in the elevation and azimuth angles maps derived from the motion-induced DW images (Fig. 6.7(b) and (d)). After using the proposed signal loss-compensation method, regular variation patterns of elevation angles on the LV were largely recovered except for the red area (Fig. 6.7(a)). The variation pattern of azimuth angles was nearly completely recovered.

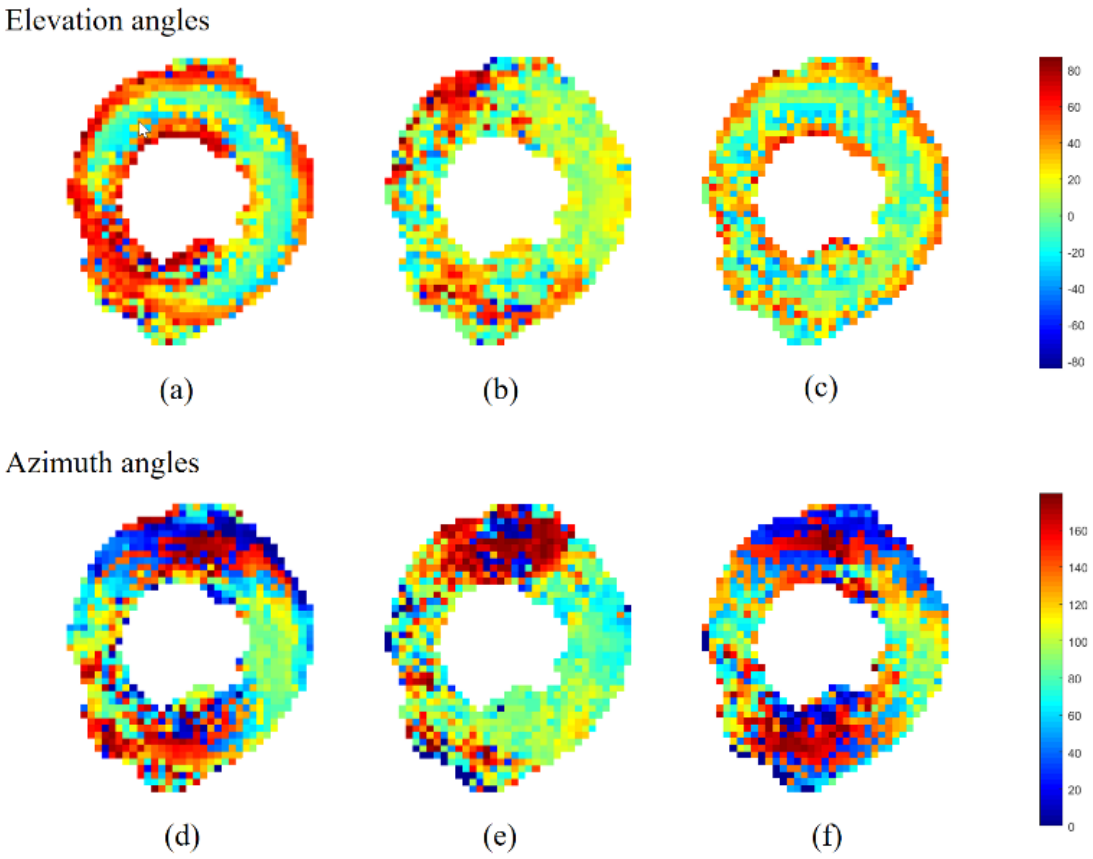

Fig. 6.7 Elevation angle maps (from (a) to (c)) and azimuth angle maps (from (d) to (f)). (a) and (d) are angle maps derived from reference DW images. (b) and (e) are angle maps derived from DW images corrupted by motion. (c) and (f) are angle maps derived from the signal loss-compensated DW images after using the proposed network.

Table 6.3 summarizes the ME ± ErrSD of elevation and azimuth obtained before and after using the proposed network on the imitated motion-induced DW images of six human hearts. Clearly, for all the six hearts, the proposed network significantly reduced the errors of estimating fiber angles. Besides, the proposed network also reduced the variation of estimation errors (see the values of ErrSD). The mean values over all the hearts (last row of Table 6.3) show that the elevation and azimuth MEs were decreased by 1.65° and 2.88°, respectively, and that these ErrSD values were 2.13° and 3.60°, respectively.

Table 6.3 Estimation errors (ME ± ErrSD) of fiber orientation computed from the imitated motion-induced and signal loss-compensated DW images for the six human hearts (H*i* denote the dataset of the *i*th human heart).

	Motion-induced		**Signal loss-compensated**	
	Elevation	Azimuth	Elevation	Azimuth
H1	5.02 ± 16.41	7.65 ± 24.66	3.01 ± 13.25	4.67 ± 19.04
H2	5.48 ± 16.98	8.66 ± 26.06	4.28 ± 14.56	7.06 ± 23.07
H3	5.04 ± 15.48	7.85 ± 23.66	3.68 ± 13.12	5.99 ± 21.43
H4	7.11 ± 19.64	10.09 ± 26.96	4.67 ± 16.20	7.74 ± 24.60
H5	9.63 ± 22.31	16.24 ± 35.95	9.14 ± 21.65	12.97 ± 32.62
H6	9.73 ± 20.67	15.18 ± 32.02	5.41 ± 15.91	8.89 ± 26.41
Mean	7.35 ± 18.58	11.43 ± 28.13	**5.70 ± 16.45**	**8.55 ± 24.53**

Table 6.4 provides a comparison of the proposed network with state-of-the art ones on the imitated *in vivo* DW images, where the ME and ErrSD values over all of the hearts were reported. The PCATMIP (Pai, et al., 2011) and WIFPCA (Wei, et al., 2014) were proposed specifically for compensation of motion-induced signal loss in *in vivo* cardiac DTI, and they relied on combining the DW images from multiple time points. The DIF-net (Jung, et al., 2020) and U2fusion-net (Xu, et al., 2020) were two representative CNN-based methods developed for image fusion tasks. Here motion-induced signal loss were added to the *ex vivo* DW images acquired at multiple time points (without signal averages) to imitate *in vivo* DW images and then we run these four methods on the imitated motion-induced images acquired at multiple time points. For the proposed method, we simply averaged the images over all time points and then generated motion-induced data for training the network. From the results, we can see that the ME and ErrSD values of both azimuth and elevation angles for WIFPCA are lower than that values for PCATMIP and the other two image fusion networks. But the proposed network gave the best results by significantly reducing estimation errors of fiber angles especially the elevation angles.

Table 6.4 Comparison results of the proposed method with state-of-the art ones. The numbers designate the ME ± ErrSD values.

	Elevation	**Azimuth**
Motion-induced	7.11 ± 19.64	10.09 ± 26.96
PCATMIP (Pai, et al., 2011)	6.84 ± 17.34	9.40 ± 26.04
WIFPCA (Wei, et al., 2014)	6.52 ± 16.86	9.07 ± 25.07
DIF-net (Jung, et al., 2020)	8.52 ± 20.37	10.40 ± 28.14
U2fusion-net (Xu, et al., 2020)	6.41 ± 19.25	9.42 ± 27.49
Proposed	**5.70 ± 16. 45**	**8.55 ± 24. 53**

Fig. 6.8 shows the elevation and azimuth angle maps obtained by running the proposed network and state-of-the art ones directly on cardiac DW images of an *in vivo* human heart. It is observed that the original elevation angles (computed from the original *in vivo* DW images) in the LV vary irregularly from epiocardium to endocardium, and that the original azimuth angles in the LV do not exhibit circularly periodic variation pattern. After signal loss compensation using the PCATMIP, WIFPCA, DIF-net, U2Fusion-net and the proposed method, the elevation angles in the LV showed a roughly linear variation from epiocardium to endocardium, and their corresponding circularly periodic variation of the azimuth angles were also recovered. In particular, regular elevation and azimuth angle variations for the WIFPCA and the proposed method were observed more clearly compared with the other methods. But angles for the proposed method showed a relatively lower noise level compared with that for the WIFPCA. Table 6.5 summarizes the mean and standard deviation (Mean ± SD) of FA and MD values obtained after using the proposed method and state-of-the art ones on the cardiac DTI datasets of all the six *in vivo* human hearts. We observe that the mean and variations of both FA and MD for WIFPCA were lower than those for PCATMIP, which is consistent with the results in (Wei, et al., 2014). In contrast, the proposed method yielded the lowest mean and variations of FA (0.42 ± 0.08) and MD (0.65 ± 0.09 × 10^{-3} mm^2/s) compared with all the other methods.

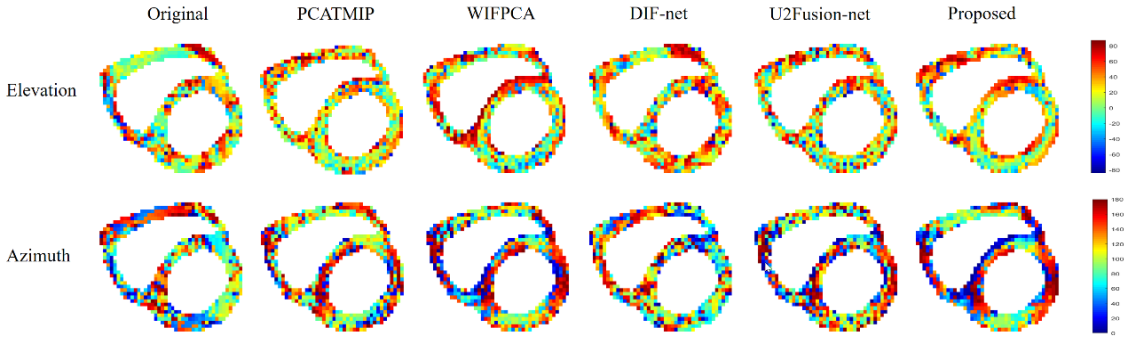

Fig. 6.8 Elevation angle maps (top row) and azimuth angle maps (bottom row) computed from the original in vivo cardiac DW images (first column on the left) and the DW images after signal loss compensation using the proposed network and state-of-the art ones (from the second to the last columns).

Table 6.5 Mean ± SD of FA and MD in the LV of six in vivo human heart. The unit of MD is 10^{-3} mm^2/s.

	FA ± SD	**MD ± SD**
Original	0.63 ± 0.14	1.68 ± 0.24
PCATMIP (Pai, et al., 2011)	0.51 ± 0.12	0.85 ± 0.11
WIFPCA (Wei, et al., 2014)	0.46 ± 0.10	0.73 ± 0.09
DIF-net (Jung, et al., 2020)	1.52 ± 0.67	2.40 ± 0.18
U2fusion-net (Xu, et al., 2020)	0.61 ± 0.16	1.42 ± 0.12
Proposed	0.42 ± 0.08	0.65 ± 0.09

6.4 Conclusions and discussions

In this chapter, we proposed a GAN-based super angular-resolution method for compensating for motion-induced signal loss in *in vivo* DTI of the human heart. The constructed network used an objective function based on adversarial learning, image similarity and angular consistency. Several pre-processing strategies of imitating lower and higher angular resolutions in *in vivo* DTI data were designed to train and test the network. The experimental results showed that the proposed method presents the ability to compensate for motion-induced signal loss in terms of accurately estimating diffusion tensors and DTI-derived parameters such as FA and MD. Moreover, the results also demonstrated that the use of angular consistency improved the network in producing DW images from which fiber orientations were more accurately derived. The proposed method then provides an interesting way to compensate for motions in *in vivo* cardiac DTI.

There are mainly two major challenges in implementing the proposed network. One is that cardiac DTI data required for training the present network to perform super angular resolution are quite difficult to obtain in practice. These data include the DW images associated with the mapping from 12 to 24 diffusion gradient directions (used as the network's input and ground-truth images, respectively) and the DW images associated with different sets of of diffusion gradient directions (used for applying the trained network to different imaging machines or schemes). The pre-processing strategies proposed in section 6.3.2 make it possible to generate enough DW images associated with these desired directions from existing *ex vivo* DW images acquired in 192 directions. The other challenge is lacking of *in vivo* cardiac DTI data for enabling the network to achieve signal loss compensation in *in vivo* DTI. The 3rd step of pre-processing strategies presented in section 6.3.2 allows us to cope with this problem by adding realistic motion-induced signal losses to *ex vivo* DW images to imitate *in vivo* DW images

corrupted by cardiac motion. A similar method for imitating *in vivo* cardiac DTI data was proposed in (Wei, et al., 2013), where motion-induced signal losses were estimated from DENSE acquisitions (Aletras, et al., 1999). In contrast, our imitation method is capable of estimating signal losses directly from a given *in vivo* DTI dataset and does not rely on any specific imaging sequence, which is more flexible and easier to implement.

During the training process, the data fed to the proposed network contain 36 DW images including 12 input images and 24 ground-truth images to learn the mapping from 12 to 24 directions. In fact, since we specified that the directions of input are a subset of output, an alternative way to perform super angular resolution is to input 12 images associated with original 12 directions to the network and let it output 12 images associated with 12 new directions. In this study, since the DW images of all 24 directions were served as network's output, the correlation between the DW images associated with the original 12 directions and new 12 directions were both enhanced after optimizing the network with adversarial learning and angular consistency. This is also an important reason contributing to the impressive performance of the proposed method. From the results in Fig. 6.7 and Fig. 6.8, it can be observed that the angular maps are noisy. This is because we directly derived these angle maps from the DW images and without using additional pre-processings such as interpolation or denoising. Even so, we also found some main features of elevation and azimuth angles, which present in normal cardiac fibers, e.g., regular variations from epicardium to endocardium of LV.

In Table 2.1, an interesting phenomenon is that the estimation errors of ME and ErrSD for the four infant hearts (from H1 to H4) are generally smaller than the errors from the two adult hearts (H5 and H6). An intuitive understanding is that this difference reflects the fact that the network compensates differently for signal loss depending on whether the myocardial structures are infant or adults. However, this phenomenon might also be caused by the differences in imaging parameters used to acquire datasets, such as angular resolution, slice thickness and number, etc, or due to the use of different DTI machines. Further study would be needed to investigate the impact of these factors on the performance of the proposed method.

The present network is dedicated to super angular resolution of motion-induced DW images acquired in 12 directions. As such it cannot take directly the data associated with other angular resolutions as input. Future work is required to modify the network so that it can be applied to DW images of new angular resolutions. Besides, the data used for training the proposed network were only obtained by adding motion-induced signal loss to *ex vivo* DW images. It would be interesting in our next step to train the proposed network using *in vivo* cardiac DTI datasets. In addition, the results of Table 6.2 also show that the network performance can be improved by constraining the diffusion tensor estimated from the generated images. From there on, another interesting attempt might be to introduce new DTI-derived parameters (e.g., FA or fiber orientation) as an additional regularized term in loss function, or redesign new sub-networks based on these parameters.

PART III

General Conclusion

Chapter 7

General conclusions and perspectives

Contents

7.1 General conclusions

DTI provides a unique medical imaging modality to noninvasively investigate the fiber architecture of the human heart, thus offering a novel tool for the diagnosis of cardiac diseases. However, clinical applications of cardiac DTI are often penalized by the limited quantity and/or quality of acquired DW images due to limited acquisition times in practice. Recent advances in deep learning, in particular in GAN hints at a new way to cope with this problem. In this thesis, we present the original work of investigating GAN-based networks for super angular resolution in cardiac DTI and compensation of motion-induced signal loss in *in vivo* cardiac DTI. Before presenting our work, the introduction to dMRI including DTI and the applications of cardiac DTI are presented in Chapter 1, and the basics and concepts of CNN and GAN are presented in Chapter 2. The conclusions of our work from Chapter 3 to Chapter 6 are summarized as follows:

In Chapter 3, we systematically studied the joint influence of angular resolutions and noise levels in cardiac DTI. The results on both synthetic and actual DW images showed that, given sufficient acquisition time and a certain noise level, the accuracy of diffusion tensors, FA and MD measurements increase as angular resolutions increase. Besides, continuing to increase the angular resolution beyond a certain value has little effect on the accuracy of diffusion tensor, FA and MD. For higher SNR (23 dB), the near-optimal diffusion tensor, FA and MD were obtained with 48 gradient directions. With lower SNR, the near-optimal diffusion tensor, FA and MD were obtained with, for example, 120 directions if SNR = 15 dB and with 156 directions if SNR = 9 dB. In addition, the results on real data show that both the mean and variance of FA or MD decrease as the angular resolution increases. When acquisition time is imposed, increasing angular resolution often comes at the cost of a reduced number of excitations used for signal averaging, thus reducing the SNR of DW images. Longer acquisition times allow the use of higher angular resolution while keeping a certain level of SNR. These findings demonstrate the validity of higher angular resolution for improving the accuracy of diffusion tensor estimation.

Based on these findings in Chapter 3, we proposed a GAN-based network for super angular resolution in cardiac DTI, which are detailed in Chapter 4. With the proposed network, DW images in additional gradient directions are estimated from given DW image sequence acquired in a limited number of gradient directions. Experiments on cardiac DTI datasets showed that the proposed network able to generate higher resolution images from existing DW images with lower angular resolution and thus improve the estimation accuracy of diffusion tensors.

In order to apply the proposed network to *in vivo* cardiac DTI data, we conducted the study of signal loss compensation in *in vivo* cardiac DTI, which are described in Chapter 5. In this study, the number of images and pixels for different degree of signal losses in *in vivo* cardiac DTI datasets were used to generate realistic motion-induced signal losses. After adding the generated signal losses to both simulated and *ex vivo* DW images, a large amount of motion-induced DW images were obtained, which can visually and distributively imitate the in vivo DW images corrupted by motion. The experiments on the generated motion-induced DW images demonstrated that the proposed methods provides an effective way to generate

realistic motion-induced DW images that can be used for develop advanced CNN-based post-processing methods in in vivo cardiac DTI.

Finally, in Chapter 6, a novel GAN-based super angular-resolution network was proposed for the compensation of motion-induced signal loss in *in vivo* DTI of the human heart. The proposed GAN network was trained using an objective function based on adversarial learning, image similarity and angular consistency. Several pre-processing strategies were used to generate enough lower and higher angular-resolutions *in vivo* DW images for training and testing the proposed network. The experimental results showed that the proposed network presents the ability to compensate for motion-induced signal loss in terms of accurately estimating diffusion tensors and DTI-derived parameters such as FA and MD. Moreover, the results also demonstrated that the use of angular consistency improved the network in producing DW images from which fiber orientations were more accurately derived. The proposed method then provides an interesting method for the compensation of motions in *in vivo* cardiac DTI.

In short, deep convolutional adversarial networks are an effective way to compensate for signal losses caused by motions in *in vivo* DTI of the human heart owing to their ability to generate a large number of DW images and improve the estimation accuracy of diffusion tensors.

7.2 Perspectives

This thesis has enabled our work to range from study of super angular resolution of cardiac DTI to the compensation of motion-induced signal loss in *in vivo* cardiac DTI. However, this work is far from complete. The study in this thesis found that it is feasible to improve the estimation accuracy of diffusion tensors in cardiac DTI. But this improvement is under certain conditions, such as sufficient acquisition time for a given SNR. The reproducibility on other conditions in practice needs to be further validated.

The networks in both Chapter 4 and Chapter 6 were dedicated to super angular resolution for several fixed angular resolutions. For example, in Chapter 6, the network was constructed to input motion-induced DW images acquired in 12 diffusion gradient directions. As such it cannot take directly the data associated with other angular resolutions as input. Future work is required to modify the network so that it can be applied to DW images of new angular resolutions. Besides, the data used for training the proposed network were only obtained by adding motion-induced signal loss to *ex vivo* DW images. It would be interesting in our next step to train the proposed network using *in vivo* cardiac DTI datasets.

In addition, the results in chapter 6 also demonstrated that it is possible to improve the network performance by constraining the diffusion tensor estimated from the generated images. From there on, another interesting attempt might be to introduce other DTI-derived parameters (e.g., FA or fiber orientation) as an additional regularized term in loss function, or redesign new sub-networks based on these parameters.

Author's publications

- **He, Y.**, Wang, L., Yang, F., Clarysse, Robini, P., Zheng, Y., Deng, Z., Robini, M., and Zhu, Y. Compensation of motion-induced signal loss in in vivo DTI of the human heart using GAN-Based super angular resolution. *IEEE Journal of Biomedical and Health Informatics* (*JBHI*), under review, 2022.
- **He, Y.**, Wang, L., Yang, F., Clarysse, P., Robini, M., and Zhu, Y. Effect of different configurations of diffusion gradient directions on accuracy of diffusion tensor estimation in cardiac DTI. *IEEE ICSP'22 & 3rd Sino-French Workshop 2022*, Beijing, China, October 21-24, submitted, 2022.
- **He, Y.**, Wang, L., Yang, F., Clarysse, P., and Zhu, Y. Study of motion compensation for in vivo DTI of the human heart. *2nd Sino-French Workshop on "Medical Image Analysis and AI"*, Virtual, France, October 25-26, 2021.
- **He, Y.**, Wang, L., Yang, F., Xia, Y., Clarysse, P., and Zhu, Y., Systematic study of joint influence of angular resolution and noise in cardiac diffusion tensor imaging. *11th International Conference on Functional Imaging and Modeling of the Heart, FIMH 2021*, Stanford, CA, USA, June 21-25, LNCS, volume 12738, 2021, p. 200–210.
- **He, Y.**, Wang, L., Yang, F., Clarysse, P., and Zhu, Y., 2020. Deep group-wise angular translation of cardiac diffusion MRI in q-space via manifold regularized GAN. *15th IEEE International Conference on Signal Processing, ICSP 2020*, Beijing, China, December 6-9, volume 1, 2020, p. 511–515.

Bibliography

Abramian, D. & Eklund, A., 2019. *Refacing: reconstructing anonymized facial features using GANs.* s.l., s.n., p. 1104–1108.

Aletras, A. H., Ding, S., Balaban, R. S. & Wen, H., 1999. DENSE: displacement encoding with stimulated echoes in cardiac functional MRI. *Journal of magnetic resonance (San Diego, Calif.: 1997),* Volume 137, p. 247.

Alzubaidi, L. et al., 2021. Review of deep learning: Concepts, CNN architectures, challenges, applications, future directions. *Journal of big Data,* Volume 8, p. 1–74.

Anderson, A. W., 2001. Theoretical analysis of the effects of noise on diffusion tensor imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine,* Volume 46, p. 1174–1188.

Arjovsky, M., Chintala, S. & Bottou, L., 2017. *Wasserstein generative adversarial networks.* s.l., s.n., p. 214–223.

Armanious, K. et al., 2019a. *Retrospective correction of rigid and non-rigid MR motion artifacts using GANs.* s.l., s.n., p. 1550–1554.

Armanious, K., Mecky, Y., Gatidis, S. & Yang, B., 2019b. *Adversarial inpainting of medical image modalities.* s.l., s.n., p. 3267–3271.

Arun, K. S., Huang, T. S. & Blostein, S. D., 1987. Least-squares fitting of two 3-D point sets. *IEEE Transactions on pattern analysis and machine intelligence,* p. 698–700.

Assaf, Y., Blumenfeld-Katzir, T., Yovel, Y. & Basser, P. J., 2008. AxCaliber: a method for measuring axon diameter distribution from diffusion MRI. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine,* Volume 59, p. 1347–1354.

Baraniuk, R. G., 2007. Compressive sensing [lecture notes]. *IEEE signal processing magazine,* Volume 24, p. 118–121.

Barrett, R. et al., 1994. *Templates for the solution of linear systems: building blocks for iterative methods.* s.l.:SIAM.

Basser, P. J., 1995. Inferring microstructural features and the physiological state of tissues from diffusion-weighted images. *NMR in Biomedicine,* Volume 8, p. 333–344.

Basser, P. J. a. P. C., 2011. Microstructural and physiological features of tissues elucidated by quantitative-diffusion-tensor MRI. *Journal of magnetic resonance,* Volume 213, pp. 560--570.

Basser, P. J., Mattiello, J. & LeBihan, D., 1994. MR diffusion tensor spectroscopy and imaging. *Biophysical journal,* Volume 66, p. 259–267.

Beers, A. et al., 2018. High-resolution medical image synthesis using progressively grown generative adversarial networks. *arXiv preprint arXiv:1805.03144.*

Berker, Y. et al., 2012. MRI-based attenuation correction for hybrid PET/MRI systems: a 4-class tissue segmentation technique using a combined ultrashort-echo-time/Dixon MRI sequence. *Journal of nuclear medicine,* Volume 53, p. 796–804.

Bermudez, C. et al., 2018. *Learning implicit brain MRI manifolds with deep learning.* s.l., s.n., p. 408–414.

Blackman, S. & Popoli, R., 1999. Design and analysis of modern tracking systems(Book). *Norwood, MA: Artech House, 1999.*.

Bloch, F., 1946. Nuclear induction. *Physical review,* Volume 70, p. 460.

Brown, R., 1828. *A brief account of microscopical observations made... on the particles contained in the pollen of plants, and on the general existence of active molecules in organic and inorganic bodies.* s.l.:s.n.

Calimeri, F., Marzullo, A., Stamile, C. & Terracina, G., 2017. *Biomedical data augmentation using generative adversarial neural networks.* s.l., s.n., p. 626–634.

Callaghan, P. T., 1993. *Principles of nuclear magnetic resonance microscopy.* s.l.:Oxford University Press on Demand.

Carlson, R. V., Boyd, K. M. & Webb, D. J., 2004. The revision of the Declaration of Helsinki: past, present and future. *British journal of clinical pharmacology,* Volume 57, p. 695–713.

Caruyer, E., Lenglet, C., Sapiro, G. & Deriche, R., 2013. Design of multishell sampling schemes with uniform coverage in diffusion MRI. *Magnetic resonance in medicine,* Volume 69, p. 1534–1540.

Catana, C. et al., 2010. Toward implementing an MRI-based PET attenuation-correction method for neurologic studies on the MR-PET brain prototype. *Journal of Nuclear Medicine,* Volume 51, p. 1431–1438.

Chartsias, A., Joyce, T., Dharmakumar, R. & Tsaftaris, S. A., 2017. *Adversarial image synthesis for unpaired multi-modal cardiac data.* s.l., s.n., p. 3–13.

Chen, G. et al., 2018. Angular upsampling in infant diffusion MRI using neighborhood matching in xq space. *Frontiers in neuroinformatics,* Volume 12, p. 57.

Cheng, J., Shen, D., Basser, P. J. & Yap, P.-T., 2015. *Joint 6D kq space compressed sensing for accelerated high angular resolution diffusion MRI.* s.l., s.n., p. 782–793.

Cheng, J., Shen, D., Yap, P.-T. & Basser, P. J., 2017. Single-and multiple-shell uniform sampling schemes for diffusion MRI using spherical codes. *IEEE transactions on medical imaging,* Volume 37, p. 185–199.

Cheng, J., Shen, D., Yap, P.-T. & Basser, P. J., 2018. Single-and multiple-shell uniform sampling schemes for diffusion MRI using spherical codes. *IEEE transactions on medical imaging,* Volume 37, p. 185–199.

Chen, J. et al., 2003. Remodeling of cardiac fiber structure after infarction in rats quantified with diffusion tensor MRI. *American Journal of Physiology-Heart and Circulatory Physiology,* Volume 285, p. H946–H954.

Chen, Y. et al., 2018. *Efficient and accurate MRI super-resolution using a generative adversarial network and 3D multi-level densely connected network.* s.l., s.n., p. 91–99.

Clark, C. A. & Le Bihan, D., 2000. Water diffusion compartmentation and anisotropy at high b values in the human brain. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine,* Volume 44, p. 852–859.

Dar, S. U. H. et al., 2018. Synergistic reconstruction and synthesis via generative adversarial networks for accelerated multi-contrast MRI. *arXiv preprint arXiv:1805.10704.*

De Santis, S., Assaf, Y., Evans, C. J. & Jones, D. K., 2011. *Improved precision in the charmed model of white matter through sampling scheme optimization and model parsimony testing.* s.l., s.n., p. 3928.

Deng, Z. et al., 2020. Investigation of in Vivo Human Cardiac Diffusion Tensor Imaging Using Unsupervised Dense Encoder-Fusion-Decoder Network. *IEEE Access*, Volume 8, p. 220140–220151.

Denton, E. L., Chintala, S., Fergus, R. & others, 2015. Deep generative image models using a laplacian pyramid of adversarial networks. *Advances in neural information processing systems*, Volume 28.

Descoteaux, M., 1999. High angular resolution diffusion imaging (HARDI). *Wiley encyclopedia of electrical and electronics engineering*, pp. 1--25.

Descoteaux, M., 2008. *High angular resolution diffusion MRI: from local estimation to segmentation and tractography*, s.l.: s.n.

Descoteaux, M., Angelino, E., Fitzgibbons, S. & Deriche, R., 2007. Regularized, fast, and robust analytical Q-ball imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, Volume 58, p. 497–510.

Descoteaux, M. et al., 2009. *Diffusion propagator imaging: using Laplace's equation and multiple shell acquisitions to reconstruct the diffusion propagator*. s.l., s.n., p. 1–13.

Donahue, J., Krähenbühl, P. & Darrell, T., 2016. Adversarial feature learning. *arXiv preprint arXiv:1605.09782*.

Dong, C., Loy, C. C., He, K. & Tang, X., 2015. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, Volume 38, p. 295–307.

Dou, J., Reese, T. G., Tseng, W.-Y. I. & Wedeen, V. J., 2002. Cardiac diffusion MRI without motion effects. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, Volume 48, p. 105–114.

Edelman, R. R. et al., 1994. In vivo measurement of water diffusion in the human heart. *Magnetic resonance in medicine*, Volume 32, p. 423–428.

Einstein, A., 1956. *Investigations on the Theory of the Brownian Movement*. s.l.:Courier Corporation.

Frindel, C., Robini, M., Croisille, P. & Zhu, Y.-M., 2009. Comparison of regularization methods for human cardiac diffusion tensor MRI. *Medical image analysis*, Volume 13, p. 405–418.

Frindel, C. et al., 2010. A graph-based approach for automatic cardiac tractography. *Magnetic Resonance in Medicine*, Volume 64, p. 1215–1229.

Gamper, U., Boesiger, P. & Kozerke, S., 2007. Diffusion imaging of the in vivo heart using spin echoes–considerations on bulk motion sensitivity. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, Volume 57, p. 331–337.

Geerts, L., Bovendeerd, P., Nicolay, K. & Arts, T., 2002. Characterization of the normal cardiac myofiber field in goat measured with MR-diffusion tensor imaging. *American Journal of Physiology-Heart and Circulatory Physiology*, Volume 283, p. H139–H145.

Goldman, M., 2004. Histoire de la RMN. *L'actualité chimique,* Volume 273, p. 57–60.

Goodfellow, I. et al., 2014. *Generative adversarial nets*. s.l., s.n., p. 2672–2680.

Greenbaum, R. A. et al., 1981. Left ventricular fibre architecture in man.. *Heart*, Volume 45, p. 248–263.

Gudbjartsson, H. & Patz, S., 1995. The Rician distribution of noisy MRI data. *Magnetic resonance in medicine*, Volume 34, p. 910–914.

Gui, J. et al., 2021. A review on generative adversarial networks: Algorithms, theory, and applications. *IEEE Transactions on Knowledge and Data Engineering.*

Hagmann, P. et al., 2006. Understanding diffusion MR imaging techniques: from scalar diffusion-weighted imaging to diffusion tensor imaging and beyond. *Radiographics,* Volume 26, p. S205–S223.

Hahn, E. L., 1950. Spin echoes. *Physical review,* Volume 80, p. 580.

Han, C. et al., 2018. *GAN-based synthetic brain MR image generation.* s.l., s.n., p. 734–738.

Heidemann, R. M. et al., 2012. k-space and q-space: combining ultra-high spatial and angular resolution in diffusion imaging using ZOOPPA at 7 T. *Neuroimage,* Volume 60, p. 967–978.

Helm, P. A. et al., 2005. Ex vivo 3D diffusion tensor imaging and quantification of cardiac laminar structure. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine,* Volume 54, p. 850–859.

Henkelman, R. M., 1985. Measurement of signal intensities in the presence of noise in MR images. *Medical physics,* Volume 12, p. 232–233.

He, X. & Niyogi, P., 2004. *Locality preserving projections.* s.l., s.n., p. 153–160.

He, Y. et al., 2020. *Deep Group-Wise Angular Translation of Cardiac Diffusion MRI in q-space via Manifold Regularized GAN.* s.l., s.n., p. 511–515.

He, Y. et al., 2021. *Systematic Study of Joint Influence of Angular Resolution and Noise in Cardiac Diffusion Tensor Imaging.* s.l., s.n., p. 200–210.

Hiasa, Y. et al., 2018. *Cross-modality image synthesis from unpaired data using CycleGAN.* s.l., s.n., p. 31–41.

Huo, Y. et al., 2018. *Adversarial synthesis learning enables segmentation without target modality ground truth.* s.l., s.n., p. 1217–1220.

Hüppi, P. S. & Dubois, J., 2006. *Diffusion tensor imaging of brain development.* s.l., s.n., p. 489–497.

Hurwitz, R., Lane, S. R., Bell, R. A. & Brant-Zawadzki, M. N., 1989. Acoustic analysis of gradient-coil noise in MR imaging.. *Radiology,* Volume 173, p. 545–548.

Ioacara, S. et al., 2020. Acute myocardial infarction mortality rates and trends in Romania between 1994 and 2017. *International Journal of Environmental Research and Public Health,* Volume 17, p. 285.

Ioffe, S. & Szegedy, C., 2015. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167.*

Isola, P., Zhu, J.-Y., Zhou, T. & Efros, A. A., 2017. *Image-to-image translation with conditional adversarial networks.* s.l., s.n., p. 1125–1134.

Jefferies, J. L. & Towbin, J. A., 2010. Dilated cardiomyopathy. *The Lancet,* Volume 375, p. 752–762.

Johansen-Berg, H. & Behrens, T. E. J., 2013. *Diffusion MRI: from quantitative measurement to in vivo neuroanatomy.* s.l.:Academic Press.

Jones, D. K., 2004. The effect of gradient sampling schemes on measures derived from diffusion tensor MRI: a Monte Carlo study. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine,* Volume 51, p. 807–815.

Jones, D. K. & Basser, P. J., 2004. “Squashing peanuts and smashing pumpkins”: how noise distorts diffusion-weighted MR data. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine,* Volume 52, p. 979–993.

Jones, D. K. & Cercignani, M., 2010. Twenty-five pitfalls in the analysis of diffusion MRI data. *NMR in Biomedicine*, Volume 23, p. 803–820.

Jones, D. K., Horsfield, M. A. & Simmons, A., 1999. Optimal strategies for measuring diffusion in anisotropic systems by magnetic resonance imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, Volume 42, p. 515–525.

Jung, H. et al., 2020. Unsupervised deep image fusion with structure tensor representations. *IEEE Transactions on Image Processing*, Volume 29, p. 3845–3858.

Karacan, L., Akata, Z., Erdem, A. & Erdem, E., 2016. Learning to generate images of outdoor scenes from attributes and semantic layouts. *arXiv preprint arXiv:1612.00215*.

Kim, J., Kwon Lee, J. & Mu Lee, K., 2016. *Deeply-recursive convolutional network for image super-resolution.* s.l., s.n., p. 1637–1645.

Kim, K. H., Do, W.-J. & Park, S.-H., 2018. Improving resolution of MR images with an adversarial network incorporating images with different contrast. *Medical physics*, Volume 45, p. 3120–3131.

Kingsley, P. B., 2006a. Introduction to diffusion tensor imaging mathematics: Part I. Tensors, rotations, and eigenvectors. *Concepts in Magnetic Resonance Part A*, Volume 28, p. 101–122.

Kingsley, P. B., 2006b. Introduction to diffusion tensor imaging mathematics: Part II. Anisotropy, diffusion-weighting factors, and gradient encoding schemes. *Concepts in Magnetic Resonance Part A*, Volume 28, p. 123–154.

Kingsley, P. B., 2006c. Introduction to diffusion tensor imaging mathematics: Part III. Tensor calculation, noise, simulations, and optimization. *Concepts in Magnetic Resonance Part A*, Volume 28, p. 155–179.

Kingsley, P. B., 2006. Introduction to diffusion tensor imaging mathematics: Part I. Tensors, rotations, and eigenvectors. *Concepts in Magnetic Resonance Part A*, Volume 28, p. 101–122.

Kingsley, P. B., 2006. Introduction to diffusion tensor imaging mathematics: Part II. Anisotropy, diffusion-weighting factors, and gradient encoding schemes. *Concepts in Magnetic Resonance Part A*, Volume 28, p. 123–154.

Kingsley, P. B., 2006. Introduction to diffusion tensor imaging mathematics: Part III. Tensor calculation, noise, simulations, and optimization. *Concepts in Magnetic Resonance Part A*, Volume 28, p. 155–179.

Kingsley, P. B., 2006. Introduction to diffusion tensor imaging mathematics: Part III. Tensor calculation, noise, simulations, and optimization. *Concepts in Magnetic Resonance Part A*, Volume 28, p. 155–179.

Knight, F. B., 1962. On the random walk and Brownian motion. *Transactions of the American Mathematical Society*, Volume 103, p. 218–228.

Kumar, R. et al., 2009. Comparative evaluation of corpus callosum DTI metrics in acute mild and moderate traumatic brain injury: its correlation with neuropsychometric tests. *Brain injury*, Volume 23, p. 675–685.

Le Bihan, D., 1990. Magnetic resonance imaging of perfusion. *Magnetic resonance in medicine*, Volume 14, p. 283–292.

Le Bihan, D. et al., 1986. MR imaging of intravoxel incoherent motions: application to diffusion and perfusion in neurologic disorders.. *Radiology*, Volume 161, p. 401–407.

Le Bihan, D. I. V. I. M., 1990. IVIM method measures diffusion and perfusion..

Le Bihan, D. et al., 2001. Diffusion tensor imaging: concepts and applications. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, Volume 13, p. 534–546.

LeCun, Y. et al., 1989. Handwritten digit recognition with a back-propagation network. *Advances in neural information processing systems*, Volume 2.

LeCun, Y., Bottou, L., Bengio, Y. & Haffner, P., 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, Volume 86, p. 2278–2324.

Lenglet, C., 2006. *Geometric and variational methods for diffusion tensor mri processing*, s.l.: s.n.

Lenglet, C. et al., 2009. Mathematical methods for diffusion MRI processing. *NeuroImage*, Volume 45, p. S111–S122.

Levitt, M. H., 2013. *Spin dynamics: basics of nuclear magnetic resonance.* s.l.:John Wiley & Sons.

Likes, R. S., 1981. *Moving gradient zeugmatography*, s.l.: s.n.

Li, Q. et al., 2018. MR-GAN: Manifold regularized generative adversarial networks. *arXiv preprint arXiv:1811.10427*.

Ljunggren, S., 1983. A simple graphical representation of Fourier-based imaging methods. *Journal of magnetic resonance*, Volume 54, p. 338–343.

Lombaert, H. et al., 2012. Human atlas of the cardiac fiber architecture: study on a healthy population. *IEEE transactions on medical imaging*, Volume 31, p. 1436–1447.

Maas, A. L., Hannun, A. Y., Ng, A. Y. & others, 2013. *Rectifier nonlinearities improve neural network acoustic models.* s.l., s.n., p. 3.

Mani, M. et al., 2015. Acceleration of high angular and spatial resolution diffusion imaging using compressed sensing with multichannel spiral data. *Magnetic resonance in medicine*, Volume 73, p. 126–138.

Mao, X. et al., 2017. *Least squares generative adversarial networks.* s.l., s.n., p. 2794–2802.

Mardani, M. et al., 2017. Deep generative adversarial networks for compressed sensing automates MRI. *arXiv preprint arXiv:1706.00051*.

Maron, B. J. & Maron, M. S., 2013. Hypertrophic cardiomyopathy. *The Lancet*, Volume 381, p. 242–255.

Maspero, M. et al., 2018. Dose evaluation of fast synthetic-CT generation using a generative adversarial network for general pelvis MR-only radiotherapy. *Physics in Medicine & Biology*, Volume 63, p. 185001.

Mathieu, M., Couprie, C. & LeCun, Y., 2015. Deep multi-scale video prediction beyond mean square error. *arXiv preprint arXiv:1511.05440.*

McGill, L.-A.et al., 2012. Reproducibility of in-vivo diffusion tensor cardiovascular magnetic resonance in hypertrophic cardiomyopathy. *Journal of Cardiovascular Magnetic Resonance*, Volume 14, p. 1–15.

Mekkaoui, C. et al., 2013. Myocardial infarct delineation in vivo using diffusion tensor MRI and the tractographic propagation angle. *Journal of Cardiovascular Magnetic Resonance*, Volume 15, p. 1–3.

Mekkaoui, C. et al., 2017. Diffusion MRI in the heart. *NMR in Biomedicine*, Volume 30, p. e3426.

Mekkaoui, C. et al., 2017. Diffusion MRI in the heart. *NMR in Biomedicine*, Volume 30, p. e3426.

Mirza, M. & Osindero, S., 2014. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784.*

Mitchell, H. H., Hamilton, T. S., Steggerda, F. R. & Bean, H. W., 1945. The chemical composition of the adult human body and its bearing on the biochemistry of growth. *Journal of Biological Chemistry*, Volume 158, p. 625–637.

Miyato, T. & Koyama, M., 2018. cGANs with projection discriminator. *arXiv preprint arXiv:1802.05637.*

Mondal, A. K., Dolz, J. & Desrosiers, C., 2018. Few-shot 3d multi-modal medical image segmentation using generative adversarial learning. *arXiv preprint arXiv:1810.12241*.

Mori, S. & Barker, P. B., 1999. Diffusion magnetic resonance imaging: its principle and applications. *The Anatomical Record: An Official Publication of the American Association of Anatomists*, Volume 257, p. 102–109.

Mori, S. & Van Zijl, P. C. M., 2002. Fiber tracking: principles and strategies–a technical review. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In Vivo*, Volume 15, p. 468–480.

Moseley, M. E. et al., 1990. Early detection of regional cerebral ischemia in cats: comparison of diffusion-and T2-weighted MRI and spectroscopy. *Magnetic resonance in medicine*, Volume 14, p. 330–346.

Munkres, J., 1957. Algorithms for the assignment and transportation problems. *Journal of the society for industrial and applied mathematics*, Volume 5, p. 32–38.

Neil, J., Miller, J., Mukherjee, P. & Hüppi, P. S., 2002. Diffusion tensor imaging of normal and injured developing human brain-a technical review. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In Vivo*, Volume 15, p. 543–552.

Nguyen, C. et al., 2014. In vivo three-dimensional high resolution cardiac diffusion-weighted MRI: a motion compensated diffusion-prepared balanced steady-state free precession approach. *Magnetic resonance in medicine*, Volume 72, p. 1257–1267.

Nie, D. a. T. R. a. L. J. a. W. L. a. P. C. a. R. S. a. W. Q. a. S. D., 2018. Medical image synthesis with deep convolutional adversarial networks. *IEEE Transactions on Biomedical Engineering*, Volume 65, pp. 2720--2730.

Nie, D. et al., 2018. Medical image synthesis with deep convolutional adversarial networks. *IEEE Transactions on Biomedical Engineering*, Volume 65, p. 2720–2730.

Nielles-Vallespin, S. et al., 2013. In vivo diffusion tensor MRI of the human heart: reproducibility of breath-hold and navigator-based approaches. *Magnetic resonance in medicine*, Volume 70, p. 454–465.

Nowozin, S., Cseke, B. & Tomioka, R., 2016. f-gan: Training generative neural samplers using variational divergence minimization. *Advances in neural information processing systems*, Volume 29.

Oksuz, I. et al., 2018. *Cardiac MR motion artefact correction from k-space using deep learning-based reconstruction.* s.l., s.n., p. 21–29.

Özarslan, E. et al., 2006. Resolution of complex tissue microarchitecture using the diffusion orientation transform (DOT). *NeuroImage*, Volume 31, p. 1086–1103.

Pai, V. a. R. S. a. K. P. a. C. P. a. W. H., 2011. PCATMIP: enhancing signal intensity in diffusion-weighted magnetic resonance imaging. *Magnetic Resonance in Medicine*, Volume 65, pp. 1611--1619.

Pai, V. M. et al., 2011. PCATMIP: enhancing signal intensity in diffusion-weighted magnetic resonance imaging. *Magnetic Resonance in Medicine,* Volume 65, p. 1611–1619.

Pan, Y. et al., 2018. *Synthesizing missing PET from MRI with cycle-consistent generative adversarial networks for Alzheimer's disease diagnosis.* s.l., s.n., p. 455–463.

Papadakis, N. G. et al., 2000. Minimal gradient encoding for robust estimation of diffusion anisotropy. *Magnetic resonance imaging,* Volume 18, p. 671–679.

Pathak, D. et al., 2016. *Context encoders: Feature learning by inpainting.* s.l., s.n., p. 2536–2544.

Peyrat, J.-M.et al., 2007. A computational framework for the statistical analysis of cardiac diffusion tensors: application to a small database of canine hearts. *IEEE Transactions on Medical Imaging,* Volume 26, p. 1500–1514.

Pierpaoli, C. & Basser, P. J., 1996. Toward a quantitative assessment of diffusion anisotropy. *Magnetic resonance in Medicine,* Volume 36, p. 893–906.

Qin, Y. et al., 2019. Pulmonary nodule segmentation with CT sample synthesis using adversarial networks. *Medical physics,* Volume 46, p. 1218–1229.

Quan, T. M., Nguyen-Duc, T. & Jeong, W.-K., 2018. Compressed sensing MRI reconstruction using a generative adversarial network with a cyclic loss. *IEEE transactions on medical imaging,* Volume 37, p. 1488–1497.

Radford, A., Metz, L. & Chintala, S., 2015. Unsupervised representation learning with deep convolutional generative adversarial networks. *arXiv preprint arXiv:1511.06434.*

Ran, M. et al., 2019. Denoising of 3D magnetic resonance images using a residual encoder–decoder Wasserstein generative adversarial network. *Medical image analysis,* Volume 55, p. 165–180.

Reed, S. et al., 2016. *Generative adversarial text to image synthesis.* s.l., s.n., p. 1060–1069.

Reed, S. E. et al., 2016. Learning what and where to draw. *Advances in neural information processing systems,* Volume 29.

Reese, T. G. et al., 1995. Imaging myocardial fiber architecture in vivo with magnetic resonance. *Magnetic Resonance in Medicine,* Volume 34, p. 786–791.

Ribeiro, M., Lazzaretti, A. E. & Lopes, H. S., 2018. A study of deep convolutional auto-encoders for anomaly detection in videos. *Pattern Recognition Letters,* Volume 105, p. 13–22.

Ronneberger, O., Fischer, P. & Brox, T., 2015. *U-net: Convolutional networks for biomedical image segmentation.* s.l., s.n., p. 234–241.

Rosen, B. & Wald, L., 2006. MR image encoding. *Harvard-MIT Division of Health Sciences and Technology.*

Sánchez, I. & Vilaplana, V., 2018. Brain MRI super-resolution using 3D generative adversarial networks. *arXiv preprint arXiv:1812.11440.*

Sara, U., Akter, M. & Uddin, M. S., 2019. Image quality assessment through FSIM, SSIM, MSE and PSNR—a comparative study. *Journal of Computer and Communications,* Volume 7, p. 8–18.

Schmid, P. et al., 2005. Ventricular myocardial architecture as visualised in postmortem swine hearts using magnetic resonance diffusion tensor imaging. *European journal of cardio-thoracic surgery,* Volume 27, p. 468–472.

Scott, A. D. et al., 2015. Optimal diffusion weighting for in vivo cardiac diffusion tensor imaging. *Magnetic resonance in medicine,* Volume 74, p. 420–430.

Shen, D., Wu, G. & Suk, H.-I., 2017. Deep learning in medical image analysis. *Annual review of biomedical engineering*, Volume 19, p. 221–248.

Shitrit, O. & Riklin Raviv, T., 2017. Accelerated magnetic resonance imaging by adversarial neural network. In: *Deep learning in medical image analysis and multimodal learning for clinical decision support*. s.l.:Springer, p. 30–38.

Sosnovik, D. E. et al., 2014. Microstructural impact of ischemia and bone marrow–derived cell therapy revealed with diffusion tensor magnetic resonance imaging tractography of the heart in vivo. *Circulation*, Volume 129, p. 1731–1741.

Sosnovik, D. E. et al., 2009. Diffusion spectrum MRI tractography reveals the presence of a complex network of residual myofibers in infarcted myocardium. *Circulation: Cardiovascular Imaging*, Volume 2, p. 206–212.

Stejskal, E. O., 1965. Use of spin echoes in a pulsed magnetic-field gradient to study anisotropic, restricted diffusion and flow. *The Journal of Chemical Physics*, Volume 43, p. 3597–3603.

Stejskal, E. O. & Tanner, J. E., 1965. Spin diffusion measurements: spin echoes in the presence of a time-dependent field gradient. *The journal of chemical physics*, Volume 42, p. 288–292.

Stoeck, C. T. et al., 2014. Dual-phase cardiac diffusion tensor imaging with strain correction. *PLoS One*, Volume 9, p. e107159.

Stoeck, C. T. et al., 2016. Second-order motion-compensated spin echo diffusion tensor imaging of the human heart. *Magnetic resonance in medicine*, Volume 75, p. 1669–1676.

Streeter Jr, D. D. & Hanna, W. T., 1973. Engineering mechanics for successive states in canine left ventricular myocardium: I. Cavity and wall geometry. *Circulation research*, Volume 33, p. 639–655.

Streeter Jr, D. D. et al., 1969. Fiber orientation in the canine left ventricle during diastole and systole. *Circulation research*, Volume 24, p. 339–347.

Strijkers, G. J. et al., 2009. Diffusion tensor imaging of left ventricular remodeling in response to myocardial infarction in the mouse. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In vivo*, Volume 22, p. 182–190.

Taylor, W. D., Hsu, E., Krishnan, K. R. R. & MacFall, J. R., 2004. Diffusion tensor imaging: background, potential, and utility in psychiatric research. *Biological psychiatry*, Volume 55, p. 201–207.

Teanby, N. A., 2006. An icosahedron-based method for even binning of globally distributed remote sensing data. *Computers & Geosciences*, Volume 32, p. 1442–1450.

Torrey, H. C., 1956. Bloch equations with diffusion terms. *Physical review*, Volume 104, p. 563.

Tournier, J.-D., Calamante, F. & Connelly, A., 2013. Determination of the appropriate b value and number of gradient directions for high-angular-resolution diffusion-weighted imaging. *NMR in Biomedicine*, Volume 26, p. 1775–1786.

Tournier, J.-D., Calamante, F., Gadian, D. G. & Connelly, A., 2004. Direct estimation of the fiber orientation density function from diffusion-weighted MRI data using spherical deconvolution. *Neuroimage*, Volume 23, p. 1176–1185.

Toussaint, N. et al., 2013. In vivo human cardiac fibre architecture estimation using shape-based diffusion tensor processing. *Medical image analysis*, Volume 17, p. 1243–1255.

Tseng, W.-Y. I., Dou, J., Reese, T. G. & Wedeen, V. J., 2006. Imaging myocardial fiber disarray and intramural strain hypokinesis in hypertrophic cardiomyopathy with MRI. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine,* Volume 23, p. 1–8.

Tseng, W.-Y.I. et al., 2000. Myocardial fiber shortening in humans: initial results of MR imaging. *Radiology,* Volume 216, p. 128–139.

Tseng, W.-Y. I., Reese, T. G., Weisskoff, R. M. & Wedeen, V. J., 1999. Cardiac diffusion tensor MRI in vivo without strain correction. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine,* Volume 42, p. 393–403.

Tuch, D. S., 2004. Q-ball imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine,* Volume 52, p. 1358–1372.

Tuch, D. S. et al., 2002. High angular resolution diffusion imaging reveals intravoxel white matter fiber heterogeneity. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine,* Volume 48, p. 577–582.

Tuch, D. S., Reese, T. G., Wiegell, M. R. & Wedeen, V. J., 2003. Diffusion MRI of complex neural architecture. *Neuron,* Volume 40, p. 885–895.

Twieg, D. B., 1983. The k-trajectory formulation of the NMR imaging process with applications in analysis and synthesis of imaging methods. *Medical physics,* Volume 10, p. 610–621.

Veraart, J. et al., 2016. Denoising of diffusion MRI using random matrix theory. *Neuroimage,* Volume 142, p. 394–406.

Vovk, U., Pernus, F. & Likar, B., 2007. A review of methods for correction of intensity inhomogeneity in MRI. *IEEE transactions on medical imaging,* Volume 26, p. 405–421.

Wang, X. & Gupta, A., 2016. *Generative image modeling using style and structure adversarial networks.* s.l., s.n., p. 318–335.

Wang, Z., Bovik, A. C., Sheikh, H. R. & Simoncelli, E. P., 2004. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing,* Volume 13, p. 600–612.

Wang, Z., She, Q. & Ward, T. E., 2021. Generative adversarial networks in computer vision: A survey and taxonomy. *ACM Computing Surveys (CSUR),* Volume 54, p. 1–38.

Wedeen, V. J. et al., 2005. Mapping complex tissue architecture with diffusion spectrum magnetic resonance imaging. *Magnetic resonance in medicine,* Volume 54, p. 1377–1386.

Wei, H. et al., 2014. Free-breathing diffusion tensor imaging and tractography of the human heart in healthy volunteers using wavelet-based image fusion. *IEEE transactions on medical imaging,* Volume 34, p. 306–316.

Wei, H. et al., 2013. Assessment of cardiac motion effects on the fiber architecture of the human heart in vivo. *IEEE transactions on medical imaging,* Volume 32, p. 1928–1938.

Wei, W. et al., 2018. *Learning myelin content in multiple sclerosis from multimodal MRI through adversarial training.* s.l., s.n., p. 514–522.

Westbrook, C. & Talbot, J., 2018. *MRI in Practice.* s.l.:John Wiley & Sons.

Wolterink, J. M. et al., 2017. *Deep MR to CT synthesis using unpaired data.* s.l., s.n., p. 14–23.

Wu, E. X. et al., 2007. MR diffusion tensor imaging study of postinfarct myocardium structural remodeling in a porcine model. *Magnetic Resonance in Medicine: An Official*

Journal of the International Society for Magnetic Resonance in Medicine, Volume 58, p. 687–695.

Wu, M.-T.et al., 2009. Sequential changes of myocardial microstructure in patients postmyocardial infarction by diffusion-tensor cardiac MR: correlation with left ventricular structure and function. *Circulation: Cardiovascular Imaging,* Volume 2, p. 32–40.

Wu, M.-T.et al., 2006. Diffusion tensor magnetic resonance imaging mapping the fiber architecture remodeling in human myocardium after infarction: correlation with viability and wall motion. *Circulation,* Volume 114, p. 1036–1045.

Wu, Y.-C. & Alexander, A. L., 2007. Hybrid diffusion imaging. *NeuroImage,* Volume 36, p. 617–629.

Xu, H. et al., 2020. U2Fusion: A Unified Unsupervised Image Fusion Network. *IEEE Transactions on Pattern Analysis and Machine Intelligence.*

Yang, F. et al., 2012. Feature-based interpolation of diffusion tensor fields and application to human cardiac DT-MRI. *Medical image analysis,* Volume 16, p. 459–481.

Yang, F. et al., 2018. Quantitative comparison of human myocardial fiber orientations derived from DTI and polarized light imaging. *Physics in Medicine & Biology,* Volume 63, p. 215003.

Yang, F. et al., 2018. Quantitative comparison of human myocardial fiber orientations derived from DTI and polarized light imaging. *Physics in Medicine \& Biology,* Volume 63, p. 215003.

Yang, F. et al., 2011. Interpolation of vector fields from human cardiac DT-MRI. *Physics in Medicine & Biology,* Volume 56, p. 1415.

Yap, P.-T., An, H., Chen, Y. & Shen, D., 2014. Uncertainty estimation in diffusion MRI using the nonlocal bootstrap. *IEEE Transactions on Medical Imaging,* Volume 33, p. 1627–1640.

Yi, X., Walia, E. & Babyn, P., 2019. Generative adversarial network in medical imaging: A review. *Medical image analysis,* Volume 58, p. 101552.

Zanjani, F. G. et al., 2018. Histopathology stain-color normalization using deep generative models.

Zhang, L., Gooya, A. & Frangi, A. F., 2017. *Semi-supervised assessment of incomplete LV coverage in cardiac MRI using generative adversarial nets.* s.l., s.n., p. 61–68.

Zhang, S. et al., 2010. The correlation of 3D DT-MRI fiber disruption with structural and mechanical degeneration in porcine myocardium. *Annals of biomedical engineering,* Volume 38, p. 3084–3095.

Zhang, Y.-L., Liu, W.-Y., Magnin, I. E. & Zhu, Y.-M., 2013. Feature-preserving smoothing of diffusion weighted images using nonstationarity adaptive filtering. *IEEE Transactions on Biomedical Engineering,* Volume 60, p. 1693–1701.

Zhang, Z., Yang, L. & Zheng, Y., 2018. *Translating and segmenting multimodal medical volumes with cycle-and shape-consistency generative adversarial network.* s.l., s.n., p. 9242–9251.

Zhan, L. et al., 2013. Angular versus spatial resolution trade-offs for diffusion imaging under time constraints. *Human brain mapping,* Volume 34, p. 2688–2706.

Zhan, L. et al., 2010. How does angular resolution affect diffusion imaging measures?. *neuroimage,* Volume 49, p. 1357–1371.

Zhou, T., Jae Lee, Y., Yu, S. X. & Efros, A. A., 2015. *Flowweb: Joint image set alignment by weaving consistent, pixel-wise correspondences.* s.l., s.n., p. 1191–1200.

Zhou, Y. & Berg, T. L., 2016. *Learning temporal transformations from time-lapse videos.* s.l., s.n., p. 262–277.

Zhu, J.-Y., Krähenbühl, P., Shechtman, E. & Efros, A. A., 2016. *Generative visual manipulation on the natural image manifold.* s.l., s.n., p. 597–613.

Zhu, J.-Y., Park, T., Isola, P. & Efros, A. A., 2017. *Unpaired image-to-image translation using cycle-consistent adversarial networks.* s.l., s.n., p. 2223–2232.

Zhukov, L. & Barr, A. H., 2003. *Heart-muscle fiber reconstruction from diffusion tensor MRI.* s.l., s.n., p. 597–602.

FOLIO ADMINISTRATIF

THESE DE L'INSA LYON, MEMBRE DE L'UNIVERSITE DE LYON

NOM: HE DATE de SOUTENANCE: 19/10/ 2022

Prénoms: Yunlong

TITRE: Generation of diffusion tensor imaging data of the human heart using deep convolutional adversarial networks

NATURE: Doctorat Numéro d'ordre : 2022ISAL0087

Ecole doctorale : Electronique, Électrotechnique, Automatique (EEA) – ED160

Spécialité: Traitement du Signal et de l'Image

RESUME: L'imagerie du tenseur de diffusion (ITD) est récemment apparue comme une modalité d'imagerie médicale unique permettant d'étudier de manière non invasive l'architecture des fibres du cœur humain, offrant ainsi un nouvel outil pour le diagnostic des maladies cardiaques. Cependant, les applications cliniques de l'ITD cardiaque sont souvent pénalisées par la quantité et/ou la qualité limitées des images pondérées en diffusion (diffusion weighted—DW) acquises en raison des temps d'acquisition limités en pratique. Les récentes avancées dans l'apprentissage profond, en particulier dans les réseaux adversariaux génératifs (Generative neural network—GAN), laissent entrevoir une nouvelle façon de faire face à ce problème. L'objectif de cette thèse est d'aider à améliorer l'efficacité de l'acquisition et l'applicabilité clinique de l'ITD cardiaque en étudiant les méthodes de génération d'images basées sur le GAN. À cette fin, une méthode de super résolution angulaire (SRA) pour les données ITD cardiaques a d'abord été proposée pour améliorer la précision de l'estimation des tenseurs de diffusion. Ensuite, la méthode SAR développée a été appliquée pour compenser la perte de signal induite par le mouvement dans l'ITD cardiaque in vivo. Notre travail comporte les quatre contributions principales suivantes : 1) Nous avons étudié systématiquement l'influence conjointe de la résolution angulaire et du bruit sur l'ITD cardiaque, démontrant ainsi la validité de la méthode SAR pour améliorer la précision de l'estimation des tenseurs de diffusion. 2) Nous avons proposé un cadre SRA basé sur le GAN, qui peut être appliqué sur des données ITD cardiaques aussi bien simulées que réelles pour améliorer la précision de l'estimation du tenseur de diffusion. 3) Nous avons étudié la perte de signal induite par le mouvement dans l'ITD cardiaque in vivo et fourni une nouvelle méthode d'ajustement des données capable de générer des images DW réalistes induites par le mouvement directement à partir des ensembles de données ITD cardiaques existants. 4) Nous avons proposé un nouveau SRA basé sur le GAN qui peut produire des images DW à haute résolution angulaire à partir d'images DW à faible résolution angulaire et corrompues par le mouvement afin de compenser la perte de signal induite par le mouvement dans l'ITD cardiaque in vivo.

MOTS-CLÉS: Apprentissage profond, imagerie par résonance magnétique, réseaux adversariens génératifs, super résolution, imagerie cardiaque.

Laboratoire (s) de recherche: CREATIS
Directeur de thèse: ZHU Yue-Min
Président de jury : CLARYSSE Patrick
Composition du jury: RUAN Su (rapporteuse de thèse), YANG Jie (rapporteur de thèse), JOUK Pierre-Simon (examinateur), USSON Yves (examinateur), CLARYSSE Patrick (examinateur), ZHU Yue-Min (directeur de thèse), WANG Lihui (co-directrice de thèse).