

# تشخیص خودکار بیماری‌های قلبی با بهره‌گیری از تحلیل همزمان فریم‌های پایان دیاستول و پایان سیستول در تصاویر ام‌آرآی قلب با استفاده از شبکه عصبی کانولوشنی

تاریخ دریافت: ۱۴۰۴/۰۶/۱۹ - تاریخ پذیرش: ۱۴۰۴/۰۸/۲۸

## خلاصه

**مقدمه:** تصویربرداری ام‌آرآی برای تشخیص بیماری‌های قلبی در سال‌های اخیر رایج شده‌است، اما این روش‌ها زمان‌بر و همراه با خطای انسانی هستند. به همین دلیل، هوش مصنوعی و به‌ویژه یادگیری عمیق به دلیل توانایی بالایشان در پردازش تصاویر پزشکی، مورد توجه قرار گرفته‌اند. هدف این مطالعه، ارائه مدلی مبتنی بر شبکه عصبی کانولوشنی برای تشخیص دقیق بیماری قلبی و کاهش نیاز به مداخلات انسانی است.

**روش کار:** در این پژوهش، مدلی مبتنی بر شبکه عصبی کانولوشنی برای تشخیص بیماری قلبی طراحی و پیاده‌سازی شد. برخلاف رویکردهای متداول که تصاویر را به صورت تک‌فریم یا موازی پردازش می‌کنند، در این مطالعه از ورودی‌های دوکاناله شامل تصاویر پایان دیاستول و سیستول استفاده گردید تا پویایی چرخه قلبی بهتر منعکس شود. معماری پیشنهادی شامل لایه‌های متعدد کانولوشن و تجمیع بیشینه بوده و برای جلوگیری از بیش‌برازش، لایه حذف تصادفی به کار گرفته شد. علاوه بر این، تنظیم هایپرپارامترها از جمله نوع بهینه‌ساز و اندازه بچ و بهره‌گیری از روش‌های پیشرفته افزایش داده، به بهبود عملکرد شبکه کمک کرد.

**نتایج:** آزمایش‌ها بر روی مجموعه داده ACDC انجام شد و عملکرد مدل با معیارهای معتبر ارزیابی گردید. نتایج نشان داد که رویکرد دوکاناله نسبت به مدل‌های تک‌فریم و موازی، بهبود قابل توجهی ایجاد می‌کند. همچنین، مقایسه با معماری‌های مرجع و الگوریتم‌های یادگیری ماشین، برتری روش پیشنهادی از نظر دقت و پایداری عملکرد را تأیید کرد.

**نتیجه گیری:** رویکرد دوکاناله می‌تواند ابزاری کارآمد در تشخیص بیماری‌های قلبی باشد. این روش با دقت بالاتر نسبت به مدل‌های مرجع و کاهش وابستگی به تفسیر انسانی، امکان جایگزینی بخشی از فرآیندهای سنتی را فراهم می‌سازد.

**کلمات کلیدی:** تصویربرداری ام‌آرآی، بیماری قلبی، یادگیری عمیق، شبکه عصبی کانولوشنی، پایان دیاستول، پایان سیستول

عبدالرضا رضاپور<sup>۱\*</sup>

مقداد فرج پور<sup>۲</sup>

<sup>۱</sup> استادیار، گروه مهندسی کامپیوتر، واحد آستانه اشرفیه،

دانشگاه آزاد اسلامی، آستانه اشرفیه، ایران

<sup>۲</sup> استادیار، گروه مدیریت، موسسه آموزش عالی مهرآستان،

آستانه اشرفیه، ایران

Email: abdoreza.rezapour@iau.ac.ir

## مقدمه

بیماری‌های قلبی عروقی یکی از اصلی‌ترین علل مرگ‌ومیر جهانی محسوب می‌شوند و سالانه میلیون‌ها نفر را تحت تأثیر قرار می‌دهند. تشخیص زودهنگام و دقیق این بیماری‌ها می‌تواند از بروز عوارض جدی مانند نارسایی قلبی، سکته‌ی قلبی و کاهش کیفیت زندگی بیماران جلوگیری کرده و نقش اساسی در بهبود پیامدهای بالینی و انتخاب درمان مناسب ایفا کند. تصویربرداری تشدید مغناطیسی قلب یا ام آر آی قلبی<sup>۱</sup> به‌عنوان یک روش استاندارد طلایی در ارزیابی غیرتهاجمی ساختار و عملکرد قلب شناخته می‌شود و به‌دلیل کنتراست بافتی بالا و قابلیت مشاهده قلب در طول چرخه قلبی، جایگاه ویژه‌ای در تشخیص اختلالات قلبی دارد (۱). با این حال، تفسیر دستی تصاویر ام آر آی اغلب زمان‌بر بوده، به مهارت بالینی نیاز دارد و ممکن است دچار خطاهای متغیر بین ناظرین شود.

در تصاویر ام آر آی متحرک<sup>۲</sup> دو فاز کلیدی، یعنی پایان دیاستول<sup>۳</sup> و پایان سیستول<sup>۴</sup>، بیشترین اطلاعات تشخیصی را ارائه می‌کنند، زیرا این دو فریم به ترتیب بیانگر بیشترین و کمترین حجم بطن‌ها هستند و تغییرات شکل، اندازه و عملکرد آن‌ها در این دو وضعیت به‌خوبی قابل ارزیابی است. با وجود ارزش این داده‌ها، بسیاری از مطالعات پیشین صرفاً به تحلیل یک فاز محدود شده‌اند و ترکیب همزمان این دو فاز کلیدی کمتر مورد توجه قرار گرفته است. این محدودیت موجب می‌شود بخشی از الگوهای دینامیکی عملکرد قلب نادیده گرفته شود و وابستگی به تفسیر دستی و خطای انسانی همچنان باقی بماند.

در سال‌های اخیر روش‌های مبتنی بر یادگیری عمیق<sup>۵</sup> و به‌ویژه شبکه‌های عصبی کانولوشنی<sup>۶</sup> پیشرفت چشمگیری در تحلیل تصاویر پزشکی داشته‌اند و توانسته‌اند با استخراج خودکار ویژگی‌ها، بسیاری از محدودیت‌های روش‌های سنتی را برطرف

کنند (۲). با این حال، بخش عمده‌ای از پژوهش‌های موجود تنها بر یک فریم یا یک مرحله از چرخه‌ی قلبی تمرکز کرده‌اند و از اطلاعات مکمل موجود در ترکیب فریم‌های پایان دیاستول و پایان سیستول بهره‌برداری کامل نشده است. این شکاف پژوهشی ضرورت توسعه‌ی رویکردی نوین را برجسته می‌سازد که بتواند با استفاده همزمان از دو فاز کلیدی، تحلیل دقیق‌تر و طبقه‌بندی قابل‌اعتمادتر بیماری‌های قلبی را فراهم کند.

اگر بخواهیم پیشینه تحقیق را هم از منظر مقالات خارجی و هم داخلی بررسی کنیم، در سطح بین‌المللی، چندین مطالعه از شبکه‌های عصبی عمیق برای تشخیص بیماری قلبی از روی تصاویر ام آر آی قلب استفاده کرده‌اند. ژنگ<sup>۷</sup> و همکاران در سال ۲۰۱۸ با یادگیری نیمه‌نظارتی و استفاده از جریان ظاهری بین فریم‌های پایان دیاستول و پایان سیستول، ویژگی‌های حرکتی قلب را استخراج و طبقه‌بندی بیماری‌ها را انجام دادند (۳). اگرچه دقت مدل قابل قبول بود، پیچیدگی محاسباتی و نیاز به پیش‌پردازش گسترده، مانع کاربرد بالینی شد. در کار دیگری یانیک<sup>۸</sup> و همکاران در سال ۲۰۲۱، با استفاده از روش‌های تبیین‌پذیری، نشان دادند که مدل‌های عمیق می‌توانند مفاهیم بالینی معنی‌دار را برای تصمیم‌گیری شناسایی کنند. با وجود این، مدل‌ها اغلب تنها روی یک فاز قلب آموزش دیده بودند و توانایی بهره‌گیری همزمان از فاز کلیدی برای افزایش دقت کمتر مورد توجه قرار گرفت (۴).

بن خلیفه<sup>۹</sup> و همکاران در سال ۲۰۲۵ در مطالعه دیگری در این حوزه، با بهره‌گیری از یادگیری انتقالی<sup>۱۰</sup> و مدل وی‌جی‌جی‌۱۶، تصاویر ماده حاجب<sup>۱۱</sup> را برای طبقه‌بندی بیماران به سه گروه سکته قلبی، میوکاردیت و حالت سالم استفاده کردند. محدودیت این مطالعه، استفاده از نوع خاصی از تصاویر بود و

<sup>7</sup> Zheng<sup>8</sup> Janik<sup>9</sup> Ben Khalifa<sup>10</sup> Transfer Learning<sup>11</sup> Late Gadolinium Enhancement<sup>1</sup> Cardiac MRI<sup>2</sup> Cine MRI<sup>3</sup> End-Diastole(ED)<sup>4</sup> End-Systole(ES)<sup>5</sup> Deep Learning<sup>6</sup> Convolutional Neural Network

توانایی تعمیم به فازهای مختلف قلب و سایر بیماری‌ها محدود بود(۵).

در مطالعه دیگر والترینک<sup>۱</sup> و همکاران در سال ۲۰۱۷ مدلی بر پایه شبکه عصبی کانولوشنی ارائه کردند که با استفاده از تصاویر پایان دیاستول و پایان سیستول، سعی در تشخیص بیماری قلبی داشت. در این روش اگرچه دقت مدل قابل قبول بود، اما شبکه به تنهایی از تفاوت‌های عملکردی بین فازهای قلبی به صورت کامل بهره نمی‌برد و نیازمند استخراج ویژگی‌های اضافی توسط الگوریتم‌های دیگر بود(۶).

در تحقیق دیگر در این زمینه، شوآب<sup>۲</sup> و همکاران در سال ۲۰۲۵ یک چارچوب کاملاً خودکار برای سگمنتیشن انفارکتوس میوکارد از تصاویر ماده حاجب ارائه کردند. نتایج نشان داد که مدل‌های شبکه عصبی کانولوشنی سه‌بعدی توانستند حجم انفارکتوس را با توافقی بالا نسبت به متخصصان انسانی محاسبه کنند. این مطالعه عمدتاً روی سگمنتیشن بافتی متمرکز است، و برای طبقه‌بندی بیماری قلبی مناسب نیست. استفاده از شبکه عصبی کانولوشنی سه‌بعدی مناسب داده‌های حجیم گادودولینویم است، اما این مدل‌ها معمولاً به داده‌های بسیار بزرگ و پر حجم نیاز دارند و در دیتاست‌هایی مانند ACDC که نسبتاً کوچک هستند، عملکردشان پایدار نیست. همچنین، روش آن‌ها بر حجم خسارت بافتی<sup>۳</sup> تمرکز دارد و نه روی تحلیل همزمان فریم‌های های کلیدی که برای عملکرد قلب حیاتی‌تر است(۷).

فون زوبن<sup>۴</sup> و همکاران در ۲۰۲۳، با استفاده از شبکه یوه هدایت‌شده‌ی آناتومیک و یادگیری انتقالی، سگمنتیشن تصاویر محور طولی قلب را انجام دادند و توانستند با داده‌های محدود دقت مناسبی به دست آورند. گرچه روش آن‌ها در شرایط داده-های محدود عملکرد مناسبی داشت، اما خروجی نهایی تنها یک ماسک سگمنتیشن است و برای تشخیص نوع بیماری قلبی طراحی نشده است. همچنین، شبکه یو ذاتاً برای استخراج ویژگی

دینامیک زمانی ام‌آر‌آی مناسب نیست مگر اینکه از توالی چندفریمی استفاده شود(۸).

به‌طور کلی، مطالعات بین‌المللی نشان می‌دهند که شبکه‌های عصبی عمیق قابلیت تشخیص بیماری قلبی از روی تصاویر ام‌آر-آی را دارند، اما محدودیت‌هایی همانند تمرکز روی یک فاز قلب پایان دیاستول یا پایان سیستول و عدم بهره‌گیری همزمان از فازها، نیاز به پیش‌پردازش پیچیده و استخراج ویژگی توسط الگوریتم‌های دیگر و محدودیت در تعمیم‌پذیری به انواع مختلف تصاویر و بیماران وجود دارد که روش ما قصد دارد آن‌ها را برطرف کند. در مدل پیشنهادی ما، تصاویر پایان دیاستول و پایان سیستول به صورت استک‌شده به شبکه وارد می‌شوند تا تفاوت‌های عملکردی و ساختاری بین این دو فاز به طور مستقیم در تصمیم‌گیری مدل استفاده شود و نیاز به استخراج ویژگی‌های اضافه کاهش یابد.

در ایران، اگرچه کاربرد هوش مصنوعی در بیماری‌های قلبی در حال رشد است، بیشتر مطالعات به داده‌های اکوکاردیوگرافی یا الگوریتم‌های سنتی یادگیری ماشین محدود بوده‌اند و کارهای مبتنی بر تصاویر ام‌آر‌آی قلب با استفاده از شبکه‌های عصبی عمیق به صورت نظام‌مند بسیار کمتر گزارش شده است. همچنین مطالعه‌ی منسجمی که ترکیب دو فریم پایان دیاستول و پایان سیستول، را به عنوان ورودی به یک مدل شبکه عصبی کانولوشنی برای تشخیص چند نوع بیماری قلبی انجام دهد، به اندازه کافی دیده نمی‌شود(۹). این شکاف پژوهشی، ضرورت توسعه یک چارچوب خودکار و دقیق برای تشخیص بیماری‌های ساختاری قلب از روی ام‌آر‌آی را بیشتر نمایان می‌کند.

هدف پژوهش حاضر، توسعه یک مدل طبقه‌بندی مبتنی بر شبکه عصبی کانولوشنی است که با استفاده از ترکیب دو فریم پایان دیاستول و پایان سیستول به صورت ورودی استک‌شده یا دوکاناله بتواند پنج کلاس تشخیصی تعریف‌شده در مجموعه داده ACDC را به صورت خودکار تشخیص دهد. این کلاس‌ها شامل وضعیت نرمال، کاردیومیوپاتی اتساعی<sup>۶</sup>

<sup>1</sup> Wolterink

<sup>2</sup> Schwab

<sup>3</sup> Infarct Volume

<sup>4</sup> Von Zuben

<sup>5</sup> U-Net

<sup>6</sup> Dilated Cardiomyopathy (DCM)

در شکل ۱ نمای کلی روش پیشنهادی ارائه شده است. ابتدا داده‌ها وارد مرحله‌ی پیش‌پردازش می‌شوند که شامل هم‌ترازی، استخراج ناحیه‌ی مورد توجه<sup>۴</sup>، تغییر اندازه به  $224 \times 224$  و نرمال‌سازی شدت است. به‌منظور کاهش خطر بیش‌برازش ناشی از اندازه‌ی کوچک مجموعه داده، یک استراتژی افزایش داده‌ی کنترل‌شده شامل چرخش خفیف، جابه‌جایی، تغییر روشنایی و وارون‌سازی محدود اعمال شد تا تنوع طبیعی تصاویر ام‌آرآی بدون تخریب ساختار اصلی قلب شبیه‌سازی شود.

پس از آماده‌سازی، فریم‌های پایان دیاستول و پایان سیستول به‌صورت دو کانال موازی استک شده و وارد معماری شبکه‌های عصبی کانولوشنی پیشنهادی می‌شوند. این ورودی دوکاناله مهم‌ترین نوآوری روش ما محسوب می‌شود، زیرا شبکه می‌تواند الگوهای عملکردی و ساختاری را به‌صورت هم‌زمان و یکپارچه بیاموزد. معماری پیشنهادی شامل سه لایه‌ی کانولوشن همراه با سه لایه‌ی تجمع بیشینه<sup>۵</sup> است. در تمام لایه‌ها از تابع فعال‌سازی خطی اصلاح شده<sup>۶</sup> استفاده شده و لایه ریزشی یا حذف تصادفی<sup>۷</sup> برای جلوگیری از بیش‌برازش در بخش میانی شبکه قرار داده شده است. خروجی لایه‌های استخراج ویژگی وارد لایه‌های کاملاً متصل<sup>۸</sup> می‌شود و نهایتاً از تابع فعال‌ساز نرم بیشینه<sup>۹</sup> پنج کلاس بیماری قلبی را دسته‌بندی می‌کند.

به‌طور خلاصه، روش پیشنهادی ما با ترکیب قدرت شبکه‌های عصبی کانولوشنی در استخراج ویژگی‌های تصویری، ورودی دوکاناله مبتنی بر پایان دیاستول و پایان سیستول، و افزایش داده-ی هدفمند، مدلی سبک و دقیق متناسب با محدودیت‌های مجموعه داده‌های ام‌آرآی قلبی ارائه می‌کند. جدول ۱ اطلاعات کامل تمام لایه‌های معماری پیشنهادی و ابعاد خروجی آنها را نمایش می‌دهد. هر یک از این لایه‌ها در ادامه و زیربخش‌های زیر توضیح داده شده‌اند.

کاردیومیوپاتی هیپرتروفیک<sup>۱</sup>، انفارکتوس میوکارد یا همان سکتة قلبی<sup>۲</sup>، کاردیومیوپاتی آریتموژنیک بطن راست<sup>۳</sup> می‌باشد. بهره‌گیری از این دو فاز کلیدی، امکان استفاده از بیشترین اطلاعات ساختاری و عملکردی را در عین حفظ سادگی و کارایی مدل فراهم می‌کند و می‌تواند به‌عنوان ابزاری کمکی برای افزایش سرعت و دقت تشخیص اختلالات قلبی در محیط‌های بالینی مورد استفاده قرار گیرد.

## روش کار

در سال‌های اخیر، شبکه‌های عصبی عمیق به‌ویژه شبکه‌های عصبی کانولوشنی به ابزار اصلی در تحلیل و پردازش تصاویر پزشکی تبدیل شده‌اند(2). شبکه‌های عصبی کانولوشنی به دلیل توانایی بالا در استخراج خودکار ویژگی‌های مکانی و بافتی، در مقایسه با روش‌های سنتی که متکی بر ویژگی‌های مهندسی شده بودند، عملکرد بسیار دقیق‌تر و پایدارتری ارائه می‌دهند. کاربرد این مدل‌ها در ام‌آرآی قلبی نیز به‌طور گسترده اثبات شده است، زیرا بسیاری از ناهنجاری‌های قلبی تنها از طریق الگوهای فضایی و بافتی ظریف در تصاویر قابل شناسایی هستند؛ الگوهایی که شبکه‌های کانولوشنی به‌طور طبیعی قادر به یادگیری آن‌ها می‌باشند(۱۰).

با این حال، استفاده از یک فریم منفرد ام‌آرآی ممکن است تمام اطلاعات عملکردی قلب را منعکس نکند، زیرا بسیاری از بیماری‌های قلبی تنها زمانی قابل مشاهده‌اند که رفتار بطن چپ در دو فاز اصلی چرخه‌ی قلب بررسی شود. بر همین اساس، در این پژوهش یک معماری شبکه‌های عصبی کانولوشنی با ورودی دوکاناله طراحی شد که هم‌زمان از تصاویر پایان دیاستول و پایان سیستول استفاده می‌کند. هدف از این طراحی، ارائه‌ی اطلاعات ساختاری و عملکردی به‌صورت ترکیبی به شبکه و فراهم کردن امکان یادگیری اختلافات کلیدی بین انقباض و انقباض قلب است؛ اختلافاتی که برای تشخیص انواع کاردیومیوپاتی بسیار حیاتی هستند.

<sup>4</sup> Region of Interest(ROI)

<sup>5</sup> Max-Pooling

<sup>6</sup> Rectified Linear Unit(ReLU)

<sup>7</sup> Dropout

<sup>8</sup> Fully Connected Layer

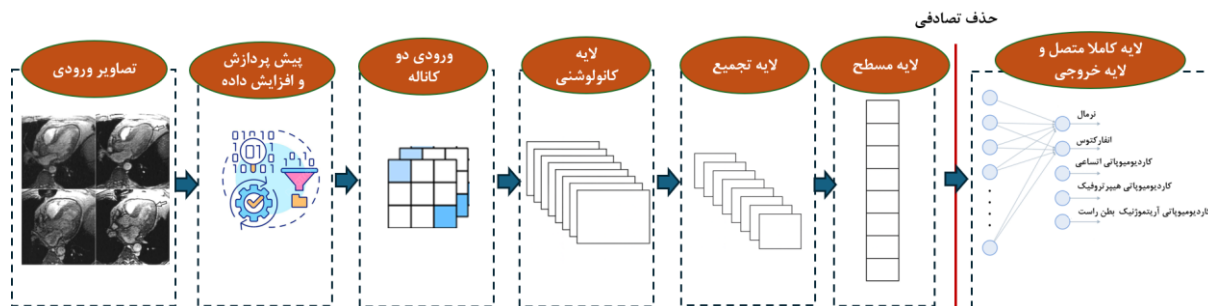
<sup>9</sup> Softmax

<sup>1</sup> Hypertrophic Cardiomyopathy (HCM)

<sup>2</sup> Myocardial Infarction (MINF)

<sup>3</sup> Abnormal Right Ventricle (ARV)





شکل ۱- نمای کلی از معماری روش پیشنهادی

جدول ۱- جزئیات کامل تمام لایه های معماری پیشنهادی برای تشخیص بیماری قلبی

| توضیح پزشکی / عملکرد                             | ابعاد خروجی  | اندازه فیلتر | تعداد فیلتر / نورون | نوع لایه              | شماره لایه |
|--------------------------------------------------|--------------|--------------|---------------------|-----------------------|------------|
| جفت فریم های پایان دیاستول و پایان سیستول        | (۲۲۴،۲۲۴،۲)  | —            | ۲ کانال             | ورودی                 | ۱          |
| استخراج ویژگی های اولیه (بافت، لبه ها)           | (۲۲۴،۲۲۴،۳۲) | ۳×۳          | ۳۲                  | کانولوشن ۱            | ۲          |
| غیرخطی سازی ویژگی ها                             | (۲۲۴،۲۲۴،۳۲) | —            | —                   | تابع فعال ساز ReLU    | ۳          |
| کاهش ابعاد و تمرکز روی نواحی مهم                 | (۱۱۲،۱۱۲،۳۲) | ۲×۲          | —                   | تجمع بیشینه           | ۴          |
| شناسایی ویژگی های ساختاری بیماری                 | (۱۱۲،۱۱۲،۶۴) | ۳×۳          | ۶۴                  | کانولوشن ۲            | ۵          |
| غیرخطی سازی ویژگی ها                             | (۱۱۲،۱۱۲،۶۴) | —            | —                   | تابع فعال ساز ReLU    | ۶          |
| کاهش ابعاد و افزایش پایداری استخراج ویژگی        | (۵۶،۵۶،۶۴)   | ۲×۲          | —                   | تجمع بیشینه           | ۷          |
| استخراج ویژگی های سطح بالا و مرتبط با نوع بیماری | (۵۶،۵۶،۱۲۸)  | ۳×۳          | ۱۲۸                 | کانولوشن ۳            | ۸          |
| غیرخطی سازی ویژگی ها                             | (۵۶،۵۶،۱۲۸)  | —            | —                   | تابع فعال ساز ReLU    | ۹          |
| کاهش نهایی ابعاد برای ورود به بخش طبقه بندی      | (۲۸،۲۸،۱۲۸)  | ۲×۲          | —                   | تجمع بیشینه           | ۱۰         |
| تبدیل نگاشت های ویژگی به بردار تحلیلی            | (۲۸،۲۸،۱۲۸)  | —            | —                   | مسطح سازی             | ۱۱         |
| کاهش بیش برآزش در مرحله طبقه بندی                | (۱۰۰۳۵۲،۱)   | نرخ ۰.۵      | —                   | حذف تصادفی            | ۱۲         |
| ترکیب ویژگی ها برای تحلیل جامع                   | (۲۵۶،۱)      | —            | ۲۵۶ نورون           | لایه کاملاً متصل ۱    | ۱۳         |
| تمرکز روی الگوهای افتراقی بیماری                 | (۱۲۸،۱)      | —            | ۱۲۸ نورون           | لایه کاملاً متصل ۲    | ۱۴         |
| پیش بینی پنج کلاسه انواع بیماری قلبی             | (۵،۱)        | —            | ۵ نورون             | لایه خروجی نرم بیشینه | ۱۵         |

یک تنسور با ابعاد (تعداد نمونه ها × تعداد کانال ها × ارتفاع × عرض) معادل  $(N \times 2 \times 224 \times 224)$  به شبکه وارد می شوند.

## ۱-۲- لایه ورودی

ورودی شبکه شامل دو تصویر دوبعدی ام آر آی قلب است که یکی مربوط به پایان دیاستول و دیگری مربوط به پایان سیستول می باشد. این دو تصویر پس از برش و تغییر اندازه به ابعاد  $224 \times 224$  پیکسل و انجام نرمال سازی شدت پیکسل ها، به صورت دو لایه روی هم قرار می گیرند. یعنی تصویر پایان دیاستول در لایه اول و تصویر پایان سیستول در لایه دوم قرار داده می شود. این فرایند که به آن استک کردن گفته می شود، باعث می شود شبکه این دو فریم را به صورت یک ورودی واحد اما با دو کانال مجزا دریافت کند. در نهایت، داده ها در قالب

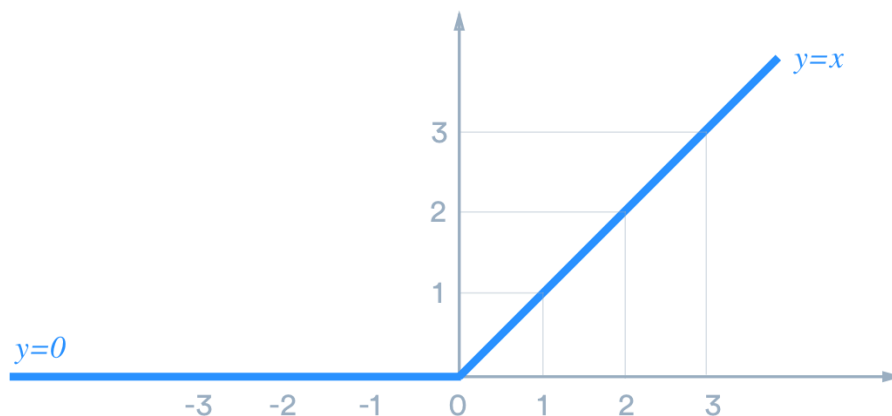
## ۲-۲- لایه کانولوشن

در معماری پیشنهادی، پس از ورود داده ها، پردازش اولیه توسط مجموعه ای از سه لایه کانولوشن انجام می شود. هر لایه کانولوشن مجموعه ای از فیلترهای قابل یادگیری را روی تصویر اعمال می کند و از طریق حرکت این فیلترها بر روی ورودی، ویژگی های مهم مانند لبه ها، بافت ها و الگوهای ساختاری قلب استخراج می شود. در لایه اول از فیلترهای بزرگ تر برای تشخیص الگوهای کلی تر استفاده شده و در لایه های بعدی تعداد فیلترها افزایش پیدا می کند تا شبکه بتواند ویژگی های

پیچیده‌تری را از تصاویر پاپان دیاستول و پاپان سیستول استخراج کند. خروجی هر لایه کانولوشن یک نگاشت ویژگی است که در آن هر کانال، بخشی از ساختار قلب را از زاویه‌ای خاص نمایش می‌دهد و این نگاشت‌ها به صورت سلسله‌مراتبی به لایه‌های بالاتر منتقل می‌شوند تا شبکه بتواند تمایز بین وضعیت طبیعی و بیمار گونه قلب را بهتر یاد بگیرد.

### ۲-۳- تابع فعال‌ساز

پس از هر لایه کانولوشن از تابع فعال‌سازی خطی اصلاح شده استفاده شد تا خروجی لایه‌ها به صورت غیرخطی تبدیل شود و



شکل ۱. \* تابع فعال‌سازی خطی اصلاح شده بکار رفته در لایه کانولوشنی

همچنین فرمول تابع فعال‌سازی خطی اصلاح شده در معادله ۱ قابل نمایش است.

$$F(x) = \max(0, x)$$

(۱)

پیش‌برازش، باعث می‌شود شبکه در لایه‌های بعدی بتواند ویژگی‌های مرتبط با اختلالات عملکرد قلب را با دقت بیشتری تحلیل کند.

### ۲-۵- لایه کاملاً متصل

پس از استخراج ویژگی‌ها در لایه‌های کانولوشن و تجمیع بیشینه، خروجی این لایه‌ها از طریق فرآیند مسطح‌سازی<sup>۱</sup> به صورت یک بردار یک‌بعدی درآمده و وارد مجموعه‌ای از لایه‌های کاملاً متصل خواهد شد. البته برای جلوگیری از پیش‌برازش، یک لایه ریزشی یا حذف تصادفی قبل از ورود به لایه‌های کاملاً متصل نیز در این بخش به کار گرفته شده است. در لایه‌های کاملاً متصل، شبکه با ترکیب تمام ویژگی‌های استخراج‌شده، ارتباط میان

### ۲-۴- لایه تجمیع بیشینه

در معماری پیشنهادی برای تشخیص بیماری قلبی، پس از هر مجموعه از لایه‌های کانولوشن، از لایه تجمیع بیشینه استفاده شد. نقش این لایه تنها کاهش اندازه داده‌ها نیست، بلکه کمک می‌کند نواحی مهم‌تر تصویر قلب که بیشترین نشانه‌های بیماری را دارند، برجسته‌تر شوند. در تصاویر ام‌آر‌آی قلب، تغییرات ساختاری مانند نازک شدن دیواره‌ها، بزرگ شدن حفرات یا بی‌نظمی در بافت، ممکن است تنها در نواحی کوچک اما معنادار ظاهر شوند. لایه تجمیع بیشینه با انتخاب بیشترین مقدار در هر بخش کوچک، همین الگوهای شاخص و برجسته را حفظ و اطلاعات کمتر دارای اهمیت را حذف می‌کند. این کار علاوه بر کاهش پیچیدگی محاسباتی و جلوگیری از

<sup>۱</sup> Flatten

اصلی چرخه قلب شامل پایان دیاستول و پایان سیستول می‌باشد. در این مجموعه داده هر بیمار یک جفت فریم پایان دیاستول و پایان سیستول دارد و داده‌ها به‌طور دقیق توسط متخصصین قلب برچسب‌گذاری شده‌اند. این مجموعه داده شامل ۵ نوع وضعیت قلبی است که شامل حالت نرمال، کاردیومیوپاتی اتساعی<sup>۳</sup>، کاردیومیوپاتی هیپرتروفیک<sup>۴</sup>، انفارکتوس میوکارد یا همان سکته قلبی<sup>۵</sup>، کاردیومیوپاتی آریتموژنیک بطن راست<sup>۶</sup> می‌باشد. این مجموعه داده به دلیل ساختار استاندارد، تنوع بیماران و برچسب‌های معتبر، یک مبنای عالی برای ارزیابی مدل‌های هوش مصنوعی در تشخیص بیماری‌های قلبی است. جدول ۲ تعداد نمونه‌های مربوط به کلاس موجود در مجموعه داده را به تفکیک مجموعه آموزش و آزمون نشان می‌دهد.

در ضمن شایان ذکر است که در این پژوهش، اگرچه مجموعه داده ACDC از نظر تعداد نمونه نسبتاً کوچک است، اما با توجه به ماهیت ساختاری و نسبتاً ثابت تصاویر ام‌آر‌آی قلب، می‌توان با بهره‌گیری از افزایش داده‌ها و برون‌اندازی، شبکه‌ی عصبی کانولوشنی را به‌صورت پایدار و قابل‌اعتماد آموزش داد. برای جلوگیری از بیش‌برازش و تقویت توان تعمیم مدل، در مرحله‌ی پیش‌پردازش از افزایش داده‌ها شامل چرخش‌های خفیف، جابه‌جایی کوچک و تغییرات جزئی شدت روشنایی استفاده شد (۱۳). این تغییرات موجب می‌شود شبکه با نمونه‌های متنوع‌تری روبه‌رو شود و وابستگی آن به الگوهای خاص کاهش یابد. همچنین در بخش طبقه‌بندی، از لایه‌ی برون‌اندازی با نرخ مناسب استفاده شد تا وابستگی مدل به نورون‌های خاص کمتر شود. ترکیب این دو روش باعث شد معماری پیشنهادی با وجود اندازه‌ی محدود داده‌ها، پایداری یادگیری و عملکرد مطلوب خود را حفظ کند.

الگوهای ساختاری قلب را در ابعاد گسترده‌تری تحلیل می‌کند. هدف این مرحله آن است که شبکه بتواند نشانه‌های پراکنده ولی مرتبط با بیماری قلبی، مانند تغییر در ضخامت دیواره بطن، تفاوت در حرکت عضله قلب بین پایان دیاستول و پایان سیستول، یا الگوهای غیرطبیعی در بافت را در کنار یکدیگر قرار داده و یک برداشت جامع از وضعیت قلب ارائه دهد. استفاده از این لایه‌ها موجب می‌شود که مدل علاوه بر تشخیص ویژگی‌های موضعی، بتواند یک جمع‌بندی کلی و تصمیم‌گیری سطح بالا را درباره سالم یا بیمار بودن نمونه انجام دهد.

## ۲-۶- لایه خروجی

در نهایت، بردار تولیدشده توسط لایه‌های کاملاً متصل وارد لایه خروجی چندکلاسه می‌شود تا وضعیت قلب را به یکی از کلاس‌های تعریف‌شده در مجموعه داده طبقه‌بندی کند. هر نورون در این لایه نمایانگر احتمال تعلق نمونه به یک کلاس مشخص است و با استفاده از تابع فعال‌ساز نرم بیشینه مقادیر نهایی به احتمال تبدیل می‌شوند. این طراحی به شبکه اجازه می‌دهد تا تغییرات عملکردی قلب بین فریم‌های پایان دیاستول و پایان سیستول را در پیش‌بینی دقیق نوع بیماری لحاظ کند و یک تصمیم‌گیری چندکلاسه قابل تفسیر برای پزشکان ارائه دهد. احتمال هر کلاس براساس تابع فعال‌ساز نرم بیشینه و معادله ۲ قابل محاسبه است.

$$\text{softmax}(x_i) = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (2)$$

که در آن  $x_i$  همان عدد بدست آمده از لایه کاملاً متصل برای هر یک از کلاس‌هاست.

## ۳- آزمایشات

### ۳-۱- مجموعه داده

مجموعه داده مورد استفاده در این پژوهش با نام ACDC<sup>۱</sup> می‌باشد (۱۲). این مجموعه داده که از مجموعه داده‌های معتبر حوزه تحلیل تصاویر قلبی است، در چالش MICCAI<sup>۲</sup> ارائه شده است. این مجموعه داده شامل تصاویر ام‌آر‌آی قلب در دو فاز

<sup>3</sup> Dilated Cardiomyopathy (DCM)

<sup>4</sup> Hypertrophic Cardiomyopathy (HCM)

<sup>5</sup> Myocardial Infarction (MINF)

<sup>6</sup> Abnormal Right Ventricle (ARV)

<sup>1</sup> Automated Cardiac Diagnosis Challenge

<sup>2</sup> <https://www.creatis.insa-lyon.fr/Challenge/acdc/>



## جدول ۲- اطلاعات مجموعه داده به تفکیک نمونه ها

| بخش   | نمونه‌ها نرمال | میوکارد | اتساعی | هیپرتروفیک | آریموژنیک | کاردیومیوپاتی |
|-------|----------------|---------|--------|------------|-----------|---------------|
|       |                |         |        |            |           |               |
| آموزش | ۱۰۰            | ۲۰      | ۲۰     | ۲۰         | ۲۰        | ۲۰            |
| آزمون | ۵۰             | ۱۰      | ۱۰     | ۱۰         | ۱۰        | ۱۰            |

## ۳-۲- تنظیم آزمایش

همانطور که در بخش قبل اشاره شده در این پژوهش، برای ارزیابی عملکرد مدل پیشنهادی در تشخیص پنج کلاسه بیماری‌های قلبی، از مجموعه داده ACDC استفاده شد. با توجه به محدود بودن حجم داده، برای استفاده کارآمد از نمونه‌ها و کاهش بیش‌برازش، روش اعتبارسنجی متقابل پنج‌قسمتی<sup>۱</sup> بر روی داده‌های آموزشی اعمال شد. در این روش، کل داده‌های آموزشی به پنج بخش تقریباً مساوی تقسیم شدند. در هر تکرار، چهار بخش برای آموزش و یک بخش برای اعتبارسنجی مورد استفاده قرار گرفت و این فرآیند پنج بار به گونه‌ای انجام شد که هر بخش دقیقاً یک بار به عنوان مجموعه اعتبارسنجی به کار گرفته شود. پس از پایان پنج تکرار، میانگین معیارهای ارزیابی برای انتخاب پیکربندی بهینه مدل محاسبه گردید. هدف از این مرحله تنظیم هایپرپارامترها و دستیابی به بهترین نسخه از مدل پیش از ارزیابی نهایی می‌باشد (۱۴). مهم‌ترین هایپرپارامترهایی که در این مرحله تنظیم شدند عبارت‌اند از: اندازه دسته<sup>۲</sup>: برای این بخش مقادیر ۸، ۱۶ و ۳۲ آزمایش شد براساس میانگین نتایج اعتبارسنجی، مقدار ۱۶ بهترین عملکرد را ارائه داد و به عنوان مقدار بهینه انتخاب شد.

نرخ یادگیری<sup>۳</sup>: چند مقدار متداول  $10^{-2}$ ،  $10^{-3}$  و  $10^{-4}$  مورد بررسی قرار گرفت و با توجه به پایداری آموزش و نرخ همگرایی مقدار  $10^{-4}$ ، به عنوان مقدار بهینه انتخاب شد.

تعداد دوره‌های آموزش<sup>۴</sup>: برای جلوگیری از بیش‌برازش، از روش توقف زود هنگام<sup>۵</sup> با صبر ۱۰ دوره و سقف حداکثر ۶۰ دوره استفاده شد.

نرخ برون اندازی: مقدارهای  $0/3$ ،  $0/5$  و  $0/6$  بررسی و مقدار  $0/5$  به دلیل ایجاد تعادل مناسب بین جلوگیری از بیش‌برازش و حفظ دقت مدل، به عنوان مقدار بهینه انتخاب شد.

بهینه‌ساز: اگوریتم آدام<sup>۶</sup> به دلیل عملکرد مناسب در مسائل تصویری و داده‌های کم حجم مورد استفاده قرار گرفت.

به منظور افزایش تنوع داده‌ها و بهبود توانایی مدل در تعمیم‌پذیری، از افزایش یا غنی سازی داده‌ها<sup>۷</sup> شامل چرخش، جابه‌جایی مکانی و تغییرات جزئی شدت روشنایی استفاده شد. لازم به ذکر است که عملیات افزایش داده صرفاً روی مجموعه آموزش و در هر بخش از تقسیم بندی به صورت مجزا انجام شد تا از نشت اطلاعات به بخش اعتبارسنجی جلوگیری شود. پس از تعیین پیکربندی بهینه مدل براساس نتایج اعتبارسنجی متقابل پنج‌قسمتی، مدل نهایی با استفاده از تمامی داده‌های آموزشی آموزش داده شد و سپس با ۵۰ نمونه موجود در بخش آزمون مجموعه داده ACDC، که از ابتدا جدا نگه داشته شده بودند و در فرآیند اعتبارسنجی دخالت نداشتند، مورد ارزیابی نهایی قرار گرفت.

## نتایج

برای ارزیابی عملکرد مدل پیشنهادی، سه دسته آزمایش طراحی و اجرا شد. در تمامی آزمایش‌ها معیارهای صحت، دقت، فراخوانی، سنجش اف و مساحت زیر منحنی حاسبه گردیدند. نتایج به صورت میانگین به همراه انحراف معیار بر اساس اعتبارسنجی ۵ قسمتی گزارش شده‌اند. در اولین آزمایش که نتایج آن در جدول ۳ مشخص است، تأثیر نوع ورودی بر عملکرد مدل بررسی شد. نتایج نشان داد که استفاده از تنها یک فریم پایان دیاستول و پایان سیستول عملکرد متوسطی دارد، در حالی که ترکیب دو فریم به صورت جداگانه باعث بهبود نسبی معیارها شد. در نهایت، روش پیشنهادی که دو فریم پایان دیاستول و پایان سیستول را به صورت استک شده وارد شبکه می‌کند، بهترین عملکرد را ارائه داد.

<sup>1</sup> 5-Fold Cross-Validation<sup>2</sup> Batch Size<sup>3</sup> Learning Rate<sup>4</sup> Epochs<sup>5</sup> Early Stopping<sup>6</sup> Adam<sup>7</sup> Data Augmentation

## جدول ۳- مقایسه عملکرد معماری پیشنهادی بر اساس نوع ورودی

| نوع ورودی                          | صحت<br>انحراف | دقت<br>انحراف | فراخوانی<br>انحراف | امتیاز اف ۱<br>انحراف | مساحت زیر منحنی<br>انحراف |
|------------------------------------|---------------|---------------|--------------------|-----------------------|---------------------------|
| فقط پایان دیاستول                  | ۰/۸۱±۰/۰۲     | ۰/۷۸±۰/۰۲     | ۰/۷۶±۰/۰۳          | ۰/۷۷±۰/۰۲             | ۰/۸۴±۰/۰۲                 |
| فقط پایان سیستول                   | ۰/۸۳±۰/۰۲     | ۰/۸۰±۰/۰۲     | ۰/۷۹±۰/۰۲          | ۰/۷۹±۰/۰۲             | ۰/۸۶±۰/۰۲                 |
| دو فریم جداگانه (دو مسیر)          | ۰/۸۶±۰/۰۱     | ۰/۸۴±۰/۰۲     | ۰/۸۳±۰/۰۲          | ۰/۸۳±۰/۰۲             | ۰/۸۹±۰/۰۱                 |
| دو فریم استک شده<br>(روش پیشنهادی) | ۰/۹۱±۰/۰۱     | ۰/۸۹±۰/۰۱     | ۰/۸۸±۰/۰۱          | ۰/۸۸±۰/۰۱             | ۰/۹۲±۰/۰۱                 |

در آزمایش دوم که نتایج آن در جدول ۴ مشخص است، مدل پیشنهادی با معماری‌های مرجع شامل شبکه عصبی کانولوشنی پایه (۱۵)، وی جی جی ۱۶، وی جی جی ۱۹<sup>۲</sup> که توسط سیمونیان و زیرمن ارائه شدند (۱۶)، رزنت ۱۸<sup>۳</sup> و رزنت ۳۴<sup>۴</sup> که نخستین بار توسط هی و همکاران معرفی شدند (۱۷)، مقایسه شد. نتایج نشان داد که اگرچه معماری‌های عمیق‌تر مانند رزنت و وی جی جی عملکرد مناسبی دارند، اما مدل پیشنهادی با طراحی سبک‌تر و استفاده از ورودی‌های فیزیولوژیک پایان دیاستول و پایان سیستول توانست از آن‌ها پیشی بگیرد.

## جدول ۴- ارزیابی مدل پیشنهادی در مقایسه با معماری‌های

یادگیری عمیق شناخته شده

| مساحت              | امتیاز اف | صحت    | دقت    | فراخوانی | زیر منحنی |
|--------------------|-----------|--------|--------|----------|-----------|
| انحراف             | انحراف    | انحراف | انحراف | انحراف   | انحراف    |
| معماری             | معماری    | معماری | معماری | معماری   | معماری    |
| مدل                | مدل       | مدل    | مدل    | مدل      | مدل       |
| شبکه عصبی کانولوشن | ۰/۸۴±     | ۰/۸۲±  | ۰/۸۱±  | ۰/۸۱±    | ۰/۸۶±     |
| ی پایه             | ۰/۰۲      | ۰/۰۲   | ۰/۰۲   | ۰/۰۲     | ۰/۰۲      |
| VGG16              | ۰/۸۷±     | ۰/۸۵±  | ۰/۸۴±  | ۰/۸۴±    | ۰/۸۹±     |
| تنظیم شده          | ۰/۰۲      | ۰/۰۲   | ۰/۰۲   | ۰/۰۲     | ۰/۰۲      |

در آزمایش سوم که نتایج آن در جدول ۵ مشخص شده است، مدل پیشنهادی با روش‌های سنتی یادگیری ماشین شامل ماشین بردار پشتیبان<sup>۵</sup> (۱۸)، درخت تصمیم<sup>۶</sup> (۱۹)، جنگل تصادفی<sup>۷</sup> (۲۰)، بیز ساده<sup>۸</sup> (۲۱) و همچنین شبکه پرسپترون چندلایه<sup>۹</sup> (۲۲) مقایسه شد. همان‌طور که انتظار می‌رفت، روش‌های سنتی عملکرد پایین‌تری نسبت به شبکه‌های کانولوشنی داشتند. در این میان، شبکه پرسپترون چندلایه نتایج نسبتاً بهتری نسبت به سایر روش‌های سنتی ارائه داد، اما همچنان از مدل پیشنهادی عقب‌تر بود.

<sup>۵</sup> Support Vector Machine<sup>۶</sup> Decision Tree<sup>۷</sup> Random Forest<sup>۸</sup> Naïve Bayes<sup>۹</sup> Multi Layer Perceptron<sup>۱</sup> VGG16<sup>۲</sup> VGG19<sup>۳</sup> ResNet18<sup>۴</sup> ResNet34

## جدول ۵- مقایسه عملکرد مدل پیشنهادی با روش‌های سنتی یادگیری ماشین

| مدل سنتی            | انحراف معیار $\pm$ میانگین | انحراف معیار $\pm$ میانگین | دقت             | فراخوانی        | امتیاز اف ۱     | مساحت زیر منحنی |
|---------------------|----------------------------|----------------------------|-----------------|-----------------|-----------------|-----------------|
| ماشین بردار پشتیبان | $0.78 \pm 0.03$            | $0.75 \pm 0.03$            | $0.74 \pm 0.03$ | $0.74 \pm 0.03$ | $0.74 \pm 0.03$ | $0.80 \pm 0.03$ |
| درخت تصمیم          | $0.71 \pm 0.03$            | $0.68 \pm 0.03$            | $0.66 \pm 0.03$ | $0.67 \pm 0.03$ | $0.67 \pm 0.03$ | $0.73 \pm 0.03$ |
| جنگل تصادفی         | $0.74 \pm 0.02$            | $0.72 \pm 0.02$            | $0.70 \pm 0.02$ | $0.71 \pm 0.02$ | $0.71 \pm 0.02$ | $0.76 \pm 0.02$ |
| بیز ساده            | $0.69 \pm 0.03$            | $0.66 \pm 0.03$            | $0.64 \pm 0.03$ | $0.65 \pm 0.03$ | $0.65 \pm 0.03$ | $0.70 \pm 0.03$ |
| چندلایه پرسپترون    | $0.80 \pm 0.02$            | $0.77 \pm 0.02$            | $0.76 \pm 0.02$ | $0.76 \pm 0.02$ | $0.76 \pm 0.02$ | $0.82 \pm 0.02$ |
| معماری پیشنهادی     | $0.91 \pm 0.01$            | $0.89 \pm 0.01$            | $0.88 \pm 0.01$ | $0.88 \pm 0.01$ | $0.88 \pm 0.01$ | $0.92 \pm 0.01$ |

## بحث

در این بخش یافته‌های پژوهش مورد بحث و تجزیه و تحلیل قرار می‌گیرند. در نخستین آزمایش که در جدول \* ارائه شد، نقش نوع ورودی در عملکرد مدل بررسی گردید. چهار حالت شامل استفاده از تصویر پایان دیاستول، استفاده از تصویر پایان سیستول، پردازش جداگانه دو تصویر در دو مسیر مستقل و ورود همزمان دو تصویر به صورت استک شده مورد ارزیابی قرار گرفت. نتایج نشان داد که روش استک شده بهترین عملکرد را در میان حالت‌های مختلف داشته است.

استفاده از تصویر منفرد، چه در پایان دیاستول و چه در پایان سیستول، تنها بخشی از اطلاعات چرخه قلب را بازتاب می‌دهد. در پایان دیاستول بیشترین حجم بطن چپ قابل مشاهده است و ارزیابی اتساع، ضخامت دیواره و نسبت حجم‌ها دقیق‌تر انجام می‌شود. در پایان سیستول کمترین حجم بطن ثبت می‌شود و الگوهای انقباضی و کاهش حرکت دیواره واضح‌تر است. بنابراین هر تصویر به تنهایی دید ناقصی از عملکرد قلب ارائه می‌کند و عملکرد مدل‌های تک تصویر محدود می‌ماند.

پردازش جداگانه دو تصویر در دو مسیر مستقل نسبت به حالت تک تصویر نتایج بهتری دارد، زیرا اطلاعات تکمیلی هر دو مرحله وارد مدل می‌شود. با این حال، ترکیب این دو مسیر در لایه‌های انتهایی انجام می‌شود و شبکه امکان یادگیری روابط

مشترک مکانی و زمانی بین پایان دیاستول و پایان سیستول را از ابتدای جریان محاسباتی به دست نمی‌آورد.

در روش استک شده، دو تصویر به صورت کانال‌های همزمان وارد شبکه می‌شوند و از همان لایه‌های اولیه، تفاوت‌ها و شباهت‌های ساختاری به طور مشترک آموخته می‌شود. به این ترتیب تغییرات حجم، حرکت دیواره و الگوهای غیرطبیعی انقباض به صورت ویژگی‌های ترکیبی یکپارچه استخراج می‌گردد و روابط بین تصویری مانند اختلاف شدت و تغییر هندسه بطن توسط کانولوشن‌ها قابل تحلیل می‌شود. این سازوکار موجب بهبود شاخص‌های عملکرد، شامل دقت، صحت، یادآوری، امتیاز اف ۱ و مساحت زیر منحنی، بر اساس میانگین به همراه انحراف معیار، در ورودی دو کاناله شده است.

در دومین آزمایش که در جدول \* ارائه شده است، عملکرد مدل پیشنهادی با معماری‌های شناخته شده یادگیری عمیق شامل وی جی جی ۱۶، وی جی جی ۱۹، رزنت ۱۸ و رزنت ۳۴ مورد مقایسه قرار گرفت. هدف این آزمایش ارزیابی توانایی این معماری‌ها در استخراج ویژگی‌های مرتبط با تصاویر ام آر آی قلب و بررسی میزان برتری مدل پیشنهادی در برابر آن‌ها بود. نتایج نشان داد که مدل پیشنهادی در مقایسه با معماری‌های ذکر شده عملکرد بهتری داشته است. هرچند معماری‌های عمیق شناخته شده عملکرد مناسبی نشان دادند و این امر با توجه به عمق زیاد و توانایی استخراج ویژگی‌های پیچیده قابل انتظار بود، اما

برتری مدل پیشنهادی دلایل مشخصی دارد. نخست آنکه معماری‌های عمومی مانند وی‌جی‌جی و رزنت عمدتاً برای تصاویر طبیعی طراحی شده‌اند، در حالی که تصاویر پزشکی از نظر کنتراست، نویز، دامنه شدت خاکستری و ساختار تفاوت‌های اساسی دارند و ویژگی‌های اولیه آموخته‌شده در این مدل‌ها با داده‌های ام‌آر‌آی قلب همخوانی کامل ندارند. دوم آنکه مدل پیشنهادی به‌طور خاص برای ورود همزمان دو تصویر پایان دیاستول و پایان سیستول طراحی شده است، در حالی که معماری‌های عمومی معمولاً ورودی تک‌تصویر دارند و روابط بین دو فریم از ابتدا در طراحی آن‌ها لحاظ نشده است. سوم آنکه شبکه پیشنهادی سبک‌تر است و با تعداد پارامتر کمتر سریع‌تر همگرا می‌شود، در حالی که معماری‌های بسیار عمیق مانند وی‌جی‌جی ۱۶ و وی‌جی‌جی ۱۹ ده‌ها میلیون پارامتر دارند و در مسائل پزشکی با داده‌های محدود احتمال بیش‌برازش در آن‌ها افزایش می‌یابد. در نهایت مهم‌ترین دلیل برتری مدل پیشنهادی امکان ترکیب مشترک دو تصویر در همان مراحل اولیه شبکه است که موجب می‌شود روابط ساختاری و زمانی بین پایان دیاستول و پایان سیستول به‌طور مستقیم آموخته شوند، در حالی که سایر معماری‌ها صرفاً یک تصویر را تحلیل می‌کنند. این ویژگی باعث برتری معنادار مدل پیشنهادی نسبت به معماری‌های شناخته‌شده با وجود عمق و پیچیدگی بیشتر آن‌ها شده است.

در سومین آزمایش که در جدول \* ارائه شده است، عملکرد مدل پیشنهادی با روش‌های سنتی یادگیری ماشین شامل ماشین بردار پشتیبان، درخت تصمیم، جنگل تصادفی، بیز ساده و پرسپترون چندلایه مورد مقایسه قرار گرفت. هدف این آزمایش ارزیابی توانایی الگوریتم‌های کلاسیک در تحلیل مستقیم داده‌های تصویری ام‌آر‌آی قلب و بررسی تفاوت آن‌ها با مدل پیشنهادی بود. نتایج نشان داد که روش‌های سنتی در مقایسه با شبکه پیشنهادی عملکرد ضعیف‌تری داشتند، که این امر با توجه به ماهیت این الگوریتم‌ها قابل انتظار است. روش‌های کلاسیک به‌طور اصولی برای داده‌های خام تصویری مناسب نیستند و نیازمند استخراج دستی ویژگی‌ها، مهندسی ویژگی و کاهش

ابعاد با روش‌هایی مانند تحلیل مؤلفه‌های اصلی هستند. از آنجا که تصاویر ام‌آر‌آی قلب دارای ویژگی‌های پیچیده‌ای همچون الگوی حرکت دیواره، تغییرات حجم در دو مرحله، ضخامت و نامتقارن بودن عضله قلب، نواحی هایپرتروفی یا آتروفی و تغییرات سیگنال در بافت هستند، استخراج این ویژگی‌ها تنها با لایه‌های چندگانه کانولوشن امکان‌پذیر است و روش‌های سنتی توانایی انجام آن را ندارند. علاوه بر این، یادگیری روابط بین دو تصویر پایان دیاستول و پایان سیستول تنها با شبکه‌های کانولوشنی قابل انجام است، زیرا الگوهای زمانی و ساختاری مانند تفاوت حجم، اختلاف شدت، تغییر هندسه بطن و کاهش حرکت موضعی تنها در قالب ورودی استک‌شده قابل کشف هستند. روش‌های سنتی فاقد توانایی درک این روابط مشترک بوده و نمی‌توانند ترکیب بهینه ویژگی‌ها را از داده خام بیاموزند. در مقابل، مدل پیشنهادی نه تنها ویژگی‌ها را استخراج می‌کند، بلکه ترکیب آن‌ها را نیز به‌طور هدفمند یاد می‌گیرد و همین امر دلیل اصلی برتری آن نسبت به الگوریتم‌های کلاسیک است. در میان روش‌های سنتی، شبکه پرسپترون چندلایه عملکرد نسبی بهتری نسبت به سایر الگوریتم‌ها داشت، زیرا توانایی یادگیری روابط غیرخطی و ترکیب ویژگی‌ها را تا حدی داراست. با این حال، این شبکه نیز فاقد سازوکار لازم برای استخراج مستقیم ویژگی‌های پیچیده تصویری و درک روابط مشترک بین دو فریم پایان دیاستول و پایان سیستول است. بنابراین، هرچند پرسپترون چندلایه نسبت به سایر روش‌های کلاسیک نتایج بهتری ارائه کرد، اما همچنان فاصله قابل توجهی با عملکرد مدل پیشنهادی دارد.

### نتیجه‌گیری

در این مطالعه، یک مدل مبتنی بر شبکه عصبی کانولوشنی دوکاناله برای تشخیص بیماری‌های قلبی ارائه شد که توانست با دقت بالا و کاهش وابستگی به مداخله انسانی، عملکرد قابل توجهی داشته باشد. نتایج آزمایش‌ها بر روی مجموعه داده ACDC نشان داد که رویکرد دوکاناله نسبت به مدل‌های تک‌فریم و موازی و همچنین معماری‌های مرجع دقت و پایداری بالاتری ارائه می‌دهد. این یافته‌ها نشان می‌دهند که

استفاده از ورودی‌های استک‌شده دوکاناله و تنظیم دقیق هاپرپارامترها، می‌تواند به بهبود عملکرد شبکه‌های عصبی کانولوشنی در کاربردهای پزشکی کمک کند. با توجه به محدودیت‌های این مطالعه، پیشنهاد می‌شود در تحقیقات آینده، از مجموعه داده‌های بزرگ‌تر و متنوع‌تر شامل بیماران با شرایط قلبی مختلف استفاده شود تا تعمیم‌پذیری مدل افزایش یابد. همچنین امکان ترکیب داده‌های چند مدالیته مانند اکوکاردیوگرافی یا سی تی اسکن قلب با تصاویر ام‌آرآی برای

بهبود دقت و تشخیص زودهنگام بررسی شود. در نهایت نیز استفاده از معماری‌های پیشرفته‌تر یادگیری عمیق یا مدل‌های ترکیبی شبکه عصبی کانولوشنی و شبکه‌های عصبی بازگشتی برای بهره‌برداری از اطلاعات زمانی چرخه قلبی مدنظر قرار گیرد. پیاده‌سازی این پیشنهادات می‌تواند کاربردهای عملی روش پیشنهادی را در محیط‌های بالینی گسترش دهد و موجب تشخیص سریع‌تر و دقیق‌تر بیماری‌های قلبی شود.

### منابع و مراجع

1. Farrant JP, Black N, Yamagata K, Soltani F, Orsborne C, Yan CK, et al. Comprehensive reference ranges for cardiovascular magnetic resonance: time to move on from single centre data? *Int J Cardiovasc Imaging* 2025 415 [Internet]. 2025 Mar 27 [cited 2025 Dec 12];41(5):913–32. Available from: <https://link.springer.com/article/10.1007/s10554-025-03370-5>
2. Mienye ID, Swart TG, Obaido G, Jordan M, Ilono P, Mienye ID, et al. Deep Convolutional Neural Networks in Medical Image Analysis: A Review. *Inf* 2025, Vol 16, [Internet]. 2025 Mar 3 [cited 2025 Dec 12];16(3). Available from: <https://www.mdpi.com/2078-2489/16/3/195>
3. Zheng Q, Delingette H, Ayache N. Explainable cardiac pathology classification on cine MRI with motion characterization by semi-supervised learning of apparent flow. *Med Image Anal* [Internet]. 2018 Nov 8 [cited 2025 Dec 12];56:80–95. Available from: <https://arxiv.org/pdf/1811.03433>
4. Janik A, Dodd J, Ifrim G, Sankaran K, Curran KM. Interpretability of a Deep Learning Model in the Application of Cardiac MRI Segmentation with an ACDC Challenge Dataset. 2021 Mar 15 [cited 2025 Dec 12];11596:111. Available from: <https://arxiv.org/pdf/2103.08590>
5. Khalifa A Ben, Mili M, Maatouk M, Abdallah A Ben, Abdellali M, Gaied S, et al. Deep Transfer Learning for Classification of Late Gadolinium Enhancement Cardiac MRI Images into Myocardial Infarction, Myocarditis, and Healthy Classes: Comparison with Subjective Visual Evaluation. *Diagnostics* 2025, Vol 15, [Internet]. 2025 Jan 17 [cited 2025 Dec 12];15(2). Available from: <https://www.mdpi.com/2075-4418/15/2/207>
6. Wolterink JM, Leiner T, Viergever MA, Išgum I. Automatic Segmentation and Disease Classification Using Cardiac Cine MR Images. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)* [Internet]. 2017 Aug 3 [cited 2025 Dec 12];10663 LNCS:101–10. Available from: <https://arxiv.org/pdf/1708.01141>
7. Schwab M, Pamminger M, Kremser C, Haltmeier M, Mayr A. Deep Learning Pipeline for Fully Automated Myocardial Infarct Segmentation from Clinical Cardiac MR Scans. *Radiol Adv* [Internet]. 2025 Feb 5 [cited 2025 Dec 12];2(4). Available from: <https://arxiv.org/pdf/2502.03272>
8. Zuben A Von, Whitt E, Viana FAC, Perotti LE. Long Axis Cardiac MRI Segmentation Using Anatomically-Guided UNets and Transfer Learning. *Lect Notes Comput Sci (including Subser Lect Notes Artif Intell Lect Notes Bioinformatics)* [Internet]. 2023 [cited 2025 Dec 12];13958 LNCS:274–82. Available from: [https://link.springer.com/chapter/10.1007/978-3-031-35302-4\\_28](https://link.springer.com/chapter/10.1007/978-3-031-35302-4_28)
9. Sandooghdar A, Yaghmaee F. Deep Learning Approach for Cardiac MRI Images. *J Inf Syst Telecommun* vol 1 61 [Internet]. 2022 Feb 19 [cited 2025 Dec 12];1(37):61. Available from: <http://jst.ir/en/Article/16121>
10. Chen C, Mat Isa NA, Liu X. A review of convolutional neural network based methods for medical image classification. *Comput Biol Med* [Internet]. 2025 Feb 1 [cited 2025 Dec 13];185:109507. Available from: <https://www.sciencedirect.com/science/article/abs/pii/S0010482524015920?via%3Dihub>
11. Szandała T. Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks. *Stud Comput Intell* [Internet]. 2021 [cited 2025 Dec 13];903:203–24. Available from: [https://link.springer.com/chapter/10.1007/978-981-15-5495-7\\_11](https://link.springer.com/chapter/10.1007/978-981-15-5495-7_11)
12. Bernard O, Lalande A, Zotti C, Cervenansky F, Yang X, Heng PA, et al. Deep Learning Techniques for Automatic MRI Cardiac Multi-Structures Segmentation and Diagnosis: Is the Problem Solved? *IEEE Trans Med Imaging*. 2018 Nov 1;37(11):2514–25.
13. Islam T, Hafiz MS, Jim JR, Kabir MM, Mridha MF. A systematic review of deep learning data augmentation in medical imaging: Recent advances and future research directions. *Healthc Anal* [Internet]. 2024 Jun 1 [cited

- 2025 Dec 12];5:100340. Available from: [https://www.sciencedirect.com/science/article/pii/S277244252400042X?utm\\_source=chatgpt.com](https://www.sciencedirect.com/science/article/pii/S277244252400042X?utm_source=chatgpt.com)
14. Raiaan MAK, Sakib S, Fahad NM, Mamun A Al, Rahman MA, Shatabda S, et al. A systematic review of hyperparameter optimization techniques in Convolutional Neural Networks. Decis Anal J [Internet]. 2024 Jun 1 [cited 2025 Dec 12];11:100470. Available from: <https://www.sciencedirect.com/science/article/pii/S2772662224000742>
15. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM [Internet]. 2017 Jun 1 [cited 2025 Dec 11];60(6):84–90. Available from: <https://dl.acm.org/doi/pdf/10.1145/3065386>
16. Simonyan K, Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition. 3rd Int Conf Learn Represent ICLR 2015 - Conf Track Proc [Internet]. 2014 Sep 4 [cited 2025 Dec 11]; Available from: <https://arxiv.org/pdf/1409.1556>
17. He K, Zhang X, Ren S, Sun J. Deep Residual Learning for Image Recognition [Internet]. 2016 [cited 2025 Dec 11]. p. 770–8. Available from: <http://image-net.org/challenges/LSVRC/2015/>
18. Pisner DA, Schnyer DM. Support vector machine. Mach Learn Methods Appl to Brain Disord [Internet]. 2020 Jan 1 [cited 2025 Dec 12];101–21. Available from: <https://www.sciencedirect.com/science/chapter/edited-volume/abs/pii/B9780128157398000067>
19. Suthaharan S. Decision Tree Learning. 2016 [cited 2025 Dec 12];237–69. Available from: [https://link.springer.com/chapter/10.1007/978-1-4899-7641-3\\_10](https://link.springer.com/chapter/10.1007/978-1-4899-7641-3_10)
20. Parmar A, Katariya R, Patel V. A Review on Random Forest: An Ensemble Classifier. Lect Notes Data Eng Commun Technol [Internet]. 2019 [cited 2025 Dec 12];26:758–63. Available from: [https://link.springer.com/chapter/10.1007/978-3-030-03146-6\\_86](https://link.springer.com/chapter/10.1007/978-3-030-03146-6_86)
21. Yang FJ. An implementation of naive bayes classifier. Proc - 2018 Int Conf Comput Sci Comput Intell CSCI 2018 [Internet]. 2018 Dec 1 [cited 2025 Dec 12];301–6. Available from: <https://ieeexplore.ieee.org/abstract/document/8947658>
22. Kruse R, Mostaghim S, Borgelt C, Braune C, Steinbrecher M. Multi-layer Perceptrons. 2022 [cited 2025 Dec 12];53–124. Available from: [https://link.springer.com/chapter/10.1007/978-3-030-42227-1\\_5](https://link.springer.com/chapter/10.1007/978-3-030-42227-1_5)

## Original Article

# Automatic Diagnosis of Cardiac Diseases Through Joint Analysis of End-Diastolic and End-Systolic MRI Frames Using a Convolutional Neural Network

Received: 10/09/2025 - Accepted: 19/11/2025

Abdoreza Rezapour<sup>1\*</sup>  
Meghdad Farajpour<sup>2</sup>

<sup>1</sup> Department of Computer Engineering,  
Ast.C., Islamic Azad University, Astaneh  
Ashrafieh, Iran

<sup>2</sup> Department of Managment, MehrAstaneh  
Institute of Higher Education, Astaneh  
Ashrafieh, Iran

Email: abdoreza.rezapour@iau.ac.ir

### Abstract

**Introduction:** Magnetic resonance imaging (MRI) has become increasingly common for the diagnosis of cardiovascular diseases in recent years. However, this technique remains time-consuming and prone to human error. These limitations highlight the need for artificial intelligence, particularly deep learning, which has demonstrated strong capabilities in processing medical images. The aim of this study is to present a convolutional neural network (CNN) model for accurate detection of heart disease while reducing reliance on human intervention.

**Methods:** In this research, a CNN-based model was designed and implemented for cardiac disease diagnosis. Unlike conventional approaches that process images as single frames or in parallel, this study employed dual-channel inputs consisting of end-diastolic and end-systolic images to better capture cardiac cycle dynamics. The proposed architecture included multiple convolutional and max-pooling layers, with dropout layers applied to prevent overfitting. Furthermore, fine-tuning of hyperparameters such as optimizer type and batch size, along with advanced data augmentation techniques, contributed to improved network performance.

**Results:** Experiments were conducted on the standard ACDC dataset, and model performance was evaluated using established metrics. Findings revealed that the dual-channel approach achieved significant improvements compared to single-frame and parallel models. Moreover, comparisons with baseline CNN architectures and traditional machine learning algorithms confirmed the superiority of the proposed method in terms of accuracy and stability.

**Conclusion:** The dual-channel approach can serve as an effective and reliable tool for cardiac disease diagnosis. By offering higher accuracy and reducing dependence on human interpretation, this method has the potential to replace parts of conventional, costly diagnostic processes.

**Keywords:** MRI imaging, Heart disease, Deep learning, Convolutional neural network, End-diastole, End-systole