
بسم الله الرحمن الرحيم

بهبود تصویرهای MRI با استفاده از روش MWACNN و مقایسه با روش U-NET

MRI Image Enhancement Using the MWACNN Method and Comparison with U-Net

معاونت پژوهش و فن آوری

با نام خدا

منشور اخلاق پژوهش

با یاری از خداوند سبحان و اعتقاد به اینکه عالم محضر خداست و همواره ناظر بر اعمال انسان و به منظور پاس داشت مقام بلند دانش و نظر به اهمیت جایگاه دانشگاه در اعتلای فرهنگ و تمدن بشری، ما دانشجویان و اعضای هیئت علمی واحدهای دانشگاه آزاد اسلامی تعهد می‌گردیم اصول زیر را ر انجام فعالیت‌های پژوهشی مد نظر قرار داده و از آن تخطی نکنیم:

۱- اصل حقیقت جویی: تلاش در راستای پی جویی حقیقت و وفاداری به آن و دوری از هرگونه پنهان سازی حقیقت.

۲- اصل رعایت حقوق: التزام به رعایت کامل حقوق پژوهشگران و پژوهیدگان (انسان، حیوان و نبات) و سایر صاحبان حق.

۳- اصل مالکیت مادی و معنوی: تعهد به رعایت کامل حقوق مادی و معنوی دانشگاه و کلیه همکاران پژوهش.

٤- اصل منافع ملی: تعهد به رعایت مصالح ملی و در نظر داشتن پیشبرد و توسعه کشور در کلیه مراحل پژوهش.

٥- اصل رعایت انصاف و امانت: تعهد به اجتناب از هرگونه جانب داری غیر علمی و حفاظت از اموال ، تجهیزات و منابع در اختیار.

٦- اصل راز داری: تعهد به صیانت از اسرار و اطلاعات محرمانه افراد، سازمان‌ها و کشور و کلیه افراد و نهادهای مرتبط در اختیار.

۷- اصل احترام: تعهد به رعایت حریم‌ها حرمت‌ها در انجام تحقیقات و رعایت جانب نقد و خودداری از هرگونه حرمت شکنی.

۸- اصل ترویج: تعهد به رواج دانش و اشاعه نتایج تحقیقات و انتقال آن به همکاران علمی و دانشجویان به غیر از مواردی که منع قانونی دارد.

۹- اصل برائت: التزام به برائت جویی از هرگونه رفتار غیر حرفه‌ای و اعلام موضع نسبت به کسانی که حوزه علم و پژوهش را شائبه‌های غیر علمی آن می‌آلایند.

دانشگاه آزاد اسلامی

واحد علوم و تحقیقات

اینجانب سید حسین طاهری دانش‌آموخته مقطع کارشناسی ارشد ناپیوسته در رشته علوم کامپیوتر گرایش محاسبات نرم و هوش مصنوعی که در تاریخ ..۱۴۰۵/۰۶/۲۱.. از پایان‌نامه / رساله خود تحت عنوان

<< بهبود تصویرهای MRI با استفاده از روش MWACNN و مقایسه با روش UNET >>

با کسب نمره ....۱۹.۷۵.... و درجه ..... دفاع نموده‌ام، بدین‌وسیله تعهد می‌شوم:

۱. این پایان‌نامه / رساله حاصل تحقیق و پژوهش انجام‌شده توسط اینجانب بوده و در مواردی که از دستاوردهای علمی و پژوهشی دیگران (اعم از پایان‌نامه، کتاب، مقاله و ...) استفاده نموده‌ام، مطابق ضوابط و رویه موجود، نام منبع مورد استفاده و سایر مشخصات آن را در فهرست مربوطه ذکر و درج کرده‌ام .

۲. این پایان‌نامه / رساله قبلاً برای دریافت هیچ مدرک تحصیلی (هم‌سطح، پایین‌تر یا بالاتر) در سایر دانشگاه‌ها و مؤسسات آموزش عالی ارائه نشده است .

۳. چنانچه بعد از فراغت از تحصیل قصد استفاده و هرگونه بهره‌برداری اعم از چاپ کتاب، ثبت اختراع و ... از این پایان‌نامه / رساله داشته باشم، از حوزه معاونت پژوهشی واحد مجوزهای مربوطه را اخذ نمایم .

٤. چنانچه در هر مقطع زمانی خلاف موارد فوق ثابت شود، عواقب ناشی از آن را می‌پذیرم و دانشگاه آزاد اسلامی واحد علوم و تحقیقات مجاز است با اینجانب مطابق ضوابط و مقررات رفتار نموده و در صورت ابطال مدرک تحصیلی، هیچ‌گونه ادعایی نخواهم داشت .

نام و نام خانوادگی : سید حسین طاهری

تاریخ و امضاء





دانشگاه آزاد اسلامی

واحد علوم و تحقیقات

دانشکده علوم و فناوری‌های همگرا، گروه علوم کامپیوتر

پایان نامه برای دریافت درجه کارشناسی ارشد در رشته علوم کامپیوتر (M.Sc)

گرایش محاسبات نرم و هوش مصنوعی

عنوان :

بهبود تصویرهای MRI با استفاده از روش MWACNN و مقایسه با روش UNET

استاد راهنما :

جناب آقای دکتر مهدی آژینی

استاد مشاور:

جناب آقای دکتر رضا کارگر

نگارش :

سید حسین طاهری

تابستان ۱۴۰۵

فهرست:

۱ فصل اول : کلیات تحقیق ۱۴

۱- ۱ مقدمه ۱۵

۱- ۲ بیان مسئله ۱۶

۱- ۳ اهمیت و ضرورت تحقیق ۱۷

۱- ۳- ۱ اهمیت علمی پژوهش ۱۷

۱- ۳- ۲ اهمیت کاربردی پژوهش ۱۷

۱ -۴ اهداف تحقیق ۱۸

۱- ۴- ۱ اهداف اصلی ۱۸

۱- ۴- ۲ اهداف اختصاصی ۱۸

۱- ۴- ۳ اهداف کاربردی ۱۹

۱- ۵ نوآوری‌های پژوهش ۱۹

۱- ۵- ۱ مقایسه نظام مند چهار معماری ۱۹

۱- ۵- ۲ بررسی هم زمان کیفیت بازسازی و کارایی محاسباتی ۲۰

۱- ۵- ۳ مقایسه معماری‌های مبتنی بر موجک و بررسی تأثیر توجه کانالی ۲۰

۱- ۵- ۴ ارزیابی تحت شرایط آموزشی یکسان ۲۰

۱- ۵- ۵ تحلیل آماری نتایج ۲۰

۱- ۶ سوالات تحقیق ۲۰

۱ -۶ -۱ سؤال اصلی ۲۰

۱ -۶ -۲ سؤال‌های فرعی ۲۱

۱- ۷ روش انجام تحقیق ۲۱

۱ -۸ ساختار پایان‌نامه ۲۲

۲ فصل دوم : مروری بر ادبیات تحقیق و پیشینه تحقیق ۲۳

۲- ۱ مقدمه ۲۴

۲- ۲ تصویر برداری تشدید مغناطیسی (MRI) ۲۴

۲- ۳ آرتیفکت‌های حرکتی در تصاویرMRI ۲۵

۲- ۴ روش‌های سنتی حذف آرتیفکت و بهبود کیفیت تصویر ۲۶

۲- ۴- ۱ فیلترهای مکانی و محدودیت در حفظ لبه‌ها ۲۷

۲ -۴ -۲ روش‌های غیرمحلی و استفاده از شباهت نواحی تصویر ۲۷
۲- ۴- ۳ روش‌های حوزه تبدیل و مدل‌های آماری ۲۸
۲ -۴ -۴ روش BM3D و جایگاه آن در کاهش نویز کلاسیک ۲۸
۲- ۴- ۵ روش‌های ترکیبی مبتنی بر موجک و بهینه‌سازی ۲۹
۲- ۵ روش‌های مبتنی بر تبدیل موجک در پردازش تصویر ۳۱
۲ -۶ یادگیری عمیق و شبکه‌های عصبی کانولوشنی در بازسازی تصویر ۳۴
۲ -۷ معماری U-Net ۳۸
۲ -۷ -۱ ساختار کلیU-Net ۳۹
۲ -۷ -۲ نقش اتصالات میانبر (Skip Connection) ۴۰
۲ -۷ -۳ مزایا و محدودیت‌هایU-Net ۴۰
۲ -۷ -۴ کاربرد U-Net در بازسازی تصاویر پزشکی ۴۱
۲ -۸ ترکیب تبدیل موجک و یادگیری عمیق ۴۱
۲ -۹ معماریWACNN ۴۵
۲ -۱۰ معماریMWACNN ۴۷
۲ -۱۱ مکانیزم توجه(Attention Mechanism) ۴۸
۲ -۱۲ حذف آرتیفکت‌های حرکتی در تصاویر MRI و معرفی مجموعه‌داده متن باز KMAR-50K ۵۰
۲ -۱۳ پیشینه پژوهش ۵۲
۲ -۱۴ جمع‌بندی ۵۶
۳ فصل سوم: روش اجرای تحقیق ۵۷
۳ -۱ مقدمه ۵۸
۳ -۲ نوع و روش تحقیق ۵۸
۳ -۳ محیط اجرای پژوهش ۵۹
۳- ۴ معماری مدل‌های مورد استفاده ۶۱
۳- ۴- ۱ مقدمه ۶۱
۳- ۴- ۲ U-Net ۶۱
۳- ۴- ۳ WACNN ۶۲
۳- ۴- ۴ MWACNN بدون مکانیزم توجه ۶۲
۳ -۴ -۵ معماری MWACNN مبتنی بر توجه کانالی(Channel Attention) ۶۳
۳ -۵ تنظیمات آموزش مدل‌ها ۶۵
۳ -۵ -۱ مقدمه ۶۵

۳ -۵ -۲ تنظیمات اصلی آموزش ۶۵
۳ -۵ -۳ انتخاب بهینه‌ساز (Optimizer) ۶۷
۳ -۵ -۴ زمان‌بند نرخ یادگیری (Learning Rate Scheduler) ۶۷
۳-۵ -۵ تابع هزینه (Loss Function) ۶۷
۳ -۵ -۶ استفاده از انباشت گرادیان(Gradient Accumulation) ۶۸
۳ -۵ -۷ دلایل انتخاب تنظیمات آموزشی ۶۹
۳ -۵ -۸ جمع‌بندی ۶۹
۳ -۶ معیارهای ارزیابی ۶۹
۳ -۶ -۱ مقدمه ۶۹
۳ -۶ -۲ نسبت سیگنال به نویز اوج(PSNR) ۷۰
۳ -۶ -۳ شاخص شباهت ساختاری (SSIM) ۷۰
۳ -۶ -۴ میانگین قدر مطلق خطا (MAE) ۷۱
۳ -۶ -۵ میانگین مربعات خطا (MSE) ۷۲
۳ -۶ -۶ ریشه دوم میانگین مربعات خطای نرمال‌شده (NRMSE) ۷۲
۳ -۶ -۷ ضریب همبستگی نرمال‌شده (NCC) ۷۳
۳ -۶ -۸ تعداد پارامترهای قابل آموزش(Trainable Parameters) ۷۳
۳ -۶ -۹ زمان استنتاج(Inference Time) ۷۳
۳ -۶ -۱۰ آزمون آماری ویلکاکسون ۷۵
۳ -۶ -۱۱ جمع‌بندی ۷۵
۳ -۷ روند اجرای پژوهش(Workflow) ۷۶
۳ -۷ -۱ روند کلی اجرای پژوهش ۷۶
۴ فصل چهارم: تجزیه و تحلیل داده‌ها (یافته‌ها) ۸۰
۴ -۱ معرفی مجموعه‌داده ۸۱
۴ -۱ -۱ معرفی مجموعه‌دادهKMAR-50K ۸۱
۴ -۱ -۲ ساختار مجموعه‌داده ۸۱
۴ -۱ -۳ نحوه انتخاب داده‌ها ۸۲
۴ -۱ -۴ دلیل انتخاب مجموعه‌داده KMAR-50K ۸۲
۴ -۲ پیش پردازش داده‌ها ۸۴
۴ -۲ -۱ مقدمه ۸۴
۴ -۲ -۲ استخراج برش‌های دوبعدی(Slice Extraction) ۸۴

۴ -۲ -۳ نرمال‌سازی تصاویر (Normalization) ۸۴
۴ -۲ -۴ استخراج وصله ۸۵
۴- ۲- ۵ استخراج تصادفی وصله‌ها ۸۶
۴- ۲- ۶ آماده‌سازی داده‌ها برای آموزش ۸۶
۴ -۲ -۷ جمع‌بندی پیش‌پردازش ۸۷
۴ -۳ نتایج و تحلیل ۸۸
۴ -۴ محیط اجرای آزمایش‌ها ۸۸
۴- ۵ نتایج کمی ۸۹
۴ -۶ تحلیل نتایج کمی ۹۲
۴- ۷ تحلیل معیارها ۹۳
۴ -۷ -۱ تحلیل نسبت سیگنال به نویز اوج ۹۳
۴ -۷ -۲ تحلیل شاخص شباهت ساختاری ۹۳
۴ -۷ -۳ تحلیل خطاهای بازسازی ۹۳
۴ -۷ -۴ تحلیل کارایی محاسباتی ۹۴
۴ -۷ -۵ جمع‌بندی نتایج کمی ۹۴
۴- ۸ تحلیل آماری نتایج با آزمون ویلکاکسون ۹۶
۴ -۹ تحلیل نتایج بصری ۹۸
۴ -۱۰ بحث ۱۰۰
۴ -۱۱ جمع‌بندی فصل ۱۰۲
۵ فصل پنجم: نتیجه‌گیری و پیشنهادها ۱۰۳
۵ -۱ مقدمه ۱۰۴
۵ -۲ نتیجه‌گیری ۱۰۴
۵ -۳ دستاوردهای پژوهش ۱۰۶
۵- ۴ محدودیت‌های پژوهش ۱۰۶
۵- ۵ مقایسه نتایج پژوهش حاضر با پژوهش‌های پیشین ۱۰۷
۵ -۶ پیشنهادهایی برای پژوهش‌های آینده ۱۰۹
۵- ۷ جمع‌بندی نهایی ۱۱۰
۵- ۸ پیشنهادهای کاربردی ۱۱۱

فهرست تصویرها

۱ ۲-۱ تصاویر همراه با آرتیفکت و واضح از MRI زانو [۲۸] ۲۶
۲ ۲-۲ معماری شبکه U-NET [۱۷] ۴۰
۳ ۲-۳ معماری شبکه WCNN ۴۵
۴ ۲-۴ معماری MWCNN [۳] ۴۸
۵ ۳-۱ روند کلی اجرای پژوهش ۷۹
۶ ۴-۱ ساختار و آماده‌سازی مجموعه‌دادهKMAR-۵۰K ۸۳
۷ ۴-۲ مقایسه کمی عملکرد چهار معماری بر اساس معیارهای ارزیابی ۹۱
۸ ۴-۳ مقایسه توازن میان دقت بازسازی (PSNR) و پیچیدگی محاسباتی مدل‌های مورد بررسی ۹۵
۹ ۴-۴ نمودارهای نتایج آزمون ویلکاکسون ۹۸
۱۰ ۴-۵ پنج نمونه منتخب از خروجی چهار معماری روی مجموعه آزمون مستقل ۱۰۰
۱۱ مقدار دهی اولیه کد ۱۱۵
۱۲ تغییر سایز عکس‌ها و انتخاب برش در کد ۱۱۶
۱۳ کد تبدیل موجکی ۱۱۷
۱۴ کد مکانیزم توجه ۱۱۸

فهرست جدول‌ها

۱ ۲-۱ مقایسه روش‌های سنتی بهبود کیفیت و کاهش نویز تصویر ۳۰
۲ ۲-۲ نقش زیرباندهای حاصل از تبدیل موجک دوبعدی در پردازش و بازسازی تصاویر ۳۲
۳ ۲-۳ جمع‌بندی رویکردهای مبتنی بر تبدیل موجک در پردازش و بازسازی تصویر ۳۴
۴ ۲-۴ مقایسه برخی از مهم‌ترین مدل‌های مبتنی بر یادگیری عمیق در بازسازی تصویر ۳۷
۵ ۲-۵ مقایسه مهمترین پژوهش‌های ترکیب تبدیل موجک و یادگیری عمیق ۴۳
۶ ۲-۶ معماری WACNN ۴۶
۷ ۲-۷ مقایسه WACNN و MWACNN ۴۸
۸ ۲-۸ مقایسه انواع مکانیزم‌های توجه در پردازش تصویر ۵۰
۹ ۲-۹ ویژگی‌های اصلی مجموعه‌داده - K۵۰KMAR ۵۲
۱۰ ۲-۱۰ معماری‌های پیشین و محدودیت‌ها ۵۵
۱۱ ۳-۱ مشخصات کلی محیط نرم‌افزاری و پردازشی ۶۰
۱۲ ۳-۲ مقایسه معماری‌های مورد استفاده در پژوهش ۶۴
۱۳ ۳-۳ تنظیمات اصلی آموزش مدل‌ها ۶۶
۱۴ ۳-۴ خلاصه معیارهای ارزیابی مورد استفاده در پژوهش ۷۴
۱۵ ۴-۱ تعداد داده‌های مورد استفاده ۸۲
۱۶ ۴-۲ تنظیمات استخراج وصله ۸۵
۱۷ ۴-۳ تنظیمات بار گذاری و آماده سازی داده‌ها ۸۷
۱۸ ۴-۴ مشخصات محیط اجرای آزمایش‌ها ۸۹
۱۹ ۴-۵ نتایج کمی مدل‌های مورد بررسی بر روی مجموعه آزمون مستقل ۹۰
۲۰ ۴-۶ جمع‌بندی مقایسه عملکرد مدل‌های مورد بررسی ۹۵
۲۱ ۴-۷ نتایج آزمون ویلکاکسون برای مقایسه مدل‌ها ۹۶
۲۲ ۵-۱ مقایسه پژوهش حاضر با مطالعات پیشین ۱۰۸

چکیده

تصویربرداری تشدید مغناطیسی یکی از مهم‌ترین روش‌های تصویربرداری پزشکی برای بررسی ساختارهای داخلی بدن است. با این حال، حرکت بیمار در هنگام تصویربرداری می‌تواند موجب ایجاد آرتیفکت‌های حرکتی و کاهش کیفیت تصاویر شود که این امر دقت تشخیص پزشکان را تحت تأثیر قرار می‌دهد. در سال‌های اخیر، استفاده از روش‌های مبتنی بر یادگیری عمیق به‌عنوان راهکاری مؤثر برای حذف آرتیفکت‌ها و بازسازی تصاویر تشدید مغناطیسی مورد توجه قرار گرفته است.

در این پژوهش، عملکرد چهار معماری مبتنی بر یادگیری عمیق شامل U-Net، WACNN، MWACNN بدون مکانیزم توجه و MWACNN مبتنی بر توجه کانالی برای حذف آرتیفکت‌های حرکتی تصاویرMRI مورد بررسی و مقایسه قرار گرفت. تمامی مدل‌ها با استفاده از مجموعه‌داده استاندارد KMAR-۵۰K و تحت شرایط آموزشی یکسان پیاده‌سازی، آموزش و ارزیابی شدند. به‌منظور سنجش عملکرد مدل‌ها، از معیارهای نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده، ضریب همبستگی نرمال‌شده، تعداد پارامترهای قابل آموزش و زمان استنتاج استفاده شد.

نتایج نشان داد که معماری U-Net با دستیابی به نسبت سیگنال به نویز اوج برابر با ۲۶.۰۵۷ دسی‌بل و شاخص شباهت ساختاری برابر با ۰.۷۱۴۱، بهترین عملکرد کلی را در میان مدل‌های مورد بررسی ارائه داد و در شش معیار اصلی کیفیت بازسازی، بهترین مقدار را به دست آورد. همچنین، U-Net کمترین زمان استنتاج را در میان مدل‌های مورد بررسی ثبت کرد. در مقابل، معماری MWACNN مبتنی بر توجه کانالی با استفاده از حدود یک‌سوم تعداد پارامترهای قابل آموزشU-Net ، در برخی معیارها، به‌ویژه نسبت سیگنال به نویز اوج و ضریب همبستگی نرمال‌شده، عملکردی نزدیک به U-Net ارائه داد. با این حال، در معیارهای شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا و ریشه دوم میانگین مربعات خطای نرمال‌شده همچنان عملکرد U-Net بهتر بود. این نتایج نشان می‌دهد که ارزیابی معماری‌های بازسازی تصویر باید بر پایه مجموعه‌ای از معیارهای کیفیت و کارایی انجام شود و کاهش تعداد پارامترها لزوماً به کاهش زمان استنتاج منجر نمی‌شود. همچنین، به‌عنوان یک تحلیل آماری تکمیلی، از آزمون رتبه علامت‌دار ویلکاکسون برای بررسی اختلاف مقادیر زوجی معیارهای ارزیابی میان مدل‌ها استفاده شد.

واژگان کلیدی :تصویربرداری تشدید مغناطیسی، حذف آرتیفکت حرکتی، بازسازی تصویر، یادگیری عمیق، U-Net، تبدیل موجک، توجه کانالی، .KMAR-۵۰K

۱ فصل اول : کلیات تحقیق

۱- ۱ مقدمه

پیشرفت فناوری‌های تصویربرداری پزشکی نقش مهمی در افزایش دقت تشخیص و برنامه‌ریزی درمان بیماری‌ها داشته است. در میان روش‌های مختلف تصویربرداری، تصویربرداری تشدید مغناطیسی (MRI[۱]) به دلیل توانایی بالا در نمایش بافت‌های نرم، قدرت تفکیک مناسب و عدم استفاده از پرتوهای یون‌ساز، به یکی از پر کاربردترین روش‌های تصویربرداری در پزشکی تبدیل شده است. [۲۸]

با وجود مزایای فراوان MRI، کیفیت این تصاویر همواره در شرایط مطلوب نمی‌باشد. یکی از مشکلات رایج در تصویربرداری MRI، ایجاد آرتیفکت‌های حرکتی[۲] است. این آرتیفکت‌ها معمولاً به علت حرکت بیمار، تنفس، ضربان قلب یا جابه‌جایی ناخواسته بدن در زمان تصویربرداری ایجاد می‌شوند. وجود آرتیفکت‌های حرکتی باعث کاهش وضوح تصویر، از بین رفتن بخشی از اطلاعات تشخیصی و دشوار شدن تفسیر تصاویر MRI می‌شود. در برخی موارد نیز کیفیت پایین تصویر انجام تصویربرداری مجدد را ضروری می‌کند. [۲۸]

ازاین‌رو، توسعه روش‌هایی که بتوانند آرتیفکت‌های حرکتی را کاهش داده و در عین حال اطلاعات تشخیصی و ساختارهای آناتومیکی تصویر را حفظ کنند، به یکی از موضوعات مهم پژوهشی در حوزه پردازش تصاویر پزشکی تبدیل شده است. روش‌های سنتی پردازش تصویر مانند فیلترهای مکانی[۳]، فیلترهای فرکانسی[۴]، تبدیل فوریه[۵] و برخی روش‌های مبتنی بر تبدیل موجک[۶] در سال‌های گذشته برای کاهش نویز و بهبود کیفیت تصاویر MRI مورد استفاده قرار گرفته‌اند. اگرچه این روش‌ها در برخی کاربردها عملکرد قابل قبولی داشته‌اند، اما معمولاً در حفظ هم‌زمان جزئیات ظریف تصویر و حذف آرتیفکت‌های پیچیده با محدودیت مواجه هستند. در مقابل، روش‌های مبتنی بر یادگیری عمیق [۷]به علت توانایی در یادگیری خودکار ویژگی‌های تصویر، عملکرد قابل توجهی در بازسازی تصاویر پزشکی و حذف آرتیفکت‌ها نشان داده‌اند. [۱۳]

در میان معماری‌های یادگیری عمیق، شبکه U-Net[۸] یکی از پرکاربردترین مدل‌ها در حوزه تصویربرداری پزشکی است. این معماری به دلیل ساختار معماری رمزگذار–رمزگشا[۹] و استفاده از اتصالات میانبر [۱۰] می‌تواند ویژگی‌های سطح پایین و سطح بالا را به‌صورت همزمان حفظ کند و در بازسازی تصاویر پزشکی عملکرد مناسبی داشته باشد [۱۷]. علاوه بر این، تبدیل موجک به عنوان یک روش تحلیل چندمقیاسی توانایی استخراج اطلاعات فرکانسی و مکانی تصویر را دارد و در حفظ جزئیات تصویر مؤثر است.([۳] ،[۲])

---

[۱] Magnetic Resonance Imaging
[۲] Motion Artifacts
[۳] Spatial Filters
[۴] Frequency Domain Filters
[۵] Fourier Transform (FT)
[۶] Wavelet Transform
[۷] Deep Learning
[۸] U-Net Architecture
[۹] Encoder-Decoder
[۱۰] Skip Connection

با توجه به پیشرفت‌های اخیر در حوزه یادگیری عمیق، مقایسه معماری‌های مختلف برای حذف آرتیفکت‌های حرکتی و ارزیابی نقاط قوت و محدودیت‌های هر یک، قابلیت کمک در انتخاب روش مناسب برای بازسازی تصاویر MRIرا دارد. از این رو، در این پژوهش عملکرد چهار معماری U-Net، WACNN[۱]، MWACNN[۲] بدون مکانیزم توجه[۳] و MWACNN مبتنی بر توجه مورد بررسی و مقایسه قرار گرفته‌اند.

۱- ۲ بیان مسئله

در کنار پیشرفت‌های قابل توجه در توسعه روش‌های بازسازی تصاویر پزشکی، حذف آرتیفکت‌های حرکتی در تصاویر MRI همچنان یکی از چالش‌های اساسی در حوزه پردازش تصاویر پزشکی محسوب می‌شود. هدف اصلی از حذف این آرتیفکت‌ها تنها افزایش کیفیت ظاهری تصویر نیست، بلکه حفظ اطلاعات تشخیصی و ساختارهای آناتومیکی[۴] است؛ زیرا از بین رفتن جزئیات تصویر یا ایجاد تغییرات مصنوعی در فرآیند بازسازی می‌تواند بر دقت تشخیص پزشک تأثیر منفی بگذارد. ازاین‌رو، توسعه روش‌هایی که بتوانند میان کاهش آرتیفکت‌ها، حفظ ساختارهای تصویر و هزینه محاسباتی تعادل برقرار کنند، همچنان یکی از مسائل باز این حوزه به شمار می‌رود.

روش‌های متعددی برای کاهش آرتیفکت‌های حرکتی و بازسازی تصاویر MRI پیشنهاد شده‌اند. به مرور زمان استفاده از معماری‌های مبتنی بر یادگیری عمیق نسبت به روش‌های کلاسیک رشد قابل توجهی داشته و نتایج امیدوارکننده‌ای در بازسازی تصاویر پزشکی ارائه کرده است. با این حال، هنوز مشخص نیست کدام نوع معماری می‌تواند ضمن حفظ کیفیت بازسازی، از نظر پیچیدگی محاسباتی نیز عملکرد مناسب‌تری داشته باشد. این موضوع به‌ویژه در کاربردهای بالینی که سرعت پردازش و منابع سخت‌افزاری اهمیت دارند، از اهمیت بیشتری برخوردار است. ([۲]،[۳]، [۱۷])

بررسی پژوهش‌های پیشین نشان می‌دهد که بخش قابل توجهی از مطالعات بر معرفی یا بهبود یک معماری خاص متمرکز بوده‌اند و اغلب عملکرد همان مدل را گزارش کرده‌اند. درمقابل، پژوهش‌هایی که چند معماری مختلف را تحت شرایط آموزشی یکسان، بر روی یک مجموعه‌داده استاندارد و با معیارهای ارزیابی یکسان مقایسه کرده باشند، محدود هستند. این موضوع باعث می‌شود تشخیص اینکه بهبود عملکرد ناشی از معماری شبکه است یا ناشی از تفاوت در داده‌ها، تنظیمات آموزشی یا معیارهای ارزیابی، دشوار باشد. ([۲]،[۳]، [۱۷]، [۲۸])

از سوی دیگر، اگرچه معماری U-Net به عنوان یکی از موفق‌ترین شبکه‌های بازسازی تصاویر پزشکی شناخته می‌شود و معماری‌های مبتنی بر تبدیل موجک مانند WACNN و MWACNN نیز با هدف استخراج بهتر ویژگی‌های چندمقیاسی توسعه یافته‌اند، اما در منابع بررسی‌شده در این پژوهش، مقایسه‌ای هم‌زمان میان چهار پیاده‌سازی مورد استفاده در تحقیق حاضر، تحت مجموعه‌داده، تنظیمات آموزشی و معیارهای ارزیابی یکسان مشاهده نشد. همچنین در بسیاری از مطالعات، تمرکز اصلی بر معیارهای کیفیت تصویر بوده و شاخص‌هایی مانند تعداد پارامترهای قابل

---

[۱] Wavelet Attention Convolutional Neural Network

[۲] Multi-level Wavelet Attention Convolutional Neural Network

[۳] Attention Mechanism

[۴] Anatomical Structures

آموزش، زمان استنتاج و کارایی محاسباتی کمتر به صورت هم‌زمان مورد بررسی قرار گرفته‌اند. در نتیجه، هنوز تصویر روشنی از نقاط قوت و محدودیت‌های هر یک از این معماری‌ها برای کاربردهای عملی مشاهده نشده است.

بر همین اساس، مسئله اصلی این پژوهش، ارزیابی و مقایسه عملکرد چهار معماری U-Net، WACNN، MWACNN و MWACNN مبتنی بر مکانیزم توجه در حذف آرتیفکت‌های حرکتی تصاویر MRI است. در این پژوهش، تمامی مدل‌ها بر روی مجموعه‌داده استاندارد KMAR-۵۰K و تحت شرایط آموزشی یکسان پیاده‌سازی و ارزیابی شده‌اند تا امکان مقایسه‌ای منصفانه میان آن‌ها فراهم شود. عملکرد مدل‌ها علاوه بر معیارهای متداول کیفیت تصویر شامل [۱]نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری[۲]، [۳]میانگین قدر مطلق خطا، میانگین مربعات خطا[۴]، [۵]ریشه دوم میانگین مربعات خطای نرمال‌شده و[۶] ضریب همبستگی نرمال‌شده، از نظر تعداد پارامترهای قابل آموزش و زمان استنتاج نیز بررسی شده است تا علاوه بر کیفیت بازسازی، کارایی محاسباتی آن‌ها نیز مورد ارزیابی قرار گیرد. نتایج این پژوهش بتواند دید روشن‌تری نسبت به مزایا و محدودیت‌های هر یک از معماری‌های مورد بررسی ارائه دهد و به انتخاب مناسب‌ترین مدل برای کاربردهای بازسازی تصاویر MRI کمک کند. [۲۸]

۱- ۳ اهمیت و ضرورت تحقیق

۱- ۳- ۱ اهمیت علمی پژوهش

حذف آرتیفکت‌های حرکتی و بازسازی تصاویر MRI همچنان از موضوعات مهم در حوزه پردازش تصاویر پزشکی و یادگیری عمیق محسوب می‌شود. با وجود توسعه معماری‌های متعدد، عملکرد مدل‌ها به عواملی مانند ساختار شبکه، نوع داده، تابع هزینه و تنظیمات آموزشی وابسته است و مقایسه نتایج گزارش‌شده در مطالعات مختلف همیشه امکان ارزیابی مستقیم معماری‌ها را فراهم نمی‌کند. ازاین‌رو، ارزیابی چند معماری تحت شرایط آموزشی مشترک می‌تواند در شناخت بهتر نقاط قوت و محدودیت‌های آن‌ها مؤثر باشد. پژوهش حاضر نیز با مقایسه چهار معماری تحت مجموعه‌داده و تنظیمات آموزشی یکسان، تلاش می‌کند ارزیابی کنترل‌شده‌تری از عملکرد آن‌ها در بازسازی تصاویر MRI ارائه دهد.. ([۲] ، [۳] ، [۱۷] ، [۲۸])

۱- ۳- ۲ اهمیت کاربردی پژوهش

از دیدگاه کاربردی، کیفیت تصاویر MRI تأثیر مستقیمی بر فرآیند تشخیص و تصمیم‌گیری درمانی دارد. بهبود کیفیت تصاویر و کاهش آرتیفکت‌های حرکتی می‌تواند موجب افزایش دقت تشخیص، کاهش احتمال بروز

---

[۱] Peak Signal-to-Noise Ratio

[۲] Structural Similarity Index Measure

[۳] Mean Absolute Error

[۴] Mean Squared Error

[۵] Normalized Root Mean Squared Error

[۶] Normalized Cross-Correlation

خطاهای ناشی از کاهش کیفیت تصویر و کاهش نیاز به انجام مجدد تصویربرداری شود.[۲۸] این موضوع علاوه بر کاهش زمان و هزینه‌های درمان، تجربه بهتری برای بیماران و کادر درمان فراهم می‌کند.

علاوه بر این، نتایج این پژوهش می‌تواند در توسعه سامانه‌های کمک‌تشخیص پزشکی(CAD)[۱] ، نرم‌افزارهای پردازش تصاویر پزشکی و سایر کاربردهای مبتنی بر هوش مصنوعی مورد استفاده قرار گیرد.[۱۳] همچنین مقایسه جامع معماری‌های مورد بررسی می‌تواند به پژوهشگران و توسعه‌دهندگان در انتخاب مدل مناسب برای کاربردهای مختلف تصویربرداری پزشکی کمک کرده و زمینه را برای انجام پژوهش‌های تکمیلی و توسعه معماری‌های پیشرفته‌تر فراهم سازد. ([۱۷] ، [۲])

۱- ۴ اهداف تحقیق

۱- ۴- ۱ اهداف اصلی

هدف اصلی این پژوهش، پیاده‌سازی، آموزش، ارزیابی و مقایسه عملکرد چهار معماری U-Net، WACNN، MWACNN بدون مکانیزم توجه و MWACNN مبتنی بر توجه کانالی در حذف آرتیفکت‌های حرکتی تصاویر MRIبا استفاده از مجموعه‌داده KMAR-۵۰K است. این مقایسه با هدف بررسی عملکرد معماری‌های مبتنی بر موجک در برابر مدل پایه و همچنین بررسی مستقیم اثر افزودن توجه کانالی به ساختار MWACNN انجام شده است.

۱- ۴- ۲ اهداف اختصاصی

۱. پیاده‌سازی و آموزش معماری U-Net به عنوان مدل مرجع برای بازسازی تصاویر MRI .
۲. پیاده‌سازی و آموزش معماری WACNN مبتنی بر تبدیل موجک .
۳. پیاده‌سازی و آموزش معماری MWACNN با هدف بررسی استخراج ویژگی‌های چندمقیاسی .
٤. پیاده‌سازی نسخه مجهز به مکانیزم توجه در معماری MWACNN و بررسی تأثیر آن بر فرآیند بازسازی تصاویر .
٥. آموزش و ارزیابی تمامی مدل‌ها بر روی مجموعه‌داده استاندارد KMAR-۵۰K تحت شرایط آموزشی یکسان .
٦. مقایسه عملکرد معماری‌های مورد بررسی از نظر کیفیت بازسازی تصاویر MRI .
۷. مقایسه عملکرد معماری‌های مبتنی بر تبدیل موجک با معماری پایه U-Net در شرایط آموزشی یکسان.
۸. بررسی تأثیر افزودن توجه کانالی به معماری MWACNN از طریق مقایسه نسخه دارای توجه و بدون توجه.

[۱] Computer-Aided Diagnosis (CAD) Systems

۹. مقایسه و تحلیل عملکرد معماری‌های مورد بررسی از نظر کیفیت بازسازی، تعداد پارامترهای قابل آموزش، زمان استنتاج.

۱- ۴- ۳ اهداف کاربردی

- کمک به بهبود کیفیت تصاویر MRI و کاهش تأثیر آرتیفکت‌های حرکتی .
- فراهم کردن اطلاعات لازم برای انتخاب معماری مناسب در بازسازی تصاویر MRI .
- کمک به توسعه سامانه‌های هوشمند پردازش و بازسازی تصاویر پزشکی .
- ایجاد زمینه مناسب برای استفاده از روش‌های مبتنی بر یادگیری عمیق در مراکز تصویربرداری پزشکی .
- فراهم کردن بستری برای انجام پژوهش‌های آینده در زمینه حذف آرتیفکت‌های حرکتی و توسعه مدل‌های پیشرفته‌تر بازسازی تصاویر پزشکی

۱- ۵ نوآوری‌های پژوهش

با وجود پژوهش‌های متعدد در زمینه بازسازی تصاویر MRI با استفاده از روش‌های مبتنی بر یادگیری عمیق، در منابع بررسی‌شده در این پژوهش، مقایسه‌ای هم‌زمان میان چهار پیاده‌سازی مورد استفاده در تحقیق حاضر تحت مجموعه‌داده، تنظیمات آموزشی و معیارهای ارزیابی یکسان مشاهده نشد. بر همین اساس، این پژوهش چهار معماری U-Net، WACNN، MWACNN بدون مکانیزم توجه و MWACNN مبتنی بر توجه کانالی را در یک چارچوب آزمایشی مشترک پیاده‌سازی، آموزش و ارزیابی کرده است.

انتخاب این چهار معماری با هدف مقایسه رویکردهای متفاوت در بازسازی تصاویر MRI انجام شده است. معماری U-Netبه‌عنوان یک معماری پایه و شناخته‌شده در پردازش تصاویر پزشکی، WACNNبه‌عنوان یک رویکرد مبتنی بر ترکیب شبکه‌های کانولوشنی و تبدیل موجک، و MWACNN به‌منظور بررسی تأثیر پردازش موجک در چند سطح انتخاب شدند. همچنین، نسخه MWACNN مبتنی بر مکانیزم توجه به‌منظور بررسی تأثیر افزودن توجه بر عملکرد این معماری در نظر گرفته شد. انتخاب این چهار مدل امکان بررسی و مقایسه تأثیر ساختار معماری، استفاده از اطلاعات چندمقیاسی موجک و مکانیزم توجه را تحت شرایط آموزشی یکسان فراهم می‌کند.

مهم‌ترین ویژگی‌های متمایز این پژوهش عبارت‌اند از:

۱- ۵- ۱ مقایسه نظام مند چهار معماری

چهار معماری U-Net، WACNN، MWACNN و MWACNN مبتنی بر مکانیزم توجه با استفاده از یک مجموعه‌داده استاندارد و تحت شرایط آموزشی یکسان پیاده‌سازی، آموزش و ارزیابی شده‌اند. این رویکرد امکان مقایسه مستقیم عملکرد مدل‌ها را فراهم کرده و اثر تفاوت‌های ناشی از تنظیمات آموزشی را به حداقل رسانده است.

۱- ۵- ۲ بررسی هم زمان کیفیت بازسازی و کارایی محاسباتی

علاوه بر معیارهای متداول کیفیت تصویر شامل نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده، شاخص‌های تعداد پارامترهای قابل آموزش و زمان استنتاج نیز مورد بررسی قرار گرفته‌اند. این موضوع امکان ارزیابی جامع‌تر معماری‌های مختلف را از نظر دقت و کارایی محاسباتی فراهم کرده است.

۱-۵-۳ مقایسه معماری‌های مبتنی بر موجک و بررسی تأثیر توجه کانالی

در این پژوهش، عملکرد معماری‌های مبتنی بر تبدیل موجک در کنار مدل پایه U-Net بررسی شده است. همچنین، با پیاده‌سازی دو نسخه MWACNN دارای ساختار اصلی یکسان و با/بدون توجه کانالی، تأثیر افزودن این مکانیزم به‌صورت مستقیم بررسی شده است.

۱- ۵- ۴ ارزیابی تحت شرایط آموزشی یکسان

تمامی معماری‌های مورد بررسی با استفاده از مجموعه‌داده KMAR-۵۰K، تنظیمات آموزشی، تابع هزینه و معیارهای ارزیابی یکسان آموزش و ارزیابی شده‌اند. این موضوع موجب شده است که تفاوت نتایج حاصل، تا حد امکان ناشی از تفاوت در معماری مدل‌ها باشد و مقایسه‌ای منصفانه میان آن‌ها انجام شود.

۱- ۵- ۵ تحلیل آماری نتایج

علاوه بر مقایسه مقادیر اصلی معیارهای ارزیابی، از آزمون رتبه علامت‌دار ویلکاکسون[۱] به‌عنوان یک تحلیل آماری تکمیلی برای بررسی اختلاف مقادیر زوجی میان مدل‌ها استفاده شد. این تحلیل به تفسیر آماری برخی تفاوت‌های مشاهده‌شده در کنار نتایج اصلی کمک می‌کند.

۱- ۶ سوالات تحقیق

۱- ۶- ۱ سؤال اصلی

عملکرد معماری‌های U-Net، WACNN، MWACNN و MWACNN مبتنی بر مکانیزم توجه در حذف آرتیفکت‌های حرکتی و بازسازی تصاویر MRI چگونه است و کدام‌یک از این معماری‌ها عملکرد مناسب‌تری ارائه می‌دهد؟

[۱] Wilcoxon Signed-Rank Test

۱- ۶- ۲ سؤال‌های فرعی

۱ .عملکرد معماری‌های مبتنی بر تبدیل موجک در مقایسه با معماری پایه U-Net در بازسازی تصاویر MRI چگونه است؟

۲ .استفاده از مکانیزم توجه چه تأثیری بر عملکرد معماری MWACNN در حذف آرتیفکت‌های حرکتی تصاویر MRIدارد؟

۳ .کدام‌یک از معماری‌های مورد بررسی از نظر معیارهای کیفیت تصویر شامل نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده عملکرد مناسب‌تری دارد؟

۴ .آیا افزایش پیچیدگی معماری و تعداد پارامترهای قابل آموزش، لزوماً موجب بهبود کیفیت بازسازی تصاویر MRI می‌شود؟

۵ .کدام‌یک از معماری‌های مورد بررسی از نظر کارایی محاسباتی، شامل زمان استنتاج و تعداد پارامترهای قابل آموزش، عملکرد مناسب‌تری دارد؟

۱- ۷ روش انجام تحقیق

پژوهش حاضر از نظر هدف، کاربردی و از نظر روش اجرا، تجربی و کمی است. به منظور حذف آرتیفکت‌های حرکتی و بهبود کیفیت تصاویرMRI ، چهار معماری مبتنی بر یادگیری عمیق شامل U-Net، WACNN، MWACNN و MWACNN مبتنی بر مکانیزم توجه مورد پیاده‌سازی، آموزش و ارزیابی قرار گرفته‌اند.

برای انجام این پژوهش، از مجموعه‌داده استاندارد KMAR-۵۰K [۲۸] شامل تصاویر MRI دارای آرتیفکت و تصاویر مرجع استفاده شده است. پس از آماده‌سازی داده‌ها و انجام مراحل پیش‌پردازش، تمامی مدل‌ها آموزش داده شدند؛ جزئیات دقیق نحوه‌ی آموزش و تنظیمات هر معماری در فصل سوم ارائه می‌شود.

در ادامه، عملکرد مدل‌های مورد بررسی با استفاده از معیارهای کمی شامل نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده ارزیابی شد. همچنین، به منظور بررسی کارایی محاسباتی مدل‌ها، تعداد پارامترهای قابل آموزش و زمان استنتاج نیز اندازه‌گیری و مقایسه شدند. علاوه بر این، برای بررسی آماری تفاوت عملکرد معماری‌ها، از آزمون رتبه علامت‌دار ویلکاکسون [۲۹] استفاده شد تا نتایج حاصل از مقایسه مدل‌ها از نظر آماری نیز مورد تحلیل قرار گیرند.

مراحل اجرای پژوهش شامل انتخاب و آماده‌سازی داده‌ها، پیاده‌سازی مدل‌ها، آموزش، ارزیابی، تحلیل نتایج و مقایسه عملکرد معماری‌های مختلف بوده است. جزئیات هر یک از این مراحل در فصل سوم به‌طور کامل تشریح شده است.

۱- ۸ ساختار پایان‌نامه

فصل اول به کلیات تحقیق اختصاص دارد و شامل مقدمه، بیان مسئله، اهمیت و ضرورت پژوهش، اهداف، سؤالات، نوآوری‌ها، روش انجام تحقیق و ساختار پایان‌نامه است.

فصل دوم به مرور ادبیات تحقیق و پیشینه پژوهش اختصاص دارد. در این فصل، مفاهیم مرتبط با تصویربرداری MRI، آرتیفکت‌های حرکتی، روش‌های پردازش تصویر، یادگیری عمیق، شبکه‌های عصبی کانولوشنی، تبدیل موجک، مکانیزم توجه و پژوهش‌های پیشین بررسی شده‌اند.

فصل سوم روش اجرای پژوهش، معماری‌های مورد استفاده، تنظیمات آموزش، معیارهای ارزیابی و روند کلی اجرای پژوهش را تشریح می‌کند.

فصل چهارم به معرفی مجموعه‌داده مورد استفاده، مراحل پیش پردازش، ارائه نتایج حاصل از آموزش و ارزیابی مدل‌ها اختصاص دارد. در این فصل، عملکرد معماری‌های مختلف بر اساس معیارهای کمی، مقایسه نتایج، تحلیل آماری و بررسی کارایی محاسباتی مدل‌ها ارائه شده است.

فصل پنجم شامل جمع‌بندی، نتیجه‌گیری، محدودیت‌های پژوهش و پیشنهادهایی برای پژوهش‌های آینده است.

۲ فصل دوم : مروری بر ادبیات تحقیق و پیشینه تحقیق

۲- ۱ مقدمه

کیفیت تصاویر پزشکی نقش مهمی در دقت تشخیص، برنامه‌ریزی درمان و تصمیم‌گیری بالینی دارد. در میان روش‌های مختلف تصویربرداری پزشکی، تصاویر حاصل از تصویربرداری تشدید مغناطیسی با توجه به توانایی بالا در نمایش بافت‌های نرم، از اهمیت ویژه‌ای برخوردار هستند. با این حال، تصاویر MRI ممکن است در مراحل تصویربرداری، ذخیره‌سازی یا بازسازی تحت تأثیر عواملی مانند نویز، حرکت بیمار و محدودیت‌های سیستم تصویربرداری قرار گیرند که نتیجه آن کاهش کیفیت تصویر، ایجاد آرتیفکت و از بین رفتن بخشی از اطلاعات تشخیصی است. این موضوع می‌تواند موجب کاهش دقت پزشک در تفسیر تصاویر و حتی نیاز به انجام مجدد تصویربرداری شود .([۳] ، [۲۸])

در سال‌های گذشته، روش‌های مختلفی برای کاهش نویز، حذف آرتیفکت و بازسازی تصاویر پزشکی ارائه شده است. این روش‌ها از تکنیک‌های کلاسیک مبتنی بر فیلترهای مکانی، فیلترهای فرکانسی، روش‌های آماری و تبدیل موجک آغاز شده و با پیشرفت یادگیری عمیق، به سمت استفاده از شبکه‌های عصبی کانولوشنی، مکانیزم توجه و معماری‌های پیشرفته حرکت کرده‌اند. ([۵]، [۷]، [۱۰]، [۱۳]، [۱۴]).با وجود پیشرفت‌های قابل توجه، دستیابی هم‌زمان به حذف مؤثر آرتیفکت‌ها، حفظ جزئیات ساختاری و کاهش پیچیدگی محاسباتی همچنان یکی از چالش‌های مهم در حوزه بازسازی تصاویر پزشکی محسوب می‌شود.

۲- ۲ تصویر برداری تشدید مغناطیسی (MRI)

تصویربرداری تشدید مغناطیسی یکی از پیشرفته‌ترین روش‌های تصویربرداری پزشکی است که با توجه به توانایی بالا در نمایش بافت‌های نرم، قدرت تفکیک مناسب و عدم استفاده از پرتوهای یون‌ساز، کاربرد گسترده‌ای در تشخیص و بررسی بیماری‌های مختلف دارد. برخلاف روش‌هایی مانند توموگرافی کامپیوتری[۱] یا تصویربرداری با اشعه ایکس[۲]، در MRI از میدان مغناطیسی قوی و امواج رادیویی برای تولید تصاویر استفاده می‌شود؛ ازاین‌رو، این روش خطرات ناشی از تابش پرتوهای یون‌ساز را ندارد و برای بسیاری از بیماران، به‌ویژه در تصویربرداری‌های مکرر، گزینه‌ای مناسب محسوب می‌شود.([۱۶]، [۱۹])

اساس عملکرد MRI بر ویژگی‌های مغناطیسی هسته اتم هیدروژن استوار است. با قرار گرفتن بدن بیمار در یک میدان مغناطیسی قوی، پروتون‌های هیدروژن در راستای میدان مغناطیسی منظم می‌شوند. سپس با اعمال پالس‌های فرکانس رادیویی، این تعادل به‌طور موقت برهم خورده و پس از قطع پالس، پروتون‌ها به وضعیت اولیه بازمی‌گردند. در این فرآیند، سیگنال‌هایی تولید می‌شود که توسط گیرنده‌های دستگاه ثبت شده و پس از پردازش، به تصاویر قابل تفسیر تبدیل می‌شوند. تفاوت در زمان بازگشت پروتون‌ها در بافت‌های مختلف، امکان ایجاد کنتراست مناسب میان ساختارهای بدن را فراهم می‌کند. ([۱۶]، [۱۸])

---

[۱] Computed Tomography (CT)

[۲] X-ray Imaging

به علت توانایی بالای MRI در نمایش جزئیات بافت‌های نرم، این روش به‌طور گسترده در تصویربرداری مغز، نخاع، ستون فقرات، مفاصل، عضلات، قلب و بسیاری از اندام‌های داخلی مورد استفاده قرار می‌گیرد. تصاویر حاصل از MRI اطلاعات ارزشمندی درباره ساختارهای آناتومیکی و تغییرات پاتولوژیک در اختیار پزشکان قرار می‌دهند و نقش مهمی در تشخیص بیماری‌ها، برنامه‌ریزی درمان و ارزیابی روند بهبود بیماران ایفا می‌کنند.([۱۳]،[۱۶]) همچنین، استفاده از شبکه‌های مبتنی بر یادگیری باقیمانده[۱] در بازسازی تصاویر MRI نیز نتایج موفقی در کاهش خطای بازسازی و بهبود کیفیت تصاویر ارائه کرده است. [۱۸]

همراه با مزایای متعدد، تصویربرداری MRI با چالش‌هایی نیز همراه است. مدت زمان نسبتاً طولانی تصویربرداری، حساسیت بالا به حرکت بیمار، وجود نویز و برخی محدودیت‌های سخت‌افزاری می‌توانند کیفیت تصاویر را تحت تأثیر قرار دهند. کاهش کیفیت تصاویر ممکن است باعث از بین رفتن بخشی از اطلاعات تشخیصی و دشوار شدن تفسیر آن‌ها شود. به همین دلیل، بهبود کیفیت تصاویر MRI و توسعه روش‌های مؤثر برای حذف نویز و آرتیفکت‌ها، یکی از موضوعات مهم و فعال در حوزه پردازش تصاویر پزشکی محسوب می‌شود.([۱۳]، [۲۸])

۲- ۳ آرتیفکت‌های حرکتی در تصاویرMRI

تصاویر MRI با وجود کیفیت بالا و توانایی مناسب در نمایش بافت‌های نرم، در برابر عوامل مختلفی که موجب کاهش کیفیت تصویر می‌شوند حساس هستند. یکی از قابل توجه ترین این عوامل، آرتیفکت‌های حرکتی است که از رایج‌ترین مشکلات در تصویربرداری تشدید مغناطیسی به شمار می‌رود. آرتیفکت‌های حرکتی معمولاً در اثر حرکت ارادی یا غیرارادی بیمار در طول فرآیند تصویربرداری ایجاد می‌شوند و می‌توانند موجب کاهش کیفیت تصویر و از بین رفتن بخشی از اطلاعات تشخیصی شوند. ([۲۸] ، [۱۹])

حرکت بیمار ممکن است به دلایل مختلفی مانند تنفس، ضربان قلب، بلع، حرکات ناخواسته سر یا اندام‌ها و حتی ناتوانی بیمار در ثابت ماندن در طول تصویربرداری رخ دهد. از آنجا که فرآیند تصویربرداری MRI نسبت به بسیاری از روش‌های دیگر زمان بیشتری نیاز دارد، احتمال وقوع این حرکات افزایش یافته و در نتیجه کیفیت تصاویر تحت تأثیر قرار می‌گیرد . ([۱۶] ، [۱۹])

آرتیفکت‌های حرکتی معمولاً به صورت تاری تصویر، ایجاد خطوط اضافی[۲]، اعوجاج ساختارهای آناتومیکی[۳]، کاهش وضوح لبه‌ها و از بین رفتن جزئیات بافتی ظاهر می‌شوند. وجود این اعوجاج‌ها می‌تواند تشخیص صحیح ضایعات یا ناهنجاری‌های کوچک را دشوار کرده و در برخی موارد موجب تفسیر نادرست تصاویر توسط پزشک شود. علاوه بر این، کاهش کیفیت تصاویر ممکن است نیاز به انجام مجدد تصویربرداری را ایجاد کند که علاوه بر افزایش هزینه و زمان، موجب ناراحتی بیشتر بیمار نیز خواهد شد . ([۱۳] ، [۲۸])

---

[۱] Residual Learning

[۲] Ghosting (Ghosting Artifact)

[۳] Distortion of Anatomical Structures

به منظور کاهش اثر آرتیفکت‌های حرکتی، تاکنون روش‌های مختلفی ارائه شده است. برخی از این روش‌ها بر بهبود فرآیند تصویربرداری و اصلاح حرکت در زمان تصویربرداری تمرکز دارند، در حالی که برخی دیگر پس از تشکیل تصویر و با استفاده از الگوریتم‌های پردازش تصویر و یادگیری عمیق، اقدام به حذف آرتیفکت‌ها و بازسازی تصویر می‌کنند. با پیشرفت شبکه‌های عصبی عمیق، استفاده از روش‌های مبتنی بر یادگیری عمیق برای حذف آرتیفکت‌های حرکتی مورد توجه گسترده پژوهشگران قرار گرفته و نتایج امیدوارکننده‌ای در بهبود کیفیت تصاویر MRI ارائه شده است. ([۱۳] ،[۱۶]، [۱۹])

با توجه به اهمیت کیفیت تصاویر MRI در تشخیص بیماری‌ها، توسعه روش‌هایی که بتوانند ضمن حذف مؤثر آرتیفکت‌های حرکتی، ساختارهای آناتومیکی و جزئیات بافتی را نیز حفظ کنند، همچنان یکی از چالش‌های مهم در حوزه پردازش تصاویر پزشکی محسوب می‌شود. بر همین اساس، در پژوهش حاضر از معماری‌های مبتنی بر یادگیری عمیق برای حذف آرتیفکت‌های حرکتی و بازسازی تصاویر MRI استفاده شده است.



۱-۲-۱ تصاویر همراه با آرتیفکت و واضح از MRI زانو [۲۸]

۲- ۴ روش‌های سنتی حذف آرتیفکت و بهبود کیفیت تصویر

پیش از ظهور روش‌های مبتنی بر یادگیری عمیق، پژوهشگران از روش‌های کلاسیک پردازش تصویر برای کاهش نویز، حذف آرتیفکت و بهبود کیفیت تصاویر استفاده می‌کردند. این روش‌ها عمدتاً بر پایه مدل‌های ریاضی، آماری و فیلترهای پردازش تصویر توسعه یافته‌اند و هدف آن‌ها کاهش اثر نویز و بازیابی اطلاعات از دست‌رفته تصویر است. سادگی پیاده‌سازی، سرعت مناسب و نیاز کم به داده‌های آموزشی از مهم‌ترین مزایای این روش‌ها به شمار می‌رود. ([۵]، [۷]، [۱۰])

در روش‌های سنتی، معمولاً نوع نویز یا مدل تخریب تصویر از پیش مشخص فرض می‌شود و سپس با استفاده از الگوریتم‌های مناسب، فرآیند حذف نویز یا بازسازی تصویر انجام می‌شود. این رویکرد اگرچه در بسیاری از مسائل عملکرد قابل قبولی دارد، اما وابستگی آن به فرضیات اولیه باعث می‌شود در تصاویر پیچیده، به‌ویژه تصاویر پزشکی، عملکرد آن کاهش یابد. همچنین بسیاری از این روش‌ها در حذف نویز شدید یا آرتیفکت‌های پیچیده، قادر به حفظ هم‌زمان جزئیات ظریف و ساختارهای آناتومیکی نیستند. ([۵]، [۱۰])

از مهم‌ترین روش‌های کلاسیک مورد استفاده در بهبود کیفیت تصاویر می‌توان به فیلترهای مکانی، فیلترهای فرکانسی، روش‌های مبتنی بر تبدیل فوریه، تبدیل موجک و روش‌های مبتنی بر شباهت نواحی تصویر اشاره کرد. هر یک از این روش‌ها با استفاده از ویژگی‌های متفاوت تصویر، تلاش می‌کنند نویز را کاهش داده و کیفیت بازسازی را بهبود بخشند. با این حال، محدودیت این روش‌ها در استخراج ویژگی‌های پیچیده و وابستگی به تنظیم دقیق پارامترها، زمینه را برای توسعه روش‌های مبتنی بر یادگیری عمیق فراهم کرده است. ([۵]، [۷]، [۱۰]، [۱۳])

در ادامه، مهم‌ترین روش‌های کلاسیک مورد استفاده در بهبود کیفیت تصاویر به‌صورت جداگانه معرفی و بررسی می‌شوند تا مبنایی برای مقایسه آن‌ها با روش‌های مبتنی بر یادگیری عمیق فراهم شود.

۲- ۴- ۱ فیلترهای مکانی و محدودیت در حفظ لبه‌ها

فیلترهای مکانی یکی از ساده‌ترین و پرکاربردترین روش‌های کلاسیک برای کاهش نویز تصویر هستند. در این روش‌ها، مقدار هر پیکسل بر اساس اطلاعات پیکسل‌های مجاور اصلاح می‌شود تا نوسانات ناخواسته کاهش یابد. فیلتر میانگین [۱] و فیلتر گوسی [۲] از متداول‌ترین نمونه‌های این گروه محسوب می‌شوند. سادگی پیاده‌سازی، سرعت مناسب و هزینه محاسباتی پایین از مهم‌ترین مزایای این فیلترها است. ([۵]، [۱۰])

همراه با این مزایا، فیلترهای مکانی معمولاً قادر به تفکیک دقیق نویز از جزئیات واقعی تصویر نیستند. در نتیجه، کاهش نویز اغلب با هموار شدن تصویر همراه بوده و موجب محو شدن لبه‌ها، کاهش کنتراست و از بین رفتن بخشی از بافت‌های ظریف می‌شود. این موضوع در تصاویر پزشکی اهمیت بیشتری دارد، زیرا حفظ مرز ساختارهای آناتومیکی و جزئیات بافتی برای تشخیص صحیح بیماری‌ها ضروری است. ([۱۰]، [۱۳])

۲- ۴- ۲ روش‌های غیرمحلی و استفاده از شباهت نواحی تصویر

برای رفع محدودیت‌های فیلترهای محلی، روش‌های غیرمحلی [۳] معرفی شدند. ایده اصلی این روش‌ها استفاده از شباهت میان نواحی مختلف تصویر است؛ به این معنا که برای بازسازی یک ناحیه، تنها از اطلاعات همسایگی آن استفاده نمی‌شود، بلکه نواحی مشابه در سایر قسمت‌های تصویر نیز در فرآیند بازسازی نقش دارند. [۵]

یکی از شناخته‌شده‌ترین روش‌های این گروه، روش میانگین غیرمحلی [۴] است. Liو Suen با ترکیب روش NLM و نظریه خاکستری، الگوریتمی برای حذف نویز ارائه کردند که از شباهت میان وصله‌های تصویر [۵] جهت افزایش کیفیت بازسازی استفاده می‌کند. [۵] این روش در مقایسه با فیلترهای محلی، توانایی بیشتری در حفظ بافت‌ها و ساختارهای تکرارشونده دارد.

با این حال، عملکرد روش‌های غیرمحلی به دقت یافتن نواحی مشابه وابسته است. در تصاویر دارای نویز شدید یا آرتیفکت‌های پیچیده، تشخیص صحیح این نواحی دشوار شده و کیفیت بازسازی کاهش می‌یابد. همچنین این روش‌ها عمدتاً برای حذف نویز طراحی شده‌اند و توانایی محدودی در بازسازی اطلاعات از دست‌رفته تصویر دارند. ([۵]، [۱۳])

---

[۱] Mean Filter

[۲] Gaussian Filter

[۳] Non-Local Methods

[۴] Non-Local Means (NLM)

[۵] Patches

۲- ۴- ۳ روش‌های حوزه تبدیل و مدل‌های آماری

در این دسته از روش‌ها، تصویر از حوزه مکانی به حوزه‌های دیگری مانند حوزه فرکانس یا نمایش‌های چندمقیاسی منتقل می‌شود. هدف از این تبدیل، جداسازی بهتر اطلاعات تصویر از نویز و استفاده از مدل‌های آماری برای بهبود کیفیت بازسازی است. ([۱۰]، [۲۰])

Rabbani و همکاران با استفاده از مدل‌سازی آماری در هرم جهت‌پذیر[۱] و بهره‌گیری از توزیع لاپلاس محلی، روشی برای حذف نویز ارائه کردند که قادر بود ساختارهای جهت‌دار تصویر را بهتر از فیلترهای ساده حفظ کند.[۱۰] همچنین Cho و Bui با استفاده از مدل‌سازی آماری چندمتغیره در حوزه تبدیل موجک، عملکرد قابل قبول در کاهش نویز تصاویر گزارش کردند. [۲۰]

اگرچه این روش‌ها نسبت به فیلترهای مکانی عملکرد بهتری دارند، اما کیفیت نتایج آن‌ها به انتخاب نوع تبدیل، سطح تجزیه، مدل آماری و پارامترهای آستانه‌گذاری وابسته است. انتخاب نامناسب این پارامترها باعث باقی ماندن نویز، بیش‌هموارسازی تصویر یا ایجاد آرتیفکت‌های جدید شود. از آنجا که تبدیل موجک نقش اساسی در پژوهش حاضر دارد، مبانی و ویژگی‌های آن به‌صورت مستقل در بخش ۵-۲ بررسی خواهد شد.

۲- ۴- ۴ روش BM۳D و جایگاه آن در کاهش نویز کلاسیک

یکی از موفق‌ترین روش‌های کلاسیک کاهش نویز تصویر، BM۳D[۲] است که توسط Dabov و همکاران معرفی شد .این روش با ترکیب ایده شباهت غیرمحلی و فیلترگذاری در حوزه تبدیل، اول وصله‌های مشابه تصویر را شناسایی کرده و سپس آن‌ها را در قالب گروه‌های سه‌بعدی پردازش می‌کند تا نویز کاهش یافته و ساختارهای مشترک تصویر حفظ شوند. [۲۱]

BM۳D سال‌ها به عنوان یکی از معیارهای اصلی مقایسه روش‌های حذف نویز شناخته می‌شد و در بسیاری از تصاویر طبیعی نتایج مطلوبی ارائه داد. با این وجود، این روش نیز بر پایه فرضیات مشخصی درباره نوع نویز و شباهت میان بلوک‌های تصویر طراحی شده است و در تصاویر دارای تخریب‌های پیچیده یا آرتیفکت‌های شدید با محدودیت‌هایی مواجه می‌شود. [۲۱]

از سوی دیگر، BM۳D یک روش یادگیرنده نیست و قادر به استخراج روابط پیچیده میان تصویر تخریب‌شده و تصویر مرجع از طریق داده‌های آموزشی نیست. این تفاوت، یکی از دلایل اصلی گرایش پژوهشگران به استفاده از مدل‌های مبتنی بر یادگیری عمیق برای بازسازی تصاویر محسوب می‌شود. ([۱۳]، [۲۱])

علاوه بر این، روش‌های مبتنی بر حفظ لبه‌ها و استفاده از اطلاعات گرادیان نیز برای بهبود کیفیت حذف نویز پیشنهاد شده‌اند. [۲۵] همچنین، برخی پژوهش‌ها به بررسی مدل‌سازی نویز و هم‌بستگی میان کانال‌های مختلف تصویر پرداخته‌اند تا دقت فرآیند حذف نویز در تصاویر پزشکی افزایش یابد. [۸]

---

[۱] Steerable Pyramid

[۲] Block Matching and ۳D Filtering

۲- ۴- ۵ روش‌های ترکیبی مبتنی بر موجک و بهینه‌سازی

در ادامه توسعه روش‌های کلاسیک، برخی پژوهشگران تلاش کردند توانایی تحلیل چندمقیاسی تبدیل موجک را با الگوریتم‌های بهینه‌سازی ترکیب کنند. Cho و همکاران روشی مبتنی بر تبدیل موجک، آستانه‌گذاری ضرایب و الگوریتم ژنتیک ارائه کردند که هدف آن انتخاب بهینه پارامترهای آستانه‌گذاری و افزایش کیفیت حذف نویز بود. [۷]

این دسته از روش‌ها نسبت به فیلترهای ساده عملکرد بهتری ارائه می‌کنند، اما همچنان وابسته به انتخاب نوع موجک، سطح تجزیه، تابع آستانه‌گذاری و پارامترهای بهینه‌سازی هستند. به همین علت، در تصاویر دارای نویزهای پیچیده یا تخریب‌های شدید، انعطاف‌پذیری محدودی دارند و قادر به یادگیری ویژگی‌های پیچیده تصویر نیستند. ([۷]، [۲۰])

به طور کلی، روش‌های سنتی نقش مهمی در توسعه الگوریتم‌های پردازش تصویر ایفا کرده‌اند و بسیاری از مفاهیم مورد استفاده در روش‌های نوین، از جمله تحلیل چندمقیاسی، استفاده از شباهت نواحی تصویر و پردازش در حوزه تبدیل، ریشه در همین روش‌ها دارند. در همین راستا، محدودیت این روش‌ها در استخراج خودکار ویژگی‌ها و بازسازی اطلاعات از دست‌رفته، زمینه را برای توسعه روش‌های مبتنی بر یادگیری عمیق فراهم کرده است. ([۵]، [۷]، [۱۰]، [۱۳]، [۲۰]، [۲۱])

در پایان، جدول ۲-۱ خلاصه‌ای از مهم‌ترین روش‌های سنتی، ایده اصلی، مزایا و محدودیت‌های هر یک را ارائه می‌دهد و مبنایی برای مقایسه آن‌ها با روش‌های مبتنی بر یادگیری عمیق در بخش‌های بعدی فراهم می‌کند.

روش /پژوهش	ایده اصلی	محدودیت	مزیت
فیلترهای میانگین گیری و گوسی	اصلاح مقدار پیکسل بر اساس همسایگی	محو شدن لبه‌ها و حذف جزئیات ظریف	ساده، سریع و قابل پیاده سازی
Suen و Li [۵]	استفاده از Non-Local Means و نظریه خاکستری	کاهش کارایی در نویزهای پیچیده و تخریب شدید	استفاده از شباهت نواحی غیر محلی تصویر
Rabbani [۱۰]	کاهش نویز در حوزه steerable pyramid با توزیع لاپلاس محلی	وابستگی به مدل‌های آماری و انتخاب پارامترها	تحلیل چند مقیاسی و جهت دار تصویر
Bui و Cho [۲۰]	مدل سازی آماری چند متغیره در حوزه موجک	حساسیت به نوع موجک و فرضیات آماری	مناسب برای تحلیل ساختارهای چند مقیاسی
Dabov /BM۳D [۲۱]	گروه بندی بلوک‌های مشابه و فیلتر گذاری سه بعدی	طراحی شده برای حذف نویز، نه بازسازی اطلاعات از دست رفته	عملکرد قوی در حذف نویز کلاسیک
Liu [۷]	وابستگی به تنظیمات دستی و معیار بهینه سازی	بهبود انتخاب پارامترها نسبت به آستانه گذاری ساده	ترکیب تبدیل موجک وآستانه گذاری

همان‌طور که در جدول ۲-۱ مشاهده می‌شود، هر یک از روش‌های سنتی معرفی‌شده دارای نقاط قوت و محدودیت‌های خاص خود هستند. این روش‌ها به‌دلیل سادگی پیاده‌سازی، توانایی تفسیر مناسب و هزینه محاسباتی پایین، سال‌ها به‌عنوان راهکارهای اصلی در بهبود کیفیت تصاویر مورد استفاده قرار گرفته‌اند. با این حال، عملکرد آن‌ها تا حد زیادی به فرضیات اولیه درباره نوع نویز، مدل تخریب تصویر و تنظیم پارامترهای الگوریتم وابسته است. در شرایط واقعی، به‌ویژه در تصاویر پزشکی، نوع نویز و آرتیفکت‌ها معمولاً پیچیده بوده و ممکن است چندین عامل به‌طور هم‌زمان بر کیفیت تصویر تأثیر بگذارند؛ ازاین‌رو، استفاده از روش‌های مبتنی بر قواعد ثابت همواره نتایج مطلوبی به همراه ندارد.([۵]، [۷]، [۱۰]، [۲۱])

علاوه بر این، در بسیاری از کاربردهای تصویربرداری پزشکی، هدف تنها کاهش نویز یا بهبود ظاهر تصویر نیست، بلکه حفظ ساختارهای آناتومیکی، جلوگیری از حذف جزئیات بافتی و بازسازی اطلاعات از دست‌رفته نیز اهمیت بالایی دارد. اگرچه روش‌های سنتی در کاهش نویز عملکرد قابل قبولی دارند، اما در مواجهه با تخریب‌های پیچیده و

ساختارهای ظریف، معمولاً با محدودیت‌هایی روبه‌رو هستند و توانایی محدودی در استخراج ویژگی‌های سطح بالا و بازسازی دقیق تصویر دارند ([۱۳]، [۲۰]، [۲۱])

با توجه به این محدودیت‌ها، پژوهشگران به سمت توسعه روش‌هایی حرکت کردند که علاوه بر استفاده از اطلاعات آماری و چندمقیاسی تصویر، توانایی یادگیری الگوهای پیچیده را نیز داشته باشند. یکی از بهترین این رویکردها، استفاده از تبدیل موجک است که در اثر توانایی در تحلیل چندمقیاسی و حفظ اطلاعات فرکانسی، جایگاه ویژه‌ای در پردازش تصاویر پیدا کرده است. ازاین‌رو، در بخش بعدی مبانی و ویژگی‌های تبدیل موجک به‌عنوان یکی از اجزای اصلی پژوهش حاضر بررسی می‌شود. ([۷]، [۲۰])

۲- ۵ روش‌های مبتنی بر تبدیل موجک در پردازش تصویر

پس از توسعه روش‌های کلاسیک کاهش نویز، استفاده از نمایش‌های چندمقیاسی به یکی از رویکردهای مهم در پردازش تصویر تبدیل شد. در این رویکرد، تصویر تنها در حوزه مکانی بررسی نمی‌شود، بلکه به مؤلفه‌هایی با مقیاس‌ها و فرکانس‌های مختلف تجزیه می‌شود .تبدیل موجک یکی از قابل قبول ترین ابزارهای این حوزه است که به‌دلیل توانایی در نمایش هم‌زمان اطلاعات مکانی و فرکانسی، کاربرد گسترده‌ای در تحلیل، فشرده‌سازی، حذف نویز و بازسازی تصاویر پیدا کرده است.([۳]، [۷]، [۲۰])

برخلاف تبدیل فوریه که تنها اطلاعات فرکانسی سیگنال را ارائه می‌دهد، تبدیل موجک علاوه بر اطلاعات فرکانسی، محل وقوع تغییرات را نیز مشخص می‌کند. این ویژگی باعث می‌شود موجک برای تحلیل تصاویر غیرایستا و تصاویر دارای لبه‌ها، بافت‌ها و جزئیات محلی مناسب‌تر باشد. در نتیجه، این تبدیل توانایی بیشتری در حفظ ساختارهای مهم تصویر در فرآیند حذف نویز و بازسازی تصویر دارد. ([۳]، [۲۰])

اهمیت تبدیل موجک در پردازش تصویر از آنجا ناشی می‌شود که نویز و اطلاعات مفید تصویر معمولاً در یک باند فرکانسی مشخص قرار ندارند. در بسیاری از تصاویر، به‌ویژه تصاویر پزشکی، بخشی از اطلاعات مربوط به لبه‌ها و ساختارهای آناتومیکی در همان نواحی فرکانسی قرار دارد که مؤلفه‌های نویزی نیز در آن حضور دارند. بنابراین، حذف مستقیم مؤلفه‌های فرکانسی بالا می‌تواند علاوه بر کاهش نویز، موجب از بین رفتن جزئیات مهم تصویر شود. تبدیل موجک با تجزیه تصویر به زیرباندهای مختلف، امکان تفکیک مناسب‌تر اطلاعات ساختاری و مؤلفه‌های نویزی را فراهم می‌کند و در نتیجه کیفیت بازسازی تصویر را بهبود می‌بخشد. ([۷]، [۲۰])

در تجزیه موجکی دوبعدی، تصویر معمولاً به چهار زیرباند اصلی شامل LL، LH، HL و HH تقسیم می‌شود. زیرباند LL اطلاعات کم‌فرکانس و ساختار کلی تصویر را در خود نگه می‌دارد، در حالی که زیرباندهای LH، HL و HH به‌ترتیب اطلاعات مربوط به جزئیات افقی، عمودی و قطری را نمایش می‌دهند. این تفکیک چندمقیاسی امکان پردازش مستقل هر بخش از اطلاعات تصویر را فراهم کرده و نقش مهمی در بسیاری از روش‌های نوین بازسازی تصویر ایفا می‌کند. ([۳]، [۲۰])

پژوهشگران از تبدیل موجک نه تنها به‌عنوان یک ابزار کلاسیک کاهش نویز، بلکه به‌عنوان بخشی از معماری‌های یادگیری عمیق نیز استفاده کرده‌اند. ترکیب توانایی تحلیل چندمقیاسی موجک با توانایی شبکه‌های عصبی در

استخراج ویژگی‌های پیچیده، زمینه توسعه مدل‌هایی مانند Wavelet-Attention CNN (WACNN) و Multi-Level Wavelet-Attention CNN (MWACNN) را فراهم کرده است. در این معماری‌ها، اطلاعات موجود در زیرباندهای موجکی به‌صورت هدفمند در فرآیند استخراج ویژگی و بازسازی تصویر مورد استفاده قرار می‌گیرند و همین موضوع می‌تواند موجب حفظ بهتر جزئیات ساختاری و افزایش کیفیت بازسازی تصاویر شود. ([۱]، [۲]، [۳]) در سال‌های اخیر، انواع پیشرفته‌تری از کانولوشن، از جمله کانولوشن پویا[۱] و کانولوشن پویای گاوسی[۲] ، برای افزایش توانایی استخراج ویژگی و بهبود عملکرد شبکه‌های کانولوشنی معرفی شده‌اند.([۶] ، [۱۲])

جدول ۲-۲ نقش زیرباندهای مختلف تبدیل موجک را در نمایش اطلاعات تصویر نشان می‌دهد.

۲ ۲-۲ نقش زیرباندهای حاصل از تبدیل موجک دوبعدی در پردازش و بازسازی تصاویر

زیرباند موجکی	نوع اطلاعات	نقش در بازسازی تصویر	اهمیت در تصاویر پزشکی
LL	مؤلفه‌های کم‌فرکانس و ساختار کلی تصویر	حفظ روشنایی، کنتراست کلی و ساختار اصلی تصویر	حفظ اطلاعات آناتومیکی و ساختار کلی بافت‌ها
LH	جزئیات افقی (تغییرات در راستای افقی)	بازیابی لبه‌ها و جزئیات افقی تصویر	حفظ مرزهای افقی بافت‌ها و ساختارهای آناتومیکی
HL	جزئیات عمودی (تغییرات در راستای عمودی)	بازسازی لبه‌ها و تغییرات عمودی	افزایش دقت در نمایش مرزهای عمودی و ساختارهای ظریف
HH	جزئیات قطری و مؤلفه‌های پرفرکانس	نمایش تغییرات قطری، بافت‌های ریز و مؤلفه‌های فرکانس بالا	تفکیک بهتر جزئیات واقعی از نویز و حفظ بافت‌های ظریف

همان‌طور که در جدول ۲-۲ مشاهده می‌شود، تمامی مؤلفه‌های پرفرکانس حاصل از تبدیل موجک را نمی‌توان به‌عنوان نویز در نظر گرفت. بخش قابل توجهی از این مؤلفه‌ها شامل اطلاعات ارزشمندی مانند لبه‌ها، مرزهای بافتی و جزئیات ظریف تصویر هستند. بنابراین، یک روش موجکی مناسب باید بتواند میان مؤلفه‌های نویزی و اطلاعات واقعی تصویر تمایز قائل شود. این موضوع در تصاویر پزشکی اهمیت بیشتری دارد، زیرا حذف نادرست مؤلفه‌های پرفرکانس می‌تواند موجب از بین رفتن اطلاعات تشخیصی و کاهش کیفیت بازسازی تصویر شود. ([۷]، [۲۰])

پژوهش‌های متعددی نشان داده‌اند که استفاده از نمایش‌های چندمقیاسی مبتنی بر تبدیل موجک می‌تواند در کاهش نویز و حفظ ساختار تصویر مؤثر باشد. ([۷]، [۱۰]، [۲۰]) تفاوت اصلی این روش‌ها با فیلترهای متداول در آن است که تصویر را تنها بر اساس همسایگی پیکسل‌ها پردازش نمی‌کنند، بلکه اطلاعات تصویر را در مقیاس‌ها و زیرباندهای مختلف مورد تحلیل قرار می‌دهند. در همین راستا، عملکرد بسیاری از روش‌های موجکی سنتی به انتخاب نوع موجک، تعداد سطوح تجزیه، مدل آماری ضرایب و روش آستانه‌گذاری وابسته است و این وابستگی می‌تواند استفاده از آن‌ها را در شرایط مختلف با محدودیت مواجه کند. ([۷]، [۲۰])

[۱] Dynamic Convolution

[۲] Gaussian Dynamic Convolution (GDC)

این محدودیت‌ها در تصاویر پزشکی اهمیت بیشتری پیدا می‌کنند، زیرا هدف تنها کاهش نویز نیست، بلکه حفظ ساختارهای آناتومیکی، مرزهای بافتی و جزئیات کم‌کنتراست نیز از اهمیت بالایی برخوردار است. اگر آستانه‌گذاری بیش از حد شدید باشد، بخشی از اطلاعات تشخیصی تصویر حذف می‌شود و در صورت انتخاب آستانه نامناسب، نویز به میزان کافی کاهش نخواهد یافت. بنابراین، اگرچه تبدیل موجک ابزار قدرتمندی برای تحلیل چندمقیاسی تصاویر محسوب می‌شود، اما به‌تنهایی قادر به بازسازی دقیق اطلاعات از دست‌رفته در تصاویر پیچیده نیست. ([۷]، [۲۰])

از این رو، این نوع از عملیات‌ها به سمت ترکیب تبدیل موجک با روش‌های مبتنی بر یادگیری عمیق حرکت کرده‌اند. در این رویکرد، تبدیل موجک اطلاعات چندمقیاسی تصویر را استخراج می‌کند و شبکه عصبی با یادگیری روابط پیچیده میان تصویر تخریب‌شده و تصویر مرجع، فرآیند بازسازی را انجام می‌دهد. این ترکیب موجب می‌شود مدل علاوه بر بهره‌گیری از اطلاعات فرکانسی و مکانی تصویر، وابستگی کمتری به تنظیمات دستی و پارامترهای از پیش تعیین‌شده داشته باشد. ([۱]، [۲]، [۳])

در همین راستا، مدل Multi-Level Wavelet CNN با هدف ترکیب تجزیه چندسطحی موجک و شبکه‌های کانولوشنی ارائه شد. [۳] در این معماری، اطلاعات استخراج‌شده از زیرباندهای مختلف موجکی به‌صورت هم‌زمان در فرآیند استخراج ویژگی مورد استفاده قرار می‌گیرند و شبکه قادر است ویژگی‌های تصویر را در مقیاس‌های مختلف یاد بگیرد. این رویکرد نشان می‌دهد که تبدیل موجک فراتر از یک مرحله پیش‌پردازش عمل کرده و به بخشی از ساختار اصلی شبکه تبدیل می‌شود.

علاوه بر این، مدل پیش‌بینی عمیق مبتنی بر موجک[۱] نیز از یادگیری ضرایب موجکی برای افزایش وضوح و بازسازی تصاویر استفاده می‌کند. [۲۴] در این روش، به‌جای یادگیری مستقیم تصویر در حوزه پیکسل، شبکه روابط میان ضرایب موجکی را یاد می‌گیرد و از آن‌ها برای بازسازی تصویر باکیفیت‌تر استفاده می‌کند. این رویکرد نشان‌دهنده ظرفیت بالای تبدیل موجک در توسعه معماری‌های نوین بازسازی تصویر است.

برای جمع‌بندی، جدول ۲-۳ روند تکامل روش‌های مبتنی بر موجک را از روش‌های کلاسیک تا معماری‌های مبتنی بر یادگیری عمیق نشان می‌دهد. همان‌طور که مشاهده می‌شود، تبدیل موجک به‌تدریج از یک ابزار مستقل برای حذف نویز به یکی از اجزای مهم معماری‌های یادگیری عمیق تبدیل شده است. این روند، زمینه توسعه مدل‌های مبتنی بر موجک مانند WACNN و MWACNN را فراهم کرده است که در بخش‌های بعدی این فصل معرفی و بررسی خواهند شد. ([۱]، [۲]، [۳])

[۱] Deep Wavelet Prediction (DWP)

مرجع	محدودیت‌ها	مزایا	ایده اصلی	روش
۷	وابستگی به نوع موجک، سطح تجزیه و مقدار آستانه	ساده، قابل پیاده‌سازی و مناسب برای کاهش نویزهای شناخته‌شده	تجزیه تصویر به زیرباندهای موجکی و حذف یا تضعیف ضرایب نویزی	تبدیل موجک همراه با آستانه‌گذاری
۱۰، ۲۰	وابستگی به مدل آماری و تنظیم پارامترها	حفظ بهتر ساختار تصویر نسبت به فیلترهای کلاسیک	مدل‌سازی آماری ضرایب موجکی و تحلیل چندمقیاسی تصویر	روش‌های آماری مبتنی بر موجک
۳	پیچیدگی بیشتر معماری و نیاز به داده آموزشی مناسب	ترکیب اطلاعات فرکانسی و مکانی با یادگیری عمیق	ادغام تبدیل موجک با شبکه‌های عصبی کانولوشنی برای استخراج ویژگی‌های چندمقیاسی	Wavelet-CNN
۴، ۲۴	تمرکز بیشتر بر افزایش وضوح تصویر (Super-Resolution) و نه حذف آرتیفکت‌هایMRI	حفظ بهتر جزئیات و مؤلفه‌های پرفرکانس، مناسب برای افزایش وضوح تصویر	یادگیری ضرایب موجکی به‌جای بازسازی مستقیم تصویر در حوزه پیکسل	Deep Wavelet Prediction

در مجموع، تبدیل موجک یکی از قابل توجه ترین ابزارهای تحلیل چندمقیاسی در پردازش تصویر محسوب می‌شود و نسبت به بسیاری از روش‌های کلاسیک، توانایی بیشتری در حفظ اطلاعات مکانی و فرکانسی تصویر دارد. این ویژگی باعث شده است که تبدیل موجک در کاربردهایی مانند حذف نویز، بهبود کیفیت تصویر و بازسازی تصاویر پزشکی به‌طور گسترده مورد استفاده قرار گیرد. در همین راستا، وابستگی عملکرد این روش به عواملی مانند نوع موجک، تعداد سطوح تجزیه، روش آستانه‌گذاری و تنظیم پارامترها، استفاده از آن را در شرایط مختلف با محدودیت‌هایی مواجه می‌کند. ([۷]، [۲۰])

این محدودیت‌ها موجب شد پژوهشگران به سمت استفاده از روش‌های مبتنی بر یادگیری عمیق حرکت کنند؛ روش‌هایی که بدون وابستگی به قواعد ثابت، قادرند ویژگی‌های پیچیده تصویر را به‌صورت خودکار از داده‌های آموزشی استخراج کرده و فرآیند بازسازی تصویر را با دقت بیشتری انجام دهند. ازاین‌رو، در بخش بعدی، مبانی یادگیری عمیق و کاربرد آن در پردازش تصاویر پزشکی بررسی خواهد شد و سپس معماری‌های مبتنی بر تبدیل موجک مورد استفاده در این پژوهش معرفی می‌شوند. ([۱]، [۳]، [۱۳])

۲- ۶ یادگیری عمیق و شبکه‌های عصبی کانولوشنی در بازسازی تصویر

با پیشرفت یادگیری عمیق، رویکردهای بهبود کیفیت و بازسازی تصویر از روش‌های مبتنی بر قواعد ثابت به سمت مدل‌های داده‌محور تغییر یافتند. در روش‌های سنتی، نوع فیلتر، مدل آماری یا پارامترهای پردازش باید به‌صورت

دستی تعیین می‌شد، در حالی که مدل‌های یادگیری عمیق قادرند رابطه میان تصویر تخریب‌شده و تصویر مرجع را مستقیماً از داده‌های آموزشی یاد بگیرند. این ویژگی موجب شده است که یادگیری عمیق در مسائل پیچیده‌ای مانند حذف نویز، حذف آرتیفکت، افزایش وضوح و بازسازی تصاویر پزشکی عملکرد قابل‌توجهی داشته باشد. [۱۳]

در میان روش‌های یادگیری عمیق، شبکه‌های عصبی کانولوشنی Convolutional Neural Networks یا CNN در اثر سازگاری با ساختار مکانی تصاویر، جایگاه ویژه‌ای در پردازش تصویر پیدا کرده‌اند. این شبکه‌ها با استفاده از عملیات کانولوشن، ویژگی‌هایی مانند لبه‌ها، بافت‌ها و ساختارهای محلی را استخراج می‌کنند. برخلاف فیلترهای کلاسیک که ضرایب آن‌ها از پیش تعیین می‌شود، فیلترهای کانولوشنی در طی فرآیند آموزش به‌صورت خودکار یاد گرفته می‌شوند و به این علت قادرند خود را با ویژگی‌های داده و نوع تخریب تصویر سازگار کنند. ([۱۳]، [۱۴]) همچنین چارچوب فریم‌لت‌های کانولوشنی عمیق[۱] به‌عنوان یکی از روش‌های مؤثر برای حل مسائل معکوس در بازسازی تصاویر معرفی شده است. [۲۲] مطالعات دیگری نیز نقش عملگر کانولوشن را در استخراج ویژگی‌های معنایی و ساختاری تصاویر به‌صورت دقیق بررسی کرده‌اند. [۲۳]

علاوه بر شبکه‌های کانولوشنی متداول، معماری‌های مبتنی بر شبکه عصبی کانولوشنی[۲] با مقادیر مختلط[۳] نیز برای پردازش تصاویر MRI پیشنهاد شده‌اند. این شبکه‌ها با بهره‌گیری از مقادیر مختلط، قادرند اطلاعات دامنه و فاز را به‌صورت هم‌زمان مدل‌سازی کنند و در برخی کاربردهای تصویربرداری پزشکی عملکرد متناسبی از خود نشان داده‌اند. [۹]

مرور جامع Tian و همکاران نشان می‌دهد که شبکه‌های عصبی کانولوشنی در بسیاری از مسائل کاهش نویز و بازسازی تصویر، عملکرد بهتری نسبت به روش‌های سنتی ارائه کرده‌اند. [۱۳] دلیل اصلی این برتری، توانایی شبکه در یادگیری نگاشت غیرخطی میان تصویر ورودی و تصویر مرجع است؛ تواناییی که امکان مدل‌سازی تخریب‌های پیچیده را فراهم می‌کند. با این حال، عملکرد این مدل‌ها همچنان به عواملی مانند کیفیت داده‌های آموزشی، طراحی معماری، تابع هزینه و روش آموزش وابسته است.

یکی از نخستین رویکردهای موفق در این زمینه، استفاده از یادگیری باقی‌مانده در شبکه‌های کانولوشنی بود. در این رویکرد، شبکه به‌جای تولید مستقیم تصویر بازسازی‌شده، اختلاف میان تصویر تخریب‌شده و تصویر مرجع را یاد می‌گیرد. این ایده موجب پایداری بیشتر فرآیند آموزش و افزایش دقت حذف نویز شد و زمینه توسعه معماری‌های پیشرفته‌تر را فراهم کرد. در همین راستا، بسیاری از این مدل‌ها تنها برای سطوح مشخصی از نویز آموزش داده می‌شدند و در مواجهه با شرایط متنوع، انعطاف‌پذیری محدودی داشتند. [۱۸]

در ادامه، Zhang و همکاران با معرفی FFDNet[۴] تلاش کردند این محدودیت را برطرف کنند. این معماری با درنظر گرفتن نقشه شدت نویز[۵] ، قادر است با سطوح مختلف نویز سازگار شده و نتیجه ی قابل قبولی در حذف نویز ارائه دهد. [۱۴] در کنار این، FFDNetهمچنان عمدتاً در حوزه پیکسل عمل می‌کند و اطلاعات چندمقیاسی یا

---

[۱] Deep Convolutional Framelets (DCF)

[۲] Convolutional Neural Network (CNN)

[۳] Complex-Valued Convolutional Neural Network (Complex-Valued CNN)

[۴] Fast and Flexible Denoising Convolutional Neural Network (FFDNet)

[۵] Noise Level Map

فرکانسی تصویر را به‌صورت صریح در فرآیند یادگیری به کار نمی‌گیرد. این موضوع یکی از انگیزه‌های توسعه معماری‌هایی بود که از تبدیل موجک در کنار شبکه‌های کانولوشنی استفاده می‌کنند.

هم‌زمان با توسعه شبکه‌های کانولوشنی، پژوهشگران تلاش کردند با استفاده از مکانیزم‌های توجه ، توانایی شبکه را در تمرکز بر ویژگی‌های مهم تصویر افزایش دهند. در این معماری‌ها فرض بر آن است که تمام نواحی یا کانال‌های تصویر اهمیت یکسانی ندارند و شبکه باید بتواند بخش‌های مهم‌تر را تشخیص داده و وزن بیشتری به آن‌ها اختصاص دهد. مطالعات مختلف نشان داده‌اند که استفاده از مکانیزم توجه توانایی موجب حفظ بهتر ساختارهای تصویر و بهبود کیفیت بازسازی را دارد. [۱]

در کنار این پیشرفت‌ها، معماری‌های متعددی برای مسائل تصویر به تصویر[۱] معرفی شدند که U-Net یکی از شناخته‌شده‌ترین آن‌ها است.[۱۷] همچنین، توسعه معماری‌های مبتنی بر تبدیل موجک مانندWACNN وMWACNN نشان داد که ترکیب تحلیل چندمقیاسی موجک با توانایی یادگیری شبکه‌های کانولوشنی می‌تواند کیفیت بازسازی تصاویر را بیش از پیش بهبود دهد. ([۲]، [۳]) ازاین‌رو، این معماری‌ها در علم نوین توجه زیادی را در حوزه بازسازی تصاویر پزشکی به خود جلب کرده‌اند.

به طور کلی، مرور پژوهش‌های انجام‌شده نشان می‌دهد که روند توسعه روش‌های بازسازی تصویر از فیلترهای ثابت و مدل‌های آماری به سمت معماری‌های عمیق، شبکه‌های مبتنی بر توجه و مدل‌های ترکیبی موجک–کانولوشنی حرکت کرده است. در ادامه، مهم‌ترین معماری‌های مرتبط با این پژوهش شامل U-Net، WACNN، MWACNN و مکانیزم توجه به‌صورت جداگانه معرفی و بررسی می‌شوند.

[۱] Image-to-Image Translation

مدل	ایده اصلی	مزایا	محدودیت‌ها
Tianو همکاران[۱۳]	مرور و تحلیل روش‌های مبتنی بر یادگیری عمیق در کاهش نویز تصویر	ارائه دیدگاه جامع از روند تحول روش‌های سنتی به مدل‌های داده‌محور	مقاله مروری است و معماری جدیدی ارائه نمی‌کند
FFDNet [۱۴]	استفاده از شبکه کانولوشنی همراه با Noise Level Map برای حذف نویز در سطوح مختلف	انعطاف‌پذیری در مواجهه با شدت‌های مختلف نویز و سرعت پردازش مناسب	پردازش عمدتاً در حوزه پیکسل و عدم استفاده صریح از اطلاعات چندمقیاسی
Attention guided CNN [۱]	استفاده از مکانیزم توجه برای تمرکز بر ویژگی‌ها و نواحی مهم تصویر	بهبود استخراج ویژگی‌های مؤثر و افزایش کیفیت بازسازی	افزایش پیچیدگی معماری و نیاز به آموزش دقیق‌تر
U-Net [۱۷]	معماری Encoder–Decoder همراه با Skip Connection	حفظ بهتر اطلاعات مکانی و عملکرد مناسب در مسائل تصویر به تصویر	عدم بهره‌گیری مستقیم از اطلاعات فرکانسی یا نمایش موجکی در ساختار پایه
Residual Dense Network (RDN) [۲۷]	استفاده از اتصال‌های Residual و Dense برای افزایش وضوح تصویر	انتقال مؤثرتر ویژگی‌ها و حفظ بهتر اطلاعات در لایه‌های عمیق	تمرکز اصلی بر افزایش وضوح تصویر (Super Resolution) و نیاز به محاسبات بیشتر

همان‌طور که در جدول ۲-۴ مشاهده می‌شود، مدل‌های مبتنی بر یادگیری عمیق نسبت به روش‌های سنتی توانایی بیشتری در یادگیری الگوهای پیچیده تصویر دارند. این مدل‌ها با استفاده از داده‌های آموزشی می‌توانند برای وظایفی مانند کاهش نویز، افزایش وضوح و بازسازی تصویر به کار گرفته شوند. با این حال، بسیاری از این معماری‌ها تصویر را عمدتاً در حوزه پیکسل پردازش می‌کنند و اطلاعات فرکانسی یا چندمقیاسی به‌صورت صریح در فرآیند یادگیری آن‌ها لحاظ نمی‌شود ([۱۳]، [۱۴])

این موضوع در بازسازی تصاویر پزشکی اهمیت بیشتری پیدا می‌کند. در تصاویر MRI، هدف تنها کاهش نویز یا بهبود ظاهر تصویر نیست، بلکه حفظ مرز بافت‌ها، ساختارهای آناتومیکی و جزئیات کم‌کنتراست نیز از اهمیت بالایی برخوردار است. ازاین‌رو، یک مدل بازسازی مناسب باید علاوه بر حذف آرتیفکت‌ها، قادر به حفظ اطلاعات ساختاری و بازیابی جزئیات از دست‌رفته نیز باشد. ([۱۶]،[۱۸]، [۱۹])

پژوهش‌های اخیر نشان می‌دهند که استفاده از داده‌های جفت‌شده شامل تصویر تخریب‌شده و تصویر مرجع، یکی از رایج‌ترین رویکردها برای آموزش مدل‌های بازسازی تصویر است. در این روش، شبکه با یادگیری نگاشت میان تصویر ورودی و تصویر مرجع، تلاش می‌کند تصویر باکیفیت را بازسازی کند. این رویکرد نسبت به روش‌های مبتنی بر قواعد ثابت، انعطاف‌پذیری بیشتری در مدل‌سازی تخریب‌های پیچیده دارد و امکان یادگیری ویژگی‌های غیرخطی تصویر را فراهم می‌کند. ([۱۳]، [۱۴])

علاوه بر پیشرفت‌های قابل توجه، مطالعات مختلف نشان داده‌اند که استفاده از شبکه‌های کانولوشنی به‌تنهایی همواره برای حفظ کامل اطلاعات ساختاری و فرکانسی کافی نیست. به همین علت، به مرور زمان نوآوران به سمت توسعه معماری‌هایی حرکت کرده‌اند که توانایی یادگیری شبکه‌های عمیق را با تحلیل چندمقیاسی تصویر ترکیب می‌کنند. این رویکرد زمینه شکل‌گیری معماری‌های مبتنی بر تبدیل موجک و مکانیزم‌های توجه را فراهم کرده است که توانایی بیشتری در استخراج ویژگی‌های چندمقیاسی و حفظ جزئیات ساختاری دارند. ([۱]، [۲]، [۳])

به طور کلی، یادگیری عمیق و شبکه‌های عصبی کانولوشنی تحول چشمگیری در حوزه بازسازی تصویر ایجاد کرده‌اند و محدودیت بسیاری از روش‌های کلاسیک را برطرف ساخته‌اند. با این حال، وابستگی به داده‌های آموزشی، هزینه محاسباتی بالا و عدم استفاده صریح از اطلاعات فرکانسی همچنان از چالش‌های مهم این معماری‌ها به شمار می‌رود. ازاین‌رو، در بخش‌های بعدی، معماری‌های U-Net، WACNN، MWACNN و مکانیزم توجه که مستقیماً با موضوع این پژوهش مرتبط هستند، به‌صورت جداگانه معرفی و بررسی می‌شوند.

۲- ۷ معماری U-Net

یکی از بهترین پیشرفت‌ها در حوزه پردازش تصاویر پزشکی، معرفی معماری U-Netتوسط Ronneberger، Fischer و Brox در سال ۲۰۱۵ بود. [۱۷] این معماری در ابتدا با هدف بخش‌بندی تصاویر زیست‌پزشکی طراحی شد، اما به علت توانایی بالا در استخراج ویژگی‌ها و بازسازی اطلاعات مکانی، به‌سرعت در بسیاری از مسائل پردازش تصویر از جمله حذف نویز، حذف آرتیفکت، بازسازی تصویر، افزایش وضوح و بازسازی تصاویر MRI مورد استفاده قرار گرفت. ([۱۷]،[۱۳])

دلیل اصلی موفقیت U-Net، ساختار متقارن رمزگذار—رمزگشا آن است. این ساختار به شبکه اجازه می‌دهد که ویژگی‌های تصویر را در سطوح مختلف استخراج کرده و سپس با استفاده از همان ویژگی‌ها، تصویر را با حفظ اطلاعات مکانی بازسازی کند. برخلاف بسیاری از شبکه‌های کانولوشنی که تنها بر استخراج ویژگی تمرکز دارند .

U-Net برای مسائل Image-to-Image طراحی شده است؛ یعنی ورودی و خروجی هر دو تصویر هستند و هدف، تبدیل یک تصویر به نسخه‌ای با کیفیت بالاتر یا دارای اطلاعات کامل‌تر است [۱۷]

۲- ۷- ۱ ساختار کلی U-Net

معماری U-Net از دو مسیر اصلی تشکیل شده است:

- مسیر فشرده‌سازی Encoder یا Contracting Path
- مسیر بازسازی Decoder یا Expanding Path

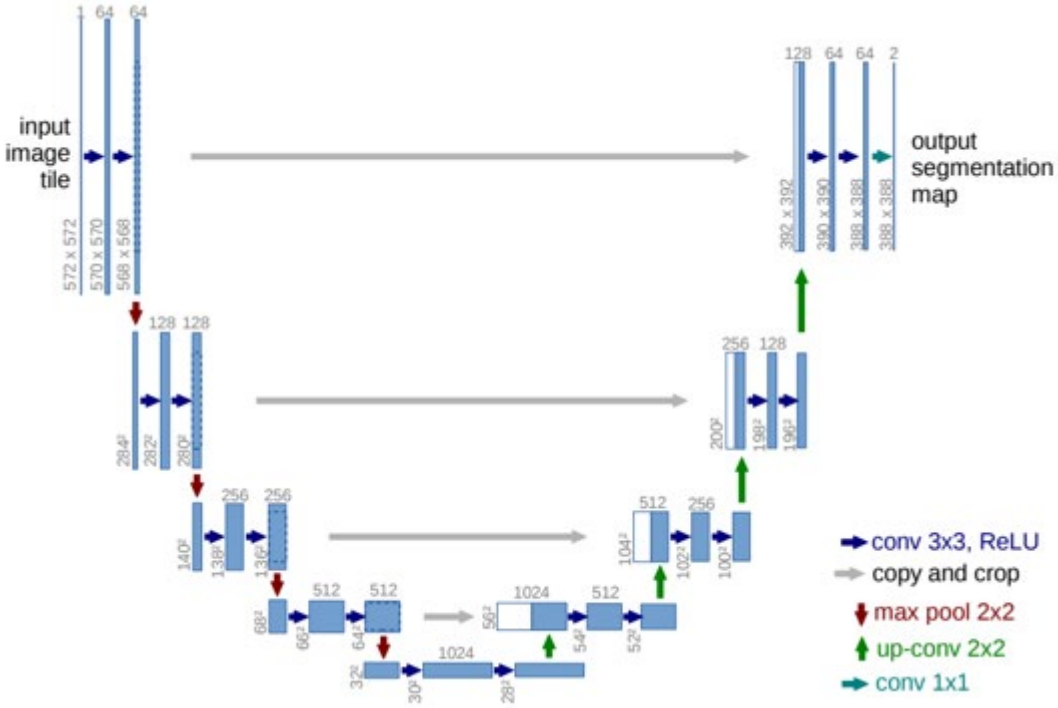
این دو مسیر توسط اتصالات میانبر به یکدیگر متصل می‌شوند و شکلی شبیه حرف U ایجاد می‌کنند که دلیل نام‌گذاری این معماری نیز همین موضوع است [۱۷]

در مسیر رمزگذار ، تصویر ورودی به‌تدریج از چندین لایه کانولوشنی و عملیات Max Pooling عبور می‌کند. در هر مرحله، ابعاد مکانی تصویر کاهش یافته و همچنین تعداد کانال‌های ویژگی افزایش پیدا می‌کند. این فرآیند باعث می‌شود شبکه بتواند ویژگی‌های سطح پایین مانند لبه‌ها و بافت‌ها و همچنین ویژگی‌های سطح بالاتر و مفهومی‌تر را استخراج کند.

در مسیر رمزگشا ، عملیات معکوس انجام می‌شود. در این مرحله، با استفاده از نمونه‌برداری افزایشی[۱] یا کانولوشن انتقال‌یافته [۲]، ابعاد تصویر دوباره افزایش یافته و اطلاعات استخراج‌شده در مرحله رمزگذار به‌تدریج برای بازسازی تصویر به کار گرفته می‌شوند. هدف این بخش، تولید خروجی‌ای است که تا حد امکان به تصویر مرجع نزدیک باشد ([۱۷]، [۱۳])

[۱] Up-sampling

[۲] Transposed Convolution



۲-۲-۲ معماری شبکه U-NET [۱۷]

۲- ۷- ۲ نقش اتصالات میانبر (Skip Connection)

یکی از تاثیر گذارترین ویژگی‌های U-Net، استفاده از اتصالات میانبر است. در بسیاری از شبکه‌های عمیق، اطلاعات مکانی اولیه تصویر در اثر چندین مرحله نمونه‌برداری کاهشی[۱] تا حدی از بین می‌روند. در نتیجه، بازسازی دقیق جزئیات تصویر دشوار می‌شود.

برای رفع این مشکل، U-Net ویژگی‌های استخراج‌شده در هر سطح از رمزگذار را مستقیماً به لایه متناظر در رمزگشا منتقل می‌کند. این انتقال باعث می‌شود اطلاعات مکانی و جزئیات تصویر در مراحل بازسازی حفظ شوند و شبکه بتواند مرزها، لبه‌ها و ساختارهای ظریف را با دقت بیشتری بازسازی کند [۱۷]

وجود اتصالات میانبر علاوه بر حفظ اطلاعات مکانی، باعث بهبود جریان گرادیان در هنگام آموزش شبکه نیز می‌شود و فرآیند یادگیری را میتواند پایدارتر می‌کند [۱۳]

۲- ۷- ۳ مزایا و محدودیت‌های U-Net

محبوبیت گسترده U-Net در پردازش تصاویر پزشکی ناشی از چند ویژگی مهم این معماری است. نخست آنکه ساختار رمزگذار–رمزگشا امکان استخراج ویژگی‌ها در مقیاس‌های مختلف را فراهم می‌کند. دوم آنکه اتصالات میانبر موجب حفظ اطلاعات مکانی و جلوگیری از حذف جزئیات مهم تصویر می‌شود. همچنین، این معماری حتی با

[۱] Down-sampling

مجموعه‌داده‌های نسبتاً کوچک نیز عملکرد قابل قبول دارد و ازین رو در بسیاری از کاربردهای پزشکی مورد استفاده قرار گرفته است ([۱۷]، [۱۳])

به علاوه ی این مزایا، U-Net نیز محدودیت‌هایی دارد. در ساختار پایه این معماری، تصویر عمدتاً در حوزه مکانی پردازش می‌شود و اطلاعات فرکانسی یا نمایش‌های چندمقیاسی به‌صورت صریح وارد فرآیند یادگیری نمی‌شوند. علاوه بر این، تمام ویژگی‌های استخراج‌شده اهمیت یکسانی دارند و مکانیزمی برای تمرکز شبکه بر نواحی یا کانال‌های مهم‌تر مشاهده نشد. این محدودیت‌ها باعث شد پژوهشگران به سمت توسعه نسخه‌های پیشرفته‌تر مبتنی بر تبدیل موجک، مکانیزم توجه و معماری‌های ترکیبی حرکت کنند. ([۱]، [۲]، [۳])

۲- ۷- ۴ کاربرد U-Net در بازسازی تصاویر پزشکی

اگرچه U-Net نخستین بار برای بخش‌بندی تصاویر زیستی معرفی شد، اما در سال‌های بعد کاربرد آن به حوزه‌های مختلف پردازش تصویر پزشکی گسترش یافت. از این معماری در حذف نویز، حذف آرتیفکت، بازسازی تصاویرMRI ، افزایش وضوح تصویر، بازسازی تصاویر CT و سایر مسائل تصویر به تصویر استفاده شده است. ([۱۳]، [۱۷])

در مسئله بازسازی تصاویرMRI ، هدف شبکه تولید تصویری است که بیشترین شباهت را به تصویر مرجع داشته باشد و در عین حال، ساختارهای آناتومیکی و جزئیات تشخیصی حفظ شوند. ساختار رمزگذار—رمزگشا و استفاده از اتصالات میانبر موجب شده است که U-Net به‌عنوان یکی از رایج‌ترین مدل‌های پایه در پژوهش‌های مرتبط با بازسازی تصاویر پزشکی شناخته شود. در همین راستا، بسیاری از معماری‌های جدید مانند مدل‌های مبتنی بر تبدیل موجک یا مکانیزم توجه، در واقع نسخه‌های توسعه‌یافته‌ای از ایده اولیه U-Net محسوب می‌شوند. ([۲]، [۳]، [۱۷])

۲- ۸ ترکیب تبدیل موجک و یادگیری عمیق

در کنار پیشرفت‌های قابل توجه در روش‌های مبتنی بر تبدیل موجک و شبکه‌های عصبی کانولوشنی، هر یک از این رویکردها به‌تنهایی با محدودیت‌هایی مواجه هستند. روش‌های موجکی اگرچه توانایی تحلیل چندمقیاسی و حفظ اطلاعات فرکانسی را دارند، اما عملکرد آن‌ها معمولاً به انتخاب نوع موجک، تعداد سطوح تجزیه و پارامترهای پردازش وابسته است. بر خلاف آن، شبکه‌های یادگیری عمیق می‌توانند رابطه میان تصویر تخریب‌شده و تصویر مرجع را از داده‌های آموزشی یاد بگیرند، اما بسیاری از آن‌ها تصویر را عمدتاً در حوزه پیکسل پردازش می‌کنند و اطلاعات فرکانسی را به‌صورت صریح در اختیار مدل قرار نمی‌دهند. ([۳]، [۱۳])

بر همین اساس، پژوهشگران به سمت ترکیب تبدیل موجک و یادگیری عمیق حرکت کرده‌اند. در این رویکرد، تبدیل موجک با تجزیه تصویر به زیرباندهای مختلف، اطلاعات ساختاری و فرکانسی را در مقیاس‌های متفاوت استخراج می‌کند و شبکه عصبی از این اطلاعات برای یادگیری فرآیند بازسازی استفاده می‌کند. به این ترتیب، شبکه علاوه بر یادگیری ویژگی‌های مکانی تصویر، به اطلاعات چندمقیاسی نیز دسترسی خواهد داشت که می‌تواند موجب حفظ بهتر لبه‌ها، بافت‌ها و جزئیات ظریف تصویر شود. ([۲]، [۳])

اهمیت این رویکرد در تصاویر پزشکی بیشتر آشکار می‌شود. در تصاویر MRI، آرتیفکت‌های حرکتی ممکن است علاوه بر ساختار کلی تصویر، مرز بافت‌ها و جزئیات کم‌کنتراست را نیز تحت تأثیر قرار دهند. بنابراین، استفاده از اطلاعات چندمقیاسی به شبکه کمک می‌کند تا میان ساختارهای واقعی تصویر و مؤلفه‌های ناشی از آرتیفکت تمایز دقیق‌تری ایجاد کند و کیفیت بازسازی را افزایش دهد. ([۲]، [۲۸])

یکی از نخستین نمونه‌های موفق این رویکرد، شبکه عصبی کانولوشنی موجک عمیق[۱] است، که برای بازسازی تصاویر MRI ارائه شد. [۲] در این معماری، اطلاعات موجکی به‌عنوان بخشی از فرآیند استخراج ویژگی، وارد شبکه می‌شوند و مدل تصویر را در مقیاس‌های مختلف تحلیل می‌کند. نتایج این پژوهش نشان داد که استفاده از نمایش موجکی در کنار شبکه‌های کانولوشنی، توانایی بهبود کیفیت بازسازی تصاویر MRI را نسبت به روش‌های مبتنی بر پردازش صرفاً پیکسلی را دارد.

در ادامه، شبکه عصبی کانولوشنی موجک چندسطحی[۲] با هدف ادغام تجزیه چندسطحی موجک و شبکه‌های کانولوشنی معرفی شد. [۳] در این معماری، تبدیل موجک تنها یک مرحله پیش‌پردازش نیست، بلکه به بخشی از ساختار شبکه تبدیل می‌شود و اطلاعات استخراج‌شده از زیرباندهای مختلف، مستقیماً در فرآیند یادگیری مورد استفاده قرار می‌گیرند. این ویژگی موجب کاهش وابستگی به آستانه‌گذاری دستی و افزایش توانایی شبکه در حفظ جزئیات ساختاری تصویر می‌شود.

علاوه بر این، مدل Deep Wavelet Prediction نیز نشان داد که یادگیری ضرایب موجکی در بازسازی جزئیات پرفرکانس و افزایش وضوح تصویر مؤثر می‌باشد. ([۴]، [۲۴]) در این رویکرد، شبکه به‌جای بازسازی مستقیم تصویر در حوزه پیکسل، روابط میان ضرایب موجکی را یاد می‌گیرد و از آن‌ها برای بازیابی اطلاعات از دست‌رفته استفاده می‌کند. این ایده بیانگر آن است که فضای موجک می‌تواند بستری مناسب برای یادگیری ویژگی‌های مؤثر در بازسازی تصویر باشد.

از سوی دیگر، چارچوب Deep Convolutional Framelets نیز مبنای نظری مناسبی برای ترکیب تبدیل موجک و شبکه‌های کانولوشنی ارائه می‌دهد. [۲۲] این چارچوب نشان می‌دهد که شبکه‌های عمیق را می‌توان از دیدگاه تجزیه و بازسازی سیگنال تحلیل کرد و استفاده از نمایش‌های چندمقیاسی فرآیند استخراج ویژگی را هدفمندتر می‌سازد. این دیدگاه، پشتوانه نظری بسیاری از معماری‌های جدید مبتنی بر موجک و یادگیری عمیق محسوب می‌شود.

به طور کلی، پژوهش‌های انجام‌شده نشان می‌دهند که ترکیب تبدیل موجک و یادگیری عمیق توانسته است بخشی از محدودیت‌های هر یک از این دو رویکرد را برطرف کند. استفاده هم‌زمان از تحلیل چندمقیاسی و یادگیری داده‌محور، امکان حفظ بهتر ساختارهای تصویر، بازیابی جزئیات ظریف و افزایش کیفیت بازسازی را فراهم می‌کند. به همین علت، بسیاری از معماری‌های جدید بازسازی تصاویر پزشکی بر پایه این ایده توسعه یافته‌اند. در بخش‌های بعدی، معماری‌های WACNNو MWACNN که مستقیماً بر این رویکرد استوار هستند، به‌صورت جداگانه معرفی و بررسی خواهند شد.

---

[۱] Deep Convolutional Wavelet Convolutional Neural Network (DC-WCNN)

[۲] Multi-Level Wavelet Convolutional Neural Network (MWCNN)

۵ ۲-۵ مقایسه مهمترین پژوهش‌های ترکیب تبدیل موجک و یادگیری عمیق

پژوهش	ایده اصلی	مزایا	محدودیت‌ها
DC-WCNN [۲]	استفاده از تبدیل موجک در کنار شبکه کانولوشنی برای بازسازی تصاویرMRI	بهره‌گیری از اطلاعات چندمقیاسی و بهبود کیفیت بازسازی تصاویرMRI	افزایش پیچیدگی معماری نسبت بهCNN های متداول
MWCNN [۳]	ادغام تجزیه چندسطحی موجک با شبکه کانولوشنی	استخراج هم‌زمان ویژگی‌های مکانی و فرکانسی و حفظ بهتر جزئیات تصویر	نیاز به طراحی مناسب معماری و هزینه محاسباتی بیشتر
Deep Wavelet Prediction [۲۴]	یادگیری و پیش‌بینی ضرایب موجکی برای بازسازی جزئیات تصویر	بهبود بازسازی مؤلفه‌های پرفرکانس و افزایش وضوح تصویر	تمرکز اصلی بر افزایش وضوح تصویر(Super-Resolution)و نه حذف آرتیفکت‌هایMRI
Deep Convolutional Framelets [۲۲]	ارائه چارچوب نظری برای تفسیر شبکه‌های عمیق بر پایه تجزیه و بازسازی سیگنال	ایجاد مبنای نظری برای ترکیب نمایش‌های چندمقیاسی و یادگیری عمیق	بیشتر یک چارچوب نظری است و نیازمند پیاده‌سازی متناسب با کاربردهای مختلف است

همان‌طور که در جدول ۲-۵ مشاهده می‌شود، پژوهش‌های انجام‌شده نشان می‌دهند که ترکیب تبدیل موجک و یادگیری عمیق توانایی برطرف کردن بخشی از محدودیت‌های هر یک از این دو رویکرد را دارد. در این معماری‌ها، تبدیل موجک اطلاعات تصویر را در مقیاس‌ها و زیرباندهای مختلف استخراج می‌کند و شبکه‌های کانولوشنی با بهره‌گیری از این اطلاعات، فرآیند بازسازی تصویر را به‌صورت داده‌محور انجام می‌دهند. در نتیجه، مدل قادر است علاوه بر کاهش نویز و آرتیفکت، ساختارهای مهم و جزئیات تصویر را نیز بهتر حفظ کند ([۲]، [۳])

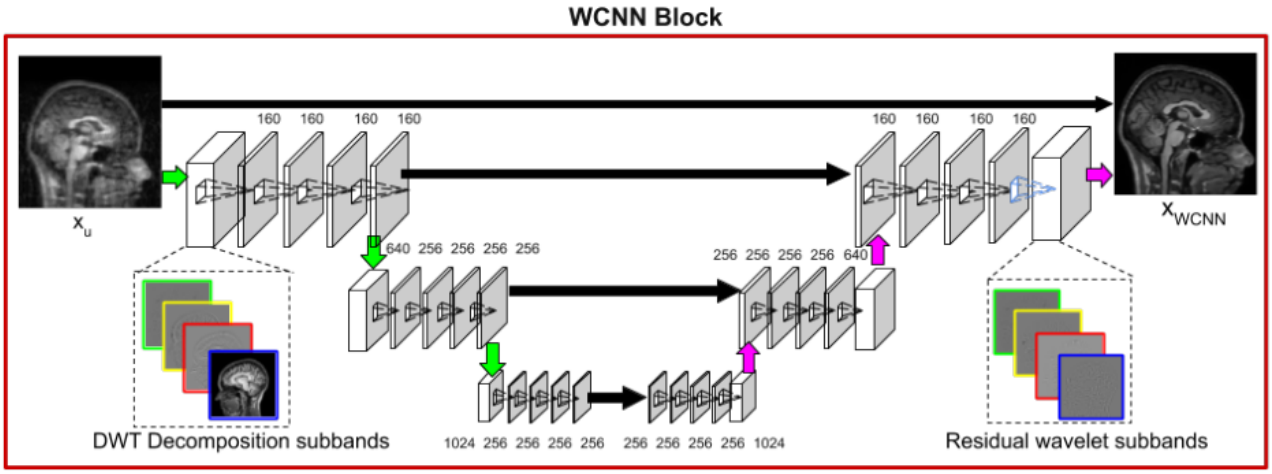
مطالعات مختلف نشان داده‌اند که استفاده از تجزیه چندسطحی موجک، امکان استخراج ویژگی‌های چندمقیاسی را فراهم می‌کند و به شبکه اجازه می‌دهد اطلاعات مربوط به ساختار کلی تصویر و جزئیات موضعی را به‌طور هم‌زمان مورد استفاده قرار دهد. [۳] این ویژگی در تصاویر پزشکی اهمیت ویژه‌ای دارد، زیرا آرتیفکت‌های حرکتی می‌توانند هم بر ساختارهای بزرگ و هم بر جزئیات کم‌کنتراست اثر بگذارند. بنابراین، تحلیل تصویر در مقیاس‌های مختلف به بهبود کیفیت بازسازی کمک می‌کند. ([۲]، [۲۸])

برخی پژوهش‌ها با استفاده از اتصال‌های باقی‌مانده[۱] و مکانیزم‌های توجه تلاش کرده‌اند عملکرد این معماری‌ها را بهبود دهند. اتصال‌های باقی‌مانده موجب تسهیل آموزش شبکه و حفظ بهتر اطلاعات تصویر می‌شوند، در حالی که مکانیزم‌های توجه با اختصاص وزن بیشتر به ویژگی‌ها یا کانال‌های مهم، توانایی شبکه را در استخراج اطلاعات مؤثر افزایش می‌دهند. ([۱]، [۴]، [۱۱]) همچنین، استفاده از توابع زیان ترکیبی که هم‌زمان خطای پیکسلی و شباهت ساختاری را در نظر می‌گیرند، کیفیت بازسازی تصاویر پزشکی را بهبود می‌دهد. [۱۵]

در کنار پیشرفت‌های قابل توجه، مقایسه مستقیم نتایج عددی گزارش‌شده در مطالعات مختلف همواره امکان‌پذیر نیست؛ زیرا این پژوهش‌ها معمولاً از مجموعه‌داده‌ها، انواع تخریب، شرایط آموزش و معیارهای ارزیابی متفاوتی استفاده کرده‌اند. ازاین‌رو، در این فصل تمرکز بر مقایسه مفهومی و معماری روش‌های موجود است و مقایسه کمی عملکرد مدل‌ها بر اساس داده‌ها و شرایط یکسان، در فصل نتایج ارائه خواهد شد.

در مجموع، ترکیب تبدیل موجک و یادگیری عمیق را می‌توان یکی از مهم‌ترین مسیرهای توسعه روش‌های نوین بازسازی تصویر دانست. تبدیل موجک با فراهم کردن نمایش چندمقیاسی، امکان تحلیل هم‌زمان ساختار کلی و جزئیات تصویر را فراهم می‌کند، در حالی که شبکه‌های یادگیری عمیق با بهره‌گیری از داده‌های آموزشی، قادر به یادگیری روابط پیچیده میان تصویر تخریب‌شده و تصویر مرجع هستند. ترکیب این دو رویکرد موجب شده است که بسیاری از محدودیت‌های روش‌های سنتی و مدل‌های مبتنی بر پردازش صرفاً پیکسلی کاهش یابد و کیفیت بازسازی تصاویر، به‌ویژه در کاربردهای پزشکی، بهبود پیدا کند. ([۲]، [۳]، [۱۳]) ازاین‌رو، معماری‌های مبتنی بر موجک و شبکه‌های کانولوشنی در مطالعات اخیر به یکی از مورد توجه ترین زمینه‌های پژوهشی در حوزه بازسازی تصاویر پزشکی تبدیل شده‌اند و مبنای توسعه بسیاری از مدل‌های پیشرفته امروزی را تشکیل می‌دهند.

[۱] Residual Connections



۳-۲-۳ معماری شبکه [۲WCNN]

۲- ۹ معماری WACNN

رویکردهای مبتنی بر ترکیب تبدیل موجک و شبکه‌های کانولوشنی با هدف استفاده هم‌زمان از اطلاعات مکانی و فرکانسی تصویر توسعه یافته‌اند. پژوهش‌هایی مانند DC-WCNN [۲] و MWCNN [۳] نشان داده‌اند که تبدیل موجک می‌تواند به‌عنوان بخشی از ساختار شبکه مورد استفاده قرار گیرد. همچنین، استفاده از مکانیزم‌های توجه در کنار نمایش موجکی در برخی معماری‌ها بررسی شده است [۴]. در پژوهش حاضر، با الهام از این رویکردها، یک معماری WACNN شامل تجزیه موجکی تک‌سطحی، بلوک‌های باقیمانده و توجه کانالی پیاده‌سازی شده است. جزئیات پیاده‌سازی در فصل سوم ارائه می‌شود.

ایده اصلی WACNN بر این فرض استوار است که همه اطلاعات تصویر از اهمیت یکسانی برخوردار نیستند. در تبدیل موجک، تصویر به زیرباندهای مختلف شامل LL، LH، HL و HH تقسیم می‌شود. زیرباند LL اطلاعات کم‌فرکانس و ساختار کلی تصویر را در خود نگه می‌دارد، در حالی که زیرباندهای LH، HL و HH شامل اطلاعات مربوط به لبه‌ها، تغییرات موضعی و جزئیات پرفرکانس هستند. این تفکیک به شبکه اجازه می‌دهد ویژگی‌های هر زیرباند را به‌صورت مستقل تحلیل کرده و نقش هر بخش را در بازسازی تصویر بهتر یاد بگیرد [۳]

در معماری WACNN، پس از انجام تجزیه موجکی، ضرایب حاصل به‌عنوان ورودی شبکه کانولوشنی مورد استفاده قرار می‌گیرند. شبکه با یادگیری روابط میان این ضرایب، تلاش می‌کند مؤلفه‌های ناشی از نویز یا آرتیفکت را از اطلاعات واقعی تصویر تفکیک کند. در پایان نیز با استفاده از تبدیل موجک معکوس[۱]، تصویر بازسازی‌شده تولید می‌شود. این فرآیند باعث می‌شود اطلاعات فرکانسی در تمام مراحل بازسازی حفظ شوند و تنها به‌عنوان یک مرحله پیش‌پردازش مورد استفاده قرار نگیرند ([۲]، [۴])

[۱] Inverse Wavelet Transform

یکی از مهم‌ترین مزایای WACNN، بهره‌گیری هم‌زمان از اطلاعات مکانی و فرکانسی تصویر است. در حالی که شبکه‌های کانولوشنی متداول عمدتاً از همسایگی پیکسل‌ها برای استخراج ویژگی استفاده می‌کنند، WACNN با استفاده از نمایش موجکی، امکان تحلیل تصویر در چند مقیاس مختلف را فراهم می‌سازد. این ویژگی به‌ویژه در تصاویر پزشکی اهمیت دارد، زیرا بسیاری از ساختارهای آناتومیکی دارای تغییرات تدریجی شدت روشنایی هستند و در کنار آن‌ها جزئیات بسیار ظریفی نیز وجود دارد که برای تشخیص پزشکی اهمیت بالایی دارند ([۲]، [۳]، [۴])

علاوه بر این مزایا، WACNN نیز محدودیت‌هایی دارد. عملکرد این معماری تا حدی به نوع موجک، تعداد سطوح تجزیه و نحوه ادغام اطلاعات موجکی با شبکه کانولوشنی وابسته است. این موضوع باعث می‌شود برخی ویژگی‌های مهم تصویر به اندازه کافی مورد توجه شبکه قرار نگیرند. ([۱]، [۳])

به همین دلیل، پژوهش‌های بعدی تلاش کردند با توسعه معماری WACNN و استفاده از تجزیه چندسطحی موجک و مکانیزم‌های توجه، محدودیت‌های آن را برطرف کنند. نتیجه این روند، معرفی معماری MWACNN بود که علاوه بر استفاده از نمایش چندمقیاسی موجک، از سازوکارهای پیشرفته‌تری برای استخراج و تقویت ویژگی‌های مهم تصویر بهره می‌گیرد. این معماری در بخش بعدی به‌طور مفصل بررسی خواهد شد.

۶-۲۶ معماری WACNN

ویژگی	توضیح
نوع معماری	Wavelet-Attached Convolutional Neural Network (WACNN)
ایده اصلی	ترکیب تبدیل موجک و شبکه کانولوشنی برای استخراج هم‌زمان ویژگی‌های مکانی و فرکانسی
مهم‌ترین مزیت	تحلیل چندمقیاسی و حفظ بهتر جزئیات تصویر
کاربردها	حذف نویز، بازسازی تصویر و پردازش تصاویر پزشکی
محدودیت‌ها	وابستگی به نوع موجک و نبود مکانیزم توجه در نسخه‌های اولیه
مرجع	([۲]، [۳])

۲- ۱۰ معماری MWACNN

یکی از معماری‌های شناخته‌شده در ترکیب تبدیل موجک چندسطحی و شبکه‌های کانولوشنی، MWCNN است که توسط Liu و همکاران معرفی شد [۳]. در این رویکرد، تجزیه موجکی در چند سطح وارد ساختار شبکه می‌شود و امکان پردازش ویژگی‌های تصویر در مقیاس‌های مختلف فراهم می‌گردد. ساختار چندسطحی مورد استفاده در MWACNN این پژوهش با الهام از این ایده پیاده‌سازی شده است، با این تفاوت که پیاده‌سازی حاضر ساختار خاص خود را دارد و در یک نسخه نیز توجه کانالی به آن افزوده شده است.

تفاوت اصلی MWACNN با WACNN در آن است که در این معماری، تبدیل موجک تنها یک بار روی تصویر اعمال نمی‌شود، بلکه تصویر در چندین سطح متوالی تجزیه می‌شود. در هر سطح، زیرباند کم‌فرکانس دوباره وارد مرحله بعدی تجزیه شده و مجموعه‌ای جدید از ضرایب موجکی تولید می‌شود. این فرآیند باعث می‌شود اطلاعات مربوط به ساختارهای بزرگ، جزئیات متوسط و ویژگی‌های بسیار ریز تصویر به‌صورت هم‌زمان در اختیار شبکه قرار گیرد. ([۳]، [۲۲])

پس از انجام تجزیه چندسطحی، ضرایب موجکی وارد بلوک‌های کانولوشنی می‌شوند. شبکه با یادگیری روابط میان این ضرایب، ویژگی‌های مرتبط با تصویر سالم را از مؤلفه‌های ناشی از نویز یا آرتیفکت تفکیک می‌کند. در پایان، با استفاده از تبدیل موجک معکوس، تصویر بازسازی‌شده تولید می‌شود. برخلاف روش‌های سنتی مبتنی بر موجک که معمولاً از آستانه‌گذاری یا قواعد از پیش تعیین‌شده استفاده می‌کنند، در MWACNN فرآیند بازسازی به‌صورت داده‌محور انجام می‌شود و پارامترهای شبکه طی فرآیند آموزش یاد گرفته می‌شوند. [۳]

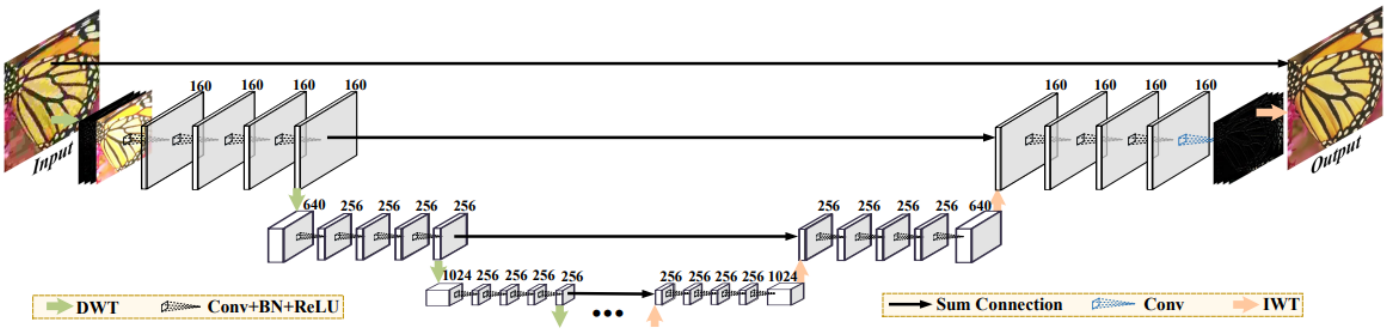
یکی از پرکاربرد ترین مزایایMWACNN ، توانایی آن در استخراج ویژگی‌های چندمقیاسی است. در تصاویر پزشکی، آرتیفکت‌های حرکتی می‌توانند هم بر ساختار کلی تصویر و هم بر لبه‌ها و بافت‌های ظریف اثر بگذارند. تجزیه چندسطحی این امکان را فراهم می‌کند که شبکه اطلاعات مربوط به هر یک از این سطوح را به‌طور مستقل پردازش کرده و در مرحله بازسازی، آن‌ها را به‌صورت مؤثرتری با یکدیگر ترکیب کند. این ویژگی موجب حفظ بهتر ساختارهای آناتومیکی و کاهش تخریب جزئیات مهم تصویر می‌شود. ([۲]، [۳])

از دیدگاه محاسباتی، استفاده از چندین سطح تجزیه موجکی موجب افزایش تعداد ویژگی‌های قابل پردازش و در نتیجه افزایش پیچیدگی شبکه می‌شود. همچنین، انتخاب تعداد مناسب سطوح تجزیه و نحوه ترکیب ضرایب موجکی می‌تواند بر عملکرد نهایی مدل تأثیرگذار باشد. اگر تعداد سطوح بیش از حد کم باشد، بخشی از اطلاعات چندمقیاسی از دست می‌رود و اگر بیش از حد زیاد باشد، هزینه محاسباتی افزایش یافته و احتمال یادگیری ویژگی‌های غیرضروری نیز بیشتر می‌شود. ([۳]، [۲۲])

به علاوه ی این محدودیت‌ها، نتایج پژوهش‌های انجام‌شده نشان می‌دهد که معماری‌های مبتنی بر تجزیه چندسطحی موجک در بسیاری از مسائل بازسازی تصویر، نسبت به معماری‌های مبتنی بر یک سطح موجک عملکرد بهتری در حفظ جزئیات و کاهش آرتیفکت‌ها دارند. ([۲]، [۳]) به همین علت، این معماری به‌عنوان یکی از پایه‌های توسعه مدل‌های جدید در بازسازی تصاویر پزشکی شناخته می‌شود.

پژوهشگران تلاش کرده‌اند قابلیت‌های MWACNN را با استفاده از مکانیزم‌های توجه نیز تقویت کنند. در این رویکرد، علاوه بر استخراج ویژگی‌های چندمقیاسی، شبکه باعث یادگیری اهمیت نسبی هر کانال یا هر ویژگی نیز

می‌شود و تمرکز بیشتری بر اطلاعات مؤثر در بازسازی تصویر داراست. این توسعه موجب شکل‌گیری نسخه‌های مبتنی بر توجه از MWACNN شده است که در بخش بعدی مورد بررسی قرار می‌گیرند. ([۱]، [۱۱])



۴-۲۴ معماری MWCNN [۲]

۷-۲۷ مقایسه WACNN و MWACNN

ویژگی	WACNN	MWACNN
تعداد سطوح تجزیه موجکی	یک سطح	دو سطح
تجزیه مجددLL	ندارد	دارد
ترکیب اطلاعات دو سطح	ندارد	دارد
بازسازی باIWT	دارد	دارد
مناسب برای تصاویر پزشکی	مناسب	مناسب‌تر برای تصاویر دارای ساختارهای پیچیده
قابلیت توسعه باAttention	محدود	بالا

۲- ۱۱ مکانیزم توجه(Attention Mechanism)

با افزایش عمق شبکه‌های عصبی، یکی از چالش‌های مهم آن بود که همه ویژگی‌های استخراج‌شده اهمیت یکسانی ندارند. در بسیاری از مسائل پردازش تصویر، تنها بخشی از ویژگی‌ها یا نواحی تصویر اطلاعات اصلی را در اختیار دارند، در حالی که سایر ویژگی‌ها ممکن است شامل اطلاعات تکراری، نویز یا جزئیات کم‌اهمیت باشند. از این رو، پژوهشگران مکانیزم توجه را معرفی کردند تا شبکه بتواند به‌صورت خودکار اهمیت نسبی ویژگی‌های مختلف را یاد بگیرد و منابع محاسباتی خود را بر اطلاعات مؤثرتر متمرکز کند.[۱]،

ایده اصلی مکانیزم توجه آن است که همه ویژگی‌های استخراج‌شده توسط شبکه ارزش یکسانی ندارند. در این روش، شبکه با یادگیری ضرایب وزن‌دهی، مشخص می‌کند که کدام ویژگی‌ها یا کانال‌ها نقش بیشتری در بازسازی تصویر دارند. در نتیجه، ویژگی‌های مهم تقویت شده و تأثیر ویژگی‌های کم‌اهمیت یا نویزی کاهش می‌یابد. این فرآیند باعث می‌شود شبکه بتواند نمایش دقیق‌تری از تصویر تولید کند و کیفیت بازسازی بهبود یابد. [۱]

در پردازش تصویر، مکانیزم‌های توجه معمولاً در سه گروه اصلی شامل توجه کانالی[۱]، توجه مکانی[۲] و توجه ترکیبی طبقه‌بندی می‌شوند. در توجه کانالی، اهمیت هر کانال ویژگی به‌طور جداگانه تعیین می‌شود، در حالی که در توجه مکانی، شبکه نواحی مهم تصویر را شناسایی کرده و وزن بیشتری به آن‌ها اختصاص می‌دهد. برخی معماری‌های جدید نیز از ترکیب این دو رویکرد برای افزایش دقت استخراج ویژگی استفاده می‌کنند.[۱]

در معماری‌های مبتنی بر تبدیل موجک، استفاده از توجه کانالی اهمیت ویژه‌ای دارد. پس از انجام تجزیه موجکی، زیرباندهای مختلف تصویر اطلاعات متفاوتی را نمایش می‌دهند؛ برای مثال، زیرباند کم‌فرکانس بیانگر ساختار کلی تصویر است، در حالی که زیرباندهای پرفرکانس حاوی اطلاعات مربوط به لبه‌ها، بافت‌ها و تغییرات موضعی هستند. اهمیت این زیرباندها در همه تصاویر یکسان نیست و ممکن است با توجه به نوع تصویر یا شدت آرتیفکت تغییر کند. مکانیزم توجه کانالی به شبکه اجازه می‌دهد تا اهمیت هر یک از این مؤلفه‌ها را به‌صورت خودکار یاد بگیرد و ویژگی‌های مؤثرتر را با وزن بیشتری در فرآیند بازسازی مورد استفاده قرار دهد. ([۱]، [۳])

در تصاویرMRI ، اهمیت استفاده از مکانیزم توجه بیشتر نمایان می‌شود. در این تصاویر، وجود آرتیفکت‌های حرکتی یا نویز می‌تواند موجب کاهش وضوح مرز بافت‌ها و از بین رفتن جزئیات ظریف شود. اگر شبکه تمام ویژگی‌ها را با اهمیت یکسان پردازش کند، ممکن است بخشی از اطلاعات تشخیصی نیز تضعیف شود. استفاده از مکانیزم توجه موجب می‌شود شبکه تمرکز بیشتری بر ساختارهای آناتومیکی و ویژگی‌های مهم تصویر داشته باشد و در نتیجه کیفیت بازسازی و حفظ جزئیات افزایش یابد. ([۱]، [۲])

یکی دیگر از مزایای مکانیزم توجه، افزایش کارایی استخراج ویژگی بدون نیاز به افزایش قابل توجه تعداد پارامترهای شبکه است. در بسیاری از معماری‌های جدید، تنها با اضافه شدن یک بلوک توجه[۳]، عملکرد شبکه در معیارهایی مانند نسبت سیگنال به نویز اوج و شاخص شباهت ساختاری بهبود یافته است، بدون آنکه افزایش چشمگیری در پیچیدگی محاسباتی ایجاد شود.[۱] به همین علت، مکانیزم توجه به یکی از اجزای متداول معماری‌های پیشرفته بازسازی تصویر تبدیل شده است.

علاوه بر این مزایا، استفاده از مکانیزم توجه نیز محدودیت‌هایی دارد. عملکرد این روش به طراحی صحیح بلوک توجه، نحوه قرارگیری آن در معماری شبکه و کیفیت داده‌های آموزشی وابسته است. همچنین، در برخی کاربردها، استفاده از بلوک‌های توجه متعدد موجب افزایش زمان آموزش و هزینه محاسباتی می‌شود. بنابراین، انتخاب نوع و محل استفاده از توجه باید متناسب با مسئله و معماری شبکه انجام گیرد.[۱]

---

[۱] Channel Attention

[۲] Spatial Attention

[۳] Residual Blocks with Channel Attention

در مجموع، مکانیزم توجه را می‌توان یکی از مهم‌ترین پیشرفت‌های معماری‌های یادگیری عمیق در سال‌های اخیر دانست. این مکانیزم با فراهم کردن امکان تمرکز انتخابی بر ویژگی‌های مهم، توانسته است کیفیت بازسازی تصاویر را در بسیاری از کاربردهای پردازش تصویر و به‌ویژه تصاویر پزشکی بهبود دهد. به همین دلیل، امروزه استفاده از توجه در کنار شبکه‌های کانولوشنی و تبدیل موجک به یکی از رویکردهای رایج در توسعه معماری‌های پیشرفته بازسازی تصویر تبدیل شده است.

۲-۸-۸ مقایسه انواع مکانیزم‌های توجه در پردازش تصویر

نوع مکانیزم توجه	ایده اصلی	مزایا	محدودیت‌ها
Channel Attention	وزن‌دهی به کانال‌های ویژگی	تقویت ویژگی‌های مهم و بهبود استخراج اطلاعات	وابستگی به طراحی مناسب بلوک توجه
Spatial Attention	تمرکز بر نواحی مهم تصویر	حفظ بهتر ساختارها و نواحی مؤثر تصویر	افزایش پیچیدگی محاسباتی
Channel + Spatial Attention	ترکیب توجه کانالی و مکانی	استخراج دقیق‌تر ویژگی‌ها و بهبود کیفیت بازسازی	افزایش تعداد پارامترها و زمان آموزش

۲- ۱۲ حذف آرتیفکت‌های حرکتی در تصاویر MRI و معرفی مجموعه‌داده متن باز KMAR-۵۰K

تصویربرداری تشدید مغناطیسی یکی از مهم‌ترین روش‌های تصویربرداری پزشکی برای بررسی بافت‌های نرم، مفاصل، مغز و اندام‌های داخلی بدن است. این روش به دلیل استفاده نکردن از پرتوهای یونیزان و توانایی بالا در نمایش جزئیات آناتومیکی، کاربرد گسترده‌ای در تشخیص بیماری‌ها و برنامه‌ریزی درمان دارد. [۲۸] با وجود این مزایا، کیفیت تصاویر MRI به عوامل مختلفی وابسته است و وجود آرتیفکت‌های تصویری دقت تشخیص را کاهش می‌دهد. در میان انواع مختلف آرتیفکت‌ها، آرتیفکت حرکتی یکی از رایج‌ترین و مهم‌ترین عوامل کاهش کیفیت تصاویر MRI محسوب می‌شود. [۲۸]

آرتیفکت‌های حرکتی معمولاً در اثر حرکت بیمار در زمان تصویربرداری ایجاد می‌شوند. حتی حرکات کوچک و غیرارادی مانند تنفس، ضربان قلب یا جابه‌جایی اندام‌ها نیز می‌توانند موجب ایجاد تاری، کشیدگی، تکرار ساختارها یا اعوجاج در تصویر شوند. این تغییرات ممکن است سبب کاهش وضوح مرز بافت‌ها، از بین رفتن جزئیات ظریف و دشوار شدن تشخیص ضایعات شوند. در برخی موارد نیز شدت آرتیفکت به اندازه‌ای است که تصویربرداری باید تکرار شود که این موضوع علاوه بر افزایش هزینه، زمان معاینه بیمار را نیز افزایش می‌دهد. [۲۸]

در میان کاربردهای مختلف MRI، تصویربرداری از مفصل زانو اهمیت ویژه‌ای دارد. مفصل زانو از ساختارهای متعددی مانند غضروف، رباط‌ها، منیسک و استخوان تشکیل شده است که بسیاری از آن‌ها دارای جزئیات ظریف و مرزهای کم‌کنتراست هستند. وجود آرتیفکت‌های حرکتی در این تصاویر توانایی مواجه با مشکل مشاهده دقیق این ساختارها را دارد و احتمال بروز خطا در تشخیص آسیب‌های مفصلی را افزایش دهد. ازاین‌رو، توسعه روش‌های مؤثر برای حذف آرتیفکت‌های حرکتی در تصاویر MRI زانو به یکی از موضوعات مهم پژوهشی در عصر حاضر تبدیل شده است. [۲۸]

یکی از چالش‌های مهم در توسعه مدل‌های یادگیری عمیق برای حذف آرتیفکت، کمبود مجموعه‌داده‌های استاندارد و جفت‌شده است. برای آموزش مدل‌های نظارت‌شده، وجود تصاویر دارای آرتیفکت و تصاویر مرجع متناظر ضروری است تا شبکه بتواند رابطه میان تصویر تخریب‌شده و تصویر مطلوب را یاد بگیرد. با این حال، تهیه چنین داده‌هایی در شرایط واقعی دشوار، زمان‌بر و پرهزینه است؛ زیرا ثبت هم‌زمان تصاویر سالم و تصاویر دارای آرتیفکت از یک بیمار در شرایط کاملاً یکسان معمولاً امکان‌پذیر نیست. [۲۸]

به‌منظور رفع این محدودیت، Xiو همکاران مجموعه‌داده KMAR-۵۰K را معرفی کردند که یکی از بزرگ‌ترین مجموعه‌داده‌های متن‌باز در زمینه حذف آرتیفکت‌های حرکتی در تصاویر MRI زانو محسوب می‌شود. [۲۸] این مجموعه‌داده شامل تصاویر جفت‌شده MRI است که در آن، برای هر تصویر مرجع، نسخه‌ای دارای آرتیفکت حرکتی نیز در دسترس قرار گرفته است. وجود این تصاویر جفت‌شده، امکان آموزش و ارزیابی منصفانه مدل‌های یادگیری عمیق را فراهم می‌کند و به همین دلیل، KMAR-۵۰K به یکی از مجموعه‌داده‌های مرجع در پژوهش‌های اخیر تبدیل شده است. [۲۸]

یکی از ویژگی‌های مهم KMAR-۵۰K، تنوع مناسب داده‌ها از نظر شدت آرتیفکت، نوع برش‌های تصویربرداری و تفاوت میان نمونه‌های مختلف است. این ویژگی موجب می‌شود مدل‌های آموزش‌دیده تنها برای یک نوع خاص از تخریب بهینه نشوند و توانایی تعمیم به شرایط متنوع‌تر را نیز داشته باشند. علاوه بر این، متن‌باز بودن این مجموعه‌داده امکان مقایسه مستقیم عملکرد مدل‌های مختلف را در شرایط یکسان فراهم کرده و نقش مهمی در استانداردسازی پژوهش‌های این حوزه ایفا کرده است. [۲۸]

در مقاله معرفی این مجموعه‌داده، عملکرد چندین معماری یادگیری عمیق از جمله U-Net، EDSR و DDPM بر روی KMAR-۵۰K مورد بررسی قرار گرفته است. [۲۸] نتایج این پژوهش نشان داد که هر یک از این معماری‌ها نقاط قوت و محدودیت‌های خاص خود را دارند. برای مثال، U-Net به خاطر ساختار رمزگذار–رمزگشا و استفاده از اتصالات میانبر عملکرد مناسبی در بازسازی تصاویر ارائه می‌دهد، در حالی که مدل‌های مبتنی بر انتشار[۱] قادرند جزئیات بیشتری را بازیابی کنند، اما معمولاً به هزینه محاسباتی بالاتر و زمان استنتاج بیشتری نیاز دارند. ([۱۷]، [۲۸])

مطالعات انجام‌شده بر روی KMAR-۵۰K نشان می‌دهد که حذف آرتیفکت‌های حرکتی تنها به کاهش نویز یا افزایش وضوح تصویر محدود نمی‌شود، بلکه هدف اصلی، بازسازی دقیق ساختارهای آناتومیکی و حفظ اطلاعات تشخیصی تصویر است. ازاین‌رو، استفاده هم‌زمان از معیارهای کمی مانند نسبت سیگنال به نویز اوج، شاخص شباهت

[۱] Diffusion Models

ساختاری، میانگین قدر مطلق خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده در کنار ارزیابی بصری، به یکی از رویکردهای رایج در بررسی عملکرد مدل‌های بازسازی تصویر تبدیل شده است [۲۸]

در مجموع، معرفی مجموعه‌داده KMAR-۵۰K گام مهمی در پیشرفت پژوهش‌های مرتبط با حذف آرتیفکت‌های حرکتی در تصاویر MRI به شمار می‌رود. وجود داده‌های جفت‌شده، تنوع مناسب نمونه‌ها و امکان مقایسه منصفانه مدل‌های مختلف، این مجموعه‌داده را به یکی از منابع استاندارد در ارزیابی روش‌های بازسازی تصویر تبدیل کرده است. به همین دلیل، ساختار جفت‌شده و متن‌باز این مجموعه‌داده امکان استفاده از آن در توسعه و ارزیابی روش‌های حذف آرتیفکت حرکتی را فراهم می‌کند. [۲۸]

۹ ۲-۹ ویژگی‌های اصلی مجموعه‌داده - K۵۰KMAR

ویژگی	توضیحات
نام مجموعه‌داده	KMAR-۵۰K
ارائه‌دهندگان	Xiو همکاران [۲۸]
نوع تصاویر	تصاویر MRI زانو
نوع داده	تصاویر جفت‌شده (دارای آرتیفکت و مرجع)
کاربرد اصلی	آموزش و ارزیابی مدل‌های حذف آرتیفکت حرکتی
مزیت اصلی	امکان مقایسه منصفانه مدل‌ها در شرایط یکسان

۲- ۱۳ پیشینه پژوهش

در سال‌های اخیر، پژوهش‌های متعددی در زمینه حذف نویز، حذف آرتیفکت و بازسازی تصاویر پزشکی با استفاده از روش‌های پردازش تصویر و یادگیری عمیق انجام شده است. در این مطالعات، روش‌های کلاسیک مبتنی بر فیلترگذاری و مدل‌های آماری مورد استفاده قرار گرفتند و سپس با پیشرفت یادگیری عمیق، معماری‌های مبتنی بر شبکه‌های عصبی کانولوشنی، تبدیل موجک و مکانیزم‌های توجه توسعه یافتند. مرور این پژوهش‌ها نشان می‌دهد که هر نسل از روش‌ها بخشی از محدودیت‌های نسل قبل را برطرف کرده است، اما همچنان چالش‌هایی مانند حفظ ساختارهای ظریف، کاهش آرتیفکت‌های پیچیده و افزایش کارایی محاسباتی باقی مانده است.

در میان روش‌های سنتی، Li و Suen با استفاده از روش Non-Local Means و نظریه خاکستری، روشی برای کاهش نویز تصویر ارائه کردند که از شباهت میان نواحی مختلف تصویر برای بهبود کیفیت استفاده می‌کرد. [۵]

همچنین Liu با ترکیب تبدیل موجک، آستانه‌گذاری و الگوریتم ژنتیک، روشی برای کاهش نویز معرفی کرد که نسبت به آستانه‌گذاری ساده عملکرد بهتری داشت، اما همچنان به تنظیم پارامترهای متعدد وابسته بود. [۷] در ادامه، Rabbani و Cho و Bui نیز با استفاده از مدل‌های آماری در حوزه‌های چندمقیاسی، تلاش کردند ساختار تصویر را بهتر حفظ کنند، هرچند عملکرد این روش‌ها به نوع مدل آماری و پارامترهای انتخابی وابسته بود. ([۱۰]، [۲۰])

با گسترش یادگیری عمیق، پژوهشگران استفاده از شبکه‌های عصبی کانولوشنی را برای بازسازی تصویر آغاز کردند. Tian در یک مطالعه مروری، روند توسعه روش‌های مبتنی بر یادگیری عمیق در کاهش نویز تصویر را بررسی کرد و نشان داد که شبکه‌های کانولوشنی نسبت به روش‌های کلاسیک، توانایی بیشتری در یادگیری روابط میان تصویر نویزی و تصویر مرجع دارند. [۱۳] همچنین FFDNet به‌عنوان یکی از شبکه‌های مؤثر در کاهش نویز، امکان پردازش تصاویر با سطوح مختلف نویز را فراهم کرد و نسبت به بسیاری از مدل‌های پیشین انعطاف‌پذیری بیشتری نشان داد. [۱۴]

در ادامه، پژوهشگران به سمت استفاده از تبدیل موجک در کنار شبکه‌های عصبی حرکت کردند. DC-WCNN از نخستین معماری‌هایی بود که تجزیه موجکی را با شبکه کانولوشنی برای بازسازی تصاویر MRI ترکیب کرد و نشان داد که استفاده از اطلاعات چندمقیاسی کیفیت بازسازی را بهبود می‌دهد. [۲] پس از آن، معماری MWCNN با استفاده از تجزیه چندسطحی موجک، توانست اطلاعات مکانی و فرکانسی را به‌صورت هم‌زمان در فرآیند یادگیری وارد کند و عملکرد مناسبی در بازسازی تصویر ارائه دهد. [۳] همچنین، پژوهش Deep Wavelet Prediction نشان داد که پیش‌بینی ضرایب موجکی به بازیابی جزئیات پرفرکانس و افزایش وضوح تصویر کمک می‌کند. ([۴]، [۲۴])

هم‌زمان با توسعه معماری‌های موجکی، استفاده از مکانیزم‌های توجه نیز گسترش یافت. Tian با معرفی شبکه عصبی کانولوشنی هدایت‌شده با مکانیزم توجه[۱] نشان داد که وزن‌دهی به ویژگی‌های مهم می‌تواند موجب بهبود کیفیت بازسازی و حفظ بهتر ساختارهای تصویر شود. [۱] پس از آن، مکانیزم‌های توجه کانالی و مکانی در بسیاری از معماری‌های بازسازی تصویر به کار گرفته شدند تا شبکه بتواند اطلاعات مهم را از میان ویژگی‌های استخراج‌شده انتخاب و تقویت کند. [۱۱]

در حوزه تصاویرMRI ، معرفی مجموعه‌داده KMAR-۵۰K توسط Xi و همکاران نقطه عطف مهمی در ارزیابی روش‌های حذف آرتیفکت حرکتی محسوب می‌شود .این مجموعه‌داده با ارائه تصاویر جفت‌شده دارای آرتیفکت و تصاویر مرجع، امکان آموزش و مقایسه منصفانه مدل‌های مختلف را فراهم کرده است. در این پژوهش، عملکرد معماری‌هایی مانند U-Net، EDSRو DDPM بر روی مجموعه‌داده KMAR-۵۰K بررسی شد و نتایج نشان داد که هر یک از این معماری‌ها نقاط قوت و محدودیت‌های خاص خود را دارند. [۲۸]

بررسی مطالعات انجام‌شده نشان می‌دهد که اگرچه روش‌های کلاسیک پایه‌های مهمی برای پردازش تصویر ایجاد کردند و شبکه‌های کانولوشنی نیز توانایی یادگیری روابط پیچیده را فراهم ساختند، اما استفاده از اطلاعات چندمقیاسی و مکانیزم‌های توجه نقش مهمی در بهبود کیفیت بازسازی تصاویر پزشکی دارد. به همین دلیل، روند پژوهش‌ها به سمت توسعه معماری‌هایی حرکت کرده است که تبدیل موجک، شبکه‌های کانولوشنی و مکانیزم‌های

[۱] Attention-Guided Convolutional Neural Network (Attention-Guided CNN)

توجه را به‌صورت هم‌زمان به کار می‌گیرند تا علاوه بر کاهش آرتیفکت، ساختارهای آناتومیکی و جزئیات تصویر نیز بهتر حفظ شوند.

پژوهش	روش	مهم‌ترین دستاورد	محدودیت
Suen وLi [۵]	Non-Local Means	استفاده از شباهت نواحی تصویر برای کاهش نویز	کاهش کارایی در تخریب‌های شدید
Liu [۷]	Wavelet + Genetic Algorithm	بهبود آستانه‌گذاری موجکی	وابستگی به تنظیم پارامترها
Rabbani [۱۰]	Steerable Pyramid	تحلیل چندمقیاسی تصویر	وابستگی به مدل آماری
Tian [۱۳]	مرور یادگیری عمیق	تبیین روند توسعه روش‌های داده‌محور	مقاله مروری
FFDNet [۱۴]	CNN	کاهش نویز با سطوح مختلف نویز	پردازش عمدتاً در حوزه پیکسل
U-Net [۱۷]	Encoder–Decoder	حفظ اطلاعات مکانی با Skip Connection	نبود تحلیل فرکانسی صریح
DC-WCNN [۲]	Wavelet + CNN	ورود اطلاعات موجکی به بازسازی MRI	افزایش پیچیدگی معماری
MWCNN [۳]	Multi-level Wavelet CNN	استخراج ویژگی‌های چندمقیاسی	هزینه محاسباتی بیشتر
Attention-guided CNN [۱]	CNN + Attention	تقویت ویژگی‌های مهم	افزایش پیچیدگی آموزش
Xi و همکاران [۲۸]	KMAR-۵۰k	ارائه مجموعه‌داده استاندارد برای حذف آرتیفکتMRI	محدود بودن بهMRI زانو

۲- ۱۴ جمع‌بندی

در این فصل، مبانی نظری و پیشینه پژوهش‌های مرتبط با بازسازی تصاویر پزشکی و حذف آرتیفکت‌های حرکتی در تصاویر MRI مورد بررسی قرار گرفت. در آغاز اصول تصویربرداری تشدید مغناطیسی و ویژگی‌های این روش تصویربرداری بیان شد و سپس مهم‌ترین عوامل ایجاد آرتیفکت‌های حرکتی و تأثیر آن‌ها بر کاهش کیفیت تصاویر پزشکی تشریح گردید. بررسی این مباحث نشان داد که آرتیفکت‌های حرکتی یکی از پرتکرار ترین چالش‌های موجود در تصویربرداری MRI هستند و می‌توانند موجب کاهش دقت تشخیص و افزایش نیاز به تصویربرداری مجدد شوند.

در ادامه، روش‌های سنتی بهبود کیفیت تصویر شامل فیلترهای مکانی، روش‌های غیرمحلی، مدل‌های آماری، روش‌های مبتنی بر تبدیل موجک و الگوریتم‌های کلاسیک کاهش نویز بررسی شدند. اگرچه این روش‌ها در کاهش نویز و بهبود کیفیت تصویر نتایج قابل قبولی ارائه کرده‌اند، اما وابستگی آن‌ها به فرضیات از پیش تعیین‌شده، تنظیم پارامترها و محدودیت در حفظ هم‌زمان جزئیات و ساختارهای تصویر، استفاده از آن‌ها را در مسائل پیچیده بازسازی تصاویر پزشکی با محدودیت مواجه کرده است.

سپس، روند توسعه روش‌های مبتنی بر یادگیری عمیق مورد بررسی قرار گرفت. نتایج پژوهش‌های مرورشده نشان داد که شبکه‌های عصبی کانولوشنی با قابلیت یادگیری مستقیم از داده‌ها، پیشرفت قابل توجهی در بازسازی تصویر ایجاد کرده‌اند. همچنین، معرفی معماری‌هایی مانند U-Net، WACNN، MWACNN و مدل‌های مبتنی بر مکانیزم توجه نشان داد که ترکیب استخراج ویژگی‌های چندمقیاسی، یادگیری داده‌محور و تمرکز بر ویژگی‌های مهم می‌تواند کیفیت بازسازی تصاویر را نسبت به روش‌های متداول بهبود دهد.

در بخش پایانی فصل، مجموعه‌داده استاندارد KMAR-۵۰K و پژوهش‌های انجام‌شده بر پایه آن بررسی شد. وجود تصاویر جفت‌شده، تنوع مناسب داده‌ها و امکان ارزیابی منصفانه مدل‌های مختلف، این مجموعه‌داده را به یکی از منابع مهم پژوهش در حوزه حذف آرتیفکت‌های حرکتی تبدیل کرده است. همچنین، مرور مطالعات انجام‌شده نشان داد که اگرچه معماری‌های مختلف عملکرد قابل قبولی در بازسازی تصاویر ارائه کرده‌اند، اما همچنان حفظ هم‌زمان جزئیات ساختاری، کاهش آرتیفکت و افزایش کارایی محاسباتی از مهم‌ترین چالش‌های این حوزه به شمار می‌رود.

با وجود پیشرفت معماری‌های مبتنی بر تبدیل موجک و مکانیزم توجه، گزارش‌های موجود در ادبیات پژوهش نشان می‌دهد که عملکرد این معماری‌ها به نوع داده، شرایط آموزش و معیارهای ارزیابی وابسته است و هنوز اجماع روشنی درباره برتری یک معماری نسبت به سایر روش‌ها مشاهده نشده است. ازاین‌رو، مقایسه این معماری‌ها در یک چارچوب یکسان همچنان یکی از موضوعات مهم پژوهشی به شمار می‌رود.

به‌طور کلی، بررسی مبانی نظری و پیشینه پژوهش نشان داد که حرکت از روش‌های کلاسیک به سمت معماری‌های مبتنی بر یادگیری عمیق و تبدیل موجک، مسیر اصلی توسعه پژوهش‌های بازسازی تصویر در سال‌های اخیر بوده است. این مطالعات زمینه علمی مناسبی برای انجام پژوهش حاضر فراهم می‌کنند و ضرورت بررسی و مقایسه معماری‌های مختلف در حذف آرتیفکت‌های حرکتی و بهبود کیفیت تصاویر MRI را نشان می‌دهند.

۳ فصل سوم: روش اجرای تحقیق

۳- ۱ مقدمه

در فصل‌های پیشین، مبانی نظری و پژوهش‌های مرتبط با تصویربرداری تشدید مغناطیسی، آرتیفکت‌های حرکتی، روش‌های بازسازی تصویر، شبکه‌های عصبی کانولوشنی، تبدیل موجک و معماری‌های مبتنی بر یادگیری عمیق بررسی شد. در این فصل، روش اجرای پژوهش به‌صورت مرحله‌به‌مرحله تشریح می‌شود. هدف اصلی فصل حاضر، بیان دقیق فرآیند انجام تحقیق از انتخاب و آماده‌سازی داده‌ها تا پیاده‌سازی، آموزش، اعتبارسنجی و ارزیابی مدل‌های مورد بررسی است.

چهار معماری شاملU-Net ، WACNN، MWACNN بدون مکانیزم توجه و MWACNN مبتنی بر توجه کانالی برای حذف آرتیفکت‌های حرکتی و بازسازی تصاویر MRI پیاده‌سازی و ارزیابی شده‌اند. در ادامه‌ی این فصل، ابتدا مجموعه‌داده و مراحل پیش‌پردازش، سپس معماری هر مدل، تنظیمات آموزش و در پایان معیارهای ارزیابی معرفی می‌شوند.

در ادامه، ابتدا نوع و رویکرد پژوهش معرفی می‌شود. سپس مجموعه‌داده مورد استفاده، نحوه تقسیم داده‌ها، مراحل پیش‌پردازش و آماده‌سازی تصاویر، معماری مدل‌ها، تنظیمات آموزش، تابع هزینه، روش انتخاب بهترین وزن، معیارهای ارزیابی، تحلیل آماری و محیط اجرای آزمایش‌ها تشریح خواهد شد. در پایان فصل نیز روند کلی اجرای پژوهش به‌صورت یکپارچه جمع‌بندی می‌شود.

۳- ۲ نوع و روش تحقیق

پژوهش حاضر از نظر هدف، در دسته پژوهش‌های کاربردی قرار می‌گیرد؛ زیرا در آن چند معماری مبتنی بر یادگیری عمیق برای حذف آرتیفکت‌های حرکتی و بهبود کیفیت تصاویر MRI پیاده‌سازی و ارزیابی شده‌اند. نتایج حاصل از این پژوهش می‌تواند در توسعه سامانه‌های پردازش تصاویر پزشکی، نرم‌افزارهای بازسازی تصویر و سامانه‌های کمک‌تشخیص مورد استفاده قرار گیرد.

از نظر نحوه اجرا، این تحقیق یک پژوهش تجربی و کمی محسوب می‌شود. چهار معماری شاملU-Net ، WACNN، MWACNN بدون مکانیزم توجه و MWACNN مبتنی بر توجه کانالی تحت شرایط آموزشی یکسان پیاده‌سازی، آموزش و ارزیابی شده‌اند. عملکرد مدل‌ها با استفاده از معیارهای کمی شامل نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده مورد مقایسه قرار گرفته است. همچنین تعداد پارامترهای قابل آموزش و زمان استنتاج هر مدل به‌منظور بررسی کارایی محاسباتی اندازه‌گیری شده‌اند.

به‌منظور بررسی دقیق‌تر اختلاف عملکرد معماری‌ها، مقادیر معیارهای ارزیابی برای هر نمونه از مجموعه آزمون مستقل محاسبه شد و از آزمون ناپارامتری آزمون رتبه علامت‌دار ویلکاکسون برای انجام مقایسه‌های زوجی میان مدل‌ها استفاده گردید. این تحلیل امکان بررسی آماری تفاوت عملکرد مدل‌ها را علاوه بر مقایسه میانگین معیارها فراهم می‌کند. [۲۹]

۳- ۳ محیط اجرای پژوهش

پیاده‌سازی مدل‌ها و اجرای آزمایش‌های این پژوهش در محیط Kaggle Notebook انجام شد. این محیط امکان استفاده از پردازنده گرافیکی، اجرای کدهای مبتنی بر یادگیری عمیق و دسترسی به کتابخانه‌های مورد نیاز را فراهم می‌کند. با توجه به محدودیت منابع پردازشی و زمان اجرای جلسات در این محیط، تنظیمات آموزش به‌گونه‌ای انتخاب شد که اجرای تمامی مدل‌ها تحت شرایط یکسان امکان‌پذیر باشد.

زبان برنامه‌نویسی مورد استفاده در این پژوهش Python و چارچوب اصلی پیاده‌سازی مدل‌ها PyTorch بود. برای خواندن فایل‌های پزشکی با قالب NIfTI[۱] از کتابخانه Nibabel و برای انجام عملیات عددی و پردازش آرایه‌ها از NumPy استفاده شد. همچنین کتابخانه‌های Pandas برای ذخیره و تحلیل نتایج، Scikit-image برای محاسبه برخی معیارهای کیفیت تصویر، Matplotlib برای ترسیم نمودارها و SciPy برای اجرای آزمون آماری ویلکاکسون مورد استفاده قرار گرفتند. به‌منظور کاهش تغییرپذیری ناشی از مقداردهی تصادفی، Seed برابر با ۴۲ برایPython ، NumPyو PyTorch در نظر گرفته شد.

برای افزایش سرعت آموزش، انتقال داده‌ها و محاسبات شبکه بر روی GPU انجام شد و در صورت در دسترس بودن پردازنده گرافیکی، از آموزش با دقت ترکیبی[۲] استفاده گردید. همچنین در صورت وجود بیش از یکGPU ، امکان استفاده از موازی‌سازی داده‌ها[۳] در کد در نظر گرفته شد.

جدول ۳-۱ مشخصات کلی محیط نرم‌افزاری و پردازشی پژوهش را نشان می‌دهد.

[۱] Neuroimaging Informatics Technology Initiative (NIfTI)

[۲] Mixed Precision Training

[۳] Data Parallelism (DataParallel)

۱۱ ۳-۱ مشخصات کلی محیط نرم‌افزاری و پردازشی

مشخصات	مؤلفه
Kaggle Notebook	محیط اجرا
Python ۳	زبان برنامه‌نویسی
PyTorch	چارچوب یادگیری عمیق
NIfTI	قالب داده‌های پزشکی
Nibabel	کتابخانه خواندن تصاویرMRI
NumPy	پردازش عددی
Pandas	تحلیل و ذخیره نتایج
Scikit-image	معیارهای کیفیت تصویر
SciPy	تحلیل آماری
Matplotlib	ترسیم نمودارها
GPU	شتاب‌دهنده سخت‌افزاری
فعال در صورت وجود GPU	آموزش با دقت ترکیبی
DataParallel در صورت دسترسی	پردازش چندGPU

کلیه‌ی آزمایش‌های این پژوهش در محیط Kaggle Notebook و با استفاده از GPU مدل Tesla T٤ انجام شدند. جزئیات تنظیمات آموزشی هر مدل - شامل تعداد Epoch، اندازه وصله وBatch در بخش ۳-۵ به‌طور کامل ارائه می‌شود. استفاده از تنظیمات آموزشی یکسان موجب شد اختلاف نتایج حاصل، ناشی از تفاوت معماری مدل‌ها باشد و امکان مقایسه دقیق و تکرارپذیر عملکرد آن‌ها فراهم شود.

۳- ۴ معماری مدل‌های مورد استفاده

۳- ۴- ۱ مقدمه

به‌منظور ارزیابی عملکرد معماری‌های مختلف در حذف آرتیفکت‌های حرکتی تصاویرMRI، در این پژوهش چهار مدل شامل U-Net، WACNN، MWACNN بدون مکانیزم توجه و MWACNN مبتنی بر توجه کانالی پیاده‌سازی و ارزیابی شدند. انتخاب این چهار معماری با هدف مقایسه‌ی سه سطح متفاوت پیچیدگی صورت گرفت: U-Net به‌عنوان معماری مرجع بدون تحلیل موجک، WACNN با یک سطح تجزیه‌ی موجکی، و دو نسخه از MWACNN برای بررسی جداگانه‌ی نقش مکانیزم توجه کانالی.

U-Net به‌عنوان مدل پایه در نظر گرفته شد و عملکرد آن با معماری‌های مبتنی بر تبدیل موجک مقایسه گردید. همچنین، به‌منظور بررسی تأثیر مکانیزم توجه کانالی، دو نسخه از معماری MWACNN شامل نسخه بدون توجه و نسخه مجهز به توجه کانالی پیاده‌سازی و با یکدیگر مقایسه شدند.

۳- ۴- ۲ U-Net

شبکه U-Net که نخستین بار توسط Ronneberger و همکاران معرفی شد، یکی از شناخته‌شده‌ترین معماری‌ها در حوزه پردازش تصاویر پزشکی است. این معماری از ساختار رمزگذار–رمزگشا تشکیل شده و با استفاده از اتصالات میانبر اطلاعات استخراج‌شده در مرحله رمزگذار را مستقیماً به مرحله رمزگشا منتقل می‌کند. این ویژگی باعث می‌شود جزئیات مکانی تصویر حفظ شده و فرآیند بازسازی با دقت بیشتری انجام شود.

ورودی شبکه، تصاویر دارای آرتیفکت و خروجی آن تصاویر بازسازی‌شده بدون آرتیفکت است. ساختار این مدل شامل لایه‌های کانولوشن، فعال‌ساز ReLU[1]، عملیات Downsampling و Upsampling است و اتصالات میانبرها نقش مهمی در حفظ اطلاعات بافتی تصاویر ایفا می‌کنند.

در پیاده‌سازی این پژوهش، مدل U-Net شامل سه سطح رمزگذار، یک بخش میانی (Bottleneck) و سه سطح رمزگشا است. در مسیر رمزگذار، تعداد کانال‌های ویژگی به‌ترتیب از ۱ به ۳۲، سپس ۶۴ و ۱۲۸ افزایش یافته و در بخش Bottleneck به ۲۵۶ می‌رسد. هر بلوک از دو لایه کانولوشنی ۳×۳، Batch Normalization و تابع فعال‌سازی ReLU تشکیل شده است. کاهش ابعاد مکانی با Max Pooling و افزایش ابعاد در مسیر رمزگشا با Transposed Convolution انجام می‌شود. ویژگی‌های متناظر رمزگذار از طریق اتصالات میانبر با ویژگی‌های رمزگشا ترکیب می‌شوند. در پایان، یک کانولوشن ۱×۱ خروجی تک‌کاناله شبکه را تولید می‌کند و مقادیر خروجی در بازه صفر تا یک محدود می‌شوند.

[1] Rectified Linear Unit (ReLU)

۳-۴-۳ WACNN

در معماری WACNN مورد استفاده در این پژوهش، تصویر ورودی ابتدا با استفاده از تبدیل موجک دوبعدی ثابت به چهار زیرباند LL، LH، HL و HH تجزیه می‌شود. فیلترهای مورد استفاده از نوع Haar-like با ابعاد ۲×۲ هستند و در طول آموزش به‌روزرسانی نمی‌شوند. ضرایب حاصل از تجزیه موجکی توسط یک لایه کانولوشنی ۳×۳ از چهار کانال به ۶۴ کانال ویژگی تبدیل شده و پس از Batch Normalization و ReLU وارد هشت بلوک باقیمانده مجهز به توجه کانالی می‌شوند.

هر بلوک شامل دو کانولوشن ۳×۳، Batch Normalization و اتصال Residual است. در WACNN پیاده‌سازی‌شده، توجه کانالی در این بلوک‌ها فعال است. علاوه بر این، پس از مجموعه هشت بلوک، یک ماژول توجه کانالی دیگر نیز به‌عنوان Gate روی ویژگی‌های نهایی اعمال می‌شود. در ادامه، ویژگی‌ها با یک لایه کانولوشنی به چهار کانال موجکی بازگردانده شده، با ضرایب موجکی اولیه جمع می‌شوند و در نهایت تصویر با تبدیل موجک معکوس بازسازی می‌شود. ([۲]، [۳])

۳-۴-۴ MWACNN بدون مکانیزم توجه

در نسخه MWACNN بدون مکانیزم توجه، ابتدا تصویر ورودی در سطح اول با استفاده از تبدیل موجک به چهار زیرباند تجزیه می‌شود. سپس مؤلفه LL سطح اول برای بار دوم تحت تبدیل موجک قرار می‌گیرد تا اطلاعات تصویر در سطح دوم نیز استخراج شود. ضرایب موجکی سطح دوم توسط یک شبکه کانولوشنی کوچک با ساختار ۴⟶۳۲⟶۴ پردازش شده و خروجی این بخش با ضرایب ورودی سطح دوم از طریق اتصال Residual ترکیب می‌شود.

در ادامه، اطلاعات پردازش‌شده سطح دوم با تبدیل موجک معکوس بازسازی و سپس مجدداً به حوزه موجک منتقل می‌شوند. در صورت وجود اختلاف ابعاد، اندازه ویژگی‌های سطح دوم با استفاده از درون‌یابی Bilinear با ضرایب سطح اول هماهنگ می‌شود. سپس اطلاعات موجکی سطح اول و دوم در امتداد کانال‌ها با یکدیگر ترکیب شده و توسط یک لایه کانولوشنی به ۶۴ کانال ویژگی تبدیل می‌شوند. ویژگی‌های حاصل از هشت بلوک باقیمانده عبور می‌کنند. در این نسخه، ماژول توجه کانالی غیرفعال است. در پایان، ویژگی‌ها به چهار کانال موجکی تبدیل شده، با ضرایب موجکی سطح اول جمع می‌شوند و تصویر نهایی با استفاده از تبدیل موجک معکوس بازسازی می‌شود. ([۳]، [۲])

در این پژوهش، علاوه بر نسخه کامل، نسخه‌ای از MWACNN بدون مکانیزم توجه نیز پیاده‌سازی شد تا تأثیر استفاده از تبدیل موجک بدون ماژول توجه به‌صورت مستقل بررسی شود

۳- ۴- ۵ معماری MWACNN مبتنی بر توجه کانالی(Channel Attention)

به‌منظور بررسی تأثیر مکانیزم توجه بر فرآیند بازسازی تصاویر MRI، نسخه‌ای از معماری MWACNN مجهز به توجه کانالی نیز پیاده‌سازی شد. ساختار اصلی این معماری، شامل تجزیه موجکی دو سطحی، نحوه ترکیب اطلاعات حاصل از دو سطح موجک و مسیر بازسازی تصویر، مشابه نسخه MWACNN بدون توجه در نظر گرفته شد. تفاوت اصلی میان دو نسخه، فعال بودن مکانیزم توجه کانالی در بخش استخراج ویژگی‌ها است. این طراحی امکان بررسی مستقیم اثر افزودن مکانیزم توجه را بدون ایجاد تغییر اساسی در سایر اجزای معماری فراهم می‌کند.

پس از ترکیب اطلاعات موجکی سطح اول و دوم، ویژگی‌های حاصل توسط یک لایه کانولوشنی به ۶۴ کانال ویژگی تبدیل می‌شوند و سپس از هشت بلوک باقیمانده مجهز به توجه کانالی عبور می‌کنند. هر بلوک شامل دو لایه کانولوشنی با کرنل ۳×۳، لایه‌های Batch Normalization، تابع فعال‌سازی ReLU و یک اتصال باقیمانده است. در نسخه مبتنی بر توجه، خروجی بخش کانولوشنی پیش از ترکیب با ورودی بلوک، توسط ماژول توجه کانالی وزن‌دهی می‌شود. در نسخه بدون توجه، همین بخش با نگاشت همانی جایگزین شده و سایر اجزای بلوک بدون تغییر باقی می‌مانند. علاوه بر توجه کانالی داخل هشت بلوک باقیمانده، یک ماژول توجه کانالی اضافی نیز پس از مجموعه این بلوک‌ها اعمال می‌شود.

در ماژول توجه کانالی، ابتدا Global Average Pooling برای هر کانال یک مقدار نماینده تولید می‌کند. سپس این مقادیر از دو لایه تمام‌متصل عبور کرده و با ReLU و در نهایت Sigmoid به ضرایب وزن‌دهی کانالی تبدیل می‌شوند. این ضرایب درfeature map های متناظر ضرب می‌شوند.

علاوه بر مکانیزم توجه موجود در هر یک از بلوک‌های باقیمانده، پس از عبور ویژگی‌ها از مجموعه هشت بلوک، یک ماژول توجه کانالی اضافی نیز بر روی ویژگی‌های نهایی اعمال می‌شود. سپس ویژگی‌های پردازش‌شده توسط یک لایه کانولوشنی مجدداً به چهار کانال متناظر با ضرایب موجکی تبدیل می‌شوند. این ضرایب با ضرایب موجکی سطح اول جمع شده و در نهایت با استفاده از تبدیل موجک معکوس، تصویر بازسازی‌شده تولید می‌شود. کد پیاده‌سازی نیز دقیقاً توالی «ترکیب اطلاعات دو سطح ← استخراج ویژگی ← بلوک‌های باقیمانده ← توجه کانالی نهایی ← بازسازی موجکی» را اجرا می‌کند.

نسخه مجهز به توجه کانالی با هدف بررسی تأثیر وزن‌دهی تطبیقی کانال‌های ویژگی بر کیفیت بازسازی تصاویر MRI پیاده‌سازی شد. از آنجا که ساختار موجکی، تعداد کانال‌های ویژگی، تعداد بلوک‌ها و سایر اجزای اصلی شبکه در نسخه‌های دارای توجه و بدون توجه یکسان در نظر گرفته شدند، مقایسه این دو مدل امکان بررسی نقش مکانیزم توجه کانالی را در چارچوب شرایط آزمایشی یکسان فراهم می‌کند. ([۱]، [۱۱])

۲-۳-۱۲ مقایسه معماری‌های مورد استفاده در پژوهش

معماری	ویژگی اصلی	مزیت	هدف استفاده در پژوهش
U-Net	ساختار Encoder–Decoder همراه با Skip Connection	حفظ بهتر اطلاعات مکانی و بازسازی تصویر	استفاده به‌عنوان مدل پایه برای مقایسه
WACNN	ترکیب شبکه کانولوشنی با تجزیه موجکی تک‌سطحی	استخراج هم‌زمان ویژگی‌های مکانی و فرکانسی	بررسی عملکرد معماری مبتنی بر تجزیه موجکی تک‌سطحی در مقایسه با U-Net
MWACNN	استفاده از تجزیه موجکی چندسطحی در معماری کانولوشنی	استخراج ویژگی‌ها در مقیاس‌های مختلف و حفظ بهتر جزئیات	بررسی عملکرد پردازش موجکی چندسطحی بدون توجه کانالی
MWACNN + Channel Attention	ترکیب تجزیه موجکی چندسطحی با مکانیزم توجه کانالی	تقویت کانال‌های ویژگی مؤثر و بهبود نمایش اطلاعات	بررسی تأثیر مکانیزم توجه کانالی در کنار تبدیل موجک

همان‌طور که در جدول ۳-۲ مشاهده می‌شود، هر یک از معماری‌های مورد بررسی با هدف ارزیابی یکی از جنبه‌های مؤثر در بازسازی تصاویر MRI انتخاب شده‌اند. U-Net به‌عنوان مدل پایه، مبنای مقایسه سایر معماری‌ها را فراهم می‌کند. WACNN نقش استفاده از تبدیل موجک را بررسی می‌کند، MWACNN تأثیر تحلیل چندمقیاسی را مورد ارزیابی قرار می‌دهد و نسخه مجهز به توجه کانالی، امکان بررسی نقش مکانیزم توجه در فرآیند بازسازی تصویر را فراهم می‌سازد. به این ترتیب، مقایسه این چهار معماری امکان تحلیل اثر هر یک از اجزای اصلی مدل پیشنهادی را به‌صورت مستقل فراهم می‌کند.

۳- ۵ تنظیمات آموزش مدل‌ها

۳- ۵- ۱ مقدمه

پس از انتخاب مجموعه‌داده مورد استفاده در پژوهش، در نخستین مرحله داده‌های مورد نیاز از مجموعه‌داده KMAR - K۵۰ دریافت شدند. از آنجا که داده‌های مورد استفاده در چند بخش قرار داشتند، پس از دانلود، بخش‌های مختلف مجموعه‌داده با یکدیگر تجمیع شدند تا ساختار واحدی از داده‌های مورد نیاز پژوهش ایجاد شود. در ادامه، فایل‌های حاصل به‌صورت منظم دسته‌بندی و مرتب شدند تا دسترسی به تصاویر دارای آرتیفکت حرکتی و تصاویر مرجع متناظر در مراحل بعدی پردازش، آموزش و ارزیابی مدل‌ها به شکل ساختاریافته انجام شود. این مرحله به‌عنوان پایه آماده‌سازی داده‌ها، پیش از اجرای مراحل پردازشی و ورود تصاویر به مدل‌ها انجام شد.

پس از آماده‌سازی و سازمان‌دهی مجموعه‌داده، چهار معماری U-Net، WACNN، MWACNN بدون مکانیزم توجه و MWACNN مبتنی بر توجه کانالی در محیط PyTorch پیاده‌سازی شدند. هدف از پیاده‌سازی این چهار مدل، بررسی عملکرد ساختارهای متفاوت در بازسازی تصاویر MRI و فراهم کردن امکان مقایسه آن‌ها تحت شرایط آموزشی یکسان بود. به همین منظور، پس از تکمیل پیاده‌سازی معماری‌ها، تنظیمات آموزشی مشترکی برای تمامی مدل‌ها در نظر گرفته شد تا تفاوت نتایج تا حد امکان ناشی از ساختار خود معماری‌ها باشد و نه تفاوت در شرایط آموزش.

تنظیمات نهایی آموزش با توجه به نتایج اجرای آزمایشی مدل‌ها و محدودیت‌های پردازشی محیط Kaggle Notebook تعیین شدند. در انتخاب این تنظیمات، عواملی مانند میزان حافظه در دسترس، زمان آموزش، پایداری فرآیند یادگیری و کیفیت بازسازی تصاویر در نظر گرفته شدند. در نهایت، تمامی مدل‌ها با مجموعه‌داده، تابع هزینه، بهینه‌ساز، زمان‌بند نرخ یادگیری و سایر شرایط آموزشی یکسان آموزش و ارزیابی شدند تا امکان مقایسه مستقیم و منصفانه عملکرد آن‌ها فراهم شود.

۳- ۵- ۲ تنظیمات اصلی آموزش

جدول ۳-۳ مهم‌ترین‌هایپر پارامترهای مورد استفاده در فرآیند آموزش مدل‌ها را نشان می‌دهد.

۳-۳-۱۳ تنظیمات اصلی آموزش مدل‌ها

هایپرپارامتر	مقدار
تعدادEpoch	۲۰
Batch Size	۸
Gradient Accumulation	۲
مؤثرBatch	۱۶
اندازهPatch	۱۲۸ × ۱۲۸
تعداد Patch از هر برش	۱
تعداد برش از هر حجم	۶ برش مرکزی
Learning Rate اولیه	۰.۰۰۰۱
حداقلLearning Rate	۰.۰۰۰۰۰۱
Optimizer	Adam
Weight Decay	۱×۱۰^{-5}
Scheduler	Cosine Annealing
Gradient Clipping	۱.۰
Mixed Precision	فعال در صورت وجودGPU

همان‌طور که در جدول ۳-۳ مشاهده می‌شود، تمامی مدل‌ها با تنظیمات آموزشی یکسان آموزش داده شدند. استفاده از Batch Size برابر با ۸ و انباشت گرادیان برابر با ۲ باعث شد اندازه مؤثر Batch به ۱۶ برسد، بدون آنکه مصرف حافظه GPU به‌طور قابل توجهی افزایش یابد. همچنین، استفاده از اندازه وصله برابر با ۱۲۸×۱۲۸ و انتخاب ۶ برش مرکزی از هر حجم MRI موجب شد فرآیند آموزش با هزینه محاسباتی مناسب و در عین حال با حفظ اطلاعات تشریحی مهم انجام شود.

۳- ۵- ۳ انتخاب بهینه‌ساز (Optimizer)

برای بهینه‌سازی پارامترهای شبکه از الگوریتم Adam استفاده شد. این الگوریتم یکی از پرکاربردترین روش‌های بهینه‌سازی در یادگیری عمیق است و با استفاده هم‌زمان از میانگین متحرک گرادیان‌ها و میانگین متحرک مربع گرادیان‌ها،میتواند فرآیند آموزش را پایدارتر و همگرایی مدل را سریع‌تر می‌کند. به همین علت، Adam در بسیاری از مسائل پردازش تصویر، به‌ویژه بازسازی تصاویر پزشکی، عملکرد مناسبی از خود نشان داده است.

در پیاده‌سازی این پژوهش، نرخ یادگیری اولیه برابر با 1×10^{-4}در نظر گرفته شد و به‌عنوان یک مؤلفه منظم‌سازی و با هدف کاهش احتمال بیش‌برازش، مقدار Weight Decay برابر با 1×10^{-5}تنظیم شد. این تنظیمات برای تمامی معماری‌های مورد بررسی یکسان در نظر گرفته شدند تا مقایسه عملکرد آن‌ها تحت شرایط آموزشی مشابه انجام شود.

۳- ۵- ۴ زمان‌بند نرخ یادگیری (Learning Rate Scheduler)

به‌منظور بهبود فرآیند آموزش، از زمان‌بند CosineAnnealingLR برای تنظیم نرخ یادگیری استفاده شد. در این روش، نرخ یادگیری از مقدار اولیه به‌صورت تدریجی و بر اساس تابع کسینوسی کاهش می‌یابد تا در پایان آموزش به مقدار حداقل 1×10^{-6}برسد. این کاهش تدریجی میتواند باعث شود فرآیند یادگیری در مراحل پایانی پایدارتر شده و مدل بتواند با نوسانات کمتر به پاسخ بهینه نزدیک شود.

استفاده از CosineAnnealingLR همچنین موجب می‌شود شبکه در مراحل ابتدایی آموزش با نرخ یادگیری بزرگ‌تر، ویژگی‌های اصلی تصویر را سریع‌تر یاد بگیرد و در مراحل پایانی، با نرخ یادگیری کوچک‌تر، پارامترهای خود را با دقت بیشتری تنظیم کند. به همین دلیل، این روش در بسیاری از پژوهش‌های مرتبط با بازسازی تصویر و آموزش شبکه‌های کانولوشنی مورد استفاده قرار می‌گیرد.

۳- ۵- ۵ تابع هزینه (Loss Function)

انتخاب تابع هزینه مناسب یکی از عوامل مهم در آموزش مدل‌های یادگیری عمیق برای بازسازی تصاویر پزشکی است. اگر فرآیند آموزش تنها بر پایه اختلاف شدت پیکسل‌ها انجام شود، ممکن است تصویر خروجی از نظر عددی به تصویر مرجع نزدیک باشد، اما بخشی از ساختارها، مرزها و جزئیات مهم آن به‌درستی بازسازی نشوند. ازین رو ، در این پژوهش از یک تابع هزینه ترکیبی[۱] استفاده شد تا علاوه بر کاهش خطای پیکسلی، شباهت ساختاری میان تصویر بازسازی‌شده و تصویر مرجع نیز حفظ شود.

تابع هزینه مورد استفاده از ترکیب L۱ Loss [۲] و SSIM Loss تشکیل شده است. بخش L۱ Loss اختلاف مطلق میان شدت پیکسل‌های تصویر خروجی و تصویر مرجع را اندازه‌گیری می‌کند و شبکه را به تولید تصاویری با خطای عددی کمتر هدایت می‌کند. در مقابل ، SSIM Loss میزان شباهت ساختاری میان دو تصویر را ارزیابی کرده و مدل

[۱] Combined Loss

[۲] L۱ Loss (Least Absolute Error Loss)

را به حفظ روشنایی، کنتراست و ساختارهای مهم تصویر ترغیب می‌کند. ترکیب این دو معیار باعث می‌شود مدل علاوه بر کاهش خطای پیکسلی، در حفظ جزئیات ساختاری نیز عملکرد بهتری داشته باشد.

در این پژوهش، تابع هزینه به‌صورت زیر تعریف شد:

$$Loss = \alpha L_{L1} + (1 - \alpha) L_{SSIM}$$

که در آن L_{L1} نشان‌دهنده خطای مطلق میان تصویر بازسازی‌شده و تصویر مرجع و L_{SSIM} بیانگر خطای مبتنی بر شاخص شباهت ساختاری است. همچنین، ضریب α میزان سهم هر یک از این دو مؤلفه را در تابع هزینه تعیین می‌کند.

بر اساس آزمایش‌های متعدد، تنظیمات نهایی مدل، مقدار $\alpha = 0.8$ انتخاب شد. در نتیجه، ۸۰ درصد از تابع هزینه به L۱ Loss و ۲۰ درصد به SSIM Loss اختصاص یافت. این ترکیب باعث شد مدل در کنار کاهش اختلاف پیکسلی، ساختارهای مهم تصویر را نیز بهتر حفظ کند؛ موضوعی که در بازسازی تصاویر MRI و حذف آرتیفکت‌های حرکتی اهمیت زیادی دارد. مؤلفه شاخص شباهت ساختاری مورد استفاده در تابع هزینه به‌صورت مشتق‌پذیر در فرآیند آموزش پیاده‌سازی شد؛ در حالی که شاخص شباهت ساختاری مرحله ارزیابی با استفاده از پیاده‌سازی کتابخانه Scikit-imageمحاسبه شد.

استفاده از تابع هزینه یکسان برای تمامی معماری‌های مورد بررسی، باعث شد فرآیند آموزش تحت شرایط مشابه انجام شود و مقایسه عملکرد مدل‌های مختلف تنها به تفاوت ساختار آن‌ها وابسته باشد، نه تفاوت در معیار آموزش.

۳- ۵- ۶ استفاده از انباشت گرادیان(Gradient Accumulation)

به‌منظور استفاده مؤثر از منابع سخت‌افزاری و در عین حال حفظ اندازه مناسب Batch،از روش انباشت گرادیان استفاده شد. در این روش، به‌جای به‌روزرسانی وزن‌های شبکه پس از هر Batch، گرادیان‌های حاصل از چند Batch متوالی با یکدیگر تجمیع شده و سپس یک‌بار عملیات به‌روزرسانی پارامترها انجام می‌شود.

مقدار انباشت گرادیان برابر با ۲ انتخاب شد. با توجه به اندازه Batch برابر با ۸، این روش باعث شد اندازه مؤثر Batchبه ۱۶ افزایش یابد، بدون نیاز به نگهداری مستقیم ۱۶ نمونه به‌صورت هم‌زمان در حافظهGPU . استفاده از این روش موجب پایداری بیشتر فرآیند آموزش و استفاده بهینه‌تر از منابع پردازشی محیط Kaggle شد.

علاوه بر این، به‌منظور جلوگیری از افزایش بیش از حد مقدار گرادیان‌ها و کاهش احتمال ناپایداری در فرآیند آموزش، از Gradient Clipping استفاده شد. در این پژوهش، حداکثر نرم گرادیان‌ها برابر با ۱.۰در نظر گرفته شد. پیش از هر مرحله به‌روزرسانی پارامترهای شبکه، در صورتی که نرم گرادیان‌ها از این مقدار بیشتر بود، مقادیر آن‌ها محدود می‌شد. این روش به حفظ پایداری فرآیند بهینه‌سازی، به‌ویژه هنگام آموزش معماری‌های عمیق‌تر، کمک می‌کند.

همچنین، برای کاهش مصرف حافظه GPU و افزایش سرعت محاسبات، در صورت در دسترس بودن پردازنده گرافیکی از آموزش با Mixed Precision استفاده شد. در این حالت، محاسبات مناسب با دقت پایین‌تر انجام شده و برای حفظ پایداری گرادیان‌ها از روش مقیاس‌بندی گرادیان استفاده شد.

۳- ۵- ۷ دلایل انتخاب تنظیمات آموزشی

تنظیمات آموزشی این پژوهش با هدف ایجاد تعادل میان کیفیت بازسازی، زمان آموزش و محدودیت‌های سخت‌افزاری انتخاب شدند. از آنجا که اجرای هم‌زمان چهار معماری مختلف بر روی مجموعه‌داده به منابع محاسباتی قابل توجهی نیاز داشت، تمامی پارامترهای آموزشی به‌گونه‌ای تنظیم شدند که ضمن حفظ کیفیت آموزش، امکان اجرای کامل آزمایش‌ها در محیط Kaggle نیز فراهم شود.

بر همین اساس، تعداد Epoch۲۰ ، اندازه وصله برابر با ۱۲۸×۱۲۸ پیکسل، Batch Size برابر با ۸ و انتخاب ۶ برش مرکزی از هر حجم MRI برای آموزش در نظر گرفته شد. همچنین، استفاده از انباشت گرادیان، آموزش با Mixed Precision و بارگذاری بهینه داده‌ها باعث شد مصرف حافظه کاهش یافته و تمامی مدل‌ها تحت شرایط آموزشی یکسان آموزش داده شوند.

انتخاب یکسان این تنظیمات برای همه معماری‌ها موجب شد مقایسه نتایج صرفاً بر اساس تفاوت ساختار مدل‌ها انجام شود و عوامل جانبی در ارزیابی عملکرد آن‌ها تأثیر حداقلی داشته باشند.

۳- ۵- ۸ جمع‌بندی

در این بخش، تنظیمات مورد استفاده برای آموزش مدل‌های مختلف معرفی شد. تمامی معماری‌های مورد بررسی با استفاده از تنظیمات آموزشی یکسان شامل بهینه‌ساز Adam، زمان‌بند CosineAnnealingLR، تابع هزینه ترکیبی، اندازه وصله ، اندازه Batch، نرخ یادگیری و سایر پارامترهای آموزشی مشابه آموزش داده شدند. همچنین، استفاده از تکنیک‌هایی مانند انباشت گرادیان و Mixed Precision امکان بهره‌گیری مؤثرتر از منابع سخت‌افزاری را فراهم کرد.

یکسان بودن شرایط آموزش برای تمامی مدل‌ها باعث شد نتایج حاصل از ارزیابی آن‌ها به‌صورت منصفانه قابل مقایسه باشد و تفاوت عملکرد مدل‌ها به ویژگی‌های معماری آن‌ها نسبت داده شود، نه به تفاوت در تنظیمات آموزشی.

۳- ۶ معیارهای ارزیابی

۳- ۶- ۱ مقدمه

به‌منظور ارزیابی عملکرد مدل‌های مورد بررسی، از مجموعه‌ای از معیارهای کمی استفاده شد. این معیارها کیفیت بازسازی تصاویر، میزان شباهت میان تصویر بازسازی‌شده و تصویر مرجع، مقدار خطای بازسازی و میزان حفظ ساختارهای تصویری را از جنبه‌های مختلف مورد بررسی قرار می‌دهند. استفاده هم‌زمان از چند معیار ارزیابی باعث می‌شود عملکرد مدل‌ها تنها بر اساس یک شاخص عددی سنجیده نشود و کیفیت خروجی از دیدگاه‌های مختلف تحلیل گردد.

معیارهای نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده برای ارزیابی کیفیت بازسازی تصاویر

MRIمورد استفاده قرار گرفتند. علاوه بر این، تعداد پارامترهای قابل آموزش[۱] و زمان استنتاج[۲] نیز به‌منظور بررسی کارایی محاسباتی مدل‌ها اندازه‌گیری شدند. تمامی این معیارها برای چهار معماری مورد بررسی و بر روی مجموعه آزمون مستقل محاسبه شدند.

۳- ۶- ۲ نسبت سیگنال به نویز اوج(PSNR)

یکی از رایج‌ترین معیارهای ارزیابی کیفیت تصاویر بازسازی‌شده، نسبت سیگنال به نویز اوج Peak Signal-to-Noise Ratio یا PSNR است. این معیار میزان اختلاف بین تصویر بازسازی‌شده و تصویر مرجع را بر اساس خطای پیکسلی اندازه‌گیری می‌کند و معمولاً بر حسب دسی‌بل (dB) بیان می‌شود. هرچه مقدار نسبت سیگنال به نویز اوج بزرگ‌تر باشد، اختلاف میان تصویر خروجی و تصویر مرجع کمتر بوده و کیفیت بازسازی بالاتر خواهد بود.

رابطه محاسبه نسبت سیگنال به نویز اوج به صورت زیر است:

$$PSNR = ۱۰\log_{۱۰}\left(\frac{MAX^{۲}}{MSE}\right)$$

که در آن:

- MAX بیشترین مقدار شدت پیکسل تصویر است .
- میانگین مربعات خطا میانگین مربعات خطا بین تصویر بازسازی‌شده و تصویر مرجع است .

از آنجا که تصاویر این پژوهش پس از پیش‌پردازش به بازه ۰ تا ۱ نرمال‌سازی شدند، مقدار MAX برابر با ۱ در نظر گرفته شد. بنابراین، مقدار نسبت سیگنال به نویز اوج تنها به میزان خطای بازسازی وابسته است. [۳۰]

نسبت سیگنال به نویز اوج به‌عنوان یکی از معیارهای اصلی ارزیابی کیفیت بازسازی مورد استفاده قرار گرفت و برای مقایسه عملکرد تمامی مدل‌ها بر روی مجموعه آزمون مستقل محاسبه شد. مقدار بیشتر این معیار نشان‌دهنده بازسازی دقیق‌تر تصویر و کاهش خطای پیکسلی نسبت به تصویر مرجع است.

۳- ۶- ۳ شاخص شباهت ساختاری (SSIM)

شاخص شباهت ساختاری Structural Similarity Index Measure یا SSIM یکی از اصلی ترین معیارهای ارزیابی کیفیت تصاویر بازسازی‌شده، به‌ویژه در تصاویر پزشکی، است. برخلاف معیار نسبت سیگنال به نویز اوج که تنها بر اختلاف شدت پیکسل‌ها تمرکز دارد، شاخص شباهت ساختاری میزان شباهت دو تصویر را از نظر روشنایی، کنتراست و ساختار بررسی می‌کند. به همین دلیل، این معیار توانایی بیشتری در ارزیابی حفظ بافت‌ها، لبه‌ها و جزئیات ساختاری تصاویر MRI دارد.

---

[۱] Trainable Parameters

[۲] Inference Time

رابطه محاسبه شاخص شباهت ساختاری به صورت زیر است:

$$SSIM(x,y) = \frac{(۲\mu_x\mu_y + C_۱)(۲\sigma_{xy} + C_۲)}{(\mu_x^۲ + \mu_y^۲ + C_۱)(\sigma_x^۲ + \sigma_y^۲ + C_۲)}$$

که در آن:

- μ_x و μ_y میانگین شدت تصاویر هستند .
- σ_x و σ_y انحراف معیار تصاویر را نشان می‌دهند .
- σ_{xy} کوواریانس میان دو تصویر است .
- $C_۱$ و $C_۲$ ثابت‌هایی برای جلوگیری از ناپایداری محاسبات و تقسیم بر صفر هستند .

مقدار شاخص شباهت ساختاری در بازه ۰ تا ۱ قرار دارد و هرچه این مقدار به ۱ نزدیک‌تر باشد، شباهت ساختاری تصویر بازسازی‌شده به تصویر مرجع بیشتر خواهد بود. [۱۵]

و شاخص شباهت ساختاری به‌عنوان یکی از معیارهای اصلی ارزیابی کیفیت بازسازی مورد استفاده قرار گرفت؛ زیرا حفظ ساختارهای آناتومیکی، مرز بافت‌ها و جزئیات ظریف در تصاویر MRI از اهمیت بالایی برخوردار است.

۳- ۶- ۴ میانگین قدر مطلق خطا (MAE)

میانگین قدر مطلق خطا Mean Absolute Error یا MAE یکی از معیارهای متداول برای اندازه‌گیری اختلاف میان تصویر بازسازی‌شده و تصویر مرجع است. این معیار میانگین اختلاف مطلق شدت پیکسل‌های متناظر در دو تصویر را محاسبه می‌کند و دید مناسبی از میزان خطای بازسازی ارائه می‌دهد.

رابطه محاسبه میانگین قدر مطلق خطا به صورت زیر است:

$$MAE = \frac{۱}{N}\sum_{i=۱}^{N} | x_i - y_i |$$

که در آن:

- x_i مقدار پیکسل متناظر در تصویر مرجع است .
- y_i مقدار پیکسل متناظر در تصویر بازسازی‌شده است .
- N تعداد کل پیکسل‌های تصویر است .

مقدار کمتر میانگین قدر مطلق خطا نشان‌دهنده اختلاف کمتر میان تصویر بازسازی‌شده و تصویر مرجع و در نتیجه عملکرد بهتر مدل است. از آنجا که میانگین قدر مطلق خطا نسبت به میانگین مربعات خطا حساسیت کمتری به خطاهای بزرگ دارد، استفاده هم‌زمان از این دو معیار توانایی ارائه ی ارزیابی دقیق‌تری از کیفیت بازسازی تصاویر را داراست. [۳۰]

همچنین میانگین قدر مطلق خطابرای مقایسه میزان خطای بازسازی تمامی مدل‌ها بر روی مجموعه آزمون مستقل محاسبه شد.

۳- ۶- ۵ میانگین مربعات خطا (MSE)

میانگین مربعات خطا Mean Squared Error یا MSE یکی از معیارهای متداول برای اندازه‌گیری میزان اختلاف میان تصویر بازسازی‌شده و تصویر مرجع است. در این معیار، اختلاف شدت پیکسل‌های متناظر ابتدا به توان دو رسانده شده و سپس میانگین آن‌ها محاسبه می‌شود. به علت مجذور کردن خطاها، میانگین مربعات خطا نسبت به خطاهای بزرگ حساس‌تر از میانگین قدر مطلق خطا است و می‌تواند تفاوت‌های قابل توجه میان دو تصویر را بهتر آشکار کند

رابطه محاسبه میانگین مربعات خطا به صورت زیر است:

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2$$

که در آن:

- x_i مقدار پیکسل متناظر در تصویر مرجع است .
- y_i مقدار پیکسل متناظر در تصویر بازسازی‌شده است .
- N تعداد کل پیکسل‌های تصویر است .

هرچه مقدار میانگین مربعات خطا کمتر باشد، اختلاف میان تصویر بازسازی‌شده و تصویر مرجع کمتر بوده و کیفیت بازسازی تصویر بهتر خواهد بود. میانگین مربعات خطا برای مقایسه میزان خطای بازسازی مدل‌های مختلف بر روی مجموعه آزمون مستقل محاسبه شد. [۳۰]

۳- ۶- ۶ ریشه دوم میانگین مربعات خطای نرمال‌شده (NRMSE)

ریشه دوم میانگین مربعات خطای نرمال‌شده Normalized Root Mean Squared Error یا NRMSE نسخه نرمال‌شده RMSE است که امکان مقایسه عملکرد مدل‌ها را در تصاویر با دامنه‌های شدت متفاوت فراهم می‌کند. این معیار مقدار خطای بازسازی را نسبت به دامنه شدت تصویر نرمال‌سازی کرده و در نتیجه مقایسه نتایج بین مجموعه‌داده‌های مختلف را آسان‌تر می‌سازد

رابطه محاسبه ریشه دوم میانگین مربعات خطای نرمال‌شده به صورت زیر است:

$$NRMSE = \frac{\sqrt{MSE}}{I_{max} - I_{min}}$$

که در آن:

- MSE میانگین مربعات خطا است .
- I_{max} و I_{min} به‌ترتیب بیشترین و کمترین شدت پیکسل در تصویر مرجع هستند .

هرچه مقدار ریشه دوم میانگین مربعات خطای نرمال‌شده کمتر باشد، کیفیت بازسازی تصویر بهتر خواهد بود. این معیار برای ارزیابی میزان خطای نرمال‌شده و مقایسه عملکرد مدل‌ها مورد استفاده قرار گرفت. [۳۰]

۳- ۶- ۷ ضریب همبستگی نرمال‌شده (NCC)

ضریب همبستگی نرمال‌شده Normalized Cross-Correlation یا NCC میزان شباهت و همبستگی میان تصویر بازسازی‌شده و تصویر مرجع را اندازه‌گیری می‌کند. این معیار نشان می‌دهد که تغییرات شدت پیکسل‌ها در دو تصویر تا چه اندازه با یکدیگر همخوانی دارند و یکی از معیارهای مناسب برای ارزیابی کیفیت بازسازی تصاویر پزشکی محسوب می‌شود .

رابطه محاسبه ضریب همبستگی نرمال‌شده به صورت زیر است:

$$NCC = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}}$$

که در آن:

- x_i و y_i مقادیر پیکسل‌های متناظر در تصویر مرجع و تصویر بازسازی‌شده هستند .
- $\bar{x}$ و $\bar{y}$ میانگین شدت پیکسل‌های دو تصویر را نشان می‌دهند .

مقدار ضریب همبستگی نرمال‌شده در بازه ۱ -تا ۱ قرار می‌گیرد و هرچه این مقدار به ۱ نزدیک‌تر باشد، همبستگی میان تصویر بازسازی‌شده و تصویر مرجع بیشتر است. ضریب همبستگی نرمال‌شده به‌عنوان یکی از معیارهای ارزیابی میزان شباهت کلی تصاویر مورد استفاده قرار گرفت. [۳۰]

۳- ۶- ۸ تعداد پارامترهای قابل آموزش(Trainable Parameters)

علاوه بر معیارهای کیفیت بازسازی، تعداد پارامترهای قابل آموزش هر مدل نیز محاسبه شد. این معیار بیانگر میزان پیچیدگی معماری شبکه و حجم پارامترهایی است که در فرآیند آموزش به‌روزرسانی می‌شوند. تعداد پارامترهای قابل آموزش یکی از شاخص‌های اندازه مدل است، اما به‌تنهایی معیار کاملی برای هزینه محاسباتی یا زمان استنتاج نیست؛ نوع عملیات و ساختار معماری نیز بر سرعت اجرا تأثیر دارند. [۳۰] در این پژوهش، تعداد پارامترهای قابل آموزش برای تمامی معماری‌های مورد بررسی محاسبه شد تا علاوه بر مقایسه کیفیت بازسازی، پیچیدگی محاسباتی مدل‌ها نیز مورد ارزیابی قرار گیرد.

۳- ۶- ۹ زمان استنتاج(Inference Time)

زمان استنتاج مدت زمانی است که یک مدل برای تولید تصویر بازسازی‌شده از یک تصویر ورودی نیاز دارد. این معیار یکی از شاخص‌های مهم در ارزیابی کارایی مدل‌های یادگیری عمیق است، زیرا در بسیاری از کاربردهای پزشکی، علاوه بر کیفیت بازسازی، سرعت پردازش نیز اهمیت بالایی دارد و می‌تواند بر قابلیت استفاده از مدل در محیط‌های بالینی تأثیرگذار باشد.

برای مقایسه سرعت اجرای مدل‌ها، یک نمونه از مجموعه آزمون مستقل به مدل داده شد و فرآیند استنتاج ۱۰۰ بار تکرار گردید. میانگین زمان هر اجرا بر حسب میلی‌ثانیه به‌عنوان زمان استنتاج گزارش شد. بنابراین، این مقدار یک شاخص مقایسه‌ای در شرایط سخت‌افزاری پژوهش است و نباید به‌عنوان زمان ثابت مدل در تمامی محیط‌های اجرایی تفسیر شود.

۴-۳ ۱۴ خلاصه معیارهای ارزیابی مورد استفاده در پژوهش

معیار	هدف ارزیابی	مقدار مطلوب
PSNR	سنجش کیفیت بازسازی بر اساس اختلاف شدت پیکسل‌ها	بیشتر
SSIM	سنجش شباهت ساختاری تصویر بازسازی‌شده و تصویر مرجع	بیشتر
MAE	اندازه‌گیری میانگین قدر مطلق خطای بازسازی	کمتر
MSE	اندازه‌گیری میانگین مربعات خطای بازسازی	کمتر
NRMSE	اندازه‌گیری خطای نرمال‌شده بازسازی	کمتر
NCC	سنجش همبستگی میان تصویر بازسازی‌شده و تصویر مرجع	بیشتر
Inference Time	ارزیابی سرعت بازسازی مدل	کمتر
Trainable Parameters	ارزیابی پیچیدگی محاسباتی مدل	کمتر (در صورت حفظ کیفیت بازسازی)

۳- ۶- ۱۰ آزمون آماری ویلکاکسون

علاوه بر مقایسه نتایج اصلی چهار معماری، از آزمون رتبه علامت‌دار ویلکاکسون به‌عنوان یک تحلیل آماری تکمیلی برای بررسی اختلاف مقادیر زوجی برخی معیارهای ارزیابی استفاده شد. این آزمون ناپارامتری برای مقایسه دو مجموعه وابسته کاربرد دارد و به فرض نرمال بودن توزیع اختلافات زوجی نیاز ندارد. [۲۹]

دلیل انتخاب آزمون ویلکاکسون در این پژوهش، ماهیت زوجی داده‌های حاصل از ارزیابی مدل‌ها و عدم نیاز این آزمون به فرض نرمال بودن توزیع اختلافات است. تمامی معماری‌های مورد بررسی بر روی مجموعه آزمون یکسان ارزیابی شده‌اند؛ بنابراین، برای هر نمونه آزمون، نتایج متناظر حاصل از دو مدل قابل مقایسه بوده و داده‌ها به‌صورت زوجی در نظر گرفته می‌شوند. در چنین شرایطی، استفاده از آزمون‌های مربوط به نمونه‌های مستقل، مانند Independent t-test، مناسب نیست؛ زیرا نتایج دو مدل از نمونه‌های مستقل به دست نیامده‌اند. همچنین، Paired t-test بر فرض نرمال بودن توزیع اختلافات زوجی تکیه دارد، در حالی که آزمون ویلکاکسون به‌عنوان یک روش ناپارامتری، به این فرض نیاز ندارد. بر این اساس، با توجه به ماهیت زوجی داده‌ها و هدف پژوهش در مقایسه عملکرد U-Netبا هر یک از مدل‌های مورد بررسی، آزمون رتبه علامت‌دار ویلکاکسون برای انجام مقایسه‌های آماری انتخاب شد.

در این پژوهش، عملکرد معماری U-Net به‌صورت جداگانه با سه معماری WACNN، MWACNN بدون توجه و MWACNN مبتنی بر توجه کانالی مقایسه شد. این مقایسه بر اساس معیارهای نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده انجام گرفت. هدف از انجام آزمون برای هر معیار، بررسی این موضوع بود که آیا اختلاف مشاهده‌شده میان عملکرد U-Net و مدل مورد مقایسه از نظر آماری معنادار است یا می‌تواند ناشی از تغییرات موجود در نتایج نمونه‌های آزمون باشد.

سطح معنی‌داری[۱] برابر با ۰.۰۵در نظر گرفته شد. بر این اساس، اگر مقدار p-value کمتر از ۰.۰۵باشد، اختلاف میان دو مدل از نظر آماری معنادار تلقی می‌شود؛ در غیر این صورت، اختلاف مشاهده‌شده از نظر آماری معنادار نخواهد بود.

نتایج حاصل از آزمون ویلکاکسون در فصل چهارم ارائه و تحلیل شده است. این نتایج در کنار معیارهای کمی مورد استفاده قرار گرفتند تا علاوه بر مقایسه مستقیم مقادیر عملکرد مدل‌ها، معناداری آماری اختلاف میان معماری‌های مورد بررسی نیز ارزیابی شود.

۳- ۶- ۱۱ جمع‌بندی

عملکرد مدل‌ها تنها بر اساس یک معیار ارزیابی نشد، بلکه مجموعه‌ای از معیارهای مکمل برای بررسی کیفیت بازسازی، حفظ ساختار تصویر، میزان خطای بازسازی و کارایی محاسباتی مورد استفاده قرار گرفت. معیارهای نسبت سیگنال به نویز اوج و شاخص شباهت ساختاری برای ارزیابی کیفیت و شباهت ساختاری تصاویر، میانگین قدر مطلق

[۱] Significance Level

خطا، میانگین مربعات خطا و ریشه دوم میانگین مربعات خطای نرمال‌شده برای سنجش میزان خطای بازسازی، ضریب همبستگی نرمال‌شده برای بررسی همبستگی میان تصویر بازسازی‌شده و تصویر مرجع، و همچنین تعداد پارامترهای قابل آموزش و زمان استنتاج برای ارزیابی پیچیدگی محاسباتی و سرعت اجرای مدل‌ها محاسبه شدند. استفاده هم‌زمان از این معیارها امکان مقایسه جامع و منصفانه معماری‌های مختلف را فراهم کرد و مبنای تحلیل نتایج ارائه‌شده در فصل چهارم قرار گرفت.

۳- ۷ روند اجرای پژوهش(Workflow)

۳- ۷- ۱ روند کلی اجرای پژوهش

روند کلی اجرای این پژوهش از آماده‌سازی مجموعه‌داده آغاز شده و تا آموزش، انتخاب بهترین وزن و ارزیابی نهایی معماری‌های مورد بررسی ادامه می‌یابد. پس از آماده‌سازی و سازمان‌دهی مجموعه‌داده -K۵۰KMAR، تصاویر MRI دارای آرتیفکت حرکتی با تصاویر مرجع متناظر جفت‌سازی شدند تا برای هر نمونه ورودی، تصویر Ground Truth صحیح در اختیار شبکه قرار گیرد. سپس زوج‌های مربوط به داده‌های آموزش و اعتبارسنجی به‌صورت تصادفی مرتب شده و با نسبت ۹۰ درصد برای آموزش و ۱۰ درصد برای اعتبارسنجی تقسیم شدند. داده‌های مربوط به مجموعه آزمون مستقل نیز به‌صورت جداگانه نگهداری شدند و در فرآیند آموزش و انتخاب مدل مورد استفاده قرار نگرفتند.

پس از تقسیم داده‌ها، مراحل پیش‌پردازش مورد نیاز برای ورود تصاویر به شبکه انجام شد. حجم‌های MRI با فرمت NIfTI بارگذاری شده و مقادیر شدت تصویر به بازه صفر تا یک نرمال‌سازی شدند. سپس برش‌های دوبعدی و وصله‌های مورد نیاز از زوج تصاویر دارای آرتیفکت و مرجع استخراج شدند تا ورودی و هدف متناظر برای آموزش، اعتبارسنجی و آزمون مدل‌ها فراهم شود. در کد، نرمال‌سازی هر حجم بر اساس حداقل و حداکثر شدت همان حجم انجام شده است.

در مرحله بعد، چهار معماری U-Net، WACNN، MWACNN بدون مکانیزم توجه و MWACNN مبتنی بر توجه کانالی مورد پیاده‌سازی و آزمایش قرار گرفتند. برای فراهم کردن امکان مقایسه منصفانه، فرآیند آموزش با تنظیمات مشترک شامل تابع هزینه ترکیبی، بهینه‌ساز Adam، زمان‌بند CosineAnnealingLR، انباشت گرادیان، Gradient Clipping و در صورت وجود GPU، Mixed Precision انجام شد. این روند آموزشی برای هر معماری به‌صورت مشابه اجرا شد تا تأثیر تفاوت ساختاری مدل‌ها بر نتایج قابل بررسی باشد. در کد، تابع آموزش نیز به‌صورت یک تابع مشترک تعریف شده و مدل موردنظر به آن ارسال می‌شود.

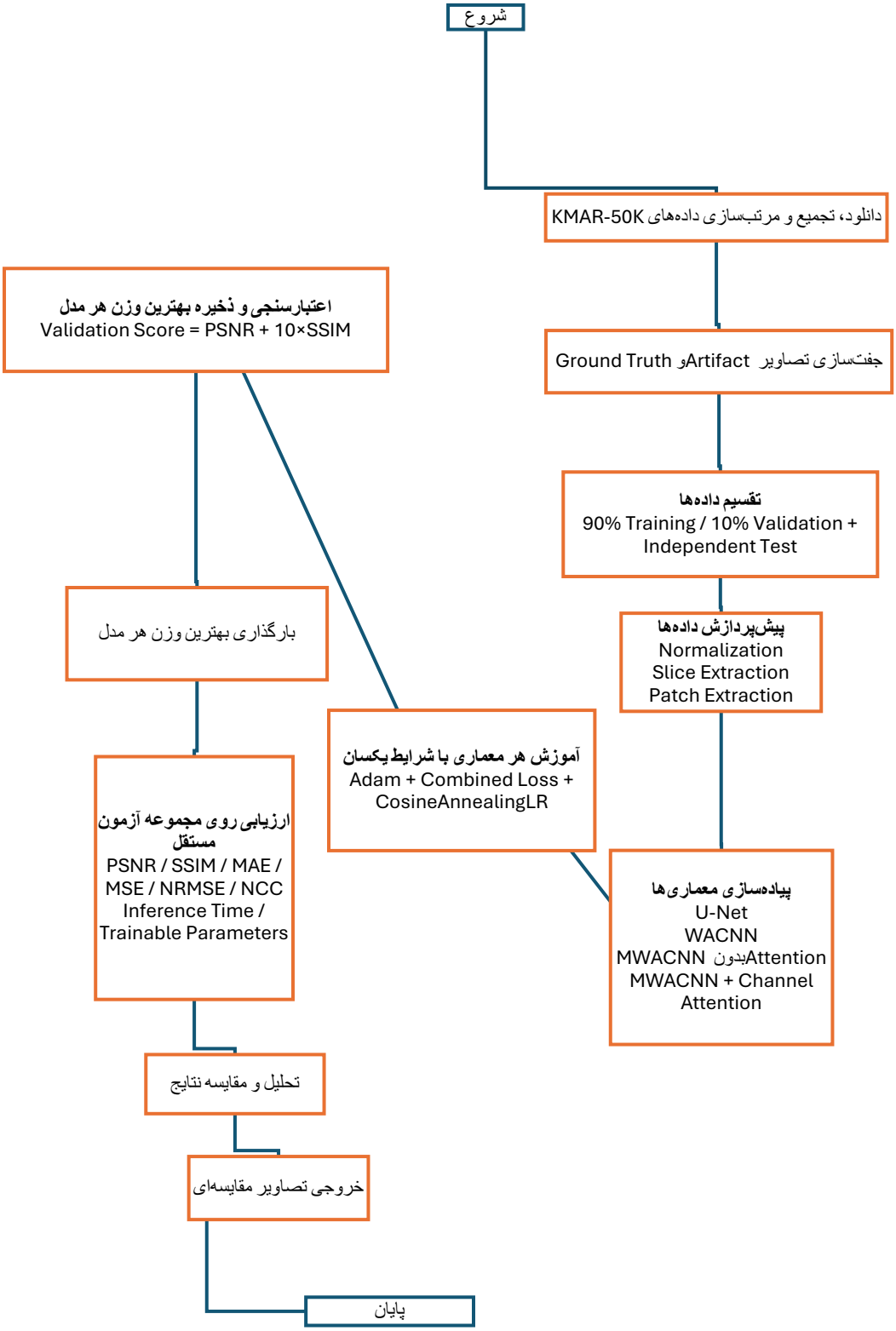
در طول فرآیند آموزش، عملکرد مدل در بازه‌های مشخص بر روی مجموعه اعتبارسنجی محاسبه شد. برای انتخاب بهترین نسخه هر مدل، امتیازی بر اساس ترکیب دو معیار نسبت سیگنال به نویز اوج و شاخص شباهت ساختاری و مطابق رابطه PSNR+ ۱۰×SSIM محاسبه شد. هر زمان این امتیاز از بهترین مقدار قبلی بیشتر بود،

وزن‌های مدل ذخیره شدند. پس از پایان آموزش نیز بهترین وزن ذخیره‌شده مجدداً بارگذاری شد و برای ارزیابی نهایی مورد استفاده قرار گرفت.

در مرحله ارزیابی نهایی، بهترین نسخه هر مدل بر روی مجموعه آزمون مستقل ارزیابی شد. معیارهای نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده برای سنجش کیفیت بازسازی و میزان خطای مدل‌ها محاسبه شدند. همچنین تعداد پارامترهای قابل آموزش هر معماری و زمان استنتاج مورد اندازه‌گیری قرار گرفت. در

محاسبه زمان استنتاج، یک نمونه از مجموعه آزمون به شبکه وارد شده و اجرای مدل ۱۰۰ بار تکرار شد؛ سپس میانگین زمان هر اجرا بر حسب میلی‌ثانیه محاسبه گردید.

در نهایت، نتایج کمی و کیفی حاصل از چهار معماری با یکدیگر مقایسه شدند تا نقاط قوت و محدودیت‌های هر مدل از نظر کیفیت بازسازی و کارایی محاسباتی مشخص شود. این فرآیند امکان بررسی تأثیر استفاده از تبدیل موجک، پردازش چندسطحی و مکانیزم توجه کانالی را در بازسازی تصاویر MRI فراهم کرد. شکل ۳-۱ روند کلی اجرای پژوهش را به‌صورت شماتیک نشان می‌دهد.



۳۵-۱ روند کلی اجرای پژوهش

٤ فصل چهارم: تجزیه و تحلیل داده‌ها (یافته‌ها)

۴- ۱ معرفی مجموعه‌داده

۴- ۱- ۱ معرفی مجموعه‌دادهKMAR-۵۰K

در این پژوهش، به‌منظور آموزش و ارزیابی مدل‌های یادگیری عمیق از مجموعه‌داده KMAR-۵۰K (Knee MRI Artifact Reduction) استفاده شد. این مجموعه‌داده یکی از مجموعه‌داده‌های استاندارد در حوزه حذف آرتیفکت‌های حرکتی تصاویر MRI بوده و به‌طور ویژه برای ارزیابی روش‌های بازسازی تصویر توسعه یافته است. استفاده از یک مجموعه‌داده استاندارد، علاوه بر افزایش قابلیت تکرارپذیری پژوهش، امکان مقایسه نتایج با سایر مطالعات را نیز فراهم می‌کند. [۲۸]

پس از دانلود بخش‌های مورد نیازKMAR-۵۰K ، فایل‌ها تجمیع و سازمان‌دهی شدند. تصاویر دارای آرتیفکت و تصاویر مرجع به‌صورت منظم دسته‌بندی شدند تا دسترسی به زوج‌های متناظر در مراحل پردازش، آموزش و ارزیابی فراهم شود.

مجموعه‌داده KMAR-۵۰K شامل حجم‌های سه‌بعدی MRI زانو در دو گروه اصلی است:

- تصاویر دارای آرتیفکت حرکتی (Artifact Images)
- تصاویر مرجع بدون آرتیفکت (Ground Truth Images)

برای هر حجم دارای آرتیفکت، یک حجم مرجع متناظر وجود دارد؛ بنابراین، امکان استفاده از یادگیری نظارت‌شده[۱] فراهم می‌شود. در این چارچوب، مدل تلاش می‌کند نگاشت میان تصویر دارای آرتیفکت و تصویر مرجع را یاد گرفته و تصویر بازسازی‌شده‌ای نزدیک به تصویر مرجع تولید کند.

این مجموعه‌داده بر پایه تصاویر واقعی MRI تهیه شده و در اثر کیفیت مناسب، ساختار استاندارد و استفاده گسترده در پژوهش‌های مرتبط با حذف آرتیفکت و بازسازی تصویر، به‌عنوان مجموعه‌داده اصلی این پژوهش انتخاب شد. [۲۸]

۴- ۱- ۲ ساختار مجموعه‌داده

پس از دریافت مجموعه‌داده، داده‌ها به سه بخش مستقل شامل مجموعه آموزش[۲] ، مجموعه اعتبارسنجی[۳] و مجموعه آزمون مستقل [۴] تقسیم شدند.

مجموعه آموزش برای یادگیری پارامترهای مدل مورد استفاده قرار گرفت. مجموعه اعتبارسنجی به‌منظور پایش روند آموزش، انتخاب بهترین وزن مدل و جلوگیری از بیش‌برازش[۵] به کار رفت. در نهایت، عملکرد مدل‌ها تنها یک‌بار و پس از پایان آموزش، بر روی مجموعه آزمون مستقل ارزیابی شد تا توانایی تعمیم مدل‌ها به داده‌های دیده‌نشده بررسی شود.

---

[۱] Supervised Learning

[۲] Training Set

[۳] Validation Set

[۴] Independent Test Set

[۵] Overfitting

تصاویر دارای آرتیفکت و تصاویر مرجع بر اساس نام فایل متناظر جفت‌سازی شدند و تنها فایل‌هایی که در هر دو مجموعه دارای نام مشترک بودند وارد فرآیند آموزش یا ارزیابی شدند. پس از این تطبیق، ۱۰۱ زوج معتبر برای مجموعه آزمون مستقل مورد استفاده قرار گرفت. پس از تقسیم‌بندی، برش‌های دوبعدی مورد نیاز برای آموزش، اعتبارسنجی و آزمون از هر حجم استخراج شدند.

تعداد نمونه‌های مورد استفاده در هر بخش در جدول ۴-۱ ارائه شده است.

۱۵ ۴-۱ تعداد داده‌های مورد استفاده

مجموعه	تعداد حجم(Volume)	تعداد برش(Slice)
آموزش	۱۲۰۵	۷۲۳۰
اعتبارسنجی	۱۳۴	۸۰۴
آزمون مستقل	۱۰۱	۶۰۶

۴- ۱- ۳ نحوه انتخاب داده‌ها

آموزش مدل‌ها بر اساس زوج داده‌های متناظر [۱] انجام شد. به این ترتیب، برای هر حجم MRI دارای آرتیفکت، یک حجم مرجع بدون آرتیفکت وجود داشت که به‌عنوان هدف آموزش مورد استفاده قرار گرفت. این ساختار امکان استفاده از یادگیری نظارت‌شده را فراهم کرد و مدل‌ها توانستند نگاشت میان تصاویر دارای آرتیفکت و تصاویر مرجع را فرا بگیرند.

تقسیم داده‌ها در سطح حجم‌های سه‌بعدی [۲] انجام شد؛ به‌گونه‌ای که هیچ حجمی به‌طور هم‌زمان در بیش از یکی از مجموعه‌های آموزش، اعتبارسنجی یا آزمون مستقل قرار نگرفت. پس از انجام این تقسیم‌بندی، برش‌های دوبعدی مورد نیاز از هر حجم استخراج شدند. این روش از ورود اطلاعات مشترک میان مجموعه‌ها جلوگیری کرده و ارزیابی عملکرد مدل‌ها را بر روی داده‌های دیده‌نشده امکان‌پذیر ساخت.

۴- ۱- ۴ دلیل انتخاب مجموعه‌داده KMAR-۵۰K

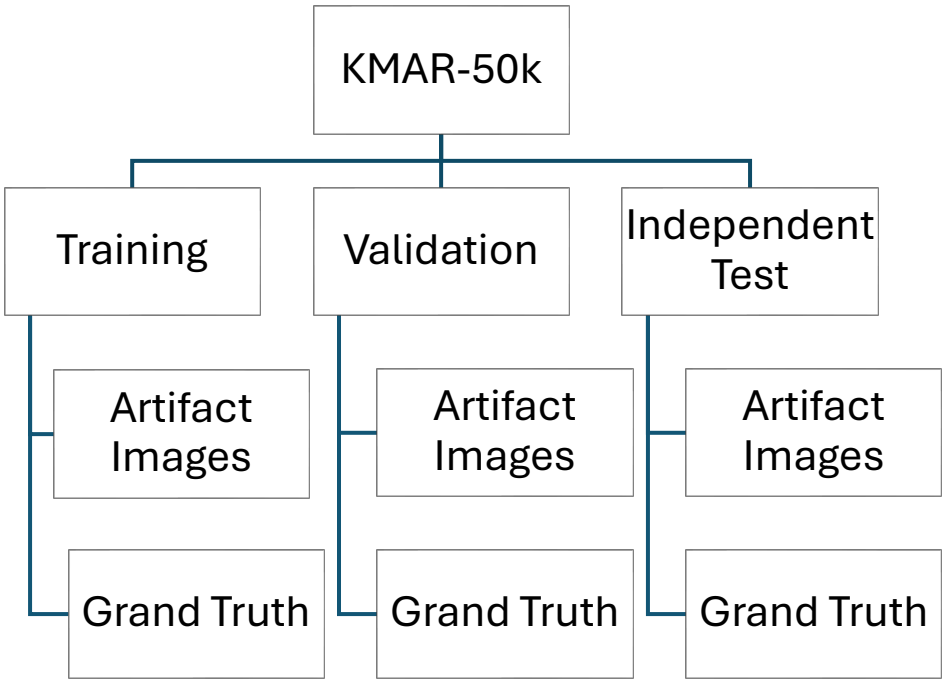
انتخاب یک مجموعه‌داده مناسب نقش مهمی در اعتبار نتایج پژوهش‌های مبتنی بر یادگیری عمیق دارد. در این تحقیق، مجموعه‌داده KMAR-۵۰K به دلیل ویژگی‌های زیر به‌عنوان مجموعه‌داده اصلی انتخاب شد:

[۱] Paired Data

[۲] Volumes

- استفاده از تصاویر واقعی MRI زانو؛
- وجود زوج تصاویر دارای آرتیفکت و تصاویر مرجع متناظر؛
- مناسب بودن برای آموزش نظارت‌شده؛
- ساختار استاندارد و استفاده در پژوهش‌های مرتبط با حذف آرتیفکت و بازسازی تصویر؛
- امکان مقایسه نتایج با مطالعات پیشین؛
- تنوع مناسب داده‌ها برای آموزش، اعتبارسنجی و آزمون مدل‌های یادگیری عمیق .

این ویژگی‌ها باعث شد مجموعه‌داده KMAR-۵۰K بستر مناسبی برای آموزش، اعتبارسنجی و ارزیابی معماری‌های مورد بررسی در این پژوهش فراهم کند [۲۸]



۴۶-۱ ساختار و آماده‌سازی مجموعه‌داده KMAR-50K

تصاویر موجود در مجموعه‌داده با فرمت (.nii.gz) NIfTI ذخیره شده‌اند که هر فایل شامل یک حجم سه‌بعدی MRI است. در مرحله آماده‌سازی داده‌ها، این حجم‌ها به برش‌های دوبعدی تبدیل شدند تا امکان استفاده از معماری‌های دوبعدی مورد استفاده فراهم شود. استفاده از برش‌های دوبعدی علاوه بر کاهش هزینه محاسباتی، امکان آموزش و مقایسه مستقیم مدل‌های U-Net، WACNN، MWACNN بدون توجه و MWACNN مبتنی بر توجه کانالی را فراهم کرد

۴- ۲ پیش پردازش داده‌ها

۴- ۲- ۱ مقدمه

کیفیت داده‌های ورودی و نحوه آماده‌سازی آن‌ها نقش مهمی در عملکرد مدل‌های یادگیری عمیق دارد. تصاویر MRI معمولاً به‌صورت حجم‌های سه‌بعدی با ابعاد نسبتاً بزرگ ذخیره می‌شوند و ممکن است دامنه شدت پیکسل‌های آن‌ها در نمونه‌های مختلف متفاوت باشد. به همین دلیل، پیش از ورود داده‌ها به شبکه‌های عصبی، مجموعه‌ای از مراحل پیش‌پردازش بر روی آن‌ها انجام شد تا ضمن کاهش هزینه محاسباتی، شرایط یکنواختی برای آموزش تمامی مدل‌ها فراهم شود.

مراحل پیش‌پردازش شامل استخراج برش‌های دوبعدی از حجم‌های سه‌بعدی، نرمال‌سازی شدت تصاویر، انتخاب برش‌های مرکزی، استخراج وصله‌های آموزشی، اعمال برش مرکزی در مرحله اعتبارسنجی و آزمون و آماده‌سازی داده‌ها برای ورود به مدل‌های یادگیری عمیق بود.

۴- ۲- ۲ استخراج برش‌های دوبعدی(Slice Extraction)

مجموعه‌داده KMAR-۵۰K به‌صورت حجم‌های سه‌بعدی MRI با فرمت (.nii.gz) NIfTIارائه شده است. از آنجا که تمامی معماری‌های مورد استفاده شامل U-Net، WACNN، MWACNN بدون توجه و MWACNN مبتنی بر توجه کانالی، بر پایه تصاویر دوبعدی طراحی شده‌اند، ابتدا هر حجم سه‌بعدی به مجموعه‌ای از برش‌های دوبعدی تبدیل شد.

از هر حجم شش برش مرکزی انتخاب شد تا ضمن تمرکز بر ناحیه مرکزی حجم، تعداد نمونه‌ها و هزینه محاسباتی در حد قابل مدیریت باقی بماند. انتخاب برش‌های مرکزی باعث شد نواحی دارای اطلاعات تشریحی بیشتر در فرآیند آموزش مورد استفاده قرار گیرند. همین روش برای تمامی مدل‌ها به‌صورت یکسان اعمال شد تا امکان مقایسه منصفانه نتایج فراهم شود.

۴- ۲- ۳ نرمال‌سازی تصاویر(Normalization)

به‌منظور کاهش اختلاف دامنه شدت تصاویر و افزایش پایداری فرآیند آموزش، شدت پیکسل‌های هر حجم MRI به‌صورت مستقل با استفاده از نرمال‌سازی Min–Max به بازه ۰ تا ۱ تبدیل شد. در این روش، کمترین و بیشترین مقدار شدت هر حجم به‌ترتیب به صفر و یک نگاشت شد و سایر مقادیر به‌صورت خطی بین این دو مقدار مقیاس‌بندی شدند.

این مرحله با هدف یکسان‌سازی دامنه شدت تصاویر و فراهم کردن شرایط عددی یکنواخت‌تر برای ورود داده‌ها به شبکه انجام شد.

۴- ۲- ۴ استخراج وصله

یکی از مراحل مهم پیش‌پردازش در این پژوهش، استفاده از آموزش مبتنی بر وصله [۱]بود. در این روش، به‌جای استفاده از کل تصویر، بخش‌های کوچکتری از تصویر به‌عنوان ورودی شبکه مورد استفاده قرار می‌گیرند. این رویکرد علاوه بر کاهش هزینه محاسباتی و مصرف حافظه، امکان تمرکز بیشتر مدل بر ویژگی‌های محلی، لبه‌ها و جزئیات ساختاری تصویر را نیز فراهم می‌کند.

از وصله‌هایی با اندازه ۱۲۸×۱۲۸ پیکسل استفاده شد. در هر دوره آموزش، برای هر برش MRI یک وصله به‌صورت تصادفی [۲] استخراج و به‌عنوان ورودی شبکه مورد استفاده قرار گرفت. این فرآیند باعث شد در دوره‌های مختلف آموزش، نواحی متفاوتی از تصاویر در اختیار مدل قرار گیرد و تنوع نمونه‌های آموزشی بدون نیاز به اعمال روش‌های متداول افزایش داده فراهم شود.

در مرحله اعتبارسنجی و آزمون، به‌منظور حفظ یکنواختی ارزیابی، از برش مرکزی[۳] با همان اندازه استفاده شد تا تمامی مدل‌ها تحت شرایط یکسان مورد ارزیابی قرار گیرند.

۱۶ ۴-۲ تنظیمات استخراج وصله

پارامتر	مقدار
اندازهPatch	۱۲۸ × ۱۲۸
روش آموزش	Random Crop
روش اعتبارسنجی	Center Crop
روش آزمون	Center Crop
تعداد Patch از هر برش	۱
تعداد برش از هر حجم	۶

[۱] Patch-Based Training
[۲] Random Crop
[۳] Center Crop

۴- ۲- ۵ استخراج تصادفی وصله‌ها

به‌منظور افزایش تنوع نمونه‌های آموزشی و استفاده مؤثر از داده‌ها، در این پژوهش از روش استخراج تصادفی وصله‌ها [۱] استفاده شد. در این روش، در هر دوره آموزش، از هر جفت تصویر شامل تصویر دارای آرتیفکت و تصویر مرجع متناظر، یک وصله با ابعاد ۱۲۸×۱۲۸ پیکسل به‌صورت تصادفی استخراج و به‌عنوان ورودی شبکه مورد استفاده قرار گرفت.

استفاده از استخراج تصادفی وصله‌ها باعث شد مدل در طول فرآیند آموزش، نواحی متفاوتی از تصاویر را مشاهده کند و بدون نیاز به روش‌های متداول افزایش داده مانند چرخش یا وارونگی تصویر، تنوع مناسبی در نمونه‌های آموزشی ایجاد شود. همچنین، استفاده از وصله‌های کوچک‌تر موجب کاهش هزینه محاسباتی، استفاده بهینه‌تر از حافظه GPU و افزایش کارایی فرآیند آموزش شد.

همچنین برای مجموعه‌های اعتبارسنجی و آزمون، به‌جای استخراج تصادفی، از Center Crop با همان ابعاد استفاده شد تا شرایط ارزیابی برای تمامی مدل‌ها یکسان و نتایج قابل مقایسه باشند.

۴- ۲- ۶ آماده‌سازی داده‌ها برای آموزش

پس از انجام مراحل پیش‌پردازش، تصاویر دارای آرتیفکت و تصاویر مرجع به‌صورت زوج‌های متناظر سازمان‌دهی شدند. سپس داده‌ها توسط DataLoader کتابخانه PyTorch برای فرآیند آموزش، اعتبارسنجی و آزمون بارگذاری شدند.

به‌منظور افزایش سرعت انتقال داده‌ها و کاهش زمان انتظارGPU ، از بارگذاری موازی داده‌ها[۲] ، حافظه سنجاق‌شده[۳] ، پیش‌بارگذاری داده‌ها [۴] و حافظه موقت[۵] استفاده شد. همچنین داده‌های آموزشی در هر دوره به‌صورت تصادفی [۶] بارگذاری شدند، در حالی که ترتیب داده‌ها در مجموعه‌های اعتبارسنجی و آزمون ثابت نگه داشته شد تا فرآیند ارزیابی به‌صورت یکنواخت انجام شود.

تنظیمات اصلی بارگذاری و آماده‌سازی داده‌ها در جدول ۴-۳ ارائه شده است.

---

[۱] Random Patch Extraction

[۲] Multi-worker Data Loading

[۳] Pinned Memory

[۴] Prefetching

[۵] Cache

[۶] Shuffle

پارامتر	مقدار
Batch Size	۸
تعداد Worker ها	۲
Prefetch Factor	۲
Cache Size	۱۶
Pin Memory	فعال
Shuffle (Training)	فعال
Shuffle (Validation/Test)	غیرفعال
Patch Size	۱۲۸ × ۱۲۸
تعداد Patch از هر Slice	۱

۴- ۲- ۷ جمع‌بندی پیش‌پردازش

پیش‌پردازش داده‌ها با هدف ایجاد شرایط مناسب برای آموزش مدل‌های یادگیری عمیق، کاهش هزینه محاسباتی و افزایش یکنواختی داده‌های ورودی انجام شد. در این فرآیند، حجم‌های سه‌بعدی MRI به برش‌های دوبعدی تبدیل شدند، شدت تصاویر با استفاده از نرمال‌سازی Min–Max به بازه صفر تا یک مقیاس‌بندی شد و از برش‌های مرکزی هر حجم برای تشکیل مجموعه داده استفاده گردید. همچنین، آموزش مدل‌ها بر پایه استخراج تصادفی وصله‌های ۱۲۸×۱۲۸ انجام شد، در حالی که در مراحل اعتبارسنجی و آزمون از Center Crop استفاده شد تا شرایط ارزیابی برای تمامی مدل‌ها یکسان باشد.

در ادامه، داده‌های آماده‌شده توسط DataLoader کتابخانه PyTorch و با استفاده از بارگذاری موازی، پیش‌بارگذاری داده‌ها و حافظه موقت در اختیار مدل‌ها قرار گرفتند. استفاده از این مراحل پیش‌پردازش و تنظیمات یکسان برای تمامی معماری‌ها، موجب شد مقایسه عملکرد مدل‌های مختلف بر پایه شرایط آموزشی مشابه انجام شود و نتایج حاصل از اعتبار و قابلیت تکرار بیشتری برخوردار باشند.

۴- ۳ نتایج و تحلیل

در این فصل، نتایج حاصل از پیاده‌سازی و ارزیابی معماری‌های مختلف برای بازسازی تصاویر MRI و حذف آرتیفکت‌های حرکتی ارائه و تحلیل می‌شود. در این پژوهش، چهار معماری شامل U-Net، WACNN، MWACNN و MWACNN مبتنی بر توجه کانالی بر روی مجموعه‌داده استاندارد KMAR-۵۰K آموزش و ارزیابی شدند. تمامی مدل‌ها با استفاده از تنظیمات آموزشی یکسان آموزش داده شدند تا امکان مقایسه منصفانه عملکرد آن‌ها فراهم شود.

به‌منظور ارزیابی عملکرد مدل‌ها، از مجموعه‌ای از معیارهای کمی شامل نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده، ضریب همبستگی نرمال‌شده، زمان استنتاج و تعداد پارامترهای قابل آموزش استفاده شد. این معیارها کیفیت بازسازی تصاویر، میزان حفظ ساختارهای تصویری، خطای بازسازی و کارایی محاسباتی هر معماری را از جنبه‌های مختلف بررسی می‌کنند.

علاوه بر ارزیابی کمی، خروجی‌های تصویری مدل‌ها نیز به‌صورت کیفی مورد بررسی قرار گرفتند تا توانایی هر معماری در حفظ جزئیات، مرزهای بافتی و کاهش آرتیفکت‌های حرکتی از دیدگاه بصری نیز تحلیل شود. در ادامه این فصل، ابتدا نتایج کمی مدل‌ها ارائه می‌شود، سپس عملکرد آن‌ها بر اساس معیارهای مختلف مورد مقایسه و تحلیل قرار می‌گیرد و در نهایت، با بررسی نمونه تصاویر بازسازی‌شده، نقاط قوت و محدودیت‌های هر معماری جمع‌بندی خواهد شد.

۴- ۴ محیط اجرای آزمایش‌ها

تمامی آزمایش‌های این پژوهش در محیط Kaggle Notebook و با استفاده از کتابخانه PyTorch انجام شدند. به‌منظور مقایسه منصفانه عملکرد مدل‌ها، تمامی معماری‌ها با تنظیمات آموزشی یکسان، شامل بهینه‌ساز، تابع هزینه، زمان‌بند نرخ یادگیری و مجموعه‌داده مشابه آموزش و ارزیابی شدند .جدول ۴-۴ مشخصات کلی محیط اجرای آزمایش‌ها را نشان می‌دهد.

۱۸ ۴-۴ مشخصات محیط اجرای آزمایش‌ها

مشخصه	مقدار
محیط اجرا	Kaggle Notebook
Framework	PyTorch
GPU	NVIDIA Tesla T٤
Optimizer	Adam
Learning Rate	۱×۱۰^{-4}
Scheduler	CosineAnnealingLR
Loss Function	Combined Loss (L۱ + SSIM)
Dataset	KMAR-٥٠K
Models	U-Net، WACNN، MWACNN، MWACNN + Channel Attention

همان‌طور که در جدول ۴-۴ مشاهده می‌شود، تمامی مدل‌ها در یک محیط سخت‌افزاری و نرم‌افزاری یکسان اجرا شدند. همچنین، استفاده از تنظیمات آموزشی مشترک باعث شد اختلاف نتایج حاصل، ناشی از تفاوت در معماری مدل‌ها باشد و تأثیر عوامل جانبی مانند شرایط اجرا یا پارامترهای آموزشی به حداقل برسد.

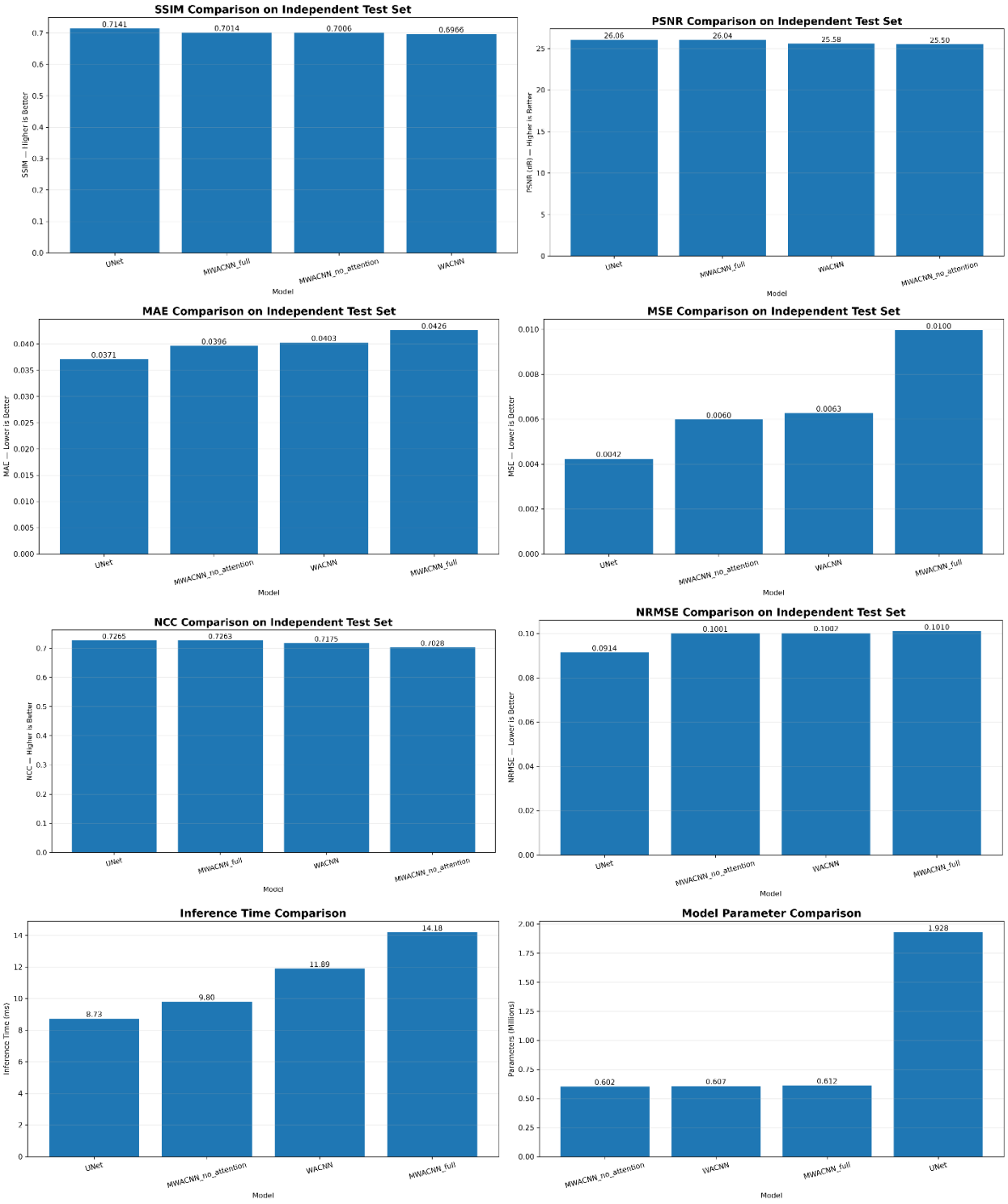
۴- ۵ نتایج کمی

به‌منظور ارزیابی عملکرد معماری‌های مختلف در حذف آرتیفکت‌های حرکتی تصاویرMRI ، چهار مدل شامل U-Net، WACNN، MWACNN (Channel Attention) و MWACNN (No Attention) بر روی مجموعه آزمون مستقل ارزیابی شدند. عملکرد این مدل‌ها با استفاده از معیارهای نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده، ضریب همبستگی نرمال‌شده، تعداد پارامترهای قابل آموزش و زمان استنتاج اندازه‌گیری شد .جدول ۴-۵ نتایج کمی حاصل از ارزیابی مدل‌ها را نشان می‌دهد.

۴-۱۹ ۵- نتایج کمی مدل‌های مورد بررسی بر روی مجموعه آزمون مستقل

Model	PSNR↑	SSIM↑	MAE↓	MSE↓	NRMSE↓	NCC↑	Parameters	Time(ms)↓
U-Net	۲۶.۰۵۷	۰.۷۱۴۱	۰.۰۳۷۰۹	۰.۰۰۴۲۴	۰.۰۹۱۴۴	۰.۷۲۶۵۲	۱,۹۲۷,۸۴۱	۸.۷۳
WACNN	۲۵.۵۸۳	۰.۶۹۶۶	۰.۰۴۰۲۶	۰.۰۰۶۲۶	۰.۱۰۰۲۴	۰.۷۱۷۵۲	۶۰۶,۹۲۴	۱۱.۸۹
MWACNN (Channel Attention)	۲۶.۰۴۲	۰.۷۰۱۴	۰.۰۴۲۶۰	۰.۰۰۹۹۵	۰.۱۰۱۰۰	۰.۷۲۶۲۶	۶۱۱,۶۴۰	۱۴.۱۸
MWACNN (No Attention)	۲۵.۵۰۴	۰.۷۰۰۶	۰.۰۳۹۶۴	۰.۰۰۵۹۸	۰.۱۰۰۰۷	۰.۷۰۲۷۹	۶۰۲,۴۲۴	۹.۸۰

و همچنین شکل ۴-۲ نمودارهای مقایسه ی معیارهای نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده، ضریب همبستگی نرمال‌شده، تعداد پارامترهای قابل آموزش و زمان استنتاج اندازه‌گیری را در هر چهار مدل نمایش می‌دهد.



۲-۴۷ مقایسه کمی‌عملکرد چهار معماری بر اساس معیارهای ارزیابی

۴- ۶ تحلیل نتایج کمی

بررسی نتایج جدول ۴-۵ نشان می‌دهد که معماری U-Net در مجموع بهترین عملکرد را در میان مدل‌های مورد بررسی ارائه کرده است. این مدل با دستیابی به نسبت سیگنال به نویز اوج برابر با ۲۶.۰۵۷ dB و شاخص شباهت ساختاری برابر با ۰.۷۱۴۱، بالاترین کیفیت بازسازی را از نظر معیارهای اصلی ارزیابی به دست آورد. همچنین، کمترین مقادیر میانگین قدر مطلق خطا، میانگین مربعات خطا و ریشه دوم میانگین مربعات خطای نرمال‌شده و بیشترین مقدار ضریب همبستگی نرمال‌شده نیز مربوط به این معماری است که بیانگر توانایی بالای آن در کاهش خطای بازسازی و حفظ ساختارهای تصویری است.

معماری MWACNN مبتنی بر توجه کانالی از نظر مقدار نسبت سیگنال به نویز اوج عملکردی بسیار نزدیک به U-Netداشت و مقدار ۲۶.۰۴۲ dB را به دست آورد. اختلاف نسبت سیگنال به نویز اوج این دو مدل تنها ۰.۰۱۵ dB است که نشان می‌دهد استفاده از تبدیل موجک و مکانیزم توجه کانالی توانسته است کیفیت بازسازی را از نظر این معیار به سطحی نزدیک به U-Net برساند. با این حال، در معیارهای میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و شاخص شباهت ساختاری عملکرد این مدل ضعیف‌تر از U-Net بود. بنابراین، اگرچه این معماری از نظر نسبت سیگنال به نویز اوج نتایج امیدوارکننده‌ای ارائه می‌دهد، اما بر اساس مجموعه معیارهای ارزیابی، همچنان عملکرد کلی آن پایین‌تر از U-Net است.

مدل WACNN درکنار بهره‌گیری از تبدیل موجک، نسبت به U-Net و MWACNN مبتنی بر توجه کانالی مقادیر پایین‌تری برای نسبت سیگنال به نویز اوج و شاخص شباهت ساختاری به دست آورد. همچنین، مقادیر میانگین قدر مطلق خطا، میانگین مربعات خطا و ریشه دوم میانگین مربعات خطای نرمال‌شده در این مدل بیشتر از U-Net بود که نشان می‌دهد دقت بازسازی آن در مقایسه با مدل پایه کمتر است. در پیاده‌سازی‌های مورد بررسی، U-Net دارای ۱,۹۲۷,۸۴۱ پارامتر قابل آموزش است، در حالی که سه معماری موجکی حدود ۶۰۰ هزار پارامتر دارند. بنابراین، مدل‌های موجکی پیاده‌سازی‌شده از نظر تعداد پارامترهای قابل آموزش کوچک‌تر از U-Net هستند.

مقایسه دو نسخه MWACNN نشان می‌دهد که افزودن توجه کانالی موجب افزایش نسبت سیگنال به نویز اوج از ۲۵.۵۰۴ به ۲۶.۰۴۲، شاخص شباهت ساختاری از ۰.۷۰۰۶ به ۰.۷۰۱۴ و ضریب همبستگی نرمال‌شده از ۰.۷۰۲۷۹ به ۰.۷۲۶۲۶ شد. با این حال، مقادیر میانگین قدر مطلق خطا، میانگین مربعات خطا و ریشه دوم میانگین مربعات خطای نرمال‌شده در نسخه دارای توجه افزایش یافتند. بنابراین، تأثیر توجه کانالی در پیاده‌سازی حاضر به معیار ارزیابی وابسته بوده و بهبود یکنواختی در تمامی شاخص‌ها مشاهده نشد.

در چهار پیاده‌سازی مورد بررسی، U-Net با وجود داشتن بیشترین تعداد پارامتر، بهترین عملکرد کلی و کمترین زمان استنتاج را ثبت کرد. بنابراین، در نتایج حاضر رابطه ساده و مستقیمی میان تعداد پارامترهای قابل آموزش و زمان استنتاج مشاهده نمی‌شود. این الگو به‌تنهایی مشخص نمی‌کند کدام تفاوت‌ها از نظر آماری معنادارند و کدام‌یک صرفاً نوسان طبیعی داده‌هاست؛ به همین علت در ادامه، ابتدا هر معیار به‌طور مجزا بررسی می‌شود و سپس در بخش ۴-۸ نتایج با آزمون **ویلکاکسون** ارزیابی می‌گردند.

۴- ۷ تحلیل معیارها

۴- ۷- ۱ تحلیل نسبت سیگنال به نویز اوج

معیار نسبت سیگنال به نویز اوج میزان شباهت تصویر بازسازی‌شده به تصویر مرجع را از نظر شدت پیکسل‌ها نشان می‌دهد و یکی از اصلی ترین شاخص‌های ارزیابی کیفیت بازسازی تصویر محسوب می‌شود. هرچه مقدار این معیار بیشتر باشد، کیفیت بازسازی تصویر بهتر خواهد بود. همان‌گونه که در جدول ۴-۵ مشاهده می‌شود، معماری U-Net با مقدار ۲۶.۰۵۷ dB بهترین عملکرد را به دست آورده است. پس از آن، MWACNN مبتنی بر توجه کانالی با مقدار ۲۶.۰۴۲ dB قرار دارد که اختلاف آن با U-Net تنها ۰.۰۱۵ dB است. این نتیجه نشان می‌دهد که استفاده از تبدیل موجک همراه با مکانیزم توجه کانالی توانسته است از نظر نسبت سیگنال به نویز اوج عملکردی بسیار نزدیک به مدل پایه ارائه دهد.

۴- ۷- ۲ تحلیل شاخص شباهت ساختاری

شاخص شباهت ساختاری میزان شباهت ساختاری میان تصویر بازسازی‌شده و تصویر مرجع را اندازه‌گیری می‌کند و نسبت به معیارهای مبتنی بر خطای پیکسلی، حساسیت بیشتری به حفظ بافت‌ها، مرزها و جزئیات تصویری دارد. نتایج نشان می‌دهد که U-Net با مقدار ۰.۷۱۴۱ بیشترین شباهت ساختاری را ایجاد کرده است. مقدار شاخص شباهت ساختاری برای MWACNN مبتنی بر توجه کانالی برابر با ۰.۷۰۱۴ و برای نسخه بدون توجه ۰.۷۰۰۶ به دست آمد. این نتایج نشان می‌دهد که استفاده از توجه کانالی در معماری پیاده‌سازی‌شده، بهبود اندکی در حفظ ساختار تصویر نسبت به نسخه بدون توجه ایجاد کرده است، هرچند همچنان عملکرد آن پایین‌تر از U-Net باقی مانده است.

۴- ۷- ۳ تحلیل خطاهای بازسازی

معیارهای میانگین قدر مطلق خطا، میانگین مربعات خطا و ریشه دوم میانگین مربعات خطای نرمال‌شده میزان خطای بازسازی تصویر را از دیدگاه‌های مختلف اندازه‌گیری می‌کنند و مقادیر کمتر در این معیارها نشان‌دهنده عملکرد بهتر مدل هستند. همان‌گونه که در جدول ۴-۵ مشاهده می‌شود، U-Net در هر سه معیار کمترین مقدار خطا را به دست آورده است که بیانگر دقت بالاتر این مدل در بازسازی تصاویر MRI است.

در میان معماری‌های مبتنی بر موجک، WACNN و MWACNN بدون توجه مقادیر خطای نزدیک‌تری به U-Net نسبت به نسخه مجهز به توجه کانالی نشان دادند. اگرچه MWACNN مبتنی بر توجه کانالی از نظر نسبت سیگنال به نویز اوج عملکرد مطلوبی داشت، اما مقادیر میانگین قدر مطلق خطا، میانگین مربعات خطا و ریشه دوم میانگین مربعات خطای نرمال‌شده آن نسبت به سایر مدل‌های موجکی بیشتر بود. این موضوع نشان می‌دهد که ارزیابی عملکرد مدل‌ها باید بر اساس مجموعه‌ای از معیارها انجام شود و اتکا به یک معیار منفرد نمی‌تواند تصویر کاملی از کیفیت بازسازی ارائه دهد.

۴- ۷- ۴ تحلیل کارایی محاسباتی

یکی از مهم‌ترین یافته‌های این پژوهش، تفاوت قابل توجه در پیچیدگی محاسباتی معماری‌ها است. در حالی که U-Net دارای ۱,۹۲۷,۸۴۱ پارامتر قابل آموزش است، سه معماری مبتنی بر تبدیل موجک همگی حدود ۶۰۰ هزار پارامتر دارند.

در اندازه‌گیری تکراری روی یک نمونه آزمون، میانگین زمان استنتاج U-Net سریع‌ترین مدل بوده و میانگین زمان پردازش آن ۸.۷۳ میلی‌ثانیه برای هر تصویر است. پس از آن، MWACNN بدون توجه با ۹.۸۰میلی‌ثانیه، WACNN با ۱۱.۸۹میلی‌ثانیه و MWACNN مبتنی بر توجه کانالی با ۱۴.۱۸میلی‌ثانیه قرار دارند. افزایش زمان استنتاج در نسخه مجهز به توجه کانالی را می‌توان به محاسبات اضافی ناشی از مکانیزم توجه نسبت داد. با این حال، این افزایش زمان در مقایسه با کاهش قابل توجه تعداد پارامترها نسبت بهU-Net ، همچنان قابل قبول است.

۴- ۷- ۵ جمع‌بندی نتایج کمی

نتایج حاصل از ارزیابی مدل‌های مورد بررسی نشان می‌دهد که U-Net در مجموع بهترین عملکرد را در بازسازی تصاویر MRI ارائه کرده است. این معماری در اغلب معیارهای ارزیابی، شامل نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده، نتایج بهتری نسبت به سایر مدل‌ها به دست آورد و در عین حال کمترین زمان استنتاج را نیز داشت. این نتایج نشان می‌دهد که ساختار رمزگذار—رمزگشا همراه با اتصالات میانبر همچنان یکی از مؤثرترین معماری‌ها برای مسائل بازسازی تصویر به شمار می‌رود.

در مقابل، معماری‌های مبتنی بر تبدیل موجک با وجود داشتن تعداد پارامترهای بسیار کمتر، عملکرد قابل قبولی ارائه کردند. به‌ویژه MWACNN مبتنی بر توجه کانالی توانست از نظر معیار نسبت سیگنال به نویز اوج به عملکردی بسیار نزدیک بهU-Net دست یابد، هرچند در برخی معیارهای دیگر مانند میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و شاخص شباهت ساختاری همچنان فاصله محسوسی با مدل پایه داشت. این موضوع نشان می‌دهد که استفاده از اطلاعات چندمقیاسی حاصل از تبدیل موجک می‌تواند کیفیت بازسازی را بهبود دهد، اما این بهبود لزوماً در همه معیارهای ارزیابی به‌صورت یکسان مشاهده نمی‌شود.

همچنین، مقایسه دو نسخه MWACNN نشان داد که افزودن مکانیزم توجه کانالی در معماری پیاده‌سازی‌شده این پژوهش، موجب بهبود برخی معیارها مانند نسبت سیگنال به نویز اوج و ضریب همبستگی نرمال‌شده نسبت به نسخه بدون توجه شده است، اما این بهبود در تمامی معیارهای ارزیابی مشاهده نشد. بنابراین، نتایج این پژوهش نشان می‌دهد که اثر مکانیزم توجه به نوع معیار ارزیابی و نحوه طراحی معماری وابسته است و نمی‌توان انتظار داشت که به‌طور هم‌زمان موجب بهبود همه شاخص‌های عملکرد شود.

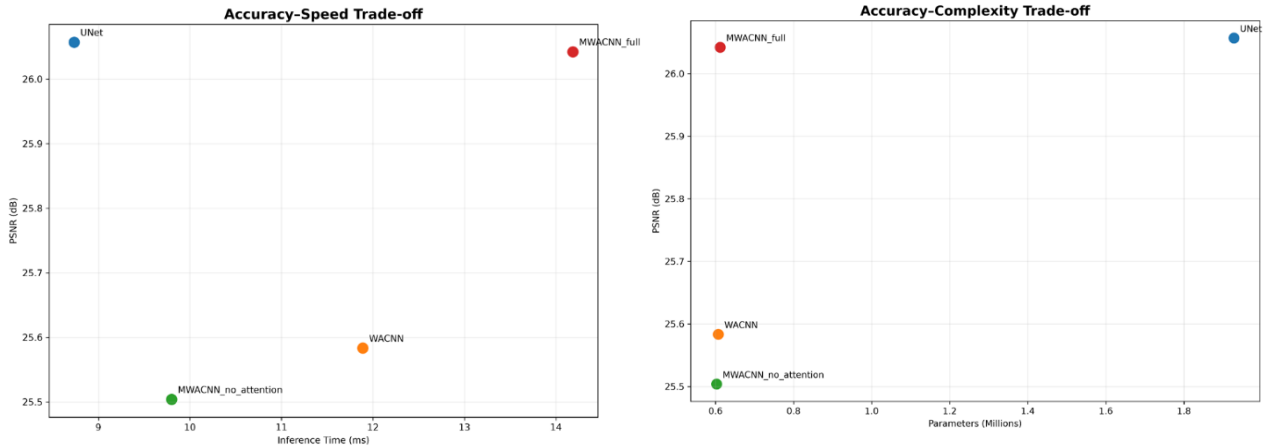
بررسی معیارها به‌صورت مجزا نشان داد که هیچ‌یک از معماری‌های موجکی در همه‌ی شاخص‌ها هم‌زمان به‌U-Net نزدیک نشدند MWACNN :با توجه کانالی در نسبت سیگنال به نویز اوج تقریباً هم‌سطح U-Net بود، اما در میانگین قدر مطلق خطا و میانگین مربعات خطا فاصله داشت؛ در مقابل، MWACNN بدون توجه و WACNN در

معیارهای خطا نزدیک‌تر عمل کردند ولی نسبت سیگنال به نویز اوج پایین‌تری داشتند. این ناهم‌خوانی بین معیارها نشان می‌دهد که رتبه‌بندی «بهترین مدل» به انتخاب معیار ارزیابی وابسته است، نه یک حکم مطلق - نکته‌ای که در بخش بعد با آزمون آماری ویلکاکسون دقیق‌تر بررسی می‌شود

۲۰ ۴-۶ جمع‌بندی مقایسه عملکرد مدل‌های مورد بررسی

معیار	بهترین مدل	دلیل
PSNR	U-Net	بیشترین کیفیت بازسازی از نظر نسبت سیگنال به نویز اوج
SSIM	U-Net	بیشترین شباهت ساختاری به تصویر مرجع
MAE	U-Net	کمترین میانگین قدر مطلق خطای بازسازی
MSE	U-Net	کمترین میانگین مربعات خطا
NRMSE	U-Net	کمترین خطای نرمال‌شده
NCC	U-Net	بیشترین همبستگی با تصویر مرجع
Trainable Parameters	MWACNN (No Attention)	کمترین تعداد پارامترهای قابل آموزش
Inference Time	U-Net	کمترین زمان استنتاج

همان‌گونه که در جدول ۴-۶ مشاهده می‌شود، معماری U-Net در اکثر معیارهای ارزیابی بهترین عملکرد را ارائه کرده است. در مقابل، MWACNN (No Attention) در کنار تعداد پارامترهای کمتر، از نظر پیچیدگی محاسباتی کارآمدتر بود. این نتایج نشان می‌دهد که اگر هدف دستیابی به بالاترین کیفیت بازسازی باشد، U-Net گزینه مناسب‌تری است، در حالی که معماری‌های مبتنی بر موجک می‌توانند با حفظ عملکرد قابل قبول، هزینه محاسباتی کمتری داشته باشند.



۳-۴ ۸ مقایسه توازن میان دقت بازسازی (PSNR) و پیچیدگی محاسباتی مدل‌های مورد بررسی

۴- ۸ تحلیل آماری نتایج با آزمون ویلکاکسون

تحلیل ویلکاکسون در این پژوهش به‌عنوان یک تحلیل آماری تکمیلی و بر اساس خروجی ارزیابی جداگانه نمونه‌های آزمون انجام شده است. مقادیر اصلی عملکرد چهار معماری در جدول ۴-۵ بر اساس خروجی نهایی چهار اجرای اصلی آموزش گزارش شده‌اند؛ ازاین‌رو، نتایج ویلکاکسون برای تفسیر تکمیلی اختلافات زوجی استفاده می‌شوند و جایگزین مقادیر اصلی گزارش‌شده در جدول ۴-۵ نیستند.

به‌منظور بررسی اینکه تفاوت عملکرد مدل‌های مورد بررسی از نظر آماری معنادار است یا خیر، از آزمون رتبه علامت‌دار ویلکاکسون استفاده شد. این آزمون یک روش ناپارامتری[۱] برای مقایسه دو مجموعه داده وابسته است. از آنجا که تمامی مدل‌ها بر روی مجموعه آزمون یکسان ارزیابی شده‌اند، استفاده از این آزمون برای مقایسه عملکرد آن‌ها مناسب است. سطح معنی‌داری آزمون برابر با ۰.۰۵در نظر گرفته شد؛ به‌گونه‌ای که اگر مقدار p-value کمتر از ۰.۰۵ باشد، اختلاف عملکرد دو مدل از نظر آماری معنادار تلقی می‌شود.

نتایج آزمون در جدول ۴-۷ ارائه شده است.

۴-۲۱ نتایج آزمون ویلکاکسون برای مقایسه مدل‌ها

Comparison	Metric	Wilcoxon statistic	p_value
UNet vs WACNN	PSNR	۸۵۱۷۶	۰.۱۱۵۶۰۷
UNet vs WACNN	SSIM	۸۱۸۴۹	۰.۰۱۹۰۲۲
UNet vs WACNN	MAE	۸۳۴۷۴.۵	۰.۰۴۹۰۵۶
UNet vs WACNN	MSE	۸۲۴۰۷	۰.۰۲۶۷۱۳
UNet vs WACNN	NRMSE	۹۰۳۳۱	۰.۷۰۵۴۹۱
UNet vs WACNN	NCC	۸۱۶۱۳	۰.۰۱۶۴۰۳
UNet vs MWACNN_no_att	PSNR	۸۸۱۲۵	۰.۳۷۳۷۱۲
UNet vs MWACNN_no_att	SSIM	۵۱۴۵۴	۵.۷۵E-۲۱
UNet vs MWACNN_no_att	MAE	۸۵۰۴۸	۰.۱۰۸۸۹۷
UNet vs MWACNN_no_att	MSE	۷۷۸۲۰	۰.۰۰۱۰۴
UNet vs MWACNN_no_att	NRMSE	۸۲۸۰۳	۰.۰۳۳۶۸۴
UNet vs MWACNN_no_att	NCC	۸۶۶۴۴	۰.۲۱۷۵۶۸
UNet vs MWACNN_full	PSNR	۹۰۴۲۸	۰.۷۲۲۲۷۳
UNet vs MWACNN_full	SSIM	۶۸۳۸۲	۴.۵۴E-۰۸
UNet vs MWACNN_full	MAE	۸۱۰۴۶	۰.۰۱۱۳۶۳
UNet vs MWACNN_full	MSE	۷۷۵۴۰	۰.۰۰۰۸۲۴
UNet vs MWACNN_full	NRMSE	۸۳۷۳۰	۰.۰۵۶۲۸۲
UNet vs MWACNN_full	NCC	۸۹۲۰۴	۰.۵۲۲۶۳

[۱] Non-Parametric Method

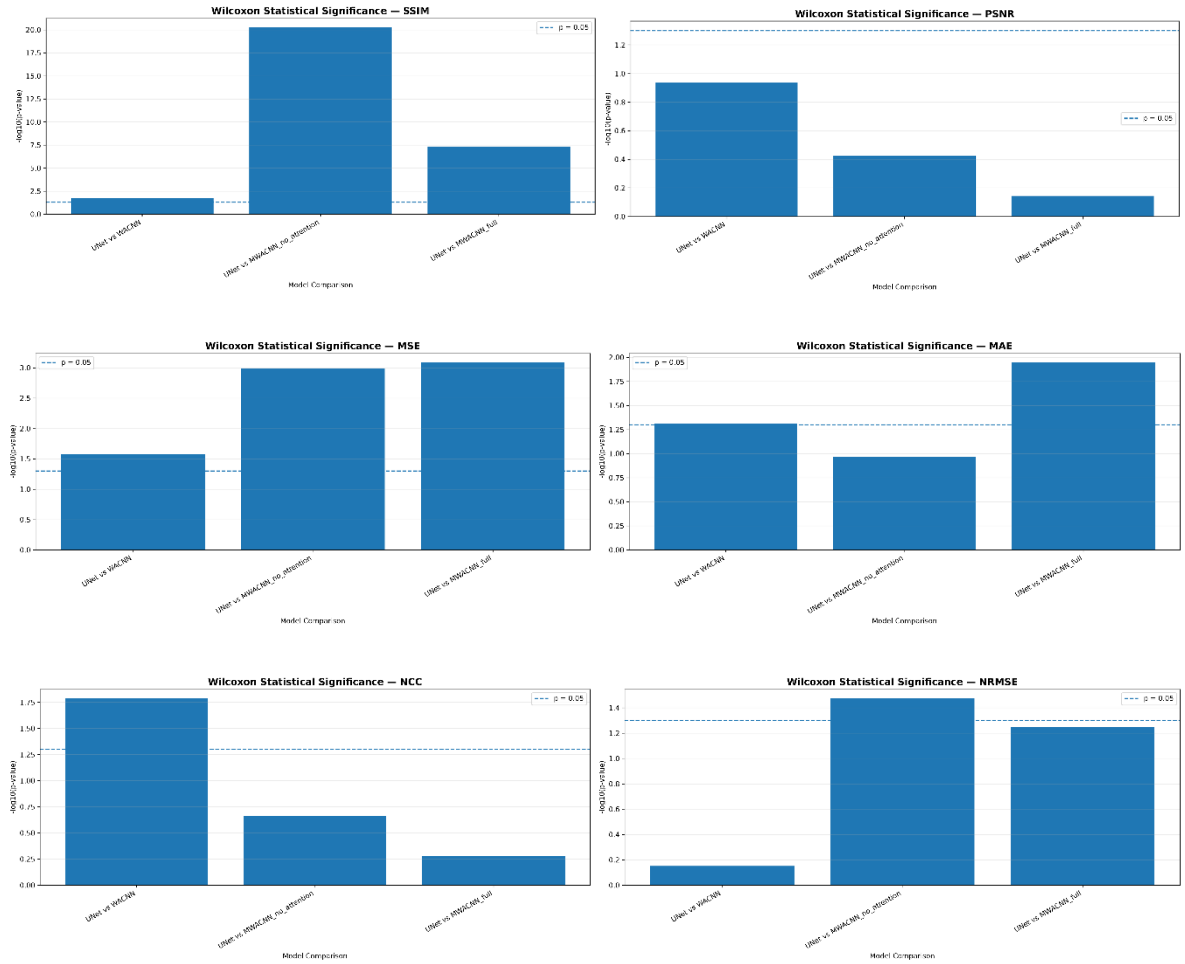
نتایج آزمون ویلکاکسون نشان داد که میزان معناداری اختلاف عملکرد مدل‌ها به معیار ارزیابی بستگی دارد.

در مقایسه U-Net و WACNN، اختلاف در معیارهای شاخص شباهت ساختاری($p = ۰,۰۱۹$) ، میانگین قدر مطلق خطا($p = ۰,۰۴۹$) ، میانگین مربعات خطا ($p = ۰,۰۲۷$) و ضریب همبستگی نرمال‌شده($p = ۰,۰۱۶$) از نظر آماری معنادار بود، در حالی که در معیارهای نسبت سیگنال به نویز اوج ($p = ۰,۱۱۶$) و ریشه دوم میانگین مربعات خطای نرمال‌شده($p = ۰,۷۰۵$) اختلاف معناداری مشاهده نشد.

در مقایسه U-Netو MWACNN بدون توجه، اختلاف در معیارهای ($p < ۰,۰۰۱$) SSIM، میانگین مربعات خطا ($p = ۰,۰۰۱$) و ریشه دوم میانگین مربعات خطای نرمال‌شده($p = ۰,۰۳۴$) معنادار بود، اما در معیارهای نسبت سیگنال به نویز اوج($p = ۰,۳۷۴$) ، میانگین قدر مطلق خطا($p = ۰,۱۰۹$) و ضریب همبستگی نرمال‌شده ($p = ۰,۲۱۸$) تفاوت آماری معناداری مشاهده نشد.

همچنین، در مقایسه U-Netو MWACNN مبتنی بر توجه کانالی، اختلاف در معیارهای شاخص شباهت ساختاری($p < ۰,۰۰۱$) ، میانگین قدر مطلق خطا ($p = ۰,۰۱۱$) و میانگین مربعات خطا($p < ۰,۰۰۱$) معنادار بود، در حالی که در معیارهای نسبت سیگنال به نویز اوج($p = ۰,۷۲۲$) ، ریشه دوم میانگین مربعات خطای نرمال‌شده ($p = ۰,۰۵۶$) و ضریب همبستگی نرمال‌شده($p = ۰,۵۲۳$) تفاوت معناداری مشاهده نشد.

به‌طور کلی، نتایج آزمون ویلکاکسون نشان می‌دهد که اگرچه U-Netدر برخی معیارهای ارزیابی نسبت به سایر معماری‌ها عملکرد بهتری داشته است، اما این برتری در همه شاخص‌ها از نظر آماری معنادار نیست. در این مقایسه‌ها شواهد آماری کافی برای تأیید وجود اختلاف در نسبت سیگنال به نویز اوج مشاهده نشد؛ این نتیجه به‌معنای اثبات برابری عملکرد دو مدل نیست. بر خلاف آن، معیارهای شاخص شباهت ساختاری و میانگین مربعات خطا در اغلب مقایسه‌ها اختلاف معناداری را نشان دادند و حساسیت بیشتری در تشخیص تفاوت عملکرد معماری‌ها داشتند.



۹ ۴-۴ نمودارهای نتایج آزمون ویلکاکسون

۴- ۹ تحلیل نتایج بصری

علاوه بر ارزیابی کمی، کیفیت بازسازی تصاویر به‌صورت بصری نیز مورد بررسی قرار گرفت. برای این منظور، خروجی مدل‌های مختلف با تصاویر دارای آرتیفکت و تصاویر مرجع (Ground Truth) مقایسه شدند تا میزان حذف آرتیفکت‌های حرکتی، حفظ مرزهای بافتی و بازسازی جزئیات ساختاری ارزیابی شود.

بررسی تصاویر بازسازی‌شده نشان می‌دهد که تمامی مدل‌های مورد بررسی توانسته‌اند بخش قابل توجهی از آرتیفکت‌های حرکتی را کاهش دهند و کیفیت تصاویر را نسبت به ورودی دارای آرتیفکت بهبود بخشند. با این حال، تفاوت‌هایی در میزان حفظ جزئیات و ساختارهای تصویری میان معماری‌ها مشاهده می‌شود.

خروجی‌های U-Netدر بیشتر نمونه‌های بررسی‌شده، مرزهای بافتی واضح‌تر، جزئیات ساختاری دقیق‌تر و شباهت بیشتری به تصاویر مرجع داشتند. این مشاهده با نتایج کمی نیز هم‌خوانی دارد، زیرا این مدل در معیارهای نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده بهترین عملکرد را به دست آورده است.

در معماری MWACNN مبتنی بر توجه کانالی نیز کیفیت بازسازی مطلوب بود و در بسیاری از نمونه‌ها ساختار کلی تصویر به‌خوبی حفظ شد. همچنین، در مقایسه با نسخه MWACNN بدون توجه، برخی جزئیات تصویر واضح‌تر به نظر می‌رسید. این مشاهده با نتایج کمی نیز سازگار است؛ به‌طوری که نسخه مجهز به توجه کانالی در معیارهایی مانند نسبت سیگنال به نویز اوج و ضریب همبستگی نرمال‌شده عملکرد بهتری نسبت به نسخه بدون توجه نشان داد. با این حال، اختلاف بصری این دو مدل در همه نمونه‌ها چشمگیر نبود و برتری نسخه مجهز به توجه کانالی در تمام معیارهای ارزیابی مشاهده نشد.

مدل WACNN نیز توانست آرتیفکت‌های حرکتی را تا حد مناسبی کاهش دهد، اما در برخی نمونه‌ها بخشی از جزئیات ریز یا مرزهای بافتی نسبت به U-Net کمتر حفظ شده بودند. این موضوع با مقادیر پایین‌تر نسبت سیگنال به نویز اوج و شاخص شباهت ساختاری این مدل نسبت به U-Net هم‌خوانی دارد.

به‌طور کلی، نتایج بصری و نتایج کمی یکدیگر را تأیید می‌کنند. اگرچه معماری‌های مبتنی بر تبدیل موجک توانسته‌اند با تعداد پارامترهای کمتر عملکرد قابل قبولی ارائه دهند، اما U-Net همچنان بهترین کیفیت بازسازی را از نظر حفظ ساختارهای تصویری و کاهش آرتیفکت‌های حرکتی نشان داد.



۱۰ ۴-۵ پنج نمونه منتخب از خروجی چهار معماری روی مجموعه آزمون مستقل

۴- ۱۰ بحث

نتایج این پژوهش نشان داد که استفاده از ساختارهای مبتنی بر تبدیل موجک و مکانیزم توجه کانالی، در شرایط آموزشی مورد استفاده، لزوماً به برتری کلی نسبت به معماری U-Net منجر نشد .U-Net در هر شش معیار اصلی کیفیت بازسازی شامل نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده بهترین مقدار را ثبت کرد. با این حال، معماری MWACNN مبتنی بر توجه کانالی در برخی معیارها عملکردی بسیار نزدیک به U-Net داشت؛ به‌گونه‌ای که مقدار نسبت سیگنال به نویز اوج برای U-Net برابر با ۲۶.۰۵۷ دسی‌بل و برای MWACNN مبتنی بر توجه کانالی برابر با ۲۶.۰۴۲ دسی‌بل بود و اختلاف میان آن‌ها تنها ۰.۰۱۵ دسی‌بل به دست آمد. مقدار ضریب همبستگی نرمال‌شده این دو مدل نیز بسیار نزدیک بود. این نتایج نشان می‌دهد که ارزیابی عملکرد معماری‌های

بازسازی تصویر باید بر اساس مجموعه‌ای از معیارهای مکمل انجام شود و یک معیار منفرد نمی‌تواند به‌تنهایی بیانگر عملکرد کلی یک مدل باشد.

از نظر اندازه مدل، تفاوت قابل‌توجهی میان U-Net و معماری‌های مبتنی بر موجک مشاهده شد U-Net .دارای ۱,۹۲۷,۸۴۱ پارامتر قابل آموزش بود، در حالی که تعداد پارامترهایWACNN ، MWACNNبدون توجه و MWACNNمبتنی بر توجه کانالی همگی در حدود ۶۰۰ هزار پارامتر قرار داشتند. با وجود این کاهش قابل‌توجه در تعداد پارامترها، زمان استنتاج معماری‌های موجکی الزاماً کمتر نبود و U-Net با زمان استنتاج ۸.۷۳ میلی‌ثانیه سریع‌ترین مدل در میان چهار معماری مورد بررسی بود. بنابراین، تعداد پارامترهای قابل آموزش و زمان استنتاج دو جنبه متفاوت از کارایی مدل هستند و کاهش اندازه مدل لزوماً به کاهش زمان اجرا منجر نمی‌شود.

مقایسه مستقیم دو نسخه MWACNN نشان داد که افزودن مکانیزم توجه کانالی موجب افزایش نسبت سیگنال به نویز اوج از ۲۵.۵۰۴ به ۲۶.۰۴۲ دسی‌بل، شاخص شباهت ساختاری از ۰.۷۰۰۶ به ۰.۷۰۱۴ و ضریب همبستگی نرمال‌شده از ۰.۷۰۲۷۹ به ۰.۷۲۶۲۶ شد. در مقابل، مقادیر میانگین قدر مطلق خطا، میانگین مربعات خطا و ریشه دوم میانگین مربعات خطای نرمال‌شده در نسخه مبتنی بر توجه کانالی افزایش یافتند. بنابراین، مکانیزم توجه کانالی در پیاده‌سازی حاضر بر برخی جنبه‌های کیفیت بازسازی اثر مثبت داشته است، اما این اثر در تمامی معیارهای ارزیابی به‌صورت یکنواخت مشاهده نشد. از آنجا که ساختار اصلی دو نسخه MWACNN مشابه بوده و تفاوت اصلی آن‌ها در فعال یا غیرفعال بودن ماژول‌های توجه کانالی است، مقایسه این دو نسخه امکان بررسی مستقیم‌تر نقش این مکانیزم را در چارچوب آزمایشی پژوهش فراهم می‌کند.

یافته‌های پژوهش حاضر از نظر مفهومی با برخی مطالعات پیشین نیز هم‌راستا است. معماری U-Net بر پایه ساختار رمزگذار–رمزگشا و استفاده از اتصالات میانبر طراحی شده است؛ این اتصالات امکان انتقال ویژگی‌های مراحل رمزگذار به مراحل متناظر رمزگشا و حفظ بهتر اطلاعات مکانی را فراهم می‌کنند [۱۷]. همچنین، پژوهش‌های مبتنی بر تبدیل موجک نشان داده‌اند که ادغام نمایش‌های موجکی با شبکه‌های کانولوشنی می‌تواند امکان استفاده از اطلاعات تصویر در مقیاس‌ها و زیرباندهای مختلف را فراهم کند [۲]، [۳]. نتایج پژوهش حاضر نیز نشان داد که معماری‌های مبتنی بر موجک، با وجود تعداد پارامترهای کمتر، قادر به ارائه نتایج رقابتی در برخی معیارها هستند؛ با این حال، در شرایط آموزشی مورد استفاده در این پژوهش، هیچ‌یک از آن‌ها از نظر عملکرد کلی بر U-Net برتری نیافتند.

به‌طور کلی، نتایج این پژوهش نشان می‌دهد که انتخاب معماری مناسب برای بازسازی تصاویر MRI باید بر اساس مجموعه‌ای از عوامل شامل کیفیت بازسازی، اندازه مدل و زمان استنتاج انجام شود. همچنین، نتایج به مجموعه‌دادهKMAR-۵۰K ، تنظیمات آموزشی مشترک، تعداد Epoch ۲۰ و منابع محاسباتی مورد استفاده در این پژوهش محدود هستند. ازاین‌رو، نمی‌توان علت دقیق تفاوت عملکرد معماری‌ها را تنها به یک جزء خاص مانند تبدیل موجک، اتصالات میانبر یا مکانیزم توجه نسبت داد. بررسی دقیق‌تر نقش هر یک از این مؤلفه‌ها نیازمند آزمایش‌های کنترلی، تنظیم مستقلHyperparameter ها برای هر معماری و ارزیابی در شرایط آزمایشی گسترده‌تر است.

۴- ۱۱ جمع‌بندی فصل

در این فصل، عملکرد چهار معماری مختلف شامل U-Net، WACNN، MWACNNو MWACNN مبتنی بر توجه کانالی برای حذف آرتیفکت‌های حرکتی تصاویر MRI مورد ارزیابی و مقایسه قرار گرفت. نتایج کمی و کیفی نشان داد که U-Netدر مجموع بهترین عملکرد را در میان مدل‌های مورد بررسی ارائه کرده است و در اغلب معیارهای ارزیابی، شامل نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده، نتایج بهتری نسبت به سایر معماری‌ها به دست آورد.

همچنین، نتایج نشان داد که معماری‌های مبتنی بر تبدیل موجک با وجود استفاده از تعداد پارامترهای کمتر، توانستند عملکرد قابل قبولی در بازسازی تصاویر ارائه دهند. به‌ویژه، MWACNN مبتنی بر توجه کانالی در معیار نسبت سیگنال به نویز اوج عملکردی بسیار نزدیک به U-Net داشت، هرچند در برخی معیارهای دیگر همچنان عملکرد مدل پایه بهتر بود. این نتایج نشان می‌دهد که استفاده از تبدیل موجک و توجه کانالی می‌تواند راهکاری مناسب برای طراحی مدل‌های سبک‌تر باشد، اما بهبود عملکرد کلی آن‌ها همچنان نیازمند توسعه بیشتر معماری و تنظیمات آموزشی است.

به‌طور کلی، یافته‌های این فصل نشان می‌دهد که انتخاب معماری مناسب تنها به کیفیت بازسازی وابسته نیست، بلکه عواملی مانند پیچیدگی محاسباتی، تعداد پارامترهای قابل آموزش و سرعت استنتاج نیز باید متناسب با نیاز کاربرد مورد نظر در نظر گرفته شوند. نتایج ارائه‌شده در این فصل مبنای بحث و نتیجه‌گیری نهایی پژوهش را در فصل پنجم فراهم می‌کنند.

۵ فصل پنجم: نتیجه‌گیری و پیشنهادها

۵- ۱ مقدمه

در فصل چهارم، نتایج حاصل از پیاده‌سازی و ارزیابی معماری‌های مختلف برای حذف آرتیفکت‌های حرکتی و بازسازی تصاویر MRI ارائه و تحلیل شد. عملکرد چهار معماری U-Net، WACNN، MWACNN و MWACNN مبتنی بر توجه کانالی با استفاده از مجموعه‌ای از معیارهای کمی و بررسی‌های کیفی مورد مقایسه قرار گرفت و نقاط قوت و محدودیت‌های هر یک مشخص شد.

در این فصل، ابتدا مهم‌ترین یافته‌ها و دستاوردهای پژوهش جمع‌بندی می‌شوند. سپس میزان تحقق اهداف پژوهش و پاسخ به پرسش‌های تحقیق مورد بررسی قرار می‌گیرد. در ادامه، محدودیت‌های پژوهش بیان شده و پیشنهادهایی برای بهبود مدل‌های ارائه‌شده و انجام پژوهش‌های آینده در زمینه حذف آرتیفکت‌های حرکتی و بازسازی تصاویر MRI ارائه خواهد شد.

۵- ۲ نتیجه‌گیری

در عصر حاضر، بازسازی تصاویر پزشکی و حذف آرتیفکت‌های حرکتی به یکی از مهم‌ترین موضوعات پژوهشی در حوزه پردازش تصویر پزشکی و یادگیری عمیق تبدیل شده است. وجود آرتیفکت‌های حرکتی در تصاویر MRI می‌تواند موجب کاهش کیفیت تصویر، از بین رفتن جزئیات ساختاری و کاهش دقت تشخیص و تفسیر پزشکان شود. ازاین‌رو، توسعه روش‌هایی که علاوه بر کاهش آرتیفکت‌ها، توانایی حفظ ساختارهای آناتومیکی و جزئیات تصویری را نیز داشته باشند، از اهمیت ویژه‌ای برخوردار است.

هدف اصلی این پژوهش، طراحی یک چارچوب منسجم برای مقایسه عملکرد چهار معماری مبتنی بر یادگیری عمیق شامل U-Net، WACNN، MWACNN و MWACNN مبتنی بر توجه کانالی در بازسازی تصاویر MRI دارای آرتیفکت حرکتی بود. برای دستیابی به این هدف، تمامی معماری‌ها بر روی مجموعه‌داده استاندارد KMAR-۵۰K و تحت شرایط آموزشی یکسان، شامل مجموعه‌داده، تابع هزینه، بهینه‌ساز، زمان‌بند نرخ یادگیری، معیارهای ارزیابی و محیط اجرا، آموزش و ارزیابی شدند تا امکان مقایسه‌ای منصفانه و قابل استناد میان مدل‌ها فراهم شود.

نتایج حاصل از ارزیابی کمی و کیفی نشان داد که تمامی مدل‌های مورد بررسی توانایی کاهش آرتیفکت‌های حرکتی و بهبود کیفیت تصاویر MRI را دارند، اما میزان موفقیت آن‌ها در حفظ ساختارهای تصویر و کاهش خطای بازسازی متفاوت است. در میان معماری‌های بررسی‌شده، U-Net در مجموع بهترین عملکرد را ارائه داد و در اغلب معیارهای ارزیابی، شامل نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده، نتایج برتری نسبت به سایر مدل‌ها به دست آورد. همچنین این معماری کمترین زمان استنتاج را در میان مدل‌های بررسی‌شده ثبت کرد. این یافته نشان می‌دهد که ساختار رمزگذار—رمزگشا همراه با اتصالات میانبر همچنان یکی از مؤثرترین معماری‌ها برای بازسازی تصاویر پزشکی و حفظ اطلاعات مکانی محسوب می‌شود.

در مقابل، معماری‌های مبتنی بر تبدیل موجک نیز عملکرد قابل قبولی از خود نشان دادند. اگرچه این مدل‌ها در مجموع به عملکرد U-Net نرسیدند، اما با استفاده از حدود یک‌سوم تعداد پارامترهای U-Net توانستند کیفیت

بازسازی مناسبی ارائه دهند. به‌ویژه، MWACNN مبتنی بر توجه کانالی از نظر معیار نسبت سیگنال به نویز اوج عملکردی بسیار نزدیک به U-Net داشت و نشان داد که ترکیب تحلیل چندمقیاسی حاصل از تبدیل موجک با مکانیزم توجه کانالی می‌تواند در بهبود کیفیت بازسازی مؤثر باشد. با این حال، این برتری در تمامی معیارهای ارزیابی مشاهده نشد و در معیارهایی مانند میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و شاخص شباهت ساختاری همچنان عملکرد U-Net بهتر بود. این موضوع نشان می‌دهد که ارزیابی روش‌های بازسازی تصویر باید بر پایه مجموعه‌ای از معیارهای مکمل انجام شود و تکیه بر یک شاخص منفرد نمی‌تواند بیانگر عملکرد واقعی یک مدل باشد.

یکی از یافته‌های مهم این پژوهش آن است که افزایش پیچیدگی معماری لزوماً به بهبود عملکرد منجر نمی‌شود . اگرچه معماری‌های مبتنی بر تبدیل موجک و توجه کانالی از نظر ساختاری پیچیده‌تر از U-Net هستند، اما نتایج نشان داد که این پیچیدگی اضافی در شرایط آزمایش این پژوهش به برتری کامل نسبت به مدل پایه منجر نشده است. در مقابل، کاهش قابل توجه تعداد پارامترها نیز الزاماً باعث افت شدید کیفیت بازسازی نشد. این نتایج بیانگر آن است که طراحی مناسب معماری، کیفیت داده‌های آموزشی، تنظیمات فرآیند آموزش و نحوه ترکیب اجزای مختلف شبکه، تأثیر تعیین‌کننده‌تری نسبت به صرف افزایش پیچیدگی ساختار مدل دارند.

از منظر کاربردی، نتایج این پژوهش نشان می‌دهد که انتخاب معماری مناسب باید متناسب با نیاز کاربرد انجام شود. در کاربردهایی که دستیابی به بالاترین کیفیت بازسازی اولویت اصلی است، U-Net گزینه مناسب‌تری محسوب می‌شود. در مقابل، در کاربردهایی که محدودیت منابع محاسباتی، حافظه یا توان پردازشی وجود دارد، معماری‌های مبتنی بر تبدیل موجک، به‌ویژه MWACNN مبتنی بر توجه کانالی، می‌توانند با حفظ کیفیت قابل قبول و استفاده از تعداد پارامترهای کمتر، گزینه‌های مناسبی باشند.

از دیگر دستاوردهای این پژوهش، ارزیابی هم‌زمان چهار معماری مختلف تحت شرایط آموزشی یکسان و با استفاده از مجموعه‌ای جامع از معیارهای ارزیابی بود. این موضوع امکان مقایسه مستقیم و منصفانه عملکرد مدل‌ها را فراهم کرد و باعث شد تفاوت‌های مشاهده‌شده عمدتاً ناشی از ویژگی‌های معماری‌ها باشد، نه تفاوت در فرآیند آموزش یا شرایط آزمایش. از این رو، نتایج حاصل از این پژوهش می‌تواند مبنای قابل اعتمادی برای مطالعات آینده در زمینه مقایسه معماری‌های بازسازی تصویر باشد.

در کنار نتایج امیدوارکننده، باید توجه داشت که یافته‌های این پژوهش در چارچوب تنظیمات آموزشی و محدودیت‌های محاسباتی موجود به دست آمده‌اند. استفاده از منابع پردازشی قدرتمندتر، افزایش تعداد دوره‌های آموزش، بررسی انواع دیگر تبدیل موجک یا توسعه مکانیزم‌های توجه می‌تواند بر عملکرد نهایی مدل‌ها تأثیرگذار باشد و زمینه را برای دستیابی به نتایج بهتر فراهم کند.

با توجه به سؤالات تحقیق، معماری‌های مبتنی بر موجک در شرایط آموزشی این پژوهش از نظر عملکرد کلی از U-Net فراتر نرفتند؛ با این حال، MWACNN مبتنی بر توجه کانالی در معیارهای **نسبت سیگنال به نویز اوج** و **ضریب همبستگی نرمال‌شده** به U-Net نزدیک شد. از نظر تعداد پارامترهای قابل آموزش، معماری‌های موجکی کوچک‌تر بودند، در حالی که U-Net کمترین زمان استنتاج را ثبت کرد. بنابراین، کارایی هر معماری به معیار مورد نظر وابسته است و نمی‌توان برتری کلی محاسباتی را به یک گروه از مدل‌ها نسبت داد.

۵- ۳ دستاوردهای پژوهش

در این پژوهش، مجموعه‌ای از دستاوردهای علمی و کاربردی در زمینه بازسازی تصاویر MRI و حذف آرتیفکت‌های حرکتی حاصل شد که مهم‌ترین آن‌ها عبارت‌اند از:

- پیاده‌سازی و ارزیابی چهار معماری مبتنی بر یادگیری عمیق شامل U-Net، WACNN، MWACNNو MWACNN مبتنی بر توجه کانالی در یک چارچوب واحد .
- مقایسه مستقیم عملکرد معماری‌های مختلف تحت شرایط آموزشی یکسان که امکان ارزیابی منصفانه و قابل اتکای نتایج را فراهم کرد .
- بررسی نقش تبدیل موجک در استخراج ویژگی‌های چندمقیاسی و تأثیر آن بر بازسازی تصاویر .MRI
- ارزیابی اثر مکانیزم توجه کانالی در معماری MWACNN و مقایسه عملکرد آن با نسخه بدون توجه .
- ارزیابی جامع مدل‌ها با استفاده از مجموعه‌ای از معیارهای کمی شامل نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده، ضریب همبستگی نرمال‌شده، تعداد پارامترهای قابل آموزش و زمان استنتاج .
- استفاده از مجموعه‌داده استاندارد KMAR-۵۰K که امکان مقایسه نتایج این پژوهش با مطالعات مشابه را فراهم می‌کند .
- ارائه تحلیلی جامع از نقاط قوت و محدودیت‌های هر معماری از نظر کیفیت بازسازی، پیچیدگی محاسباتی و کارایی مدل .
- نشان دادن این نکته که افزایش پیچیدگی معماری لزوماً به بهبود عملکرد منجر نمی‌شود و معماری‌های سبک‌تر مبتنی بر تبدیل موجک می‌توانند با تعداد پارامتر کمتر، عملکرد قابل قبولی ارائه دهند .
- فراهم کردن بستری مناسب برای توسعه و بهبود معماری‌های مبتنی بر موجک و یادگیری عمیق در پژوهش‌های آینده .

۵- ۴ محدودیت‌های پژوهش

مهم‌ترین محدودیت این پژوهش، محدودیت منابع محاسباتی در محیط اجرای آزمایش‌ها بود. در اثر محدودیت زمان اجرای جلسات و منابع سخت‌افزاری محیط Kaggle Notebook، امکان آموزش مدل‌ها با تنظیمات ایده‌آل فراهم نبود. در نتیجه، تعداد Epochها به ۲۰، اندازه وصله برابر با ۱۲۸×۱۲۸و تعداد محدودی از برش‌های هر حجم MRI برای آموزش انتخاب شد. این تنظیمات با هدف مدیریت مصرف حافظه، کاهش زمان آموزش و فراهم کردن امکان اجرای کامل تمامی آزمایش‌ها انتخاب شدند.

اگرچه تمامی معماری‌ها با استفاده از تنظیمات آموزشی یکسان آموزش داده شدند و مقایسه‌ای منصفانه میان آن‌ها انجام گرفت، انتظار می‌رود استفاده از منابع پردازشی قدرتمندتر، افزایش تعدادEpoch ها، استفاده از وصله‌های بزرگ‌تر و به‌کارگیری تمامی برش‌های هر حجم MRI بتواند موجب بهبود عملکرد مدل‌ها و افزایش قابلیت تعمیم نتایج شود.

تمامی معماری‌ها باHyperparameter های مشترک آموزش داده شدند. این رویکرد امکان مقایسه کنترل‌شده‌تر را فراهم کرد، اما ممکن است تنظیمات بهینه هر معماری متفاوت باشد. بنابراین، نتایج حاضر لزوماً بیشینه عملکرد قابل دستیابی هر مدل را نشان نمی‌دهند.

زمان استنتاج بر اساس اجرای تکراری یک نمونه در محیط سخت‌افزاری پژوهش اندازه‌گیری شد؛ بنابراین، مقادیر گزارش‌شده باید به‌عنوان یک شاخص مقایسه‌ای در همان شرایط اجرا تفسیر شوند.

در تحلیل آماری ویلکاکسون، واحد تحلیل در سطح برش‌های دوبعدی در نظر گرفته شد. از آنجا که چند برش استخراج‌شده از یک حجم MRI دارای منشأ مشترک هستند، استقلال کامل میان تمامی مشاهدات تضمین نمی‌شود. بنابراین، نتایج آزمون آماری در این پژوهش به‌عنوان تحلیل تکمیلی تفسیر می‌شوند. در مطالعات آینده، انجام تحلیل آماری پس از تجمیع معیارهای برش‌ها در سطح حجم MRI می‌تواند ارزیابی محافظه‌کارانه‌تری ارائه دهد.

از دیگر محدودیت‌های این پژوهش می‌توان به موارد زیر اشاره کرد:

- استفاده از تنها یک مجموعه‌داده استاندارد (KMAR-۵۰K) برای آموزش و ارزیابی مدل‌ها .
- تمرکز پژوهش بر تصاویر MRI زانو و بررسی نشدن سایر اندام‌ها یا پروتکل‌های تصویربرداری .
- بررسی تنها آرتیفکت‌های حرکتی و ارزیابی نشدن سایر انواع آرتیفکت‌های رایج در تصاویرMRI .
- استفاده از معماری‌های دوبعدی (۲D) و بررسی نشدن مدل‌های سه‌بعدی که می‌توانند از اطلاعات بین برش‌های متوالی نیز بهره ببرند .
- محدود بودن ارزیابی به معیارهای کمی و تحلیل بصری و استفاده نشدن از ارزیابی تخصصی توسط رادیولوژیست‌ها یا متخصصان تصویربرداری پزشکی .
- بررسی نشدن میزان تعمیم‌پذیری مدل‌ها بر روی مجموعه‌داده‌های مستقل یا تصاویر تهیه‌شده با دستگاه‌ها و مراکز درمانی متفاوت.

۵- ۵ مقایسه نتایج پژوهش حاضر با پژوهش‌های پیشین

به‌منظور بررسی جایگاه نتایج این پژوهش در میان مطالعات پیشین، عملکرد مدل‌های مورد بررسی با مهم‌ترین پژوهش‌های مرتبط در حوزه بازسازی تصاویر پزشکی، شبکه‌های کانولوشنی و معماری‌های مبتنی بر تبدیل موجک مقایسه شد .جدول ۵-۱ خلاصه‌ای از این مقایسه را نشان می‌دهد.

پژوهش	معماری / روش	ویژگی اصلی	محدودیت	ارتباط با پژوهش حاضر
Ronneberger et al. (۲۰۱۵) [۱۷]	U-Net	استفاده از ساختار Encoder–Decoder و Skip Connection	عدم استفاده از تحلیل چندمقیاسی مبتنی بر Wavelet	مدل پایه و مبنای اصلی مقایسه
Liu et al. (۲۰۱۸) [۳]	MWCNN	ترکیب شبکه کانولوشنی با تبدیل موجک	عدم استفاده از مکانیزم توجه	مبنای توسعه معماری‌های موجکی
Sriprabha et al. (۲۰۲۰) [۲]	DC-WCNN	استفاده از Wavelet در بازسازی تصاویر MRI	ارزیابی محدود و عدم مقایسه با چند معماری	مرجع مهم در بازسازی MRI
Xi et al. (۲۰۲۵) [۲۸]	U-Net، EDSR، DDPM	ارائه مجموعه‌داده استاندارد KMAR-k۵۰	عدم بررسی معماری‌های مبتنی بر Wavelet	منبع مجموعه‌داده و مرجع اصلی ارزیابی
پژوهش حاضر	U-Net، WACNN، MWACNN، MWACNN + Channel Attention	مقایسه چهار معماری تحت شرایط آموزشی یکسان و ارزیابی جامع با معیارهای کمی و محاسباتی	محدودیت منابع پردازشی و تعداد Epoch ها	ارائه مقایسه مستقیم میان معماری‌های مبتنی بر CNN و Wavelet

جایگاه U-Net به‌عنوان مدل مرجع در این پژوهش با یافته‌های Ronneberger و همکاران [۱۷] هم‌راستاست، اما آنچه پژوهش حاضر اضافه می‌کند مقایسه‌ی مستقیم این معماری با نسخه‌های موجکی تحت دقیقاً یک مجموعه‌داده و یک تنظیم آموزشی است - چیزی که در مطالعات Sriprabha و همکاران [۲] یا Liu و همکاران [۳]، به دلیل تفاوت در مجموعه‌داده و شرایط آزمایش، به‌صورت مستقیم قابل استنتاج نبود.

بر خلاف آن، پژوهش‌های مبتنی بر تبدیل موجک نشان داده‌اند که استفاده از تحلیل چندمقیاسی می‌تواند به استخراج بهتر ویژگی‌های تصویری و کاهش پیچیدگی مدل کمک کند. نتایج این پژوهش نیز نشان داد که معماری‌های

WACNNو MWACNNبا وجود استفاده از تعداد پارامترهای بسیار کمتر نسبت بهU-Net ، عملکرد قابل قبولی در بازسازی تصاویر MRI ارائه می‌کنند.

همچنین، مقایسه دو نسخه MWACNN نشان داد که افزودن توجه کانالی موجب بهبود برخی معیارهای ارزیابی، به‌ویژه نسبت سیگنال به نویز اوج و ضریب همبستگی نرمال‌شده، نسبت به نسخه بدون توجه شده است؛ هرچند این بهبود در تمامی معیارهای ارزیابی مشاهده نشد. این یافته نشان می‌دهد که تأثیر مکانیزم توجه به نحوه طراحی معماری و تنظیمات آموزش وابسته است و استفاده از آن به‌تنهایی تضمین‌کننده بهبود همه شاخص‌های عملکرد نیست.

در مقایسه با مطالعه Xi و همکاران [۲۸]، پژوهش حاضر نیز از مجموعه‌داده استاندارد KMAR-۵۰K استفاده کرده است، اما تفاوت اصلی آن در مقایسه هم‌زمان چهار معماری مختلف تحت شرایط آموزشی یکسان و استفاده از مجموعه‌ای جامع از معیارهای ارزیابی است. این رویکرد امکان تحلیل دقیق‌تر نقاط قوت و محدودیت‌های هر معماری را فراهم کرده و مقایسه‌ای منصفانه میان مدل‌های مبتنی بر CNN و Wavelet ارائه داده است.

جدول ۵-۱ نشان می‌دهد که هیچ‌یک از مطالعات پیشین هم‌زمان چهار معماری را در چارچوب یکسانی از تنظیمات آموزشی مقایسه نکرده بودند؛ نقطه‌ی تمایز اصلی این پژوهش دقیقاً همین است، نه معرفی یک معماری جدید. این رویکرد مقایسه‌ای می‌تواند الگویی برای پژوهش‌های آینده باشد که به‌جای مقایسه‌ی نتایج گزارش‌شده در مقالات مختلف (که با مجموعه‌داده‌ها و معیارهای متفاوت به دست آمده‌اند)، ارزیابی هم‌زمان و منصفانه را در دستور کار قرار دهند.

۵- ۶ پیشنهادهایی برای پژوهش‌های آینده

با توجه به نتایج به‌دست‌آمده و محدودیت‌های این پژوهش، پیشنهادهای زیر برای توسعه مطالعات آینده در زمینه بازسازی تصاویر MRI و حذف آرتیفکت‌های حرکتی ارائه می‌شود:

۱. آموزش مدل‌ها با تنظیمات آموزشی کامل‌تر، شامل افزایش تعداد Epochها، استفاده از وصله‌های بزرگ‌تر، به‌کارگیری تمامی برش‌های هر حجم MRI و اجرای آزمایش‌ها بر روی سرورهای مجهز به GPUهای قدرتمند، تا تأثیر این عوامل بر کیفیت بازسازی و قابلیت تعمیم مدل‌ها به‌صورت دقیق بررسی شود .

۲. ارزیابی مدل‌ها بر روی چندین مجموعه‌داده استاندارد و همچنین تصاویر واقعی مراکز درمانی مختلف، به‌منظور بررسی میزان تعمیم‌پذیری و پایداری عملکرد معماری‌ها در شرایط متفاوت تصویربرداری .

۳. توسعه و ارزیابی معماری‌های سه‌بعدی (۳D CNN) یا مدل‌های مبتنی بر حجم کامل تصاویرMRI ، به‌گونه‌ای که اطلاعات بین برش‌های متوالی نیز در فرآیند بازسازی مورد استفاده قرار گیرد .

٤. بررسی استفاده از مدل‌های مبتنی بر Diffusion در حذف آرتیفکت‌های حرکتی و مقایسه عملکرد آن‌ها با معماری‌های کانولوشنی و مبتنی بر موجک .

۵. مطالعه و پیاده‌سازی روش‌های یادگیری خودنظارتی[۱] و یادگیری نیمه‌نظارتی[۲] به‌منظور کاهش وابستگی مدل‌ها به داده‌های برچسب‌گذاری‌شده و افزایش قابلیت استفاده از داده‌های بدون برچسب .
۶. بررسی ترکیب تبدیل موجک با معماری‌های جدید مبتنی بر Transformerیا سایر مدل‌های هیبریدی، به‌منظور بهره‌گیری هم‌زمان از تحلیل چندمقیاسی و توانایی مدل‌های مبتنی بر مکانیزم توجه .همچنین مقایسه عملکرد مدل‌های پیشنهادی با معماری‌های مبتنی بر Transformer مانند SwinIR می‌تواند مسیر مناسبی برای پژوهش‌های آینده باشد. [۲۶] همچنین بررسی معماری‌های مبتنی بر Residual Dense Network (RDN) نیز می‌تواند در توسعه مدل‌های سبک و دقیق برای بازسازی تصاویر MRI مؤثر باشد. [۲۷]
۷. طراحی و ارزیابی مکانیزم‌های توجه پیشرفته‌تر، مانند Spatial Attention، Dual Attention یا ترکیب توجه کانالی و مکانی، و مقایسه آن‌ها با مکانیزم توجه کانالی استفاده‌شده در این پژوهش .
۸. استفاده از توابع هزینه ترکیبی یا معیارهای ادراکی[۳] در کنار معیارهای پیکسلی و ساختاری، به‌منظور بهبود هم‌زمان کیفیت عددی و کیفیت بصری تصاویر بازسازی‌شده .
۹. انجام ارزیابی بالینی نتایج با مشارکت متخصصان رادیولوژی و مقایسه نتایج کمی با ارزیابی‌های تخصصی، به‌منظور بررسی کاربردپذیری مدل‌ها در محیط‌های واقعی درمانی .
۱۰. بررسی امکان بهینه‌سازی معماری‌های پیشنهادی برای کاربردهای بلادرنگ[۴] و استقرار آن‌ها در سامانه‌های کمک‌تشخیص پزشکی، به‌ویژه در شرایطی که محدودیت منابع محاسباتی وجود دارد .

۵- ۷ جمع‌بندی نهایی

پژوهش حاضر با هدف ارزیابی و مقایسه معماری‌های مختلف یادگیری عمیق برای حذف آرتیفکت‌های حرکتی و بازسازی تصاویر MRI انجام شد. برای این منظور، چهار معماری شامل U-Net، WACNN، MWACNN و MWACNN مبتنی بر توجه کانالی بر روی مجموعه‌داده استاندارد KMAR-۵۰K پیاده‌سازی، آموزش و ارزیابی شدند.

نتایج این پژوهش نشان داد که استفاده از روش‌های مبتنی بر یادگیری عمیق می‌تواند کیفیت تصاویرMRI را بهبود داده و آرتیفکت‌های حرکتی را کاهش دهد. در میان مدل‌های مورد بررسی، U-Net بهترین عملکرد کلی را از نظر معیارهای نسبت سیگنال به نویز اوج، شاخص شباهت ساختاری، میانگین قدر مطلق خطا، میانگین مربعات خطا، ریشه دوم میانگین مربعات خطای نرمال‌شده و ضریب همبستگی نرمال‌شده ارائه داد. در مقابل ، معماری‌های مبتنی بر تبدیل موجک علاوه بر تعداد پارامترهای کمتر، توانستند عملکرد قابل قبولی ارائه دهند. به‌ویژه، MWACNN

[۱] Self-Supervised Learning (SSL)
[۲] Semi-Supervised Learning
[۳] Perceptual Loss
[۴] Real-Time

مبتنی بر توجه کانالی از نظر نسبت سیگنال به نویز اوج عملکردی بسیار نزدیک به U-Net داشت، هرچند این برتری در تمامی معیارهای ارزیابی مشاهده نشد.

یکی از مهم‌ترین یافته‌های این پژوهش آن است که افزایش پیچیدگی معماری، لزوماً به بهبود عملکرد منجر نمی‌شود و در مقابل، کاهش تعداد پارامترهای شبکه نیز الزاماً باعث افت شدید کیفیت بازسازی نمی‌شود. این نتیجه نشان می‌دهد که طراحی مناسب معماری و انتخاب صحیح اجزای شبکه، نقش مهم‌تری نسبت به صرف افزایش پیچیدگی مدل در دستیابی به عملکرد مطلوب دارند.

در مجموع، نتایج این پژوهش نشان می‌دهد که در انتخاب معماری برای کاربردهای بالینی، «بهترین مدل» یک مفهوم مطلق نیست؛ بسته به اینکه اولویت با حداکثر کیفیت تشخیصی باشد یا با اجرا روی سخت‌افزار محدود، پاسخ درست می‌تواند U-Net یا یکی از نسخه‌های MWACNN باشد. این چارچوب انتخاب، فارغ از هر پیشرفت معماری خاص، می‌تواند راهنمای تصمیم‌گیری در پژوهش‌های آتی این حوزه باشد.

۵- ۸ پیشنهادهای کاربردی

یافته‌های این پژوهش می‌تواند در حوزه‌های مختلف پردازش تصاویر پزشکی و سامانه‌های کمک‌تشخیص مورد استفاده قرار گیرد. مهم‌ترین کاربردهای بالقوه نتایج این تحقیق عبارت‌اند از:

- استفاده از مدل‌های بازسازی تصویر به‌عنوان مرحله پیش‌پردازش در سامانه‌های کمک‌تشخیص مبتنی بر هوش مصنوعی، به‌منظور کاهش اثر آرتیفکت‌های حرکتی و بهبود کیفیت تصاویر ورودی .
- به‌کارگیری مدل‌های بازسازی برای افزایش کیفیت تصاویر MRI پیش از بررسی توسط رادیولوژیست‌ها، با هدف کمک به مشاهده دقیق‌تر ساختارهای آناتومیکی و کاهش احتمال خطا در تفسیر تصاویر .
- توسعه نسخه‌های سبک‌تر از معماری‌های مبتنی بر تبدیل موجک برای استفاده در تجهیزات پزشکی یا سامانه‌هایی با محدودیت منابع محاسباتی، بدون کاهش قابل توجه کیفیت بازسازی .
- استفاده از نتایج این پژوهش به‌عنوان مبنایی برای توسعه سامانه‌های هوشمند پردازش تصاویر پزشکی و یکپارچه‌سازی آن‌ها با سایر ابزارهای مبتنی بر یادگیری عمیق در حوزه تشخیص و تحلیل تصاویرMRI .
- بهره‌گیری از معماری‌های بررسی‌شده در پژوهش‌های آینده برای طراحی مدل‌های هیبریدی که بتوانند تعادل بهتری میان کیفیت بازسازی، سرعت پردازش و پیچیدگی محاسباتی برقرار کنند.

منابع:

[۱] C. Tian, Y. Xu, Z. Li, W. Zuo, L. Fei, and H. Liu, "Attention-guided CNN for image denoising," *Neural Networks*, vol. ۱۲٤, pp. ۱۱۷–۱۲۹, ۲۰۲۰.

[۲] R. Sriprabha, M. Balamurali, R. Keerthi, and M. Sivaprakasam, "DC-WCNN: A deep cascade of wavelet-based CNNs for MR image reconstruction," *arXiv preprint arXiv:2001.02397*, ۲۰۲۰.

[۳] P. Liu, H. Zhang, K. Zhang, L. Lin, and W. Zuo, "Multi-level wavelet-CNN for image restoration," in *CVPR Workshops*, ۲۰۱۸.

[٤] X. Zhao, P. Huang, and X. Shu, "Wavelet-Attention CNN for image classification," *Multimedia Systems*, vol. ۲۸, pp. ۹۱٥–۹۲٤, ۲۰۲۲.

[٥] H. Li and C. Y. Suen, "A novel non-local means image denoising method based on grey theory," *Pattern Recognition*, vol. ٤۹, pp. ۲۳۷–۲٤۸, ۲۰۱٦.

[٦] Y. Li, Y. Chen, X. Dai, M. Liu, D. Chen, Y. Yu, L. Yuan, Z. Liu, M. Chen, and N. Vasconcelos, "Revisiting dynamic convolution via matrix decomposition," *arXiv preprint arXiv:2103.08756*, ۲۰۲۱.

[۷] Y. Liu, "Image denoising method based on threshold, wavelet transform and genetic algorithm," *IJSP, Image Processing, Pattern Recognition*, vol. ۸, pp. ۲۹–٤۰, ۲۰۱٥.

[۸] S. Nam, Y. Hwang, Y. Matsushita, and S. J. Kim, "A holistic approach to cross-channel image noise modeling and its application to image denoising," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. ۱٦۸۳–۱٦۹۱, ۲۰۱٦.

[۹] Y. Quan, Y. Chen, Y. Shao, H. Teng, Y. Xu, and H. Ji, "Image denoising using complex-valued deep CNN," *Pattern Recognition*, vol. ۱۱۱, Art. no. ۱۰۷٦۳۹, ۲۰۲۱.

[۱۰] H. Rabbani, "Image denoising in steerable pyramid domain based on a local Laplace prior," *Pattern Recognition*, vol. ٤۲, pp. ۲۱۸۱–۲۱۹۳, ۲۰۰۹.

[۱۱] J. Hu, L. Shen, and G. Sun, "Squeeze-and-Excitation Networks," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. ۷۱۳۲–۷۱٤۱, ۲۰۱۸.

[۱۲] X. Sun, C. Chen, X. Wang, J. Dong, H. Zhou, and S. Chen, "Gaussian dynamic convolution for efficient single-image segmentation," *IEEE Transactions on Circuits and Systems for Video Technology*, ۲۰۲۱.

[۱۳] C. Tian, L. Fei, W. Zheng, Y. Xu, W. Zuo, and C. W. Lin, "Deep learning on image denoising: An overview," *Neural Networks*, vol. ۱۳۱, pp. ۲٥۱–۲۷٥, ۲۰۲۰.

[۱٤] K. Zhang, W. Zuo, and L. Zhang, "FFDNet: Toward a fast and flexible solution for CNN-based image denoising," *IEEE Transactions on Image Processing*, vol. ۲۷, pp. ٤٦۰۸–٤٦۲۲, ۲۰۱۸.

[۱٥] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. ۱۳, pp. ٦۰۰–٦۱۲, ۲۰۰٤.

[۱٦] S. Wang, Z. Su, L. Ying, X. Peng, S. Zhu, F. Liang, D. Feng, and D. Liang, "Accelerating magnetic resonance imaging via deep learning," in *IEEE International Symposium on Biomedical Imaging (ISBI)*, ۲۰۱٦.

[١٧] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, ٢٠١٥.

[١٨] D. Lee, J. Yoo, and J. C. Ye, "Deep residual learning for compressed sensing MRI," in *IEEE International Symposium on Biomedical Imaging (ISBI)*, ٢٠١٧.

[١٩] C. M. Hyun, H. P. Kim, S. M. Lee, S. Lee, and J. K. Seo, "Deep learning for undersampled MRI reconstruction," *Physics in Medicine & Biology*, ٢٠١٨.

[٢٠] D. Cho and T. D. Bui, "Multivariate statistical modeling for image denoising using wavelet transforms," *Signal Processing: Image Communication*, ٢٠٠٥.

[٢١] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse ٣-D transform-domain collaborative filtering," *IEEE Transactions on Image Processing*, ٢٠٠٧.

[٢٢] J. Ye, Y. Han, and E. Cha, "Deep convolutional framelets: A general deep learning framework for inverse problems," *SIAM Journal on Imaging Sciences*, ٢٠١٨.

[٢٣] P. Wang, P. Chen, Y. Yuan, D. Liu, Z. Huang, X. Hou, and G. Cottrell, "Understanding convolution for semantic segmentation," in *IEEE Winter Conference on Applications of Computer Vision (WACV)*, ٢٠١٨.

[٢٤] T. Guo, H. S. Mousavi, T. H. Vu, and V. Monga, "Deep wavelet prediction for image super-resolution," in *CVPR Workshops*, ٢٠١٧.

[٢٥] P. Li, J. Liang, M. Zhang, W. Fan, and G. Yu, "Joint image denoising with gradient direction and edge-preserving regularization," *Pattern Recognition*, ٢٠٢٢.

[٢٦] J. Liang, J. Cao, G. Sun, K. Zhang, L. Van Gool, and R. Timofte, "SwinIR: Image restoration using Swin Transformer," in *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, ٢٠٢١.

[٢٧] Y. Zhang, Y. Tian, Y. Kong, B. Zhong, and Y. Fu, "Residual dense network for image super-resolution," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, ٢٠١٨.

[٢٨] Y. Xi, F. Wang, F. Shi, Q. Zhou, R. Li, Y. Li, and Y. Wang, "A Multi-view Open-access Dataset of Paired Knee MRI for Motion Artifact Removal," *Scientific Data*, vol. ١٢, Art. no. ١١٧٣, ٢٠٢٥, doi: ١٠,١٠٣٨/s٤١٥٩٧-٠٢٥-٠٥٤٣٩-١.

[٢٩] F. Wilcoxon, "Individual Comparisons by Ranking Methods," *Biometrics Bulletin*, vol. ١, no. ٦, pp. ٨٠–٨٣, Dec. ١٩٤٥.

[٣٠] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, ٤th ed. Pearson, ٢٠١٨.

پیوست:

نمونه ای از کدها :

```python
SEED = 42

# Fix random seeds for reproducibility
random.seed(SEED)
np.random.seed(SEED)
torch.manual_seed(SEED)

# Improve GPU performance for fixed-size input images
torch.backends.cudnn.benchmark = True

# Select GPU if available, otherwise use CPU
device = torch.device("cuda" if torch.cuda.is_available() else "cpu")

print("Device:", device)
print("GPU count:", torch.cuda.device_count())

# Folder where results, models, plots, and JSON files will be saved
RESULT_DIR = "/kaggle/working/results"
os.makedirs(RESULT_DIR, exist_ok=True)

#Final publication settings
PATCH = 128 # Best 256
EPOCHS = 20 # Best 100
BATCH_SIZE = 8
LR = 1e-4

N_PATCHES_PER_SLICE = 1
MAX_SLICES_PER_VOLUME = 6 # Best None

NUM_WORKERS = 2
PREFETCH_FACTOR = 2
CACHE_SIZE = 16
ACCUM_STEPS = 2
```

۱۱ مقدار دهی اولیه کد

```python
def resize_match(noisy_slice, clean_slice):
    # If noisy and clean slices have the same shape, return directly
    if noisy_slice.shape == clean_slice.shape:
        return noisy_slice, clean_slice

    # Resize clean slice to match noisy slice size
    h, w = noisy_slice.shape
    clean_img = PILImage.fromarray((clean_slice * 255).astype(np.uint8))
    clean_img = clean_img.resize((w, h), PILImage.BILINEAR)
    clean_slice = np.array(clean_img).astype(np.float32) / 255.0

    return noisy_slice, clean_slice


def center_crop(img, patch_size):
    # Pad image if it is smaller than patch size
    h, w = img.shape

    if h < patch_size:
        img = np.pad(img, ((0, patch_size - h), (0, 0)), mode="reflect")
    if w < patch_size:
        img = np.pad(img, ((0, 0), (0, patch_size - w)), mode="reflect")

    # Crop from the center
    h, w = img.shape
    top = (h - patch_size) // 2
    left = (w - patch_size) // 2

    return img[top:top + patch_size, left:left + patch_size]
```

۱۲ تغییر سایز عکس‌ها و انتخاب برش در کد

```python
class WaveletTransform(nn.Module):
    def __init__(self):
        super().__init__()

        # Haar-like wavelet filters
        ll = torch.tensor([[1, 1], [1, 1]], dtype=torch.float32) / 4
        lh = torch.tensor([[-1, -1], [1, 1]], dtype=torch.float32) / 4
        hl = torch.tensor([[-1, 1], [-1, 1]], dtype=torch.float32) / 4
        hh = torch.tensor([[1, -1], [-1, 1]], dtype=torch.float32) / 4

        # Register filters as buffers so they move with model.to(device)
        self.register_buffer("ll", ll[None, None])
        self.register_buffer("lh", lh[None, None])
        self.register_buffer("hl", hl[None, None])
        self.register_buffer("hh", hh[None, None])

    def forward(self, x):
        # Input shape: B, C, H, W
        B, C, H, W = x.shape

        # Apply wavelet filters independently for each channel
        x = x.reshape(B * C, 1, H, W)

        out = torch.cat([
            F.conv2d(x, self.ll, stride=2),
            F.conv2d(x, self.lh, stride=2),
            F.conv2d(x, self.hl, stride=2),
            F.conv2d(x, self.hh, stride=2)
        ], dim=1)

        # Output shape: B, 4C, H/2, W/2
        return out.reshape(B, C * 4, H // 2, W // 2)
```

۱۳ کد تبدیل موجکی

```python
class ChannelAttention(nn.Module):
    def __init__(self, channels, reduction=8):
        super().__init__()

        hidden = max(channels // reduction, 4)

        # Global average pooling extracts channel-level statistics
        self.pool = nn.AdaptiveAvgPool2d(1)

        # Fully connected layers learn channel importance weights
        self.fc = nn.Sequential(
            nn.Linear(channels, hidden, bias=False),
            nn.ReLU(True),
            nn.Linear(hidden, channels, bias=False),
            nn.Sigmoid()
        )

    def forward(self, x):
        B, C, _, _ = x.shape

        weights = self.pool(x).view(B, C)
        weights = self.fc(weights).view(B, C, 1, 1)

        return x * weights
```

۱۴ کد مکانیزم توجه

Abstract

Magnetic Resonance Imaging (MRI) is one of the most important medical imaging modalities for visualizing internal body structures. However, patient motion during image acquisition may introduce motion artifacts that degrade image quality and adversely affect clinical diagnosis. In recent years, deep learning-based approaches have become effective solutions for MRI image reconstruction and motion artifact removal.

In this study, four deep learning architectures, namely U-Net, WACNN, MWACNN without attention, and MWACNN with Channel Attention, were implemented and comparatively evaluated for MRI motion artifact removal. All models were trained and tested on the KMAR-٥٠K benchmark dataset under identical training conditions. Model performance was evaluated using PSNR, SSIM, MAE, MSE, NRMSE, NCC, the number of trainable parameters, and inference time.

The experimental results demonstrated that U-Net achieved the best overall performance, obtaining a PSNR of ٢٦,٠٥٧ dB and an SSIM of ٠,٧١٤١, while also producing the lowest reconstruction errors and the fastest inference time among the evaluated models. In contrast, MWACNN with Channel Attention achieved results close to U-Net in selected metrics, particularly PSNR and NCC, while requiring approximately one-third of U-Net's trainable parameters. However, U-Net achieved better results in SSIM and the reconstruction error metrics. Overall, the findings indicate that deep learning-based architectures, particularly U-Net, can effectively reduce motion artifacts and improve the quality of MRI images. In addition, a supplementary Wilcoxon Signed-Rank analysis was conducted to examine paired differences between model evaluation results.

Keywords: Magnetic Resonance Imaging (MRI), Motion Artifact Removal, Image Reconstruction, Deep Learning, U-Net, Wavelet Transform, Channel Attention, KMAR-٥٠K.

Islamic Azad University

Science and Research Branch

Faculty of Convergent Sciences and Technologies, Department of Computer Science

Major:
Soft Computing and Artificial Intelligence

Title:
MRI Image Enhancement Using the MWACNN Method and Comparison with U-Net

Supervisor:
Dr. Mehdi Azhini

Advisor:
Dr. Reza Kargar

By:
Seyed Hossein Taheri

Summer 2026

MRI Image Enhancement Using the MWACNN Method and Comparison with U-Net