



دانشگاه صنعتی امیرکبیر

(پلی‌تکنیک تهران)

دانشکده مهندسی کامپیوتر

گزارش پروژه کارشناسی

تشخیص بیماری‌های قلبی با استفاده از تصاویر  
سینماتیک ام‌آرآی ۴ بعدی از طریق یادگیری عمیق

نگارش

سودا ایمانی

استاد راهنما

دکتر مریم امیرمزلقانی

خرداد ۱۴۰۳

بِسْمِ اللّٰهِ الرَّحْمٰنِ الرَّحِيْمِ

# صفحه فرم ارزیابی و تصویب پایان نامه- فرم تأیید اعضاء کمیته دفاع

در این صفحه فرم دفاع یا تایید و تصویب پایان نامه موسوم به فرم کمیته دفاع- موجود در پرونده آموزشی- را قرار دهید.

## نکات مهم:

- نگارش پایان نامه/رساله باید به **زبان فارسی** و بر اساس آخرین نسخه دستورالعمل و راهنمای تدوین پایان نامه های دانشگاه صنعتی امیرکبیر باشد.(دستورالعمل و راهنمای حاضر)
- رنگ جلد پایان نامه/رساله چاپی کارشناسی، کارشناسی ارشد و دکترا باید به ترتیب مشکی، طوسی و سفید رنگ باشد.
- چاپ و صحافی پایان نامه/رساله بصورت **پشت و رو(دورو)** بلامانع است و انجام آن توصیه می شود.



دانشگاه صنعتی امیرکبیر  
(پلی‌تکنیک تهران)

به نام خدا

تاریخ: خرداد ۱۴۰۳

## تعهدنامه اصالت اثر

اینجانب سودا ایمانی متعهد می‌شوم که مطالب مندرج در این پایان‌نامه حاصل کار پژوهشی اینجانب تحت نظرارت و راهنمایی استاد دانشگاه صنعتی امیرکبیر بوده و به دستاوردهای دیگران که در این پژوهش از آنها استفاده شده است مطابق مقررات و روال متعارف ارجاع و در فهرست منابع و مأخذ ذکر گردیده است. این پایان‌نامه قبلاً برای احراز هیچ مدرک هم‌سطح یا بالاتر ارائه نگردیده است.

در صورت اثبات تخلف در هر زمان، مدرک تحصیلی صادر شده توسط دانشگاه از درجه اعتبار ساقط بوده و دانشگاه حق پیگیری قانونی خواهد داشت.

کلیه نتایج و حقوق حاصل از این پایان‌نامه متعلق به دانشگاه صنعتی امیرکبیر می‌باشد. هرگونه استفاده از نتایج علمی و عملی، واگذاری اطلاعات به دیگران یا چاپ و تکثیر، ترجمه و اقتباس از این پایان‌نامه بدون موافقت کتبی دانشگاه صنعتی امیرکبیر ممنوع است. نقل مطالب با ذکر مأخذ بلامنع است.

سودا ایمانی

امضا

# سپاس‌گزاری

تقدیم به پرورا مادر غیرزی و هم‌بانم که در سختی‌ها و شواری‌های زندگی همواره کنارم بوده‌اند.

سودا ایانی

خرداد ۱۴۰۳

## چکیده

یادگیری عمیق، یکی از زیرمجموعه‌های هوش مصنوعی، با قابلیت خود در استخراج الگوها و بینش‌های پیچیده از داده‌های حجمی، دگرگونی‌های بسیاری را در زمینه‌های متعددی همچون بهداشت و درمان، رانندگی خودکار، خدمات مالی و نظارت بر محیط زیست به ارمغان آورده است. این فناوری بهویژه در زمینه تصویربرداری پزشکی تأثیر شگرفی نشان داده است. پروژه حاضر به طور خاص به بررسی کاربرد یادگیری عمیق در تشخیص بیماری‌های قلبی و عروقی از طریق تصاویر تشدید مغناطیسی قلب می‌پردازد. هدف این تحقیق آشکارسازی توانمندی‌های یادگیری عمیق در تشخیص این دسته از بیماری‌ها می‌باشد. با استفاده از معناری‌های پیشرفته و قدرت محاسباتی بالای مدل‌های یادگیری عمیق، این پروژه به ارائه رویکرد جدیدی در تشخیص بیماری‌های قلبی می‌پردازد و امکان تعمیم این تکنولوژی را به سایر زمینه‌های پزشکی نشان می‌دهد.

## واژه‌های کلیدی:

یادگیری عمیق، بیماری‌های قلبی و عروقی، تصاویر تشدید مغناطیسی قلب.

# فهرست مطالب

صفحه	عنوان
۵	۲ ادبیات مسئله
۶	۱-۲ تصویربرداری رزونانس مغناطیسی قلب
۶	۲-۲ تصاویر سینماتیک ام آر آی
۷	۳-۲ ام آر آی قلبی محور کوتاه
۸	۳ پیشینه موضوع
۱۲	۴ رویکرد پروژه
۱۳	۱-۴ مقدمه
۱۴	۲-۴ معماری کدگذار-کدگشا
۱۵	۳-۴ کدگذارهای خودکار تغییراتی
۱۶	۱-۳-۴ تابع هزینه
۱۷	۴-۴ مدل یونت
۱۸	۵-۴ ترانسفورمرها
۲۰	۶-۴ مدل ترانسفورمر کاملاً کانولوشنی
۲۲	۵ پیاده‌سازی
۲۳	۱-۵ مجموعه داده
۲۴	۲-۵ معماری مدل پیاده‌سازی شده
۲۶	۱-۲-۵ آزمایشات با دیگر مدل‌های پیچیده برای استخراج ویژگی‌های یادگیری عمیق
۲۷	۳-۵ استخراج ویژگی‌های دستی
۲۸	۴-۵ نرمال‌سازی داده‌ها
۲۹	۵-۵ معیارهای ارزیابی
۲۹	۱-۵-۵ ارزیابی مدل تقسیم‌بندی با معیار دایس
۲۹	۲-۵-۵ ارزیابی مدل دسته‌بندی با دقت، دقت تفکیکی، حساسیت و امتیاز F1
۳۰	۶-۵ کتابخانه‌ها و چارچوب‌های مورد استفاده

۳۱	۶	نتایج و ارزیابی
۳۲	۶-۱	تقسیم‌بندی
۳۴	۶-۲	کدگذار خودکار تغییراتی
۳۷	۶-۳	طبقه‌بندی
۳۹	۷	پیاده‌سازی سامانه نهایی
۴۰	۱-۷	Streamlit
۴۰	۲-۷	طراحی برنامه
۴۰	۱-۲-۷	تقسیم‌بندی
۴۲	۲-۲-۷	استخراج ویژگی
۴۳	۳-۲-۷	طبقه‌بندی
۴۴	۴-۲-۷	ارزیابی
۴۵	۸	جمع‌بندی و نتیجه‌گیری
۴۷		کتاب‌نامه

# فهرست تصاویر

صفحه

شکل

۱-۱	مثالی از قطعه‌بندی یک تصویر تشدید مغناطیسی قلب [۱]	۳
۱-۲	نحوه برش‌گیری قلب در تصاویر ام‌آرآی محور کوتاه	۷
۱-۳	معماری مدل استفاده شده در مقاله [۲]	۱۰
۲-۳	معماری مدل استفاده شده در مقاله [۱]	۱۱
۱-۴	معماری مدل‌های کدگذار-کدگشا	۱۴
۲-۴	معماری مدل‌های کدگذار خودکار تغییراتی	۱۶
۳-۴	معماری مدل UNet	۱۸
۴-۴	معماری مدل ترانسفورمرهای بینایی [۳]	۱۹
۵-۴	معماری مدل ترانسفورمر کاملاً کانولوشنی	۲۰
۱-۵	نمونه‌ای از تصاویر محور کوتاه قلب به همراه تصاویر قطعه‌بندی شده	۲۳
۲-۵	معماری مدل‌های استفاده شده در این پروژه	۲۵
۳-۵	روش تخمین ضخامت دیواره میوکارد در هر برش محور کوتاه در مقاله [۴]	۲۸
۱-۶	نمودار میله‌ای مقادیر دایس برای دسته‌های پنج گانه در فاز آموزش	۳۲
۲-۶	نمودار میله‌ای مقادیر دایس برای دسته‌های پنج گانه در فاز آزمایش	۳۲
۳-۶	نمونه‌ای از تصویر تقسیم‌بندی شده توسط مدل FCT و مقایسه آن با ماسک تقسیم‌بندی شده توسط کارشناسان بالینی	۳۳
۴-۶	نمودار سمت چپ کاهش ضرر کلی مدل کدگذار خودکار تغییراتی، نمودار سمت راست کاهش ضرر بخش کلاس‌بندی در مدل کدگذار خودکار تغییراتی	۳۴
۵-۶	نمودار خطی دقت کلاس‌بندی در فاز آموزش	۳۴
۶-۶	نمودار سمت چپ دقت کلاس‌بندی در فاز اعتبارسنجی، نمودار سمت راست کاهش ضرر کلی مدل کدگذار خودکار تغییراتی در فاز اعتبارسنجی	۳۵
۷-۶	نمایش بصری از خروجی‌های مدل VAE در طول ایپاک‌های مختلف	۳۵
۹-۶	ماتریس درهم‌ریختگی در داده‌های آزمایشی	۳۷

۳۷	۸-۶ ماتریس درهم‌ریختگی در داده‌های آموزشی
۳۷	۱۰-۶ نمودار یادگیری در فاز آموزش
۴۱	۱-۷ صفحه اول سامانه تحت وب
۴۲	۲-۷ صفحه دوم سامانه تحت وب
۴۳	۳-۷ صفحه سوم سامانه تحت وب

## فهرست جداول

صفحه

جدول

۱۱ .....	۱-۳ مقایسه عملکرد سه روش بررسی شده .....
۳۳ .....	۱-۶ میانگین ضرایب دایس برای بطن راست RV، بطن چپ LV، و میوکارد MYO در دو فاز قلبی، انبساط و انقباض در فاز آموزش .....
۳۳ .....	۲-۶ میانگین ضرایب دایس برای بطن راست RV، بطن چپ LV، و میوکارد MYO در دو فاز قلبی، انبساط و انقباض در فاز آزمایش .....
۳۸ .....	۳-۶ معیارهای عملکرد طبقه‌بندی برای مجموعه داده‌های آزمایش .....

مقدمه

فصل اول

## ۱-۱ مقدمه

بیماری‌های قلبی و عروقی<sup>۱</sup> در سال‌های اخیر به عنوان یکی از علل مهم مرگ و میر در سطح جهان ظاهر شده‌اند. به دلیل علائم اولیه ناشناخته‌ای که به تدریج بدتر می‌شود، تشخیص زودهنگام و تشخیص بیماری‌های قلبی و عروقی از اهمیت بالایی برخوردار است [۵]. روش‌های بالینی نظیر آزمایش خون، سیگنال‌های نوار قلب<sup>۲</sup> و تصویربرداری پزشکی از جمله تصویربرداری تشدید مغناطیسی قلب<sup>۳</sup> ابزارهای مؤثری برای تشخیص بیماری‌های قلبی و عروقی هستند. یکی از مزایای اصلی تصویربرداری تشدید مغناطیسی قلب قابلیت ارائه تصاویر با وضوح بالا و جزئیات دقیق از قلب است. این تصاویر به واقع اطلاعات مفیدی درباره شکل‌های ساختاری و اندازه و عملکرد سیستم قلبی به پزشکان ارائه می‌دهند [۶].

تصویربرداری تشدید مغناطیسی قلب نقش حیاتی در کمی‌سازی شاخص‌های قلبی دارد که برای تشخیص بیماری‌های قلبی و عروقی ضروری هستند. پارامترهایی مانند حجم بطن<sup>۴</sup>، کسر جهشی<sup>۵</sup>، جرم عضله قلبی<sup>۶</sup> و انقباض ناحیه‌ای<sup>۷</sup> می‌توانند به دقت از تصاویر تشدید مغناطیسی قلب استخراج شوند [۷]. این شاخص‌های کمی فهمی ارزشمند از کلیت عملکرد قلبی ارائه می‌دهند که در تشخیص، ارزیابی خطر و برنامه‌ریزی درمانی بیماری‌های قلبی و عروقی مختلف موثر است. با این وجود، علیرغم این مزایا، تحلیل تصاویر تشدید مغناطیسی قلب برای تشخیص بیماری‌های قلبی و عروقی هنوز یک فرآیند پیچیده و زمانبر است. تفکیک دستی ساختارهای قلبی زحمتبر و حساس است و به خطاهای انسانی منجر می‌شود. علاوه بر این، حجم بزرگ داده‌های تشدید مغناطیسی قلب تولید شده برای هر بیمار نیازمند تکنیک‌های تجزیه و تحلیل کارا و خودکار برای مقابله با بار کاری روزافزون در بیمارستان‌ها است.

به منظور غلبه بر چالش‌های مرتبط با تفکیک دستی و تفسیر ذهنی تصاویر تشدید مغناطیسی قلب، یکی از راهکارهای جالب واگذاری به روش‌های یادگیری عمیق است. مدل‌های یادگیری عمیق در بسیاری از کاربردهای تصویربرداری پزشکی به خوبی جواب داده‌اند [۸]. الگوریتم‌های یادگیری عمیق به طور خودکار و با دقت بالا قابلیت یادگیری و شناسایی الگوها و ویژگی‌های جزئی و پیچیده را از حجم بزرگی

<sup>1</sup>Cardiovascular diseases (CVDs)

<sup>2</sup>Electrocardiography (ECG) signals

<sup>3</sup>Cardiac magnetic resonance (CMR)

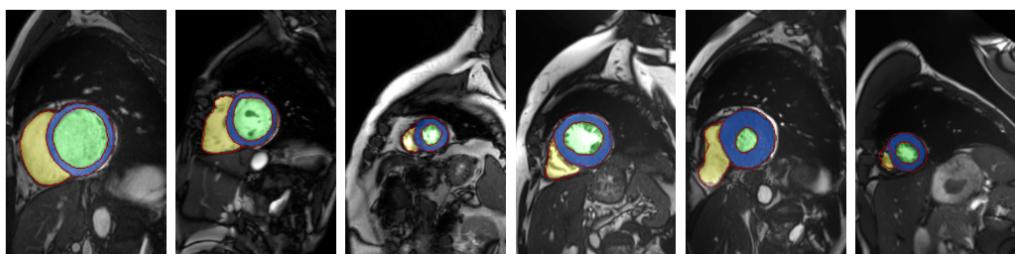
<sup>4</sup>Ventricular volume

<sup>5</sup>Ejection fraction

<sup>6</sup>Myocardial mass

<sup>7</sup>Regional contractility

از داده‌ها دارند. این قابلیت آن‌ها را به مناسب‌ترین ابزار برای تجزیه و تحلیل تصاویر تشید مغناطیسی قلب که اغلب ساختارهای آناتومیکی پیچیده، تفاوت‌هایی در شدت تصویر و بیماری‌های قلبی متنوع دارند، تبدیل کرده است [۵]. با آموزش شبکه‌های عصبی بر روی مجموعه داده‌های بزرگ و با اطلاعات برچسب‌گذاری شده، این مدل‌ها می‌توانند ساختارهای قلبی مانند بطن چپ، بطن راست و عضله قلب را با دقت و کارآیی بالا تقسیم کنند.



شکل ۱-۱: مثالی از قطعه‌بندی یک تصویر تشید مغناطیسی قلب [۱]

شکل ۱-۱: نمونه‌هایی از بخش‌بندی‌های به دست آمده توسط الگوریتم‌های یادگیری عمیق در شش بیمار مختلف، که حفره بطن چپ را به رنگ سبز، حفره بطن راست را به رنگ زرد، و عضله قلب را به رنگ آبی نشان می‌دهد. خطوط مرتع با رنگ قرمز نشان داده شده است.

علاوه بر این، الگوریتم‌های یادگیری عمیق قادرند ویژگی‌های جامع و قابل اطمینان‌تری را از تصاویر تشید مغناطیسی قلب نسبت به روش‌های سنتی دستی استخراج کنند. با بهره‌گیری از قدرت محاسباتی بزرگ و معماری‌های پیشرفته مدل‌های یادگیری عمیق، پروژه پیشنهادی ممکن است محدودیت‌های مرتبط با استخراج ویژگی دستی را برطرف کند و بتواند تحلیل دقیق‌تر و واقع‌گرایتری از تصاویر تشید مغناطیسی قلب ارائه دهد. الگوریتم‌های یادگیری عمیق همچنین توانایی تعمیم خوبی را نسبت به جمعیت‌های بیمار مختلف دارند. با استفاده از تکنیک‌های مناسب افزایش داده<sup>۸</sup> و معماری‌های قوی شبکه<sup>۹</sup>، مدل‌های یادگیری عمیق می‌توانند به تغییراتی مانند تفاوت‌های آناتومی بیمار، پروتکل‌های اسکن تصویر و کیفیت داده، سازگار شوند. این قابلیت ارزشمند است و در محیط‌های بالینی که با گروه‌های بیماری متنوع مواجه هستند، تضمین می‌کند که تکنیک‌های خودکار تفکیک و طبقه‌بندی بیماری‌های قلبی و عروقی ارائه شده می‌توانند به طور موثر در موارد متنوعی استفاده شوند. علاوه بر این، رویکرد مبتنی بر یادگیری عمیق پروژه پیشنهادی باعث امکان تجزیه و تحلیل قابل اعتماد و مقیاس‌پذیر می‌شود و برای پردازش کارآمد حجم عظیمی از تصاویر تشید مغناطیسی قلب امکان‌پذیر می‌شود و

<sup>8</sup>Data augmentation

<sup>9</sup>Robust network architectures

جريان کار در بیمارستان را ارتقا می‌بخشد. مخصوصاً بدلیل شیوع افرایشی بیماری‌های قلبی و عروقی و نیاز به تشخیص به موقع و دقیق این موضوع اهمیت زیادی دارد. با کمک به بهبود سرعت و دقت تشخیص بیماری‌های قلبی و عروقی از طریق تفکیک خودکار و طبقه‌بندی بیماری، پروژه پیشنهادی می‌تواند تأثیری متحول کننده در بیمارستان‌ها داشته باشد و باعث بهبود مدیریت کارآمد بیمار و احتمالاً کاهش تأخیر در درمان شود.

با استفاده از الگوریتم‌های یادگیری عمیق، پروژه پیشنهادی این امکان را دارد که تشخیص بیماری‌های قلبی و عروقی مبتنی بر تصاویر تشید مغناطیسی قلب را بهبود بخشد. با بهره‌گیری از مزایای یادگیری عمیق، شامل توانایی یادگیری الگوهای پیچیده، تعمیم در میان جمعیت‌های متنوع و بهبود فرآیندهای تجزیه و تحلیل تصویر، این پروژه می‌تواند تشخیص دقیق، بی‌طرف و کارآمد را به دست دهد و در نهایت در دنیای سلامت قلبی-عروقی به همه پزشکان و بیماران بهره‌مندی بیاندازد.

## فصل دوم

### ادبیات مسئله

## ۱-۲ تصویربرداری رزونانس مغناطیسی قلب

تصویربرداری رزونانس مغناطیسی قلب، که به آن Cardiovascular magnetic resonance imaging (CMR) نیز گفته می‌شود، یک تکنولوژی تصویربرداری غیرتهاجمی است که از میدان‌های مغناطیسی قدرتمند و امواج رادیویی برای تولید تصاویر دقیق از ساختار و عملکرد قلب استفاده می‌کند. قدرت CMR در توانایی آن برای ارائه تصاویر واضح از عضله قلب و بافت‌های اطراف بدون نیاز به تابش یونیزه کننده است، که آن را به یک گزینه امن برای ارزیابی‌های مکرر تبدیل می‌کند.

از جمله مزایای اصلی CMR، توانایی آن در ارائه تصاویر دقیق از دیواره‌ها و حفره‌های قلبی است که برای تشخیص و درمان بیماری‌های قلبی حیاتی است. این تکنیک می‌تواند انواع مختلف بافت‌های قلبی را تفکیک کرده و تغییرات بیماری‌زا مانند فیبروز<sup>۱</sup>، ادما<sup>۲</sup> یا نفوذ مواد خارجی را شناسایی کند.

همچنین، تصویربرداری رزونانس مغناطیسی قلب قادر است جریان خون را در داخل قلب و عروق اصلی اندازه‌گیری و تصویربرداری کند، که در ارزیابی بیماری‌های پیچیده مادرزادی قلب و مشکلات دریچه‌ای مفید است. همراه با تمام مزایای داده‌های CMR، تشخیص بیماری‌های قلبی عروقی برای پزشکان به دلیل وجود تعداد زیادی از برش‌های داده‌ای، کنتراست پایین و غیره، چالش‌برانگیز است. برای رفع این مشکلات، تکنیک‌های یادگیری عمیق برای تشخیص بیماری‌های قلبی و عروقی با استفاده از داده‌های CMR به کار گرفته شده‌اند و تحقیقات زیادی در حال حاضر در این زمینه در حال انجام است.

## ۲-۲ تصاویر سینماتیک ام‌آرآی

سینماتیک ام‌آرآی<sup>۳</sup> یک نوع تکنیک تصویربرداری رزونانس مغناطیسی قلب است که تصاویر متحرک قلب را در حین کار کردن آن نشان می‌دهد، و این امکان را می‌دهد تا ساختار و عملکرد قلب را به صورت زنده مشاهده کنیم. سینماتیک ام‌آرآی در طول چرخه قلبی، تصاویر متعددی را در زمان‌های مختلف ثبت می‌کند که این امکان را می‌دهد تا حرکت ساختارهای قلبی در هر ضربان به صورت فریم به فریم تجزیه و تحلیل شود.

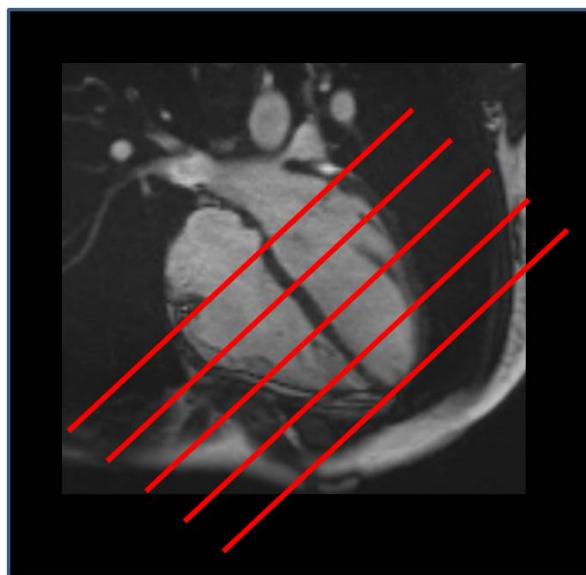
<sup>1</sup>Fibrosis

<sup>2</sup>Edema

<sup>3</sup>Cine MRI

### ۳-۲ ام آر آی قلبی محور کوتاه

تصاویر محور کوتاه در ام آر آی قلبی به تصاویر مقطعی از قلب اشاره دارند که عمود بر محور طولانی بدن گرفته شده‌اند. این جهت‌گیری برای ارزیابی قلب بسیار مفید است زیرا امکان مشاهده جامع تمام حفره‌های قلب، ضخامت دیواره و حرکت میوکارد در طول چرخه قلبی را فراهم می‌کند. استفاده از تصاویر محور کوتاه برای اندازه‌گیری حجم‌های بطنی، ارزیابی کارکرد دیواره‌های قلبی و تعیین کارایی قلب در پمپاژ خون مفید است. این دیدگاه به پزشکان کمک می‌کند تا با دقیق بیشتری به تشخیص و درمان اختلالات قلبی بپردازنند، زیرا تصاویر محور کوتاه اطلاعات بسیار دقیقی از ساختار داخلی قلب و عملکرد آن ارائه می‌دهند.



شکل ۲-۱: نحوه برش گیری قلب در تصاویر ام آر آی محور کوتاه

برش گیری به این صورت است که قلب معمولاً از پایه (نzdیک به دهلیزها) تا نوک (نوك بطن چپ) در یک سری تصاویر موازی تصویربرداری می‌شود.

## فصل سوم

### پیشینه موضوع

تشخیص کمکی کامپیوترا<sup>۱</sup> برای بیماری‌های قلبی با استفاده از فناوری‌های پیشرفته پردازش تصویر و یادگیری ماشین، از جمله موضوعات پژوهش در عرصه پزشکی مدرن می‌باشد. در این راستا، تحقیقات متعددی به منظور بهبود دقت تشخیص و کارایی درمانی از طریق تجزیه و تحلیل دقیق داده‌های تصویری بالینی انجام شده است. در ادامه، به بررسی سه مقاله برجسته در حوزه تشخیص بیماری‌های قلبی با استفاده از تکنیک‌های برش‌نگاری و دسته‌بندی تصاویر قلبی پرداخته‌ایم.

ستین و همکارانش [۹] از یک روش تقسیم‌بندی نیمه‌خودکار استفاده کردند تا به صورت دستی مرزهای ساختارهای قلبی را استخراج کنند. برای استخراج ویژگی، روش آن‌ها بر استخراج مجموعه‌ای جامع از ویژگی‌های رادیومیک، شامل ویژگی‌های آماری، شکلی و بافتی از تصاویر cine-MRI مرکز است. این ویژگی‌ها تغییرات جزئی آناتومیک و عملکردی مرتبط با انواع مختلف بیماری‌های قلبی و عروقی را ثبت می‌کنند.

- ویژگی‌های آماری: مانند میانگین و انحراف معیار که اطلاعات کلی در مورد توزیع شدت در ساختارهای قلبی را فراهم می‌کنند.

- ویژگی‌های مبتنی بر شکل: مانند حجم، کرویت و مساحت سطح که ویژگی‌های هندسی ساختارهای قلب را توصیف می‌کنند.

- ویژگی‌های بافتی: توصیف کننده‌های پیشرفته‌ای که از ماتریس‌های تجزیه و تحلیل بافت مانند ماتریس هموقوع سطح خاکستری<sup>۲</sup> و ماتریس طول دویدن سطح خاکستری<sup>۳</sup> استخراج می‌شوند، این ویژگی‌ها الگوهای تغییرات شدت پیکسل و زبری بافت را ضبط می‌کنند.

در نهایت ۵۶۷ ویژگی بدست آمد که به دلیل تعداد زیاد ویژگی‌های رادیومیک، انتخاب ویژگی مناسب برای جلوگیری از بیش‌برازش مدل ضروری بود. در این مطالعه، از روش انتخاب ویژگی پیشروی ترتیبی<sup>۴</sup> استفاده شد تا مرتبطترین ویژگی‌ها برای طبقه‌بندی بیماری‌های قلبی و عروقی شناسایی شوند. سپس، برای طبقه‌بندی بیماران، از ماشین بردار پشتیبانی<sup>۵</sup> استفاده شد. SVM با استفاده از ویژگی‌های انتخاب شده، مدل طبقه‌بندی را آموزش داد تا بتواند بین زیرگروه‌های مختلف بیماران و افراد عادی تمایز قائل شود.

<sup>1</sup>Computer-aided diagnosis (CAD)

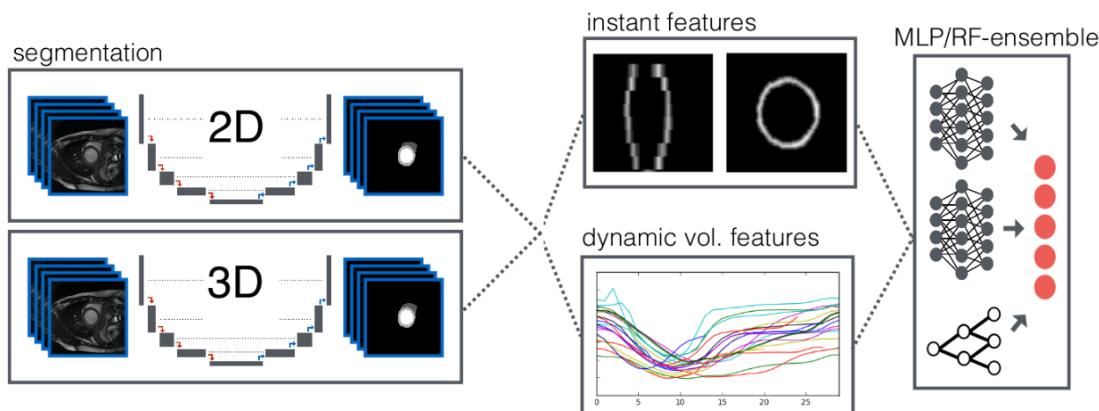
<sup>2</sup>Gray Level Co-occurrence Matrix (GLCM)

<sup>3</sup>Gray Level Run Length Matrix (GLRLM)

<sup>4</sup>sequential forward feature selection

<sup>5</sup>SVM

مقاله "ارزیابی بیماری قلبی خودکار بر روی سینماتیک ام آر آی از طریق برش نگاری سری زمانی و ویژگی های مخصوص دامنه" اثر ایزننسی و همکاران [۲]، بر دو جنبه اصلی تمرکز دارد: تقسیم‌بندی دقیق ساختارهای قلبی در طول کل چرخه قلبی و استخراج ویژگی های دقیق برای دسته‌بندی بیماری های قلبی. این تحقیق معماری های متعدد از شبکه های U-Net<sup>۶</sup> دو بعدی و سه بعدی را با استفاده از اتصالات باقی‌مانده<sup>۷</sup> در لایه های اپ‌سمپلینگ<sup>۸</sup> پیاده‌سازی کردند که برای تقسیم‌بندی دقیق ساختارهای قلبی در تمام نقاط زمانی چرخه قلبی به کار رفته است. این شبکه ها قادر به تشخیص و جداسازی حفره های بطی نیز هستند. پس از تقسیم‌بندی، ویژگی های مختلفی از چپ و راست و میوکارد در هر مرحله از دوره قلبی هستند. پس از تقسیم‌بندی، ویژگی های مختلفی از داده های تصویری استخراج می شوند که برای ارزیابی های بالینی قلب طراحی شده اند. این ویژگی ها شامل اطلاعات مربوط به ضخامت، دوره، و حجم ساختارهای قلبی در حالت های مختلف می شوند. همچنین، ویژگی های پویایی نظیر تغییرات حجمی در طول دوره قلبی نیز مورد بررسی قرار می گیرند. بر اساس ویژگی های استخراج شده، یک مجموعه از دسته‌بندان، شامل شبکه های عصبی چند لایه<sup>۹</sup> و جنگل<sup>۱۰</sup> تصادفی<sup>۱۱</sup>، آموزش داده شده اند.



شکل ۳-۱: معماری مدل استفاده شده در مقاله [۲]

ولترینک و همکارانش [۱] یک روش کاملاً خودکار که از یادگیری عمیق برای هم برش نگاری و هم دسته‌بندی بیماری های قلبی از تصاویر سینماتیک ام آر آی استفاده می کند را توسعه داده اند. در این مطالعه، یک شبکه عصبی کانولوشنی<sup>۱۰</sup> را برای تقسیم‌بندی خودکار بطن های چپ و راست و میوکارد در فازهای پایان انبساط و پایان انقباض تصاویر استفاده کرده اند. این شبکه به گونه‌ای طراحی شده است

<sup>6</sup>Residual connections

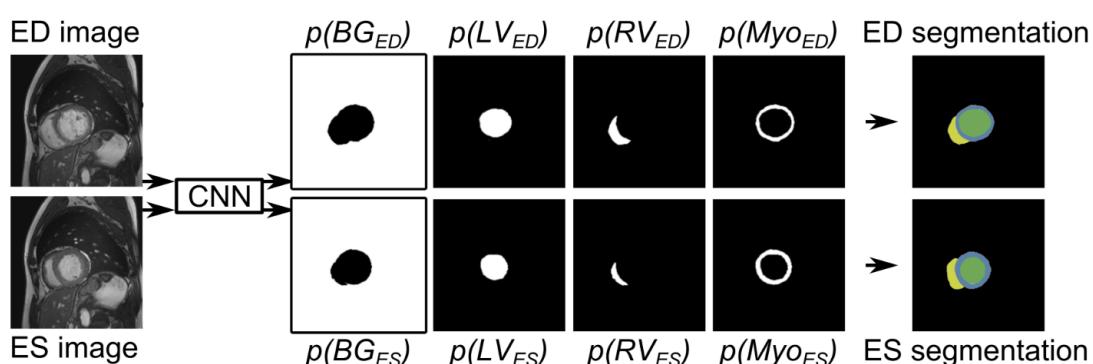
<sup>7</sup>Upsampling

<sup>8</sup>Multilayer perceptron

<sup>9</sup>RandomForest

<sup>10</sup>Convolutional neural network

که بتواند با دقت بالا ساختارهای قلبی را در فازهای مختلف چرخه قلبی شناسایی و جدا کند. پس از تقسیم‌بندی دقیق ساختارهای قلبی، ویژگی‌های کمی مختلفی از تصاویر تقسیم‌بندی شده استخراج می‌شوند. این ویژگی‌ها شامل حجم‌های بطنی، کسری جهشی، و جرم میوکارد می‌باشند. استخراج این ویژگی‌ها امکان مقایسه دقیق بین نمونه‌های بیمار و سالم را فراهم می‌آورد و به تشخیص دقیق‌تر بیماری‌های قلبی کمک می‌کند. سپس با استفاده از ویژگی‌های استخراج شده، یک دسته‌بند جنگل تصادفی برای تشخیص و دسته‌بندی بیماری‌های قلبی به کار گرفته می‌شود.



شکل ۲-۳: معماری مدل استفاده شده در مقاله [۱]

جدول ۱-۳ یک نمای کلی از عملکرد طبقه‌بندی سه روش بررسی شده ارائه می‌دهد.

Accuracy	Architectures	Authors
۹۲.۰	SVM	[۹] Cetin et al.
۹۲.۰	Random Forest	[۱۲] Isensee et al.
۸۶.۰	Random Forest	[۱۱] Wolterink et al.

جدول ۱-۳: مقایسه عملکرد سه روش بررسی شده

## فصل چهارم

### رویکرد پروژه

## ۱-۴ مقدمه

در بخش قبلی، مشاهده کردیم که اکثر پروژه‌های مربوط به تشخیص بیماری‌های قلبی عروقی بر ویژگی‌هایی که به صورت دستی از داده‌های برش‌داده شده استخراج می‌شوند، تکیه دارند. هرچند این رویکرد مؤثر است، اما ممکن است از پتانسیل کامل مدل‌های پیشرفته محاسباتی برای کشف الگوهای عمیق‌تر و پیچیده‌تر در داده‌ها که می‌توانند برای تشخیص حیاتی باشند، به طور کامل استفاده نکند. در این پروژه، تلاش شده است که روش‌های مرسوم با افزودن ویژگی‌های یادگیری عمیق به ویژگی‌های دستی ارتقاء داده شوند. این رویکرد ترکیبی با هدف بهره‌گیری از قوتهای هر دو تخصص دستی و پیچیدگی یادگیری ماشین برای بهبود دقت و قابلیت اطمینان تشخیص طراحی شده است.

برای تقسیم‌بندی تصاویر قلبی، از یک ترانسفورمر کاملاً کانولوشنی<sup>۱</sup> استفاده خواهد شد. این معماری، ترکیبی از شبکه‌های عصبی کانولوشنی<sup>۲</sup> و ترانسفورمرها را ارائه می‌دهد. مدل FCT با هدف درک و تفسیر ساختارهای پیچیده آناتومیک قلب طراحی شده است تا با جذب ویژگی‌های محلی از طریق کانولوشن‌ها و وابستگی‌های بلند مدت از طریق مکانیسم‌های توجه، دقت تقسیم‌بندی را افزایش دهد. پس از اتمام تقسیم‌بندی، یک اتوانکودر متغیر<sup>۳</sup> را برای تولید ویژگی‌های یادگیری عمیق از تصاویر تقسیم‌بندی شده آموزش داده می‌شود. VAE طراحی شده است تا داده‌های ورودی را به یک فضای کمینه فشرده کدگذاری کند، که تغییرات اساسی ساختارهای قلبی را که برای تشخیص بیماری‌های قلبی و عروقی بسیار مرتبط هستند، به دست آورد. با استفاده از قدرت این مدل تولیدی، می‌توانیم ویژگی‌های معناداری را استخراج کنیم که از طریق روش‌های دستی به راحتی قابل دسترس نیستند. مدل تشخیصی سپس این ویژگی‌های یادگیری عمیق را از فضای کمینه VAE با ویژگی‌های استخراج شده دستی ترکیب خواهد کرد. این مجموعه ویژگی‌ها برای آموزش یک دسته‌بند نهایی برای تشخیص بیماری‌های قلبی استفاده خواهد شد. با ترکیب ویژگی‌های به دست آمده به صورت دستی با ویژگی‌های آموخته شده، این مدل به دنبال دستیابی به عملکرد برتر در شناسایی انواع پاتولوژی‌ها است، که ممکن است منجر به تشخیص دقیق‌تر و زودهنگام بیماری‌های قلبی و عروقی شود.

این بخش به تشریح روش‌های دقیق استفاده شده در این پروژه، از جمله فرایندهای کسب داده‌ها، جزئیات معماری مدل برای هر سه بخش تقسیم‌بندی، استخراج ویژگی و تشخیص و راهبردها برای ارزیابی تشخیص بیماران می‌پردازد.

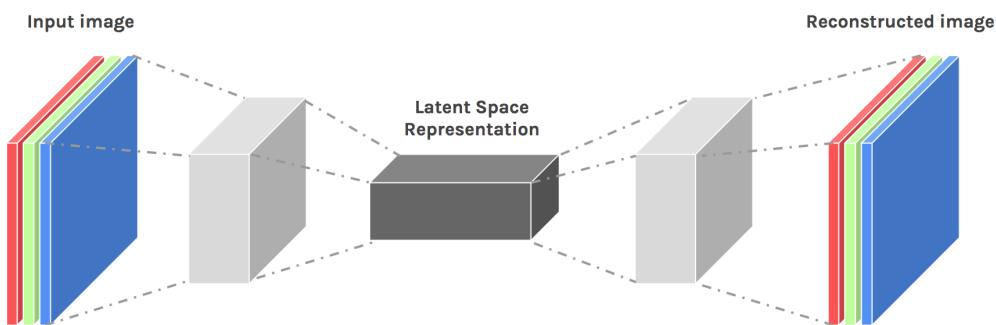
<sup>1</sup>Fully Convolutional Transformer (FCT)

<sup>2</sup>Convolutional Neural Networks (CNNs)

<sup>3</sup>Variational Autoencoder (VAE)

## ۲-۴ معماری کدگذار-کدگشا

معماری رمزگذار-رمزگشا یک چارچوب بسیار کاربردی در پردازش تصویر و بینایی کامپیوتری است که بخصوص برای کارهایی مانند تقسیم‌بندی تصویر، افزایش وضوح، و انتقال سبک مناسب است. این معماری به طور گسترده در یادگیری توالی به توالی کاربرد دارد و از دو قسمت اصلی رمزگذار و رمزگشا تشکیل شده است. بخش رمزگذار، تصویر ورودی را پردازش می‌کند و آن را به یک نمایش فشرده و متراکم به نام نقشه ویژگی یا فضای کمینه تبدیل می‌کند. این فرآیند شامل کاهش ابعاد مکانی تصویر (ارتفاع و عرض) در حالی که عمق (تعداد کانال‌های ویژگی) افزایش می‌یابد، است. تصویر از طریق چندین لایه، که اغلب لایه‌های کانولوشنی هستند، پردازش می‌شود که به طور پیشونده با استفاده از تکنیک‌هایی مانند پولینگ یا کانولوشن‌های پله‌ای<sup>۴</sup> تصویر را کاهش می‌دهند. خروجی از رمزگذار یک نمایش ویژگی فشرده است که اطلاعات حیاتی لازم برای پردازش‌های بعدی را در بر دارد.



شکل ۱-۴: معماری مدل‌های کدگذار-کدگشا

در طرف دیگر، رمزگشا ویژگی‌های رمزگاری شده را گرفته و آن‌ها را به ابعاد اصلی تصویر یا به ابعاد متفاوت بسته به کار خاص بازمی‌گرداند. به عنوان مثال، در تقسیم‌بندی تصویر، خروجی یک ماسک پیکسل به پیکسل از تصویر خواهد بود، اما در افزایش وضوح، خروجی ممکن است تصویری با ابعاد بزرگتر از ورودی باشد. رمزگشا با استفاده از لایه‌هایی که انجام افزایش نمونه‌برداری را دارند، کار می‌کند، که می‌تواند از طریق روش‌هایی مانند کانولوشن‌های معکوس، افزایش مقیاس دنبال شده توسط کانولوشن‌ها، یا عملیات تغییر ترتیب پیکسل‌ها انجام شود. در حین این فرآیند، عمق معمولاً کاهش می‌یابد در حالی که ابعاد مکانی افزایش می‌یابد.

<sup>4</sup>Strided convolutions

### ۳-۴ کدگذارهای خودکار تغییراتی

کدگذار خودکار تغییراتی<sup>۵</sup> [۱۰] نوعی خودرمنگار است که از تکنیک‌های احتمالاتی و آماری برای یادگیری نمایش‌های پنهان<sup>۶</sup> داده‌ها استفاده می‌کند. در حالی که خودرمنگارهای استاندارد داده‌ها را در یک فضای با بعد کمتر فشرده‌سازی می‌کنند و سپس خروجی را مستقیماً بازسازی می‌کنند، مدل‌های VAE با استفاده از یک رویکرد احتمالاتی برای توصیف توزیع داده‌ها در فضای پنهان عمل می‌کنند. این امر باعث می‌شود که VAE‌ها نه تنها در فشرده‌سازی داده‌ها موثر باشند بلکه در تولید داده‌های جدید که شبیه به داده‌های آموزشی هستند، نیز موثر باشند.

یک VAE شامل دو جزء اصلی رمزگذار<sup>۷</sup> و رمزگشا<sup>۸</sup> است. با این حال، رمزگذار در یک نقطه‌ای تکی در فضای پنهان تولید نمی‌کند؛ بلکه توزیعی از مقادیر ممکن را تولید می‌کند. رمزگذار در VAE داده‌های ورودی  $x$  را به دو چیز مختلف نگاشت می‌کند: میانگین  $\mu$  و واریانس  $\sigma^2$  (یا واریانس لگاریتمی  $\log \sigma^2$ )، که توزیع گاووسی<sup>۹</sup> در فضای پنهان را تعریف می‌کند. به طور ریاضی، رمزگذار پارامترهای توزیع احتمال  $(z|x)$  را تعریف می‌کند، که در آن  $\phi$  پارامترهای شبکه رمزگذار و  $z$  متغیر پنهان است.

$$q_\phi(z|x) = \mathcal{N}(z; \mu(x), \sigma^2(x)I)$$

در اینجا،  $I$  ماتریس هویت و  $\mathcal{N}$  نمایانگر توزیع گاووسی است. استفاده از توزیع گاووسی به دلیل خواص ریاضی خوب آن، محاسبات را ساده می‌کند.

برای تولید یک نمونه  $z$  از فضای پنهان، VAE از تکنیکی به نام "ترفند پارامترزنی مجدد"<sup>۱۰</sup> استفاده می‌کند تا امکان انتشار عقب‌رو<sup>۱۱</sup> از طریق نمونه‌برداری تصادفی فراهم شود. به جای نمونه‌برداری مستقیم  $z$  از  $(\mu, \sigma^2)$ ،  $\epsilon$  را از توزیع نرمال استاندارد  $(0, 1)$  نمونه‌برداری می‌کنیم و محاسبه می‌کنیم:

$$z = \mu + \sigma \odot \epsilon$$

<sup>5</sup>Variational Autoencoder (VAE)

<sup>6</sup>Latent representations

<sup>7</sup>encoder

<sup>8</sup>decoder

<sup>9</sup>Gaussian distribution)

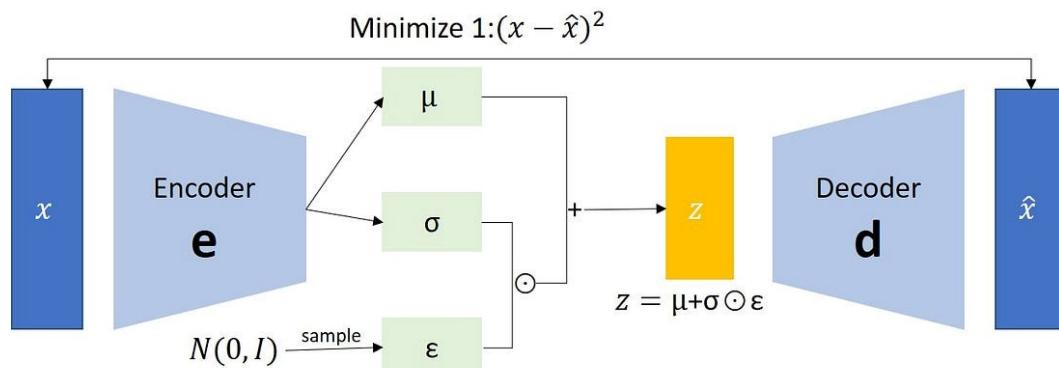
<sup>10</sup>Reparameterization trick

<sup>11</sup>backpropagation

این گام اطمینان می‌دهد که شبکه قابلیت مشتق‌پذیری داشته باشد و بنابراین از طریق گرادیان تصادفی کاهشی قابل آموزش باشد.

بخش رمزگشا در یک VAE هدفش بازسازی ورودی  $x$  از نمونه پنهان  $z$  است. این بخش توزیع احتمال داده‌های مشاهده شده را با توجه به نمایش پنهان  $(x|z)$  مدل می‌کند، که در آن  $\theta$  پارامترهای شبکه رمزگشا است.

برای سادگی، اغلب فرض می‌شود که  $p_\theta(x|z)$  نیز یک توزیع گاووسی است.



$$\text{Minimize 2: } \frac{1}{2} \sum_{i=1}^N (\exp(\sigma_i) - (1+\sigma_i) + \mu_i^2)$$

شکل ۲-۴: معماری مدل‌های کدگذار خودکار تغییراتی

### ۱-۳-۴ تابع هزینه

تابع هزینه در VAE‌ها شامل دو عبارت است:

۱. **هزینه بازسازی**: این میزان اثربخشی رمزگشا در بازسازی داده‌های ورودی را اندازه‌گیری می‌کند و در این پروژه از تابع هزینه mean squared error (MSE) استفاده شده است.

۲. **اختلاف KL**: این عبارت به عنوان یک تنظیم‌کننده عمل می‌کند و اختلاف کولبک-لیبلر<sup>۱۲</sup> بین توزیع تقریبی پنهان  $q_\phi(z|x)$  و توزیع اولیه متغیرهای پنهان است که عموماً فرض می‌شود یک نرمال استاندارد  $\mathcal{N}(0, 1)$  باشد.

<sup>12</sup>Kullback-Leibler divergence

$$\text{Loss} = \text{MSE}(\mathbf{x}, \hat{\mathbf{x}}) - \text{KL}[q_\phi(z|x)\|p(z)]$$

در این پروژه چون هدف اصلی از این آموزش این مدل یادگیری ویژگی‌هایی است که در تشخیص بیماری‌های قلبی و عروقی کمک‌کننده باشند، در تابع هزینه بالا یک بخش نیز برای کلاس‌بندی ویژگی‌ها استفاده کردیم تا مدل را قادر کنیم ویژگی‌های موثر در تشخیص را نیز یاد بگیرد. برای این کار، خروجی نقطه انسداد به یک مدل طبقه‌بندی MLP وارد می‌شود. در نهایت خواهیم داشت:

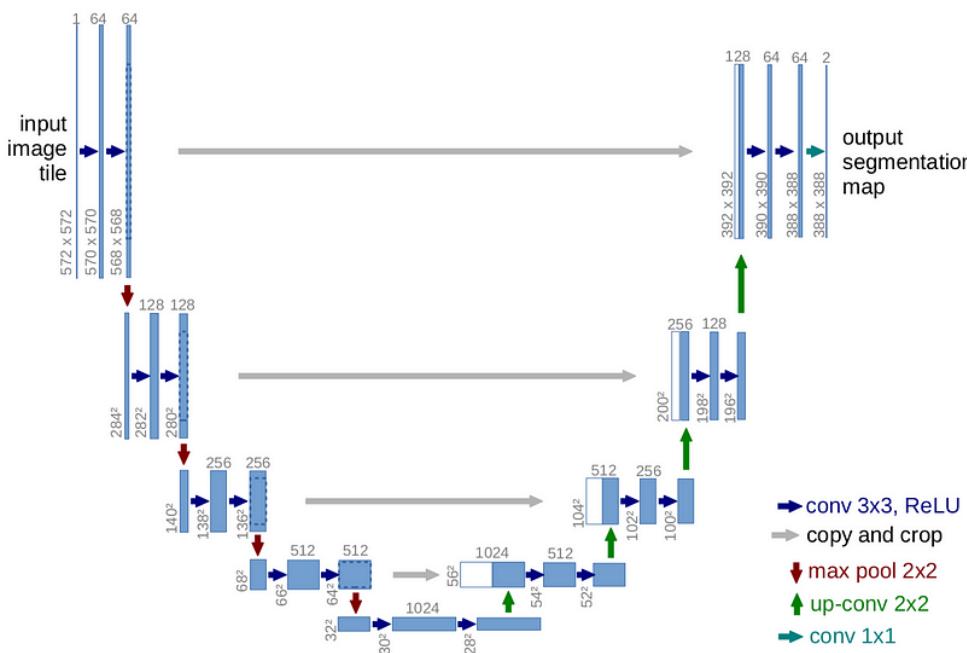
$$\text{Loss} = \text{MSE}(\mathbf{x}, \hat{\mathbf{x}}) - \text{KL}[q_\phi(z|x)\|p(z)] + \text{Categorical Cross-Entropy}$$

## ۴-۴ مدل یونت

یونت [۱۱] یک شبکه عصبی کانولوشنی است که به طور خاص برای تقسیم‌بندی تصاویر پزشکی طراحی شده است. این معماری توسط اولاف رونبرگرو همکارانش در مقاله‌ای با عنوان "یونت: شبکه‌های کانولوشنی برای تقسیم‌بندی تصاویر پزشکی" در سال ۲۰۱۵ معرفی شد. این معماری برای دقت بالا در تقسیم‌بندی تصاویر که در آن اشیاء مورد نظر اشکال و اندازه‌های متفاوتی دارند، شناخته شده است. معماری یونت به خاطر ساختار U شکل خود بارز است که شامل دو بخش اصلی رمزگذار و رمزگشا می‌باشد. مسیر رمزگذار شبیه به معماری معمول شبکه‌های کانولوشنی است و برای درک زمینه تصویر کاربرد دارد. این مسیر از واحدهای تکرار شونده تشکیل شده که هر کدام شامل دو کانولوشن  $3 \times 3$  هستند که هر کدام به دنبال یک واحد فعال سازی خطی اصلاح شده<sup>۱۳</sup> قرار دارند و پس از آن یک عملیات پولینگ حداقل  $2 \times 2$  برای کاهش نمونه وجود دارد. با پیشروی در این مسیر، تعداد کanal‌های ویژگی دو برابر می‌شود که این امکان را می‌دهد تا شبکه ویژگی‌های پیچیده‌تری را در هر سطح یاد بگیرد.

در مقابل، مسیر رمزگذار با هدف دقت مکانی با استفاده از کانولوشن‌های معکوس برای بالا بردن نمونه ویژگی‌ها و ادغام آن‌ها با نقشه‌های ویژگی مربوطه بریده شده از مسیر انقباضی عمل می‌کند. این مرحله شامل یک سری کانولوشن‌های بالا برnde و پس از آن کانولوشن‌های  $2 \times 2$  که تعداد کanal‌های ویژگی را نصف می‌کند، می‌باشد. این مسیر توسط دو کانولوشن  $3 \times 3$  دنبال می‌شود که هر کدام به دنبال

<sup>13</sup>ReLU



شکل ۴-۳: معماری مدل UNet

یک ReLU قرار دارند که به تصفیه ویژگی‌ها برای تقسیم‌بندی دقیق‌تر کمک می‌کند. لایه نهایی یونت یک کانولوشن  $1 \times 1$  است که برای نگاشت کanal‌های ویژگی عمیق به تعداد کلاس‌های مورد نیاز برای وظیفه تقسیم‌بندی مشخص استفاده می‌شود، که نقشه تقسیم‌بندی نهایی را فراهم می‌کند. معماری یونت به دلیل توانایی آن در دقت مکانی بالا، برای تحلیل تصویر پزشکی ضروری است. این معماری در تقسیم‌بندی دقیق حتی با مقدار کمی داده عملکرد بسیار خوبی دارد. این قابلیت در کاربردهای پزشکی که داده‌های تقسیم‌بندی شده با کیفیت بالا کمیاب است، بسیار ارزشمند است.

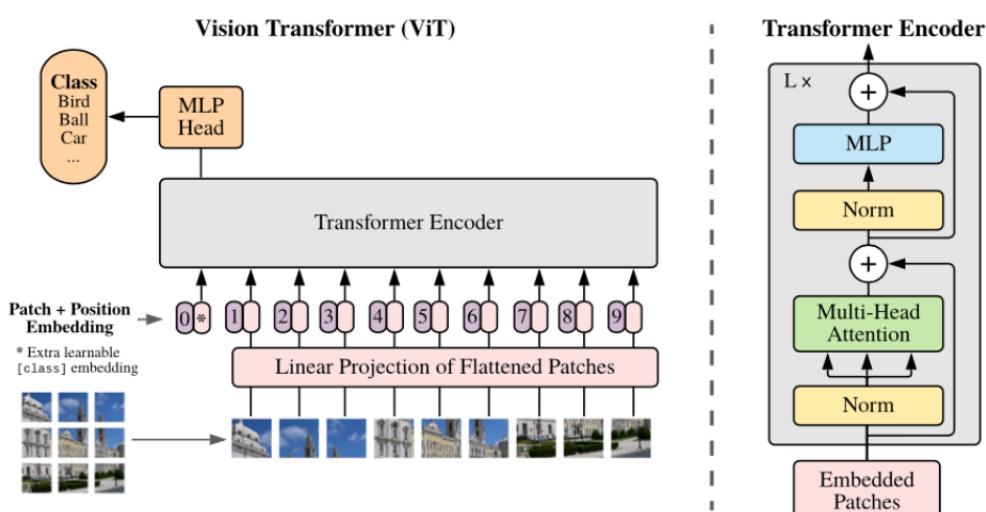
## ۵-۴ ترانسفورمرها

معماری ترانسفورمر معماري ترانسفورمرها، که برای اولین بار توسط وسوانی و همکارانش [۱۲] در مقاله‌ی "تنها به توجه نیاز است"<sup>۱۴</sup> در سال ۲۰۱۷ معرفی شد، انقلابی در نحوه پردازش داده‌های دنباله‌ای مانند متن، داده‌های سری زمانی و حتی تصاویر به وجود آورد. در قلب این معماری، مکانیزم توجه<sup>۱۵</sup> قرار دارد که به مدل امکان می‌دهد تا بر روی بخش‌های مختلف داده‌های ورودی تمرکز کند و برخی از قسمت‌ها را بر اساس اهمیت‌شان در انجام وظیفه‌ی مورد نظر، بیشتر در نظر بگیرد.

<sup>14</sup>Attention is All You Need

<sup>15</sup>Attention Mechanism

مکانیزم توجه با محاسبه‌ی امتیازاتی که اهمیت هر بخش از داده‌ها را نسبت به سایر بخش‌ها اندازه‌گیری می‌کند کار می‌کند. به عنوان مثال در زمینه‌ی پردازش جمله، این مکانیزم به مدل اجازه می‌دهد تا تمام ورودی‌ها را یکجا در نظر بگیرد و هنگام پیش‌بینی کلمه‌ی بعدی در یک دنباله، به کلمات مرتبط بیشتر توجه کند. این ویژگی برای درک وابستگی‌های درازمدت بسیار مفید است.



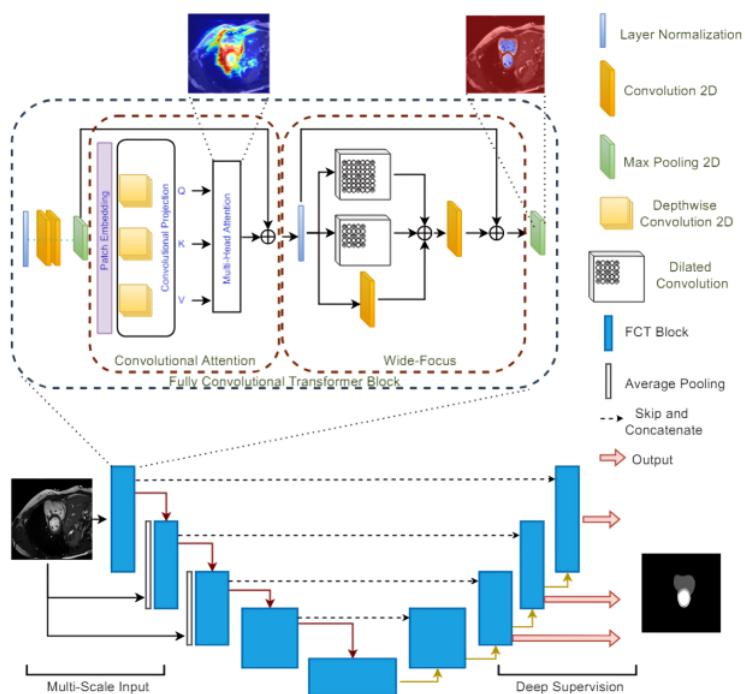
شکل ۴-۴: معماری مدل ترنسفورمرهای بینایی [۳]

برای سازگاری ترنسفورمرها با پردازش تصاویر به عنوان دنباله‌ای از برش‌ها<sup>۱۶</sup> در نظر گرفته می‌شود. هر برش مشابه یک کلمه در جمله است و ترنسفورمر با استفاده از مکانیزم توجه خود، اهمیت هر برش را در ارتباط با سایر برش‌ها سنجیده و هم ویژگی‌های محلی و هم زمینه‌ی وسیع‌تر تصویر را درک می‌کند. این رویکرد در وظایفی مانند طبقه‌بندی تصاویر، تشخیص اشیاء و حتی تولید تصاویر جدید، اثربخش ثابت شده است.

<sup>۱۶</sup>Patches

## ۶-۴ مدل ترانسفورمر کاملاً کانولوشنی

مدل ترانسفورمر کاملاً کانولوشنی<sup>۱۷</sup> [۱۳]، به چالش‌های تقسیم‌بندی تصاویر پزشکی با ترکیب نقاط قوت شبکه‌های عصبی کانولوشن<sup>۱۸</sup> و ترانسفورمرها پرداخته است. مدل‌های سنتی بر CNN مانند UNet در گرفتن ویژگی‌های محلی عالی عمل می‌کنند اما به دلیل محلی بودن عملیات کانولوشن، در درک وابستگی‌های بلندمدت مشکل دارند. ترانسفورمرها، در حالی که در گرفتن وابستگی‌های بلندمدت عالی هستند، اغلب قادر به استفاده مؤثر از زمینه فضایی نیستند. FCT طراحی شده است تا با ادغام عملیات کانولوشنی با قابلیت‌های ترانسفورمر، این شکاف را پر کند. FCT اولین مدل ترانسفورمر کاملاً کانولوشنی در ادبیات تصویربرداری پزشکی است که هدف آن ترکیب مزایای هر دو CNN و ترانسفورمر بدون معایب آنها است.



شکل ۴-۵: معماری مدل ترانسفورمر کاملاً کانولوشنی

بلاک اصلی در مدل FCT ، لایه ترانسفورمر کاملاً کانولوشنی (لایه FCT) است. این لایه از دو بخش اصلی مازمول توجه کانولوشنی و مازمول Wide-Focus تشکیل شده است

<sup>17</sup>Fully Convolutional Transformer (FCT)

<sup>18</sup>CNN

• **ماژول توجه کانولوشنی:** از کانولوشن‌های عمقی به جای پروجکشن‌های خطی سنتی استفاده می‌کند تا زمینه فضایی را حفظ کند و نیاز به کدگذاری موقعیت را از بین ببرد.

• **ماژول Wide-Focus:** ماژول Wide-Focus برای بهبود درک مدل از تصاویر طراحی شده است. این ماژول از کانولوشن‌های دیلاته استفاده می‌کند. کانولوشن‌های دیلاته نوعی عملیات ریاضی هستند که به مدل اجازه می‌دهند به نواحی بزرگتری از تصویر نگاه کند بدون اینکه جزئیات را از دست بدهد. به طور ساده، این کانولوشن‌ها نقاطی از تصویر را که فاصله بیشتری دارند، بررسی می‌کنند تا مدل بتواند الگوهای بزرگ‌تر و ساختار کلی تصویر را ببیند.

ساختار مدل از معماری رمزگذار-رمزگشا پیروی می‌کند. رمزگذار شامل چهار لایه FCT است که هر کدام مسئول استخراج و انتشار ویژگی‌ها هستند. رمزگذار از ورودی تصویر چند مقیاسی به سبک هرم استفاده می‌کند تا کلاس‌ها و نواحی مورد علاقه را در مقیاس‌های مختلف برجسته کند. رمزگشا به صورت متقارن با رمزگذار، رمزگشا نقشه‌های تقسیم‌بندی را از نمایش فشرده دوباره نمونه‌برداری می‌کند. از اتصالات پرش از لایه‌های رمزگذار متناظر برای حفظ ارتباط متنی استفاده می‌کند و نقشه‌های تقسیم‌بندی میانی را برای نظارت اضافی خروجی می‌دهد.

مدل FCT به طور قابل توجهی عملکرد بهتری نسبت به مدل‌های مبتنی بر ترانسفورمر موجود در چندین مجموعه داده تقسیم‌بندی تصاویر پزشکی دارد. این مدل با داشتن پارامترهای کمتر، عملکرد بهتری را نشان می‌دهد و طراحی فشرده و کارآمد خود را ثابت می‌کند.

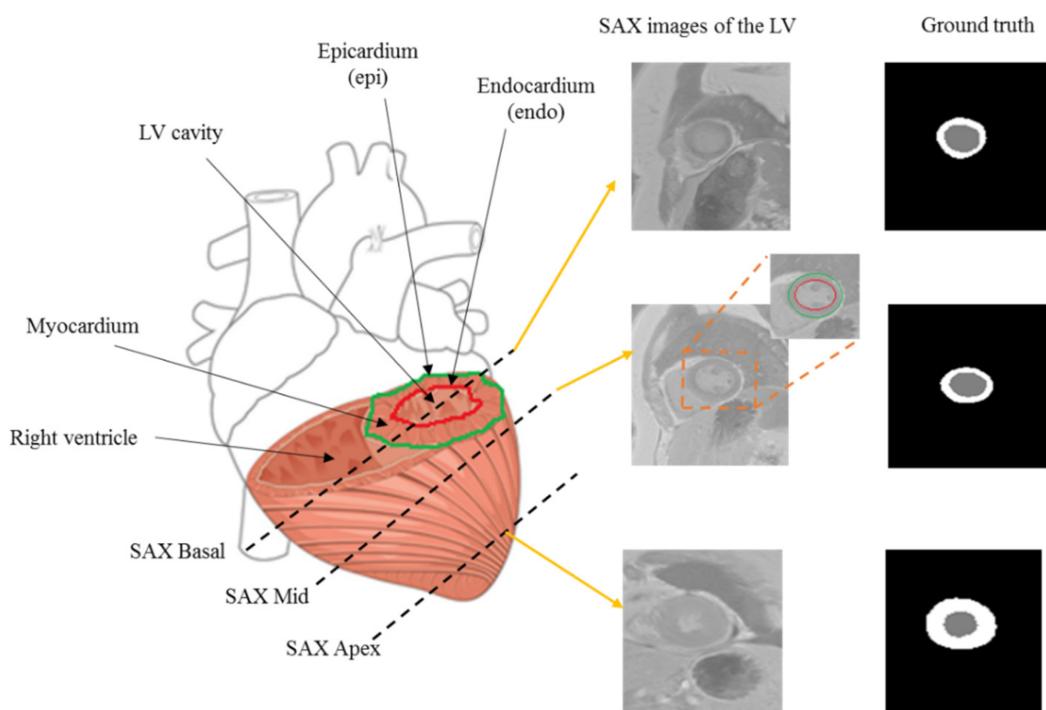
## فصل پنجم

### پیاده‌سازی

## ۱-۵ مجموعه داده

این پروژه بر روی داده‌های سینماتیک ام‌آرآی قلبی<sup>۱</sup> از چالش ACDC انجام شده است [۱۴]. مجموعه داده‌های ACDC شامل تصاویر تشدید مغناطیسی قلب با نمای کوتاه محور از ۱۵° بیمار است که در قالب NIfTI ضبط شده‌اند. این مجموعه داده به دو بخش آموزشی و آزمایشی تقسیم شده است که به ترتیب شامل ۱۰۰ و ۵۰ اسکن<sup>۲</sup> بعدی cine-CMR می‌باشند و ۲۸ تا ۴۰ استپ زمانی به طور کامل یا جزئی چرخه قلب را پوشش می‌دهد و هر استپ زمانی شامل یک اسکن کامل از تصاویر محور کوتاه قلب است که با توجه به اندازه قلب هر بیمار تعداد اسلاسیس‌های متفاوتی دارد. برای مثال دیتای ۴ بعدی بیمار با شناسه ۲۱، شامل ۳۰ استپ زمانی است که هر استپ تمام قلب را با ۱۰ اسلاسیس ۲۲۴ در ۲۵۶ پوشش می‌دهند.

تصاویر با استفاده از دو اسکنر Siemens Trio Tim (3T) و CMR: Siemens Area (1.5T)، در طی شش سال جمع‌آوری شده‌اند. ماسک‌های تقسیم‌بندی شده توسط کارشناسان بالینی برای میوکارد، بطن چپ و بطن راست در مراحل انتهای انقباض و انتهای انبساط ارائه شده است.



شکل ۱-۵: نمونه‌ای از تصاویر محور کوتاه قلب به همراه تصاویر قطعه‌بندی شده

<sup>1</sup>cine-MRI

این مجموعه داده شامل توزیع متعادلی از بیماران با تشخیص‌های سکته قلبی قبلی<sup>۲</sup>، نارسایی قلبی<sup>۳</sup>، بیش‌پروردگی ماهیچه قلب<sup>۴</sup>، بطن راست غیرطبیعی<sup>۵</sup> و دادگان طبیعی<sup>۶</sup> می‌باشد. بیماران با تشخیص سکته قلبی قبلی، دارای کسری جهشی<sup>۷</sup> بطن چپ کمتر از ۴۰٪ و چندین بخش میوکارد با انقباض غیر طبیعی هستند. در بیماران نارسایی قلبی، حجم بطن چپ در مرحله انقباض بالاتر از ۱۰۰ میلی‌لیتر بر متر مربع و کسری جهشی بطن چپ کمتر از ۴۰٪ است. بیماران با بیش‌پروردگی ماهیچه قلب، جرم قلبی بطن چپ آن‌ها بالاتر از ۱۱۰ گرم بر متر مربع است. این بیماران در چندین بخش میوکارد، ضخامت بیشتر از ۱۵ میلی‌متر در مرحله انقباض دارند و کسری جهشی آن‌ها طبیعی است. و در نهایت، بیماران با بطن راست غیرعادی، حجم حفره بطن راست آن‌ها بالاتر از ۱۱۰ میلی‌لیتر بر متر مربع است و یا کسری جهشی بطن راست آن‌ها کمتر از ۴۰٪ می‌باشد.

این ترکیب متعادل، نمای کاملی از شرایط مختلف قلبی را فراهم می‌کند که به استحکام تحلیل‌های تجربی کمک می‌کند.

## ۲-۵ معماری مدل پیاده‌سازی شده

ابتدا داده‌های هر بیمار که شامل اسلایس‌های قلب در دو مرحله انقباض و انبساط هستند برای آموزش مدل تقسیم‌بند استفاده می‌شوند. مدل تقسیم‌بندی مورد استفاده در این پروژه، مدل The Fully Convolutional Transformer (FCT) است. FCT، به عنوان اولین مدل ترانسفورمر کاملاً کانولوشنی در تصویربرداری پزشکی، با ادغام قابلیت‌های CNN و ترانسفورمر طراحی شده است تا هم در استخراج ویژگی‌های محلی توانمند باشد و هم وابستگی‌های بلندمدت را به خوبی درک کند. پس از تقسیم‌بندی توسط FCT، داده‌ها به دو مسیر برای استخراج ویژگی تقسیم می‌شوند: یکی از طریق کدگذار خودکار تغییراتی<sup>۸</sup> و دیگری از طریق استخراج ویژگی‌های دستی.

<sup>2</sup>MINF

<sup>3</sup>DCM

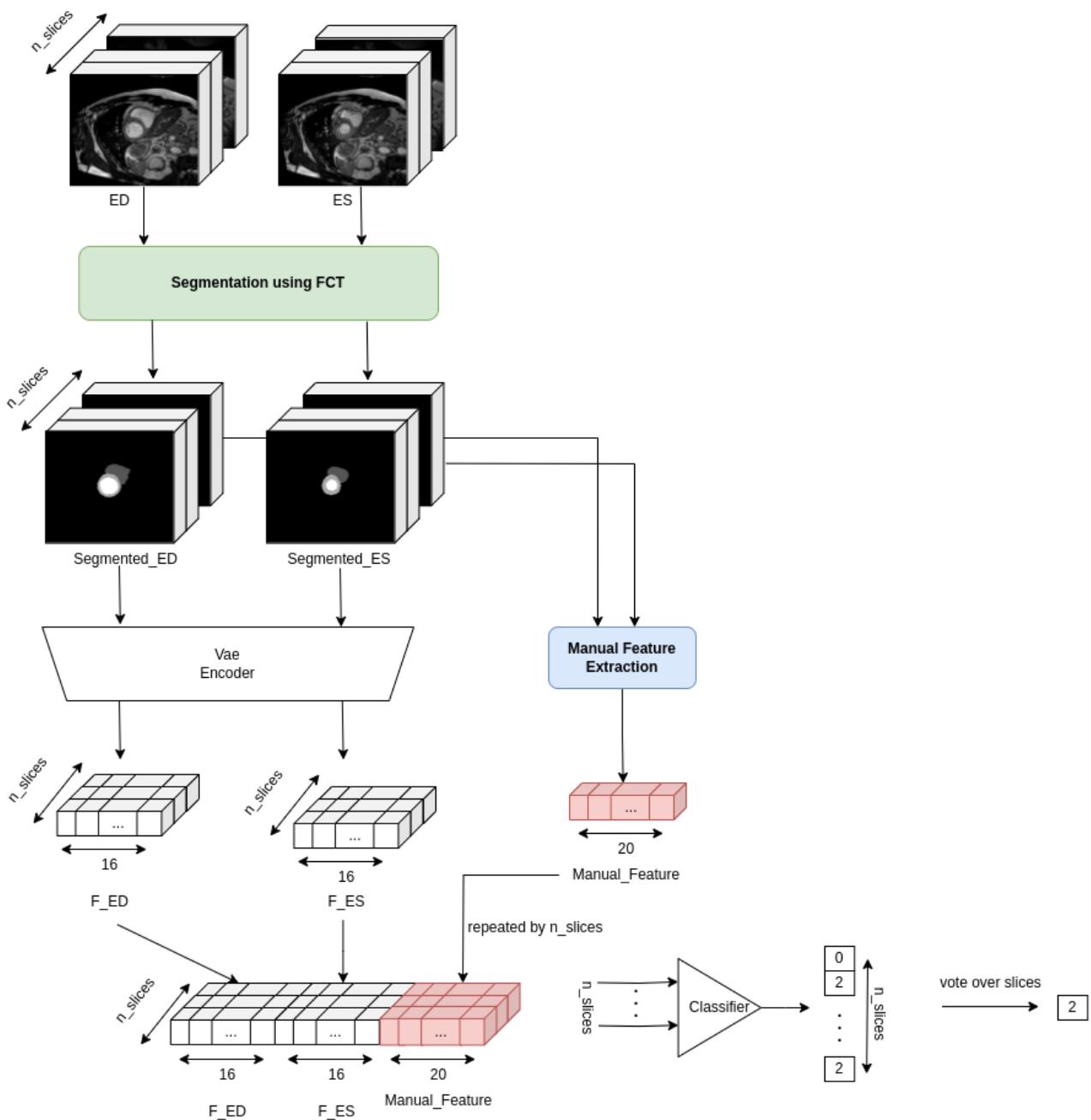
<sup>4</sup>HCM

<sup>5</sup>RV

<sup>6</sup>NOR

<sup>7</sup>Ejection fraction

<sup>8</sup>Variational Autoencoder



شکل ۲-۵: معماری مدل‌های استفاده شده در این پروژه

برای آموزش کدگذار خودکار تغییراتی از داده‌های تقسیم‌بندی شده از خروجی مدل FCT استفاده می‌شود. همانند مدل FCT، اسلایس‌های هر دو استپ زمانی ED و ES جداگانه وارد مدل می‌شوند. از تمام اسلایس‌های تقسیم‌بندی شده دو استپ زمانی ED و ES مجموعه‌ای از ویژگی‌های دستی برای هر بیمار بدست می‌آید.

در طرف دیگر برای استخراج ویژگی‌های یادگیری عمیق، به دلیل حجم کم داده‌ها، نمی‌توان از تمام

اسلایس‌های هر دو مرحله یکجا استفاده کرد. از این رو برای هر بیمار اسلایس‌های متناظر از دو مرحله VAE وارد مدل ES شده و ویژگی‌های نقطه انسداد این مدل به عنوان ویژگی‌های یادگیری عمیق استخراج می‌شود. که در نهایت  $2 \times 16$  ویژگی بدست می‌آید.

برای مدل طبقه‌بند، ۳۲ ویژگی یادگیری عمیق و ۲۰ ویژگی دستی با هم ترکیب شده و مدل را آموزش می‌دهند. این مدل همانطور که در بالا توضیح داده شد روی اسلایس‌های متناظر ED و ES آموزش می‌بیند نه بر روی بیمارها. از این رو در هنگام تست داده هر بیمار که شامل تعدادی اسلایس متناظر در دو استپ زمانی ED و ES است وارد طبقه‌بند شده و بین خروجی‌ها آن کلاسی که بیشترین تشخیص را دارد به عنوان کلاس خروجی انتخاب می‌شود.

مدل طبقه‌بندی یک یک رأی‌گیری کلاس‌بند<sup>۹</sup> است. این کلاس‌بند با ترکیب تصمیمات حاصل از یک پرسپترون چند لایه<sup>۱۰</sup>، یک جنگل تصادفی<sup>۱۱</sup> و یک کلاس SVM، که با استفاده از GridSearch بهینه‌سازی شده‌اند، تصمیم نهایی را می‌گیرد تا با افزایش دقت و استحکام پیش‌بینی‌ها، بهترین برچسب کلاس را برای تصویر تقسیم‌بندی شده تعیین کند. برای تست مدل طبقه‌بند دیتای تست که هر کدام از تعداد اسلایس‌های متفاوتی تشکیل شده وارد طبقه‌بند شده و بین خروجی اسلایس‌ها آن کلاسی که بیشترین تشخیص را دارد به عنوان کلاس خروجی انتخاب می‌شود.

## ۱-۲-۵ آزمایشات با دیگر مدل‌های پیچیده برای استخراج ویژگی‌های یادگیری عمیق

در این پژوهش، تلاش‌هایی برای به کارگیری از دیگر مدل‌های پیشرفته‌ی یادگیری عمیق از جمله Swin، VGG16، Transformer، شبکه‌های عصبی کانولوشنی سنتی (CNNs) و نقطه انسداد مدل‌های همچون UNet و ترانسفورمر کاملاً کانولوشنی (FCT) انجام شد تا ویژگی‌های استخراج شده توسط این مدل‌ها در تشخیص بیماری‌ها بکار گرفته شوند. با این حال، موانع قابل توجهی در استفاده از این مدل‌های عمیق در مجموعه داده‌های کوچک مورد استفاده در این پژوهه مشاهده شد. این مدل‌ها به دلیل پیچیدگی‌های محاسباتی بالا و نیاز به داده‌های زیاد برای آموزش مؤثر، در شرایطی که تنها تعداد محدودی داده در دسترس بود، نتوانستند به خوبی عمل کنند. نتیجه این شد که مدل‌های مذکور دچار بیش‌برازش شده و توانایی تعمیم‌دهی به داده‌های جدید را نداشتند.

<sup>9</sup>Voting Classifier

<sup>10</sup>Multilayer perceptron (MLP)

<sup>11</sup>Random Forest

در مقابل، استفاده از ویژگی‌های استخراج شده از فضای نهانی VAE نتایج بسیار بهتری ارائه داد. این مدل توانست ۱۶ ویژگی کلیدی را از تصاویر ارائه دهد که نه تنها فشرده و معنadar بودند، بلکه توانایی بالایی در آموزش دقیق‌تر طبقه‌بندها داشتند.

### ۳-۵ استخراج ویژگی‌های دستی

پس از تقسیم‌بندی ساختارهای قلبی از تصاویر ام‌آرآی، فرآیند استخراج ویژگی‌ها آغاز می‌شود. در این پروژه از روش استخراج ویژگی گفته شده در مقاله [۴] استفاده می‌شود. فرایند استخراج ویژگی‌ها به دو دسته ویژگی‌های اولیه و ویژگی‌های مشتق شده تقسیم می‌شود. ویژگی‌های اولیه مستقیماً از تصاویر تقسیم‌بندی شده بدست می‌آیند و شامل حجم بطن چپ<sup>۱۲</sup> در مراحل انبساط<sup>۱۳</sup> و انقباض<sup>۱۴</sup>، حجم بطن راست<sup>۱۵</sup> و جرم و حجم میوکارد<sup>۱۶</sup> هستند. علاوه بر این، ضخامت دیواره میوکارد در هر برش نیز از این بخش‌ها استخراج می‌شود.

ویژگی‌های مشتق شده از ترکیب ویژگی‌های اولیه حاصل می‌شوند و شامل کسر جهشی برای LV و RV، و نسبت‌های مختلف مانند نسبت LV به RV و MYO به LV در مراحل ED و ES هستند. همچنین پروفایل تغییرات ضخامت دیواره میوکارد (MWT) در محورهای کوتاه (SA) نیز مورد بررسی قرار می‌گیرد. این فرآیند شامل بخش‌بندی میوکارد برای تعریف مرزهای لازم برای محاسبه MWT است. از تکنیک‌های تشخیص لبه Canny برای تعیین دقیق کانتورهای داخلی و خارجی میوکارد استفاده می‌شود. MWT محاسبه کوتاه‌ترین فاصله اقلیدسی بین پیکسل‌ها در کانتورهای داخلی و خارجی برای هر برش از میوکارد است. که به صورت ریاضی به شکل زیر بیان می‌شود:

$$MWT|_{SA} = \left\{ \min_{e \in E} d(i, e) : i \in I \right\}$$

که در آن  $I$  و  $E$  مجموعه‌هایی از پیکسل‌ها روی کانتورهای داخلی و خارجی هستند، و  $d(i, e)$  نشان‌دهنده فاصله بین پیکسل‌های  $i$  و  $e$  است.

از تجزیه و تحلیل آماری MWT از جمله محاسبه میانگین و انحراف معیار در سراسر برش‌ها و فازهای

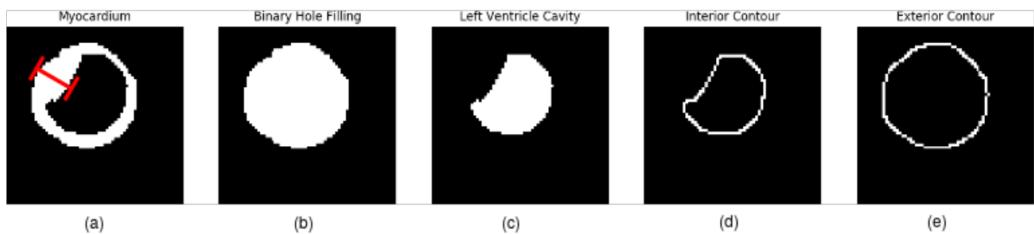
<sup>12</sup>LV

<sup>13</sup>End Diastolic phase (ED)

<sup>14</sup>End Systolic phase(ES)

<sup>15</sup>RV

<sup>16</sup>MYO



شکل ۵-۳: روش تخمین ضخامت دیواره میوکارد در هر برش محور کوتاه در مقاله [۴]

قلبی (فاز پایانی انبساط و فاز پایانی انقباض) هشت ویژگی برای سنجش پروفایل تغییرات MWT در طول چرخه قلبی استخراج می‌شود. این ویژگی‌ها نشان‌دهنده تغییرات یکنواختی MWT در برش‌ها و فازهای مختلف هستند و بینش‌های حیاتی در مورد سلامت و عملکرد میوکارد ارائه می‌دهند. این تجزیه و تحلیل امکان شخصیت‌پردازی دقیق ساختار میوکاردی را فراهم می‌کند، که برای تشخیص و درک شرایط مختلف قلبی ضروری است.

## ۴-۵ نرمال‌سازی داده‌ها

نرمال‌سازی داده‌ها یکی از مراحل اساسی در پیش‌پردازش داده‌ها برای مدل‌های یادگیری ماشین است. این کار به خصوص زمانی که ویژگی‌ها دارای مقیاس‌های متفاوت هستند اهمیت بیشتری دارد. در این پروژه از RobustScaler برای نرمال‌سازی ویژگی‌های دستی<sup>۱۷</sup> که شامل داده‌های پزشکی هستند، استفاده شده است.

RobustScaler ویژگی‌ها را به صورتی نرمال‌سازی می‌کند که تأثیر نقاط داده‌ای دور از مرکز<sup>۱۸</sup> کمتر می‌شود. این کار با استفاده از میانه<sup>۱۹</sup> و فاصله بین کوارتیل‌ها<sup>۲۰</sup> به جای میانگین و انحراف معیار انجام می‌گیرد. فرمول نرمال‌سازی به شکل زیر است:

$$X_{\text{scaled}} = \frac{X - \text{Median}(X)}{\text{IQR}(X)}$$

که در آن:  $X$  مقدار اصلی ویژگی،  $\text{Median}(X)$  میانه مقادیر ویژگی و  $\text{IQR}(X)$  فاصله بین کوارتیل اول  $Q1$  و کوارتیل سوم  $Q3$  است، یعنی  $\text{IQR} = Q3 - Q1$

<sup>17</sup>Manual features

<sup>18</sup>Outliers

<sup>19</sup>Median

<sup>20</sup>Interquartile Range (IQR)

استفاده از RobustScaler برای نرمال‌سازی ویژگی‌ها به دلیل مقاومت آن در برابر داده‌های پرت بسیار محبوب است. در داده‌های پزشکی، جایی که اغلب با داده‌های دارای توزیع نامتقارن و داده‌های پرت مواجه هستیم، این نوع نرمال‌سازی می‌تواند به مدل‌های یادگیری ماشین کمک کند تا دقیق‌تر بیشتری داشته باشند. علاوه بر این، RobustScaler با کاهش تأثیر داده‌های پرت بر محاسبات، می‌تواند به حفظ توزیع طبیعی‌تر ویژگی‌ها در مدل کمک کند، بنابراین به بهبود قابلیت تعمیم مدل‌ها منجر می‌شود.

## ۵-۵ معیارهای ارزیابی

### ۱-۵-۵ ارزیابی مدل تقسیم‌بندی با معیار دایس

معیار دایس<sup>۲۱</sup> برای سنجش دقیق تقسیم‌بندی تصاویر استفاده می‌شود. این معیار از طریق محاسبه دو برابر تعداد پیکسل‌های مشترک بین پیش‌بینی و برچسب‌های واقعیت به تعداد کل پیکسل‌های موجود در هر دو مجموعه محاسبه می‌شود:

$$\text{Dice} = \frac{2 \times |X \cap Y|}{|X| + |Y|}$$

که در آن  $X$  مجموعه پیکسل‌های پیش‌بینی شده و  $Y$  مجموعه پیکسل‌های واقعی است. نتایج این معیار در پروژه نشان داد که مدل بخش‌بندی به خوبی توانسته است ساختارهای قلبی را شناسایی و دسته‌بندی کند.

### ۲-۵-۵ ارزیابی مدل دسته‌بندی با دقیقیت، دقیقیت تفکیکی، حساسیت و امتیاز F1

در مدل دسته‌بندی، معیارهای زیر برای ارزیابی استفاده شده‌اند:

- دقیقیت (Accuracy): نسبت تعداد پیش‌بینی‌های صحیح به کل پیش‌بینی‌ها.

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Population}}$$

<sup>21</sup>Dice Coefficient

- دقت تفکیکی (Precision): نسبت پیش‌بینی‌های صحیح مثبت به کل پیش‌بینی‌های مثبت.

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}}$$

- حساسیت (Recall): نسبت پیش‌بینی‌های صحیح مثبت به کل موارد واقعی مثبت.

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}}$$

- امتیاز (F1 Score): میانگین هارمونیک دقت تفکیکی و حساسیت.

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}}$$

این معیارها به ترتیب نشان دهنده توانایی مدل در تشخیص دقیق داده‌های مثبت و تعادل بین دقت تفکیکی و حساسیت هستند. نتایج حاصل از این معیارها نشان می‌دهد که مدل دسته‌بندی به خوبی توانسته است بیماری‌های قلبی مختلف را با دقت بالا شناسایی کند.

## ۶-۵ کتابخانه‌ها و چارچوب‌های مورد استفاده

پیاده‌سازی این پروژه بر روی چارچوب پایتورچ<sup>۲۲</sup> انجام شده است. از مهم ترین کتابخانه‌های استفاده شده می‌توان به لایتنینگ<sup>۲۳</sup>، نامپای<sup>۲۴</sup>، پانداز<sup>۲۵</sup> و مت‌پلات لیب<sup>۲۶</sup> اشاره کرد. برای کار با داده‌های سینماتیک ام‌آرآی از کتاب خانه نیبابل<sup>۲۷</sup> استفاده شده است که ویژگی‌های زیادی در اختیار توسعه دهنده می‌دهد.

<sup>22</sup>Pytorch

<sup>23</sup>Lightning

<sup>24</sup>Numpy

<sup>25</sup>Pandas

<sup>26</sup>Matplotlib

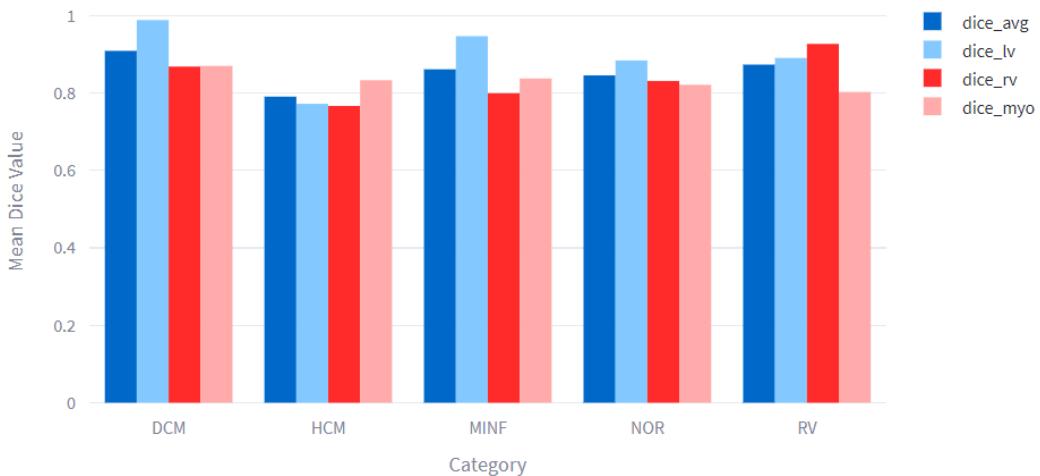
<sup>27</sup>Nibabel

## فصل ششم

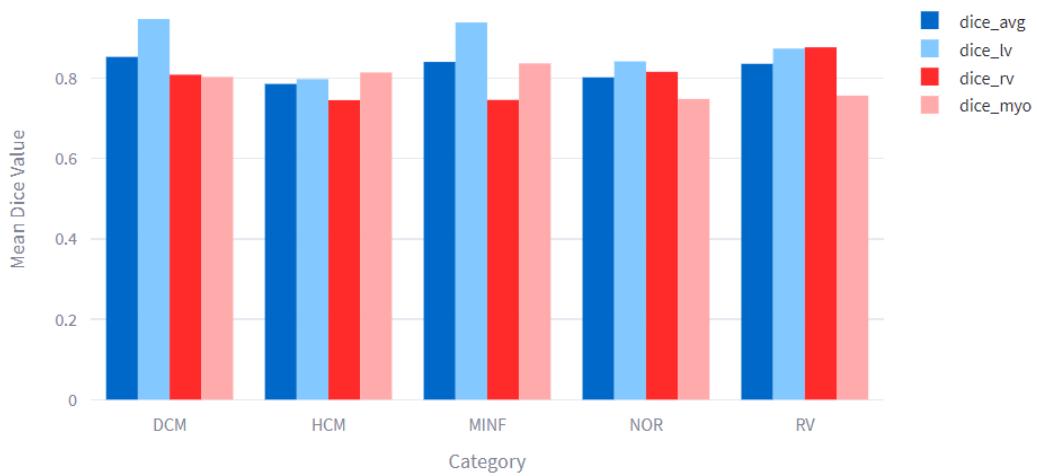
## نتایج و ارزیابی

## ۱-۶ تقسیم‌بندی

عملکرد مدل در تقسیم‌بندی با استفاده از ضرایب شbahت دایس برای پنج گروه از بیماری‌های قلبی عروقی در دادگان پروژه در فازهای آموزش و آزمایش به صورت کمی ارزیابی شد.



شکل ۱-۶: نمودار میله‌ای مقادیر دایس برای دسته‌های پنج‌گانه در فاز آموزش



شکل ۲-۶: نمودار میله‌ای مقادیر دایس برای دسته‌های پنج‌گانه در فاز آزمایش

نمودارهای میله‌ای شکل ۱-۶ و ۲-۶ به‌وضوح مقادیر دایس را برای هر دسته نشان می‌دهند که از عملکرد نسبتاً یکنواختی در تقسیم‌بندی ساختارهای مختلف قلبی خبر می‌دهد.

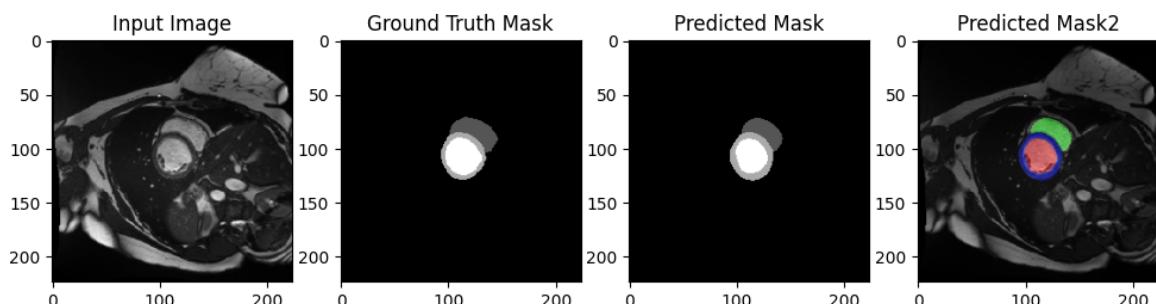
phase	dice_myo	dice_rv	dice_lv
ES	۸۱.۰	۸۲.۰	۸۶.۰
ED	۸۵.۰	۸۹.۰	۹۸.۰

جدول ۶-۱: میانگین ضرایب دایس برای بطن راست RV، بطن چپ LV، و میوکارد MYO در دو فاز قلبی، انبساط و انقباض در فاز آموزش

phase	dice_myo	dice_rv	dice_lv
ES	۷۸.۰	۷۲.۰	۸۶.۰
ED	۷۸.۰	۷۴.۰	۹۳.۰

جدول ۶-۲: میانگین ضرایب دایس برای بطن راست RV، بطن چپ LV، و میوکارد MYO در دو فاز قلبی، انبساط و انقباض در فاز آزمایش

میانگین ضرایب دایس برای دو فاز قلبی، انبساط (ES) و انقباض (ED) گزارش شده‌اند که دقت بالاتری در فاز ED نسبت به فاز ES نشان می‌دهند. این تفاوت ممکن است به ویژگی‌های قابل تمایز قلب در این فازها نسبت داده شود.

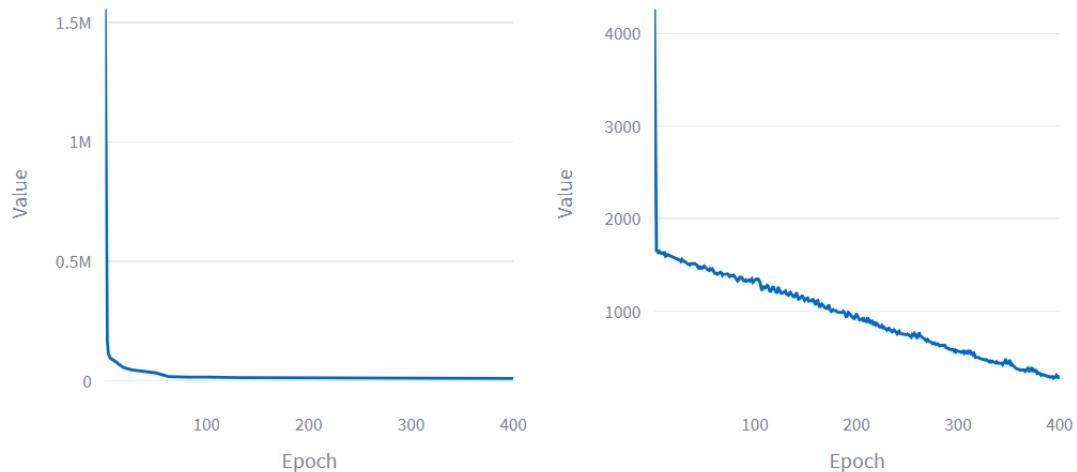


شکل ۶-۳: نمونه‌ای از تصویر تقسیم‌بندی شده توسط مدل FCT و مقایسه آن با ماسک تقسیم‌بندی شده توسط کارشناسان بالینی

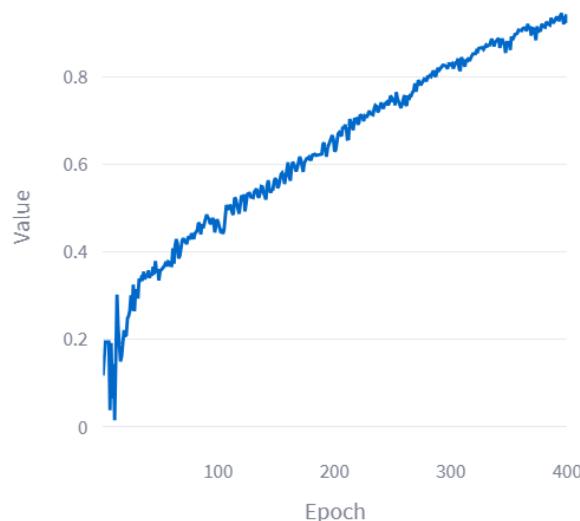
همچنین نمونه‌ای از تصویر تقسیم‌بندی شده توسط مدل FCT و مقایسه آن با ماسک تقسیم‌بندی شده توسط کارشناسان بالینی در شکل ۶-۳ قابل مشاهده است.

## ۲-۶ کدگذار خودکار تغییراتی

فاز آموزش کدگذار خودکار تغییراتی از طریق نمودارهای خطی که کاهش کلی ضرر در هر ایپاک<sup>۱</sup> را در طول دوره‌های آموزشی نشان می‌دهند، تفصیل داده شده است (شکل ۴-۶).

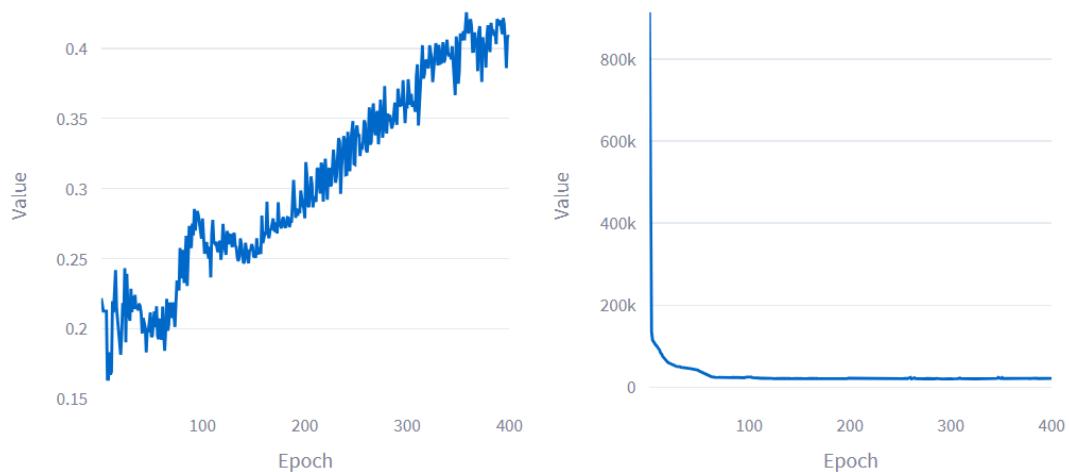


شکل ۴-۶: نمودار سمت چپ کاهش ضرر کلی مدل کدگذار خودکار تغییراتی، نمودار سمت راست کاهش ضرر بخش کلاس‌بندی در مدل کدگذار خودکار تغییراتی



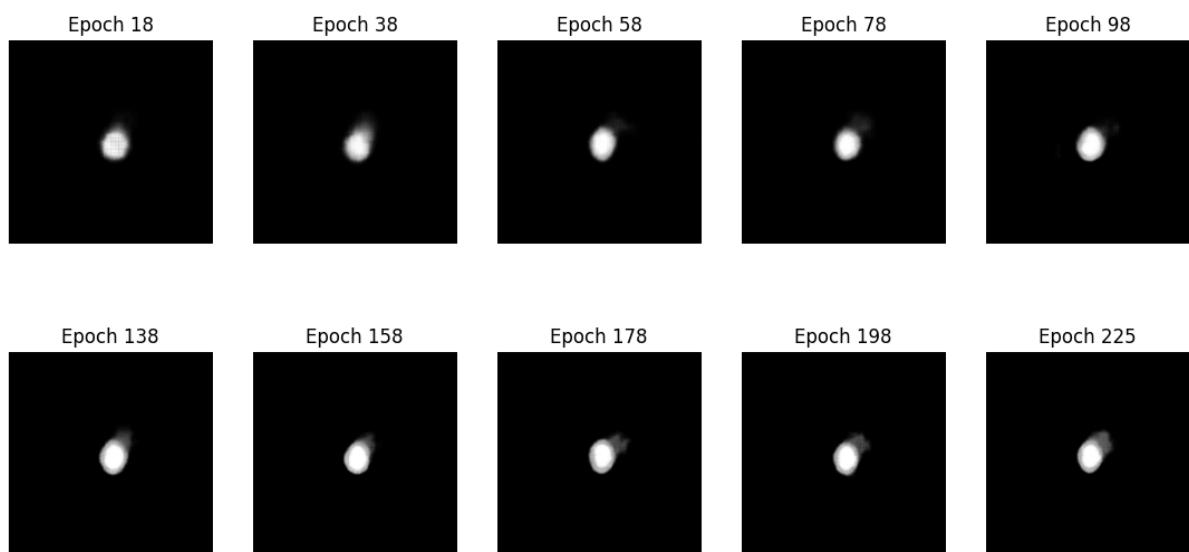
شکل ۵-۶: نمودار خطی دقت کلاس‌بندی در فاز آموزش

<sup>1</sup>Epoch



شکل ۶-۶: نمودار سمت چپ دقت کلاس‌بندی در فاز اعتبارسنجی، نمودار سمت راست کاهش ضرر کلی مدل کدگذار خودکار تغییراتی در فاز اعتبارسنجی

این نمودارها نشان‌دهنده کاهش تدریجی ضرر هستند که یادگیری مؤثر و همگرایی مدل را نشان می‌دهد. ضرر اعتبارسنجی نیز نشان داده شده است (شکل ۶-۶) که روندی موازی با هزینه آموزشی دارد و توانایی‌های تعمیم مدل بر روی داده‌های دیده‌نشده را تأیید می‌کند.

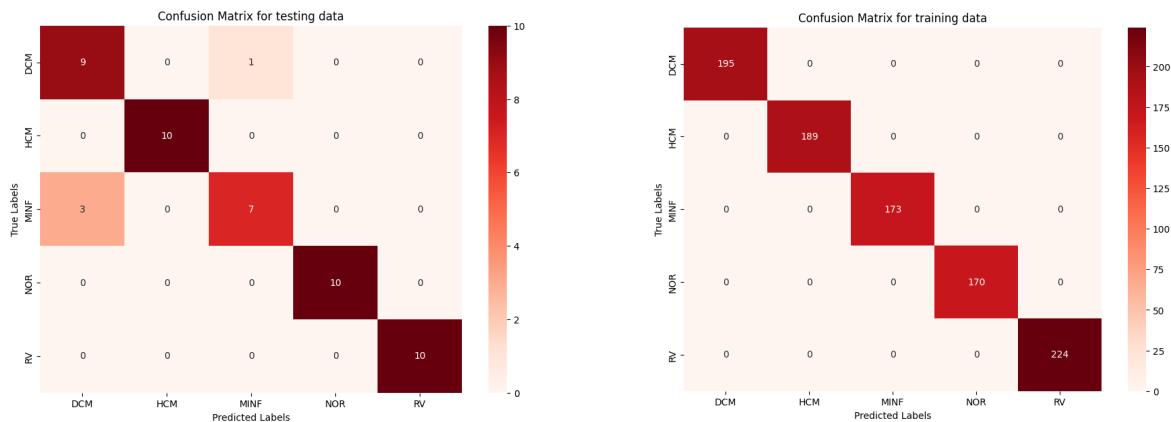


شکل ۷-۶: نمایش بصری از خروجی‌های مدل VAE در طول ایپاک‌های مختلف

همچنین، نمایش‌های بصری از خروجی‌های مدل در طول ایپاک‌های مختلف (شکل ۶-۷) بهبود تدریجی آن در بازسازی دقیق تصاویر قلبی را نشان می‌دهد و اثربخشی رمزگذاری فضای پنهان در به دست آوردن ویژگی‌های ضروری را تأیید می‌کند.

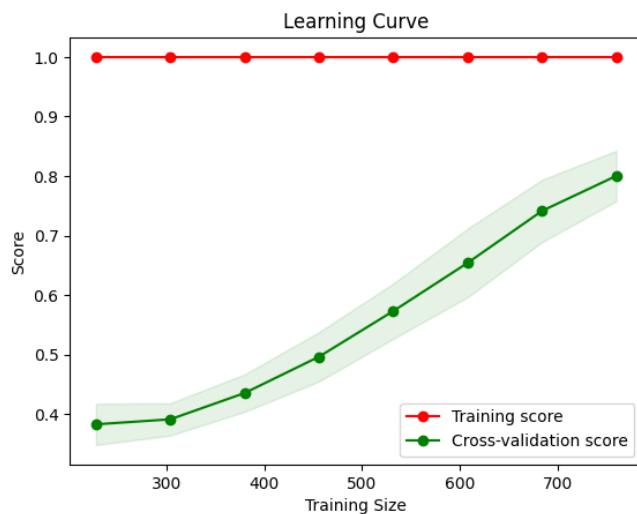
## ۳-۶ طبقه‌بندی

مرحله طبقه‌بندی با استفاده از ماتریس درهم‌ریختگی<sup>۲</sup> برای داده‌های آموزشی و آزمایشی ارزیابی شد (شکل ۶-۸ و ۶-۹) که نمایی دقیق از عملکرد مدل در برابر دسته‌های مختلف فراهم می‌کند.



شکل ۶-۹: ماتریس درهم‌ریختگی در داده‌های آزمایشی

شکل ۶-۸: ماتریس درهم‌ریختگی در داده‌های آموزشی



شکل ۶-۱۰: نمودار یادگیری در فاز آموزش

نمودار یادگیری<sup>۳</sup> برای فاز آموزش (شکل ۶-۱۰) کاهش پیوسته ضرر را با افزایش داده‌های آموزشی نشان می‌دهند.

<sup>2</sup>Confusion matrix

<sup>3</sup>Learning curve

همچنین دقت، دقت تفکیکی، حساسیت و مقادیر F1 برای مجموعه داده آزمایش در جدولی ارائه شده‌اند (جدول ۳-۶) که نمای کاملی از عملکرد مدل را ارائه می‌دهد.

Class	Precision	Recall	F1-Score	Support
DCM	۷۵.۰	۹۰.۰	۸۲.۰	۱۰
HCM	۱۰۰	۱۰۰	۱۰۰	۱۰
MINF	۸۸.۰	۷۰.۰	۷۸.۰	۱۰
NOR	۱۰۰	۱۰۰	۱۰۰	۱۰
RV	۱۰۰	۱۰۰	۱۰۰	۱۰
Accuracy			۹۲.۰	۵۰
Macro Avg	۹۳.۰	۹۲.۰	۹۲.۰	۵۰
Weighted Avg	۹۳.۰	۹۲.۰	۹۲.۰	۵۰

جدول ۳-۶: معیارهای عملکرد طبقه‌بندی برای مجموعه داده‌های آزمایش

## فصل هفتم

### پیاده‌سازی سامانه نهایی

در مرحله پیاده‌سازی نهایی این پایان‌نامه، یک برنامه کاربردی وب محور توسعه یافت تا یک رابط کاربری تعاملی برای تشخیص بیماری‌های قلبی-عروقی با استفاده از مدل‌های یادگیری عمیق فراهم آورد. انتخاب یک برنامه وب محور با توجه به افزایش حضور چنین پلتفرم‌هایی در محیط‌های بهداشتی و درمانی به دلیل دسترسی آسان و کاربردی بودن آن‌ها صورت گرفت. برای ساخت این برنامه از Streamlit استفاده شد، که یک چارچوب برنامه کاربردی منبع باز است و به سرعت امکان تبدیل اسکریپت‌های داده‌ای به برنامه‌های وب قابل به اشتراک‌گذاری را فراهم می‌آورد. سیستم نهایی به عنوان ابزار جامعی برای تشخیص و تجزیه و تحلیل بیماری‌های قلبی-عروقی عمل می‌کند، که تکنیک‌های پیشرفته یادگیری عمیق را در یک برنامه وب محور کاربردی گنجانده است، و نمونه‌ای را ارائه می‌دهد که دسترسی به ابزارهای تشخیصی پیشرفته را بهبود می‌بخشد.

## Streamlit ۱-۷

برای تیم‌های یادگیری ماشین و علم داده طراحی شده است تا بتوانند به سرعت برنامه‌های سیک برای تحلیل داده‌ها ایجاد کنند. این چارچوب به توسعه‌دهندگان اجازه می‌دهد تا یک اسکریپت بنویسند و به طور تکراری آن را به روزرسانی کنند، با این حال برنامه به طور خودکار همزمان با تغییرات اسکریپت به روز می‌شود. این فریمورک توسعه رابط کاربری جلویی را تسهیل می‌بخشد و پیوند بین بکاند<sup>۱</sup> و یک رابط جذاب را ساده می‌سازد، که آن را برای این پروژه ایده‌آل می‌سازد.

## ۲-۷ طراحی برنامه

رابط کاربری مبتنی بر Streamlit این سیستم به چهار بخش اصلی تقسیم شده است که هر کدام برای تعامل بی‌درنگ با داده‌های کاربر طراحی شده‌اند و بینش‌هایی در مورد جنبه‌های مختلف فرآیند تشخیص قلبی-عروقی ارائه می‌دهند:

## ۱-۲-۷ تقسیم‌بندی

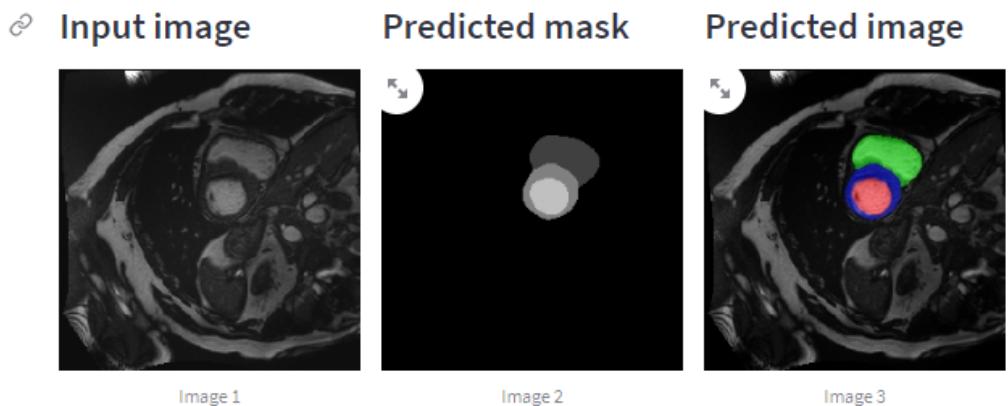
کاربران با بارگذاری فایل NIfTI خود شروع می‌کنند. سیستم این فایل را از طریق دو مدل تقسیم‌بندی یادگیری عمیق: FCT و UNet پردازش می‌کند.

<sup>1</sup> Back-end

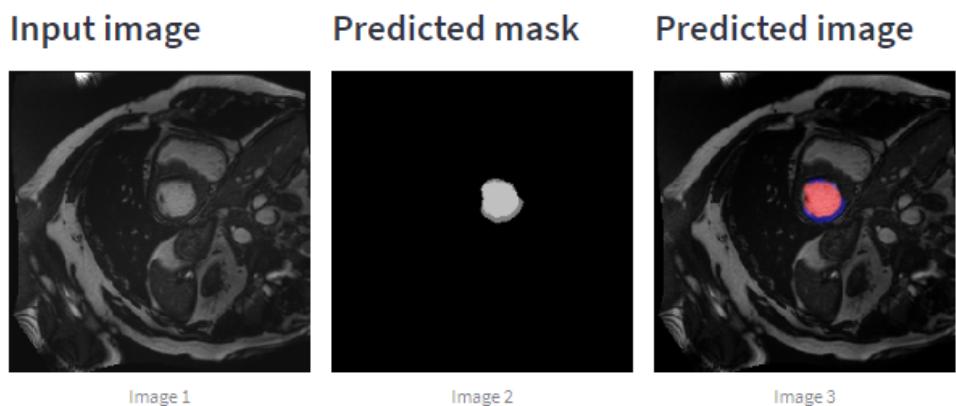
## Cine MRI Segmentation



### FCT model results:



### Unet model results:



شکل ۱-۷: صفحه اول سامانه تحت وب

رابط کاربری خروجی‌های تقسیم‌بندی شده از هر دو مدل را برای مقایسه کنار هم نمایش می‌دهد. توجه شده است که مدل FCT عملکرد به مراتب بهتری نسبت به مدل UNet دارد و نتایج تقسیم‌بندی واضح‌تر و دقیق‌تری را ارائه می‌دهد.

## ۲-۲-۷ استخراج ویژگی

کاربران دو فایل NIfTI که نمایانگر فازهای انبساط و انقباض از سینماتیک ام آر آی هستند را بارگذاری می‌کنند. سیستم ابتدا این تصاویر را تقسیم‌بندی کرده و سپس ویژگی‌های دستی مرتبط با مطالعه را استخراج می‌کند.

### Extract Cardiac Features

**Features dataframe:**

	Patient	mean_normal	std_normal
ED[vol(LV)]	223535.433506	99787.754702	30017.699022
ES[vol(LV)]	194206.983021	72027.665621	19491.544064
ED[vol(RV)]	111005.260597	119786.416804	30706.432784
ES[vol(RV)]	94857.901463	92222.532530	23140.633535
ED[mass(MYO)]	202980.451115	98418.959751	22663.323115
ES[vol(MYO)]	214093.510700	99765.489359	22411.988014
EF(LV)	13.120269	25.581513	14.264661
EF(RV)	14.546481	21.670710	14.063248
ED[vol(LV)/vol(RV)]	2.013737	0.834269	0.142996
ES[vol(LV)/vol(RV)]	2.047346	0.786322	0.102440
ED[mass(MYO)/vol(LV)]	0.908046	1.024432	0.200090
ES[vol(MYO)/vol(LV)]	1.102399	1.417453	0.223474
ES[max(mean(MWT SA) LA)]	12.065944	10.653821	1.180061
ES[stdev(mean(MWT SA) LA)]	2.244901	1.016220	0.472467
ES[mean(stdev(MWT SA) LA)]	1.410800	1.459389	0.263772
ES[stdev(stdev(MWT SA) LA)]	0.349633	0.365614	0.138745
ED[max(mean(MWT SA) LA)]	11.095671	8.883283	0.968156
ED[stdev(mean(MWT SA) LA)]	1.665670	0.800573	0.277635
ED[mean(stdev(MWT SA) LA)]	1.443342	1.400169	0.231311
ED[stdev(stdev(MWT SA) LA)]	0.264829	0.276529	0.121037

شکل ۲-۷: صفحه دوم سامانه تحت وب

در کنار استخراج ویژگی، سیستم میانگین و انحراف استاندارد مجموعه‌های داده نرمال را محاسبه می‌کند. سپس ویژگی‌های غیرعادی در داده‌های کاربر را بر جسته می‌کند، که به عنوان ویژگی‌هایی

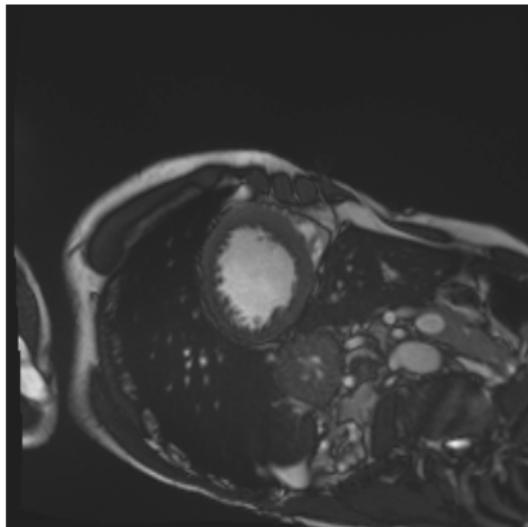
تعریف می‌شوند که از میانگین دو برابر انحراف استاندارد فاصله دارند.

### ۳-۲-۷ طبقه‌بندی

مشابه بخش استخراج ویژگی، کاربران دو فایل NIfTI را بارگذاری می‌کنند. سیستم از این داده‌ها برای تشخیص نوع بیماری قلبی که داده‌ها به آن تعلق دارند استفاده می‌کند، که شرایط را به یکی از پنج دسته مشخص شده طبقه‌بندی می‌کند.

## Diagnose Cardiac Abnormality

	ED[vol(LV)]	ES[vol(LV)]	ED[vol(RV)]	ES[vol(RV)]	ED[mass(MYO)]	ES[vol(MYO)]	EF(LV)	EF(RV)
Patient	223,535.4335	194,206.983	111,005.2606	94,857.9015	202,980.4511	214,093.5107	13.1203	14.1203



Predicted class: DCM

شکل ۳-۷: صفحه سوم سامانه تحت وب

نتایج تشخیص به صورت گرافیکی نمایش داده می‌شوند، که تفسیر آسان را برای کاربر فراهم می‌آورد.

## ۴-۲-۷ ارزیابی

این بخش آمار مختلف و مقایسه‌های تصویری را برای ارزیابی عملکرد مدل‌های پیاده‌سازی شده در بخش‌های مختلف پژوهش ارائه می‌دهد.

شامل معیارهای عملکرد برای هر مدل است، به علاوه خروجی‌های بصری که نواحی تقسیم‌بندی شده و نتایج طبقه‌بندی را نشان می‌دهند.

## فصل هشتم

### جمع‌بندی و نتیجه‌گیری

در این پژوهش، کاربرد یادگیری عمیق در تشخیص بیماری‌های قلبی عروقی با استفاده از تصاویر تشید مغناطیسی قلب بررسی شد. در فصل سوم، مطالعه‌ای دقیق بر روی کارهای مرتبط انجام گرفت. در بخش چهارم با معماری‌های مختلف برای تشخیص بیماری‌ها آشنا شدیم و به بررسی پیشینه و مفاهیم مرتبط با یادگیری عمیق پرداخته شد. همچنین رویکرد این پژوهش مورد توضیح قرار گرفت که شامل استفاده از مدل‌های پیشرفته یادگیری عمیق است. این پژوهه منجر به طراحی یک سامانه شد که به کاربران امکان می‌دهد تا خروجی‌های مختلف مدل را با داده‌های خود تست و ارزیابی کنند. نتایج به دست آمده نشان داد که یادگیری عمیق قابلیت ارزشمندی در تحلیل دقیق تصاویر قلبی دارد و می‌تواند به عنوان یک رویکرد مکمل در کنار سایر روش‌های تشخیصی به کار رود. برای کارهای آتی، توسعه و بهینه‌سازی مدل‌های یادگیری عمیق با استفاده از حجم بیشتری از داده‌های بالینی، تلاش برای کاهش خطاهای تشخیصی و بهبود قابلیت‌های تعمیم دهنده به نمونه‌های متنوع‌تر پیشنهاد می‌شود.

# كتاب نامه

- [1] Wolterink, Jelmer M, Leiner, Tim, Viergever, Max A, and Isgum, Ivana. Automatic segmentation and disease classification using cardiac cine mr images. in Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers 8, pp. 101–110. Springer, 2018.
- [2] Isensee, Fabian, Jaeger, Paul F, Full, Peter M, Wolf, Ivo, Engelhardt, Sandy, and Maier-Hein, Klaus H. Automatic cardiac disease assessment on cine-mri via time-series segmentation and domain specific features. in Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers 8, pp. 120–129. Springer, 2018.
- [3] Dosovitskiy, Alexey, Beyer, Lucas, Kolesnikov, Alexander, Weissenborn, Dirk, Zhai, Xiaohua, Unterthiner, Thomas, Dehghani, Mostafa, Minderer, Matthias, Heigold, Georg, Gelly, Sylvain, et al. An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929, 2020.
- [4] Khened, Mahendra, Kollerathu, Varghese Alex, and Krishnamurthi, Ganapathy. Fully convolutional multi-scale residual densenets for cardiac segmentation and

automated cardiac diagnosis using ensemble of classifiers. *Medical image analysis*, 51:21–45, 2019.

- [5] Jafari, Mahboobeh, Shoeibi, Afshin, Khodatars, Marjane, Ghassemi, Navid, Moridian, Parisa, Alizadehsani, Roohallah, Khosravi, Abbas, Ling, Sai Ho, Delfan, Niloufar, Zhang, Yu-Dong, et al. Automated diagnosis of cardiovascular diseases from cardiac magnetic resonance imaging using deep learning models: A review. *Computers in Biology and Medicine*, p. 106998, 2023.
- [6] Janik, Adrianna, Dodd, Jonathan, Ifrim, Georgiana, Sankaran, Kris, and Curran, Kathleen. Interpretability of a deep learning model in the application of cardiac mri segmentation with an acdc challenge dataset. in *Medical Imaging 2021: Image Processing*, vol. 11596, pp. 861–872. SPIE, 2021.
- [7] Assomull, Ravi G, Pennell, Dudley J, and Prasad, Sanjay K. Cardiovascular magnetic resonance in the evaluation of heart failure. *Heart*, 93(8):985–992, 2007.
- [8] Jamart, Kevin, Xiong, Zhaohan, Maso Talou, Gonzalo D, Stiles, Martin K, and Zhao, Jichao. Mini review: Deep learning for atrial segmentation from late gadolinium-enhanced mrис. *Frontiers in Cardiovascular Medicine*, 7:86, 2020.
- [9] Cetin, Irem, Sanroma, Gerard, Petersen, Steffen E, Napel, Sandy, Camara, Oscar, Ballester, Miguel-Angel Gonzalez, and Lekadir, Karim. A radiomics approach to computer-aided diagnosis with cardiac cine-mri. in *Statistical Atlases and Computational Models of the Heart. ACDC and MMWHS Challenges: 8th International Workshop, STACOM 2017, Held in Conjunction with MICCAI 2017, Quebec City, Canada, September 10-14, 2017, Revised Selected Papers* 8, pp. 82–90. Springer, 2018.
- [10] Kingma, Diederik P, Welling, Max, et al. An introduction to variational autoencoders. *Foundations and Trends® in Machine Learning*, 12(4):307–392, 2019.

- [11] Ronneberger, Olaf, Fischer, Philipp, and Brox, Thomas. U-net: Convolutional networks for biomedical image segmentation. in Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18, pp. 234–241. Springer, 2015.
- [12] Vaswani, Ashish, Shazeer, Noam, Parmar, Niki, Uszkoreit, Jakob, Jones, Llion, Gomez, Aidan N, Kaiser, Łukasz, and Polosukhin, Illia. Attention is all you need. Advances in neural information processing systems, 30, 2017.
- [13] Tragakis, Athanasios, Kaul, Chaitanya, Murray-Smith, Roderick, and Husmeier, Dirk. The fully convolutional transformer for medical image segmentation. in Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 3660–3669, 2023.
- [14] Bernard, Olivier, Lalande, Alain, Zotti, Clement, Cervenansky, Frederick, Yang, Xin, Heng, Pheng-Ann, Cetin, Irem, Lekadir, Karim, Camara, Oscar, Ballester, Miguel Angel Gonzalez, et al. Deep learning techniques for automatic mri cardiac multi-structures segmentation and diagnosis: is the problem solved? IEEE transactions on medical imaging, 37(11):2514–2525, 2018.

# Abstract

Deep learning, a subset of artificial intelligence, with its ability to extract complex patterns and insights from large datasets, has brought about transformations in various fields including healthcare, autonomous driving, financial services, and environmental monitoring. This technology has particularly shown remarkable effects in medical imaging. The current project specifically investigates the application of deep learning in diagnosing cardiovascular diseases through cardiac magnetic resonance imaging (MRI). The aim of this research is to reveal the capabilities of deep learning in diagnosing these types of diseases. Utilizing advanced architectures and the high computational power of deep learning models, this project introduces a new approach to the diagnosis of cardiac diseases and demonstrates the potential to generalize this technology to other medical fields.

## Key Words:

Deep learning, Cardiovascular diseases, Cardiac magnetic resonance imaging.