

بکارگیری مدل مبتنی بر یادگیری عمیق بر روی تصویر چاپ شده سیگنال‌های ECG برای تشخیص بیماری‌های قلبی عروقی

نسیم بیگزاده^۱ و عبدالحسین فتحی^{۲*}

چکیده

تشخیص بیماری‌های قلبی از روی تصاویر چاپ شده نوارهای الکتروکاردیوگرام اهمیت ویژه‌ای دارد، زیرا این نوارها به‌عنوان یک استاندارد جهانی در تمامی مراکز درمانی قابل تفسیر و استفاده هستند. استفاده از یادگیری عمیق برای تحلیل خودکار این نوارها می‌تواند دقت و سرعت تشخیص را به‌طور چشمگیری بهبود دهد. در این پژوهش یک مدل مبتنی بر یادگیری عمیق برای طبقه‌بندی تصاویر چاپ شده سیگنال‌های الکتروکاردیوگرام به پنج کلاس ارائه شده است. برای این منظور ابتدا پیش‌پردازش‌هایی روی تصاویر بدست آمده از سیگنال‌های الکتروکاردیوگرام اعمال شده است. این پیش‌پردازش‌ها برای نرمال‌سازی و تقویت داده‌های تعلیم می‌باشد. سپس با ارتقای مدل از پیش آموزش داده شده ResNet-18 از طریق ترکیب آن با ماژول‌های فشرده‌سازی و تحریک شبکه، سعی شده یک مدل پیشرفته به اسم SEResNet-18 برای بهبود ویژگی‌های استخراج شده و تشخیص دقیق‌تر بیماری‌های قلبی بکار گرفته شود. مدل پایه ResNet-18 و مدل پیشرفته SEResNet-18 روی مجموعه داده PTB-XL پیاده‌سازی و ارزیابی شده‌اند که در نهایت مدل پایه ResNet-18 توانست به صحت ۸۱/۴۵٪ و مساحت زیر منحنی ۹۵/۵۴٪ دست پیدا کند در حالیکه در مدل پیشرفته با افزودن ماژول فشرده‌سازی و تحریک، عملکرد مدل به شکل قابل توجهی بهبود یافت، به طوری که صحت مدل پیشرفته به ۸۵/۱۶٪ و مساحت زیر منحنی به ۹۶/۶۸٪ افزایش پیدا کرد. این کارایی در مقایسه با روش‌های موجود نیز برتر می‌باشد.

کلید واژه‌ها

الکتروکاردیوگرام، بیماری‌های قلبی، مدل ResNet-18، ماژول فشرده‌سازی و تحریک، یادگیری عمیق

۱- مقدمه

بررسی وضعیت قلب بیماران به کار می‌روند. پزشکان از این نوارها برای شناسایی مشکلاتی مثل آریتمی‌ها، انسداد رگ‌ها، و سایر اختلالات قلبی استفاده می‌کنند. البته تشخیص دقیق این بیماری‌ها به دقت و تجربه پزشک وابسته است و امکان خطا وجود دارد. حتی کوچک‌ترین اشتباه در تفسیر این نوارها می‌تواند برای بیمار پیامدهای خطرناکی داشته باشد. از این رو، نیاز به روش‌های خودکار و فناوری‌های جدید که بتوانند این سیگنال‌ها را با دقت و سرعت بیشتری تحلیل کنند، کاملاً حس می‌شود تا علاوه بر کاهش احتمال خطا، تشخیص بیماری‌های قلبی نیز بهبود یابد. مدل‌های یادگیری عمیق به‌عنوان روشی جدید و کارآمد برای تفسیر خودکار و دقیق این سیگنال‌ها، توجه بسیاری را به خود جلب کرده‌اند [۱]. این مدل‌ها با قدرت بالا در استخراج و تحلیل

الکتروکاردیوگرام از ابزارهای کلیدی و پراهمیت در تشخیص بیماری‌های قلبی به شمار می‌رود. این سیگنال‌های الکتریکی قلب که به صورت نوارهای چاپی در دسترس قرار می‌گیرند، به‌عنوان یک استاندارد جهانی شناخته شده، در تمامی مراکز درمانی برای

این مقاله در آبان‌ماه سال ۱۴۰۳ دریافت شد؛ در دی‌ماه بازنگری و سپس پذیرفته شد.

^۱ دانشجوی کارشناسی ارشد مهندسی کامپیوتر - هوش مصنوعی، گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه رازی، کرمانشاه

رایانامه: n.beigzadeh@stu.razi.ac.ir

^۲ دانشیار گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشگاه رازی، کرمانشاه

رایانامه: a.fathi@razi.ac.ir

نویسنده مسئول: عبدالحسین فتحی

اضافی، تصویر ECG را به عنوان ورودی دریافت کرده و با دقت بالایی COVID-19 را شناسایی کرده است. Mathunjwa و همکاران [۹] یک روش جدید برای طبقه‌بندی اختلالات قلبی با استفاده از ۲D-CNN و روش نمودار بازگشتی ارائه داده‌اند. در این روش، سیگنال‌های ECG به تصاویر دوبعدی تبدیل شده‌اند که نشان‌دهنده الگوهای بازگشتی در سیگنال‌های قلبی هستند. این تصاویر به CNN داده شده تا انواع مختلف بیماری قلبی را تشخیص دهد Safdar. و همکاران [۱۰] نیز از ترکیب ۲D-CNN و طیف‌نگاری مبتنی بر تبدیل فوریه کوتاه مدت برای طبقه‌بندی سیگنال‌های ECG استفاده کردند. Li و همکاران [۱۱] به بررسی یک روش یادگیری عمیق برای طبقه‌بندی بیماری‌های قلبی با استفاده از تجزیه حالت تجربی (EMD) برای استخراج ویژگی‌های سیگنال‌های ECG پرداخته‌اند. آنها داده‌های ECG را پس از پردازش با EMD به ویژگی‌های تصویری دوبعدی تبدیل کرده و سپس توسط یک ۲D-CNN طبقه‌بندی نموده‌اند. Pal و همکاران [۱۲] با بهره‌گیری از مدل DenseNet پیش‌آموزش دیده‌بر روی مجموعه داده ImageNet، سیگنال‌های ECG را به تصاویر دوبعدی دیجیتال تبدیل کرده و از این تصاویر برای آموزش و ارزیابی مدل نهایی استفاده کرده‌اند Narotamo. و همکاران [۱۳] از دو رویکرد برای پردازش سیگنال‌های یک‌بعدی نیز تبدیل این سیگنال‌ها به تصاویر دوبعدی جهت طبقه‌بندی با استفاده از CNN استفاده کرده‌اند. آنها همچنین با روش هم‌جوشی چندحالتی، ترکیبی از شبکه‌های سیگنال‌های یک‌بعدی و تصاویر دوبعدی برای افزایش دقت به کار گرفته‌اند Liu. و همکاران [۱۴] نیز از ترکیب شبکه‌های با ورودی‌های چند حالتی (شامل سیگنال یک‌بعدی، اطلاعات مورفولوژی و داده دوبعدی وضعیت نرمال) در تحلیل و تشخیص بیماری‌های قلبی استفاده کردند. Brisk و همکاران [۱۵] نیز سیگنال‌های ECG مجموعه داده‌های AF PhysioNet را به تصاویر دوبعدی تبدیل و با معماری ۳۴ لایه‌ای CNN تحلیل کرده‌اند.

انور و ذاکر [۱۶] به بررسی تأثیر روش‌های تقویت تصویر برای طبقه‌بندی بیماری‌ها با استفاده از تصاویر ECG و مدل یادگیری عمیق EfficientNet B3 پرداختند. آنها چهار نوع تقویت تصویر شامل برش، چرخش، تغییر کنتراست و تغییر گاما را بر روی تصاویر اعمال کردند و نشان دادند که استفاده از تقویت تصویر در برخی موارد باعث کاهش دقت و معیار F1 شده و بدون اعمال تقویت تصاویر، دقت بالاتری بدست آوردند Khan. و همکاران [۱۷] به توسعه یک مدل شبکه عصبی عمیق با استفاده از معماری SSD MobileNet V2 روی تصاویر ECG برای تشخیص چهار نوع اختلال قلبی پرداختند Zhang. و همکاران

ویژگی‌های پیچیده سیگنال‌های الکتروکاردیوگرام (ECG)، امکان تشخیص خودکار و دقیق اختلالات قلبی را فراهم می‌کنند [۲]. در سال‌های اخیر مطالعات متعددی برای استفاده از شبکه عصبی پیچشی یک بعدی که برای پردازش داده‌های تک‌بعدی مانند سیگنال‌ها و سری‌های زمانی استفاده می‌شوند، جهت تشخیص اختلالات قلبی از روی سیگنال‌های ECG صورت گرفته است [۳-۶]. Jahmunah و همکاران [۳] به توسعه مدل‌های یادگیری عمیق DenseNet و شبکه عصبی پیچشی برای تشخیص سگته قلبی از طریق سیگنال‌های ECG پرداختند Han. و همکاران [۴] یک روش تشخیص سگته قلبی را با استفاده از شبکه DenseNet و گراف دانش برای پیش‌بینی زمان و شدت سگته قلبی ارائه دادند. در این روش، با استفاده از ویژگی‌های سیگنال‌های ECG، نمودار دانش بالینی ایجاد شده است که روابط بین ویژگی‌ها و ساختار دانش مرتبط با سگته قلبی را استخراج کرده است. احمد و همکاران [۵] به بررسی روش‌های تشخیص اختلالات قلبی از طریق سیگنال‌های ECG روی مجموعه داده MIT-BIH با استفاده از شبکه ۱D-CNN برای طبقه‌بندی بیماری‌های قلبی پرداخته‌اند. Wang و همکاران [۶] نیز با استفاده از شبکه ۱D-CNN و ویژگی‌های محاسباتی مانند آنتروپی و پیچیدگی موج QRS به بهبود دقت طبقه‌بندی سیگنال‌های ECG روی مجموعه داده PTB_XL پرداخته‌اند. نتایج نشان می‌دهد که ویژگی‌های مبتنی بر آنتروپی و پیچیدگی موج QRS، دقت مدل‌ها را بهبود می‌بخشد و این روش برای تشخیص بیماری‌های قلبی مؤثر است. همچنین از شبکه حافظه طولانی کوتاه مدت (LSTM) به همراه تابع اتلاف کانونی برای بهبود تشخیص بیماری‌های قلبی در داده‌های نامتوازن استفاده شده است [۲].

در برخی پژوهش‌های دیگر برای اینکه محققان به تحلیل دقیق الگوهای پیچیده و ویژگی‌های موجود در سیگنال‌های ECG دست یابند، آن‌ها را به یک نمایش بصری قابل فهم تبدیل کرده‌اند. Safdar و همکاران [۷] روشی جدید برای افزایش داده‌های سیگنال ECG جهت بهبود دقت مدل‌های طبقه‌بندی معرفی کرده‌اند. آنها سیگنال‌های اصلی ECG را به قطعات کوچک تقسیم کرده و سپس ترتیب آن‌ها را تغییر داده‌اند تا سیگنال‌های مصنوعی جدید با ساختاری متفاوت اما ویژگی‌های مشابه ایجاد کنند. در این تحقیق، از یک مدل CNN چهارلایه‌ای برای ارزیابی دقت طبقه‌بندی استفاده شده است Sakr. و همکاران [۸] یک مدل عمیق به نام ECG-COVID برای تشخیص COVID-19 از طریق سیگنال‌های ECG ارائه داده‌اند. این مدل بدون نیاز به مراحل

□□□□□□□□□□□□□□□□ (□□□)

1-Dimensional Convolutional Neural Network (1D-CNN)

Convolutional Neural Network (CNN)

□□□□ □□□□□□□□□□□□□□ (□□□□)

Focal Loss

Recurrence Plot (RP)

Short-Time Fourier Transform (STFT)

Empirical Mode Decomposition (EMD)

بزرگترین مجموعه داده‌های در دسترس است ابتدا به فرم تصویر تبدیل می‌شوند سپس جهت تقویت داده عملیات مختلفی مانند تغییر اندازه، چرخش تصادفی و تنظیمات رنگ روی تصاویر آموزش اعمال می‌شود. این موضوع باعث می‌شود دقت طبقه‌بندی مدل افزایش یابد و از بیش‌برازش شدن آن جلوگیری شود. همچنین برای تمام داده‌ها پیش‌پردازش استاندارد سازی Z-Score استفاده می‌شود. سپس برای طبقه‌بندی سیگنال‌های ECG به پنج کلاس مختلف، از معماری بهینه‌شده ResNet-18 که قبلاً آموزش داده شده است استفاده می‌شود. در مدل بهینه شده وزن‌های لایه‌های نهایی، ویژگی‌های پیچیده‌ای که مدل طی آموزش بر روی مجموعه داده بزرگ ImageNet بدست آورده، حفظ می‌شوند، در حالیکه لایه‌های ابتدایی به‌روزرسانی می‌شوند تا بتوانند ویژگی‌های با الگوهای ساده و محلی متناسب با سیگنال‌های ECG را استخراج کنند. علاوه بر این برای ارتقا و بهبود مدل ResNet-18، ماژول فشرده سازی و تحریک (SE) در لایه‌های انتهایی آن تعبیه می‌شود. این ماژول با تمرکز بر روی ویژگی‌های خاص و حساس ECG که مرتبط با تشخیص بیماری‌های قلبی هستند، آن‌ها را با دقت بیشتری استخراج می‌کنند و باعث می‌شوند مدل عملکرد بهتری در طبقه‌بندی نهایی داشته باشد با این تنظیمات شبکه SEResNet-18 پیشنهادی به دقت بالا و عملکرد بهتر در طبقه‌بندی سیگنال‌های ECG دست می‌یابد و می‌تواند در توسعه سیستم‌های پشتیبانی تصمیم‌گیری پزشکی نقش بسزایی داشته باشد.

نو آوری‌های پژوهش عبارتند از:

- استفاده از تصاویر سیگنال ECG به عنوان ورودی مدل که به دلیل استاندارد بودن و دسترس پذیری بالا در مراکز درمانی، می‌تواند در طیف وسیع‌تری از موارد پزشکی و تشخیصی به کار گرفته شود و به متخصصان بیشتری کمک کند.
- از روش تقویت داده با هدف افزایش تنوع داده، بهبود دقت مدل و همچنین جلوگیری از بیش‌برازش شدن استفاده شده است.
- بدست آوردن یک شبکه ResNet-18 بهبود یافته با بکارگیری ماژول SE در برخی از لایه‌های این شبکه که قادر است بر روی مجموعه داده PTB_XL طبقه‌بندی با دقت بالاتری را انجام دهد.

۲- روش پیشنهادی

روش پیشنهادی برای طبقه‌بندی تصاویر ECG با استفاده از مدل یادگیری عمیق در شکل ۱ نشان داده شده است. این پژوهش شامل ۴ مرحله اصلی است: تبدیل سیگنال ECG به تصاویر ECG، بخش تقویت داده، بخش پیش‌پردازش داده و در نهایت آموزش دو شبکه ResNet18 و SEResNet-18 برای طبقه‌بندی تصاویر ECG.

[۱۸] از ترکیب مکانی و زمانی گراف پیچشی مبتنی بر شبکه ResNet برای تحلیل سیگنال‌های ECG استفاده کردند. Fatema و همکاران [۱۹] ابتدا چندین روش پیش‌پردازش، شامل حذف نویز و افزایش کنتراست بر روی تصاویر اعمال کرده تا کیفیت داده‌ها افزایش یابد. سپس پنج مدل یادگیری عمیق شامل VGG19، MobileNetV2، ResNet50، InceptionV3 و DenseNet201 را ارزیابی کردند. در نهایت، یک مدل ترکیبی به نام InRes-106 با ترکیب InceptionV3 و ResNet50 ایجاد کرده که باعث بهبود دقت نسبت به سایر مدل‌ها شده بود. Mhamdi و همکاران [۲۰] نیز ترکیب مدل‌های MobileNet V2 و VGG16 را برای طبقه‌بندی ریتم‌های نرمال و غیرنرمال قلبی بکار گرفته اند. Shi و همکاران [۲۱] از ترکیب مدل‌های مولد (generative) و contrastive برای بهبود تشخیص بیماری‌های قلبی در مدل‌های عمیق استفاده کردند. Strodtthoff و همکاران [۲۲] یک Ensemble از مدل‌های عمیق برای تشخیص بیماری‌های قلبی استفاده کردند. Dissanayake و همکاران [۲۳] از ترکیب شبکه دیکانولوشن و LSTM برای تشخیص بیماری‌های قلبی استفاده کردند. Zhang و همکاران [۲۴] و Qiang و همکاران [۲۵] از چکانش دانش برای غلبه بر پیچیدگی محاسباتی مدل‌های عمیق بخصوص مدل‌های ترکیبی استفاده کردند.

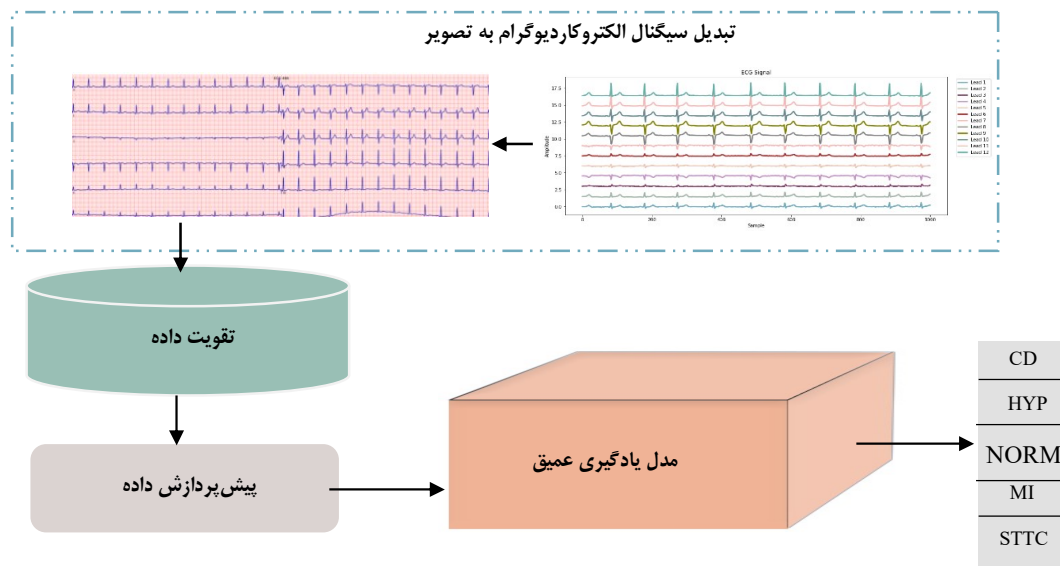
استفاده از ترکیب مدل‌های عمیق با ویژگی‌های استخراج شده دستی نیز برای تحلیل و طبقه‌بندی سیگنال‌های ECG در سال‌های اخیر مورد توجه قرار گرفته است. Eleyan و همکاران [۲۶] از ترکیب ویژگی‌های استخراج شده از تبدیل فوریه، هیستوگرام گرادیان جهت دار (HOG) و الگوهای دودویی محلی (LPB) با شبکه عمیق GRU برای طبقه‌بندی بیماری‌های قلبی استفاده کردند. Ramkumar و همکاران [۲۷] از ترکیب الگوهای دودویی محلی (LPB) با یک Ensemble از شبکه‌های CNN، LSTM و RRHOS برای تشخیص بیماری‌های قلبی استفاده کردند. Houssein و همکاران [۲۸] نیز از ترکیب ویژگی‌های مختلف استخراج شده با تبدیل موجک، الگوهای دودویی محلی (LBP) و ویژگی‌های آماری و ممانی با شبکه پیچشی عمیق برای تحلیل تصاویر ECG استفاده کردند.

این رویکردها نشان دادند که تبدیل داده‌های تک‌بعدی ECG به تصاویر دوبعدی به بهبود دقت و شناسایی ویژگی‌های مهم سیگنال کمک می‌کند. اگر تصاویر سیگنال‌های ECG که به عنوان ابزار استاندارد در دسترس عموم در مراکز درمانی موجود هستند را به عنوان ورودی شبکه‌های یادگیرنده عمیق در نظر بگیریم، تشخیص بیماری‌ها کار بردی تری و قابل‌درک تر خواهد شد. این رویکرد به پزشکان کمک می‌کند تا با استفاده از استانداردها و تجربیات موجود، به نتایج قابل اعتمادتر و سریع‌تری دست یابند. در این پژوهش با توجه به اهمیت استفاده از تصاویر چاپ شده ECG جهت قابلیت اجرا و بهره‌برداری، سیگنال‌های مجموعه داده PTB_XL که یکی از

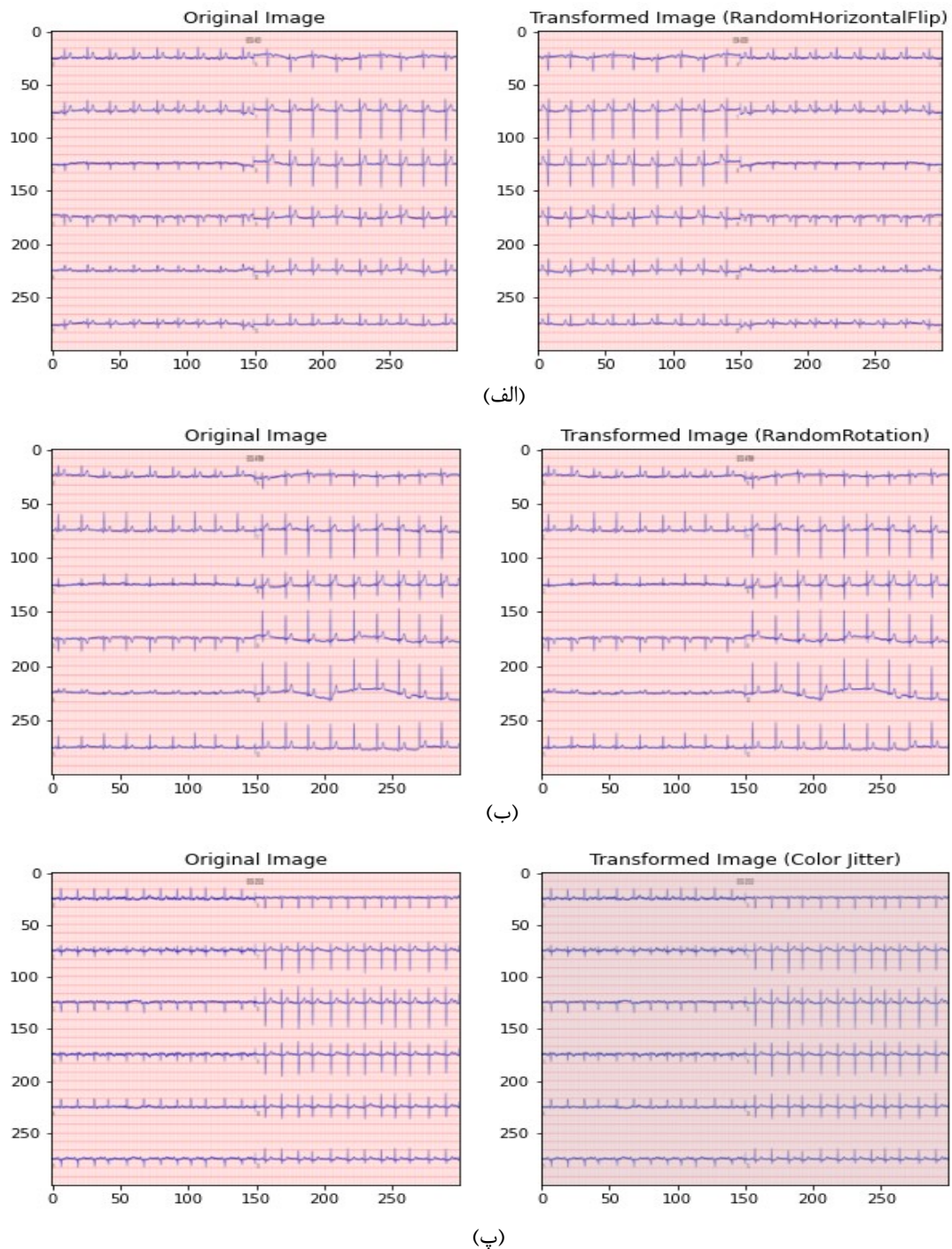
۲_۱_ تقویت داده:

در این پژوهش، از روش‌های متنوع تقویت داده استفاده شده تا مدل بتواند عملکرد بهتری در طبقه‌بندی تصاویر ECG داشته باشد و در برابر تغییرات و نویزهای مختلف مقاومت بیشتری پیدا کند. این روش‌ها امکان دسترسی مدل به نمونه‌های گوناگون‌تری از داده را فراهم می‌کنند و باعث می‌شوند مدل در شرایط متنوع‌تری به درستی عمل کند. در ابتدا، همه‌ی تصاویر به اندازه 300×300 پیکسل تغییر مقیاس داده شدند تا تمامی تصاویر مجموعه داده ابعاد یکسانی داشته باشند. سپس به منظور تقویت داده بر روی داده‌ی آموزشی چرخش افقی تصادفی، چرخش تصادفی و تنظیمات روشنایی، کنتراست، اشباع و رنگ اعمال می‌شوند که به ترتیب باعث می‌شوند، تصاویر با احتمال ۵۰ درصد به صورت افقی برگردانده شوند این کار به مدل کمک می‌کند ویژگی‌ها را از زوایای مختلف یاد بگیرد و نسبت به جهت‌گیری‌های متفاوت تصاویر حساسیت کمتری داشته باشد. سپس تصاویر با زاویه‌ی کوچکی (حداکثر بین ۱۰ تا ۱۰- درجه) به‌طور تصادفی چرخانده شدند تا مدل نسبت به چرخش‌های جزئی مقاوم‌تر شود. در نهایت روشنایی، کنتراست، اشباع و رنگ تصاویر به‌صورت تصادفی تغییر داده شد تا مدل بتواند در شرایط مختلف نوری و رنگی نیز به درستی عمل کند و نسبت به تغییرات محیطی مقاوم‌تر باشد. در شکل ۲ تغییراتی از تصاویر در هنگام اعمال تقویت داده نمایش داده شده است.

ابتدا سیگنال‌های دیجیتال ECG که به صورت داده‌ی سری زمانی در دسترس هستند، با استفاده از کد پایتون به تصاویری با ابعاد 300×300 پیکسلی تبدیل می‌شوند که شبیه به نوارهای ECG چاپ شده بر روی کاغذهای سنتی هستند. برای ایجاد تصاویر با شباهت بیشتر به نوارهای چاپی ECG، عناصر گرافیکی همچون خطوط شبکه‌ای پس‌زمینه، مقیاس‌بندی دقیق و نمای کلی کاغذی به تصاویر اضافه می‌شوند. این فرایند باعث می‌شود تصاویر حاصل نه تنها داده‌های سیگنال را به شکل گرافیکی نمایش دهند، بلکه ویژگی‌های ظاهری نوارهای واقعی ECG را نیز شبیه‌سازی کنند، تا مدل بتواند روی داده‌های تصویری مشابه نوارهای کاغذی آموزش ببیند. در مرحله تقویت داده، روش‌های مختلفی به کار گرفته می‌شوند تا تنوع تصاویر ECG افزایش یابد و مدل بتواند در برابر تغییرات مختلف مقاوم شود و دقت بیشتری در طبقه‌بندی داشته باشد. سپس در قسمت پیش‌پردازش برای توزیع مناسب داده و کمک به یادگیری بهتر شبکه از روش استاندارد سازی Z-Score بر روی تصاویر ECG استفاده می‌شود. در مرحله نهایی، مدل ResNet-18 به گونه‌ای تنظیم می‌شود که برای طبقه‌بندی تصاویر ECG آموزش داده شود. همچنین ماژول SE که توانایی استخراج ویژگی‌های مهم از تصاویر ECG را دارد در بین بعضی لایه‌های مدل ResNet-18 تعبیه می‌شود که این فرآیند آموزشی به مدل‌ها کمک می‌کند تا با دقت بالا، الگوهای بیماری‌های قلبی را در تصاویر ECG شناسایی و طبقه‌بندی کنند و به عنوان یک ابزار دقیق برای کمک به تشخیص‌های پزشکی عمل کنند.



شکل ۱. مراحل روش پیشنهادی، تبدیل سیگنال ECG به تصاویر ECG، تقویت داده و پیش‌پردازش داده و طبقه‌بندی با استفاده از مدل یادگیری عمیق.



شکل ۲. (الف) به صورت تصادفی تصویر را به صورت افقی (چپ به راست) با احتمال ۵۰ درصد برمی گرداند. (ب) تصویر چرخش یافته به صورت تصادفی. (پ) تغییرات در روشنایی، کنتراست، اشباع و رنگ را نمایش می دهد.

۲_۲_ پیش پردازش:

در این پژوهش، پیش پردازش داده با استفاده از استاندارد سازی Z-Score به منظور آماده سازی تصاویر ECG برای ورود به مدل انجام می شود. در روش استاندارد سازی، هر پیکسل تصویر به گونه ای تغییر می کند که داده ها به یک توزیع نرمال با میانگین صفر و انحراف معیار یک تبدیل شوند. برای هر کانال $\mathbf{c} = \{R, G\}$ ، میانگین کل هر کانال از طریق جمع میانگین پیکسل های هر

تصویر در کانال مورد نظر و سپس تقسیم آن بر تعداد کل نمونه ها به صورت فرمول ۱ محاسبه می شود:

$$\mu_c = \frac{1}{N} \sum_{i=1}^N \text{mean}(X_i^c) \quad (1)$$

سپس انحراف معیار هر کانال $\mathbf{c} = \{R, G\}$ به طریق مشابه بصورت زیر محاسبه می شود:

$$\sigma_c = \frac{1}{N} \sum_{i=1}^N \text{std}(X_i^c) \quad (2)$$

کرنل 3×3 به تصویر اضافه می‌شود که به کاهش بیشتر ابعاد تصویر کمک می‌کند. در لایه اول، دو بلوک قرار دارد که هر بلوک شامل دو لایه پیچشی با 3×3 کرنل و گام ۱ است. در هر یک از این بلوک‌ها، پس از هر لایه پیچشی، لایه نرمالسازی دسته‌ای و تابع فعال‌سازی ReLU به منظور بهبود یادگیری افزوده شده‌اند. در لایه دوم، هر بلوک شامل دو لایه پیچشی با ۱۲۸ فیلتر، نرمالسازی دسته‌ای، تابع فعال‌سازی ReLU، یک مسیر میان‌بر و کاهش نمونه برای تنظیم ابعاد است.

لایه سوم شامل دو بلاک پیچشی با ۲۵۶ کرنل و اندازه 3×3 است. در اولین بلوک، برای کاهش ابعاد، گام ۲ در کرنل پیچشی آن اعمال شده، یک مسیر میان‌بر و کاهش نمونه به کار گرفته شده تا ابعاد ورودی و خروجی هماهنگ شوند. هر لایه پیچشی بانرمال‌سازی دسته‌ای و تابع فعال‌سازی ReLU همراه است.

در لایه چهارم هر بلوک دارای دو لایه پیچشی با ۵۱۲ کرنل با اندازه 3×3 می‌باشد. یک مسیر میان‌بر کاهش نمونه جهت هماهنگی ابعاد استفاده شده است. مشابه لایه‌های قبلی، این بلوک‌ها هم از نرمال‌سازی دسته‌ای و ReLU برای بهبود کارایی یادگیری بهره می‌برند. وزن‌های این دو لایه ثابت نگه‌داشته شده‌اند تا به عنوان ویژگی‌های ثابت و عمومی برای مدل عمل کنند. در نهایت از یک تجمع میانگین تطبیقی استفاده می‌کند که به‌طور خودکار ابعاد و ویژگی‌های خروجی را به یک اندازه ثابت کاهش می‌دهد.

در این پژوهش برای جلوگیری از بیش‌برازش در لایه کاملاً متصل روش حذف تصادفی با احتمال ۵۰ درصد در نظر گرفته شده است. برای تنظیم بهتر وزن‌های لایه‌های مختلف این معماری و افزایش دقت و سرعت یادگیری مدل ارائه شده، بجای استفاده از وزن‌ها و تنظیمات اولیه که بصورت تصادفی انتخاب شده باشند، از روش انتقال یادگیری بهره گرفته شده است. برای این منظور از مدل ResNet-18 که قبلاً روی تصاویر مجموعه داده حجیم ImageNet آموزش داده شده استفاده کرده و ساختار لایه کاملاً متصل آن متناسب با داده‌ها و کلاس‌های بیماری‌های قلبی تغییر داده شده است. در تنظیم مجدد لایه‌ها هنگام استفاده از تکنیک انتقال یادگیری، باید ویژگی‌های سطح بالا را مجدداً تولید کرد و این ویژگی‌ها مربوط به لایه‌های آخر است و لایه‌های اول ثابت فرض می‌شود. اما با توجه به اینکه در این مساله تشابهی بین ورودی‌های مدل از پیش آموزش دیده شده با ImageNet و داده‌های ECG وجود ندارد، تنظیم مجدد ویژگی‌های لایه‌های اول مهمتر به نظر می‌رسد. از طرفی استخراج ویژگی‌های محلی مناسب داده‌های تصویری ECG (که در لایه‌های اول انجام می‌شود) به مدل

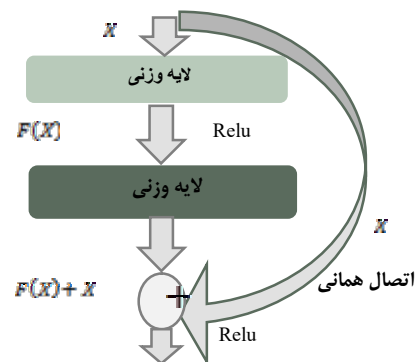
در نهایت استاندارد سازی هر کانال از پیکسل‌های تصویر با نماد Z نشان داده شده است و از مقادیر کانال پیکسل‌های تصویر $Image$ به دست می‌آید، به شکل زیر قابل انجام است:

$$Z_{(c,t)} = \frac{Image_{(c,t)} - \mu_c}{\sigma_c} \quad (3)$$

این پیش‌پردازش به مدل کمک می‌کند تا داده‌ی ورودی متعادلی را دریافت کند و تأثیر مقادیر بسیار بزرگ یا بسیار کوچک در مقایسه با دیگر مقادیر کاهش یابد که به بهبود یادگیری مدل کمک خواهد کرد.

۲_۳ مدل Resnet_18 :

شبکه ResNet-18 یکی از معماری‌های پرکاربرد در شبکه‌های عصبی عمیق است که برای حل مشکل ناپایداری گرادینان‌ها در شبکه‌های بسیار عمیق و افزایش دقت مدل طراحی شده است. این شبکه از بلوک‌های رسوبی که در شکل ۳ نمایش داده شده است استفاده می‌کند، که به مدل اجازه می‌دهد اطلاعات را از طریق مسیرهای میان‌بر به لایه‌های عمیق‌تر انتقال دهد. این ویژگی باعث می‌شود اگر برخی از ویژگی‌ها در طول مسیر پردازش کامل یاد گرفته نشوند یا اهمیت کمتری پیدا کنند، از طریق مسیر میان‌بر به لایه‌های عمیق‌تر منتقل شوند تا مدل سریع‌تر همگرا شود و عملکرد بهتری داشته باشد [۲۹] و [۳۰]. شبکه ResNet-18 شامل ۱۷ لایه پیچشی و ۱ لایه کاملاً متصل است که به ۴ بلوک یا لایه رسوبی تقسیم شده‌اند. معماری کلی این شبکه ResNet-18 در شکل ۴ نمایش داده شده است.



شکل ۳. ساختار بلوک‌های رسوبی.

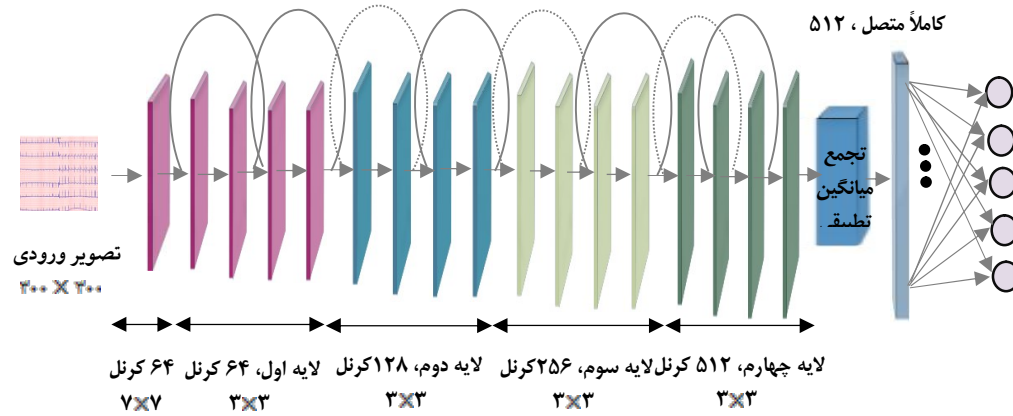
در قسمت ابتدایی شبکه ResNet-18، یک لایه پیچشی با $64 \times 7 \times 7$ و گام ۲ استفاده شده که ابعاد تصویر را کاهش می‌دهد. سپس یک لایه نرمالسازی دسته‌ای برای بهبود همگرایی و یک تابع فعال‌سازی ReLU برای یادگیری ویژگی‌های غیرخطی بکار گرفته شده‌اند. در پایان این بخش، یک لایه تجمع بیشینه با

Downsample
AdaptiveAverage Pooling
Dropout
Transfer Learning

Residual Blocks
Stride
Batch Normalization
Max Pooling

به انتهای هر کدام دو لایه آخر یک ماژول SE اضافه شده تا مدل بتواند ویژگی‌های نهایی سطح بالا را (که در لایه های آخر انجام می شود) با دقت بالایی جهت شناسایی بیماری‌های قلبی استخراج کند.

کمک می‌کند تا پایه‌ای قوی برای یادگیری جزئیات کوچک که نشانه‌های مهمی برای تحلیل و شناسایی ویژگی‌های قلبی هستند فراهم کند، لذا در روش پیشنهادی وزن‌های دو لایه اول با استفاده از داده‌های تصویری ECG مجدد آموزش داده می‌شوند. مضافاً،

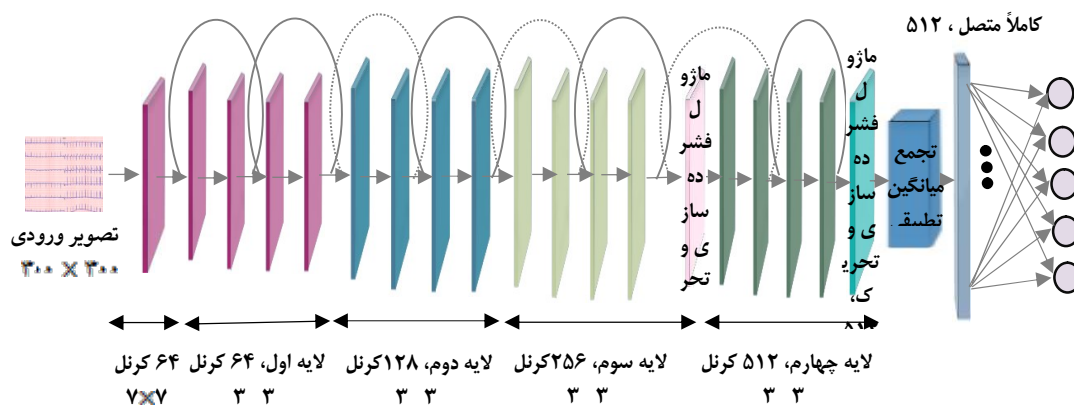


شکل ۴. شبکه ResNet-18.

البته برای تمرکز بیشتر روی ویژگی‌های مهم تصاویر ECG، ماژول SE بعد از لایه های سوم و چهارم قرار داده شده است. برای این منظور خروجی لایه های پیشگی با خروجی مسیر میان بر جمع می‌شود، سپس بر روی اطلاعات بدست آمده ماژول SE اعمال می‌شود تا به ویژگی‌های مهم‌تر وزن بیشتری اختصاص دهد و ویژگی‌های غیرضروری را کم‌اهمیت‌تر کند. به این ترتیب مدل می‌تواند هم ویژگی‌های محلی مناسب تصاویر ECG را در لایه‌های اول بهتر استخراج کند و هم تمرکز بیشتری بر اطلاعات کلیدی تصاویر ECG در لایه‌های آخر داشته باشد.

۲_۴ مدل SEResNet-18:

ساختار مدل ResNet-18 به همراه ماژول SE در شکل ۵ نشان داده شده است. در این مدل، شبکه ResNet-18 به گونه‌ای تنظیم شده که بتواند توسط لایه‌های پیشگی ابتدایی، ویژگی‌های جزئی تصاویر ECG را در طی فرآیند آموزش مجدد یاد بگیرد و استخراج کند. همچنین، از ویژگی‌های از پیش آموزش دیده شده قبلی در لایه‌های انتهایی ResNet-18 استفاده شده تا نتایج مطلوبی به دست آید.



شکل ۵. ساختار مدل ResNet-18 به همراه ماژول SE.

۳- نتایج آزمایشات:

در این پژوهش، ارزیابی‌ها و پیاده‌سازی‌ها بر روی سیستم عامل Windows 10 و سیستمی با سخت‌افزار شامل پردازنده Intel(R) Core(TM) i9، حافظه RAM با ظرفیت ۱۲۸ گیگابایت و کارت گرافیک Nvidia RTX 3090 انجام شده است. شبکه پایه-ResNet-18 در طول ۱۵۰ اپیاک روی مجموعه داده ImageNet آموزش داده شده، در حالی که نسخه اصلاح شده آن با ماژول SE تنها طی ۴۵ اپیاک به نتایج مطلوب دست یافت. برای هر دو شبکه، نرخ یادگیری ۰/۱، مقدار کاهش برابر ۰/۰۰۰۰۱، اندازه دسته‌ی برابر ۱۲۸، بهینه‌ساز گرادیان کاهشی تصادفی حذف تصادفی با مقدار ۰/۵ برای لایه‌های تمام‌متصل تنظیم شده است.

برای ارزیابی عملکرد مدل‌ها معیارهای مختلف شامل دقت^{۱۱} بازخوانی^{۱۲}، معیار-F1، ویژگی^{۱۳}، مساحت زیر منحنی ویژگی عملکرد گیرنده^{۱۴} و صحت‌لار نظر گرفته شده است. چهار پارامتر اصلی، شامل مثبت واقعی^{۱۵}، منفی واقعی^{۱۶}، مثبت کاذب^{۱۷} و منفی کاذب^{۱۸} برای محاسبه این معیارها به کار گرفته شده‌اند.

$$\text{Recall} = \frac{TP}{TP+FN} \quad (7)$$

$$\text{Precision} = \frac{TP}{TP+FP} \quad (8)$$

$$F_1 = \frac{2 \times (\text{Precision} \times \text{Recall})}{(\text{Precision} + \text{Recall})} \quad (9)$$

$$\text{Specificity} = \frac{TN}{TN+FP} \quad (10)$$

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (11)$$

$$\text{AUROC} = \int_0^1 \text{CurveROC} \quad (12)$$

جدول ۱. تعداد داده‌های هر کلاس از مجموعه داده PTB_XL

Data Class	Train	Test	Validation
NORM	7243	912	914
CD	1353	184	171
HYP	415	56	64
STTC	1903	242	255
MI	2043	256	233
Total	12957	1650	1637

Batch Size

Precision

Recall

Specificity

Area Under the Receiver Operating Characteristic Curve

(AUROC)

Accuracy

True Positive (TP)

True Negative (TN)

False Positive (FP)

False Negative (FN)

مرحل پیاده سازی ماژول SE که در شکل ۶ نمایش داده شده است شامل دو مرحله هست: مرحله فشرده‌سازی و تحریک^{۱۹} [۳۱]. ورودی تصویر ECG که $X \in R^{H \times W}$ است به شبکه ResNet-18 داده می‌شود بعد از عبور از بلوک‌های رسوبی ویژگی استخراج می‌شود. در مرحله ی تحریک از تجمع میانگین سراسری برای فشرده‌سازی هر کانال از ویژگی‌های $U = [u_1, u_2, \dots, u_C] \in R^{H' \times W'}$ استفاده می‌شود. که در اینجا ، و به ترتیب هر کدام ارتفاع، عرض و تعداد کانال‌های ویژگی است و نتیجه این مرحله به صورت فرمول ۴ نمایش داده می‌شود:

$$z_k = \frac{1}{H' \times W'} \sum_{i=1}^{H'} \sum_{j=1}^{W'} u_k(i, j), \text{fork} = 1, 2, \dots, C' \quad (13)$$

بردار $z \in R^C$ که نمایانگر ویژگی‌های فشرده شده هر کانال است از یک شبکه عصبی کوچک عبور داده می‌شود تا اهمیت هر کانال را یاد بگیرد. این شبکه شامل دو لایه کاملاً متصل است که ابتدا یک لایه کاملاً متصل $FC_1 \in R^C$ روی آن اعمال می‌شود و خروجی بدست آمده از تابع فعال‌ساز ReLU که با نماد نمایش داده شده است عبور می‌کند. سپس از لایه کاملاً متصل دوم $FC_2 \in R^C$ عبور داده شده و تابع فعال‌ساز سیگموئید که با نماد نشان داده شده است، به آن اعمال می‌شود. در این لایه ها برابر ۱۶ در نظر گرفته شده است و نتیجه نهایی آنها به صورت معادله ۵ نمایش داده شده است:

$$s = F_{ex}(z, FC) = \sigma(g(z, FC)) \sigma(FC_2) \sigma(FC_1) \quad (14)$$

بردار که حاوی وزن‌های توجه برای هر کانال است به صورت عنصر به عنصر با ویژگی‌های استخراج شده از لایه پیش‌بینی شبکه پایه ResNet-18 ضرب می‌شوند و بردار ویژگی وزن دار^{۲۰} $\tilde{X} = [\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_C]$ بدست آید. که فرمول آن به شکل زیر است:

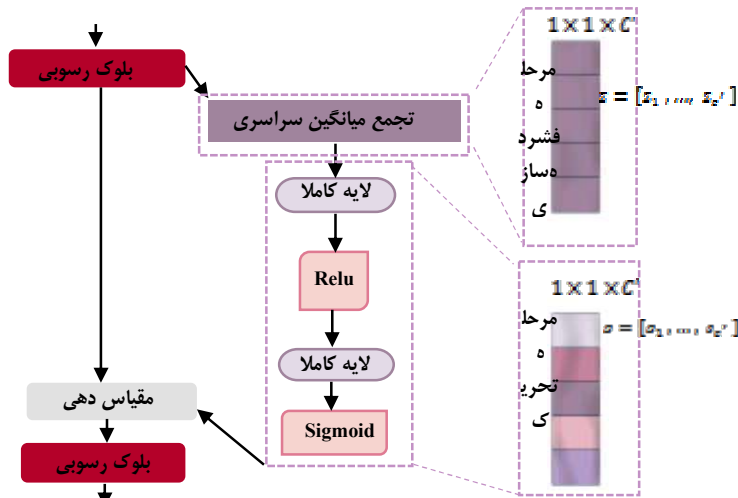
$$\tilde{x}_k = F_{scale}(u_k, s_k) = u_k \cdot s_k, \text{fork} = 1, 2, \dots, C' \quad (15)$$

به این ترتیب، کانال‌هایی که وزن توجه بیشتری دارند تقویت می‌شوند و کانال‌هایی با وزن کمتر تأثیر کمتری خواهند داشت. این فرآیند کمک می‌کند که مدل، تمرکز بیشتری بر روی ویژگی‌های مهم‌تر هر تصویر داشته باشد و به بهبود دقت و عملکرد در استخراج ویژگی‌های تصاویر ECG کمک کند.

Squeeze

Excitation

Global Average Pooling

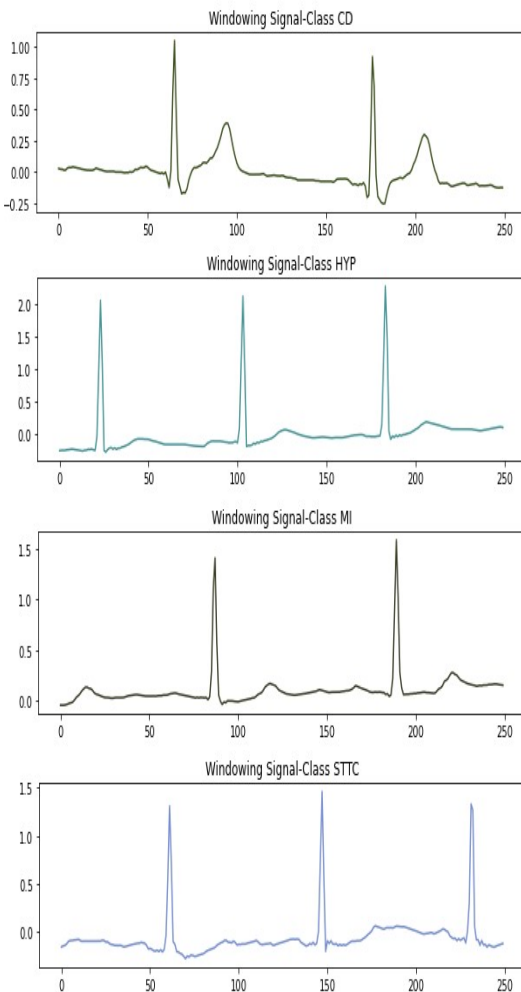


شکل ۶. مراحل پیاده‌سازی مازول SE.

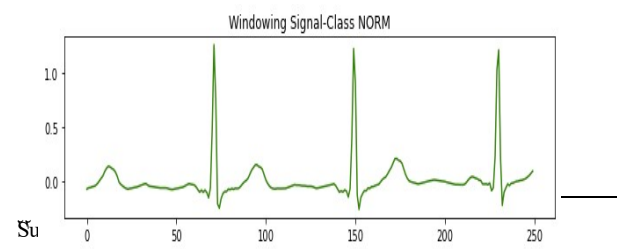
۱_۳_ مجموعه داده:

مجموعه داده PTB-XL، با داشتن مجموعه‌ای بزرگ از سیگنال‌های ECG و تنوع بالای بیماران، به عنوان یکی از منابع اصلی و ارزشمند برای تحقیقات در زمینه تشخیص خودکار بیماری‌های قلبی شناخته می‌شود. این مجموعه شامل ۲۱۷۹۹ نمونه سیگنال ECG از ۱۸۸۶۹ بیمار است که هر نمونه، نوار ECG با ۱۲ لید و مدت ۱۰ ثانیه با فرکانس‌های ۵۰۰ و ۱۰۰ هرتز را در بر می‌گیرد. در این پژوهش از دسته‌بندی ابرتشخیصی آبا پنج کلاس اصلی استفاده شده است: تغییرات ST/T^۳، حمله قلبی (MI)، طبیعی^۴ (NORM)، اختلال در هدایت (CD) و بزرگ‌شدگی عضله قلب^۵ (HYP). یک نمونه از سیگنال‌های این اختلالات را در شکل ۷ می‌توانید مشاهده نمایید.

علاوه بر این برای بهبود کارایی پردازش و کاهش پیچیدگی محاسبات، سیگنال‌های ۱۰۰ هرتز نیز مورد تحلیل قرار گرفت. طبق مرجع [۳۲]، بخش‌های ۱ تا ۸ داده برای آموزش و بخش‌های ۹ و ۱۰ به ترتیب برای اعتبارسنجی و تست در نظر گرفته شده است. تعداد نمونه‌های این مجموعه داده در جدول ۱ نمایش داده شده است.



شکل ۷. یک نمونه از سیگنال‌های مربوط به اختلالات قلبی



ST/T Change (STTC)

Norm

Conduction Disturbance (CD)

Hypertrophy (HYP)

Validation

جدول ۲. ارزیابی مدل ResNet-18 برای طبقه‌بندی تصاویر سیگنال‌های ECG

ResNet-18	Precision	Recall	F1_score	Specificity	AUROC
NORM	0.8625	0.9703	0.9133	0.8089	0.9835
CD	0.8098	0.7173	0.7608	0.9788	0.9702
HYP	0.571	0.2857	0.3810	0.9924	0.9358
STTC	0.7113	0.7231	0.7172	0.9495	0.9582
MI	0.7272	0.5312	0.6140	0.9634	0.9294
Average	0.7363	0.6455	0.6777	0.9386	0.9554

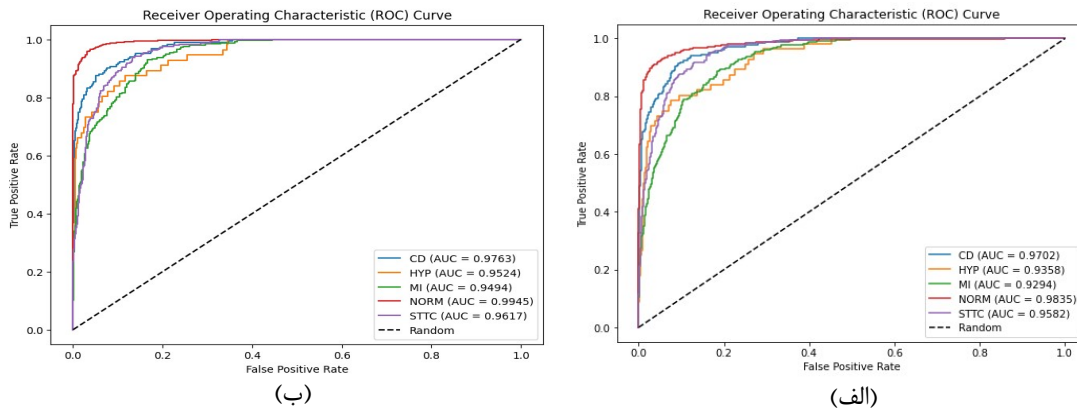
۳-۲_ ارزیابی‌ها:

SE در انتهای لایه‌های مختلف و البته آموزش مجدد لایه‌های موجود مورد ارزیابی قرار گرفت. جدول ۴ نتایج ارزیابی میانگین کلاس‌های مختلف مدل‌های را نمایش می‌دهد. ابتدا برای ارزیابی میزان کارایی استفاده از روش انتقال یادگیری و بکارگیری مدل از پیش آموزش داده شده (با توجه به عدم شباهت بین داده‌های ورودی دو مدل) حالتی که وزنها و تنظیمات اولیه مدل که بصورت تصادفی انتخاب شدند (ردیف ۱ جدول ۴) با حالتی که وزنها و تنظیمات اولیه از قبل آموزش دیده شدند (ردیف ۲ جدول ۴) پیاده‌سازی و مقایسه شده‌اند. در این ارزیابی‌ها، مدل ResNet-18 بدون استفاده از انتقال یادگیری و بدون ماژول SE (مدل ۱ جدول ۴) به Precision برابر با ۰/۷۳۶۳، Recall برابر با ۰/۶۴۵۵، F1 Score برابر با ۰/۶۷۷۷، Accuracy برابر با ۰/۸۱۴۵، Specificity برابر با ۰/۹۳۸۶ و AUROC برابر با ۰/۹۵۵۴ دست یافته است. همچنین مدل ResNet-18 که فقط از انتقال یادگیری استفاده کرده و از ماژول SE استفاده نکرده (مدل ۲ جدول ۴) به Precision برابر با ۰/۷۳۶۳، Recall برابر با ۰/۶۴۵۵، F1 Score برابر با ۰/۶۷۷۷، Accuracy برابر با ۰/۸۱۴۵، Specificity برابر با ۰/۹۳۸۶ و AUROC برابر با ۰/۹۵۵۴ دست یافته است. همانطور که مشاهده می‌کنید علیرغم عدم تشابه بین ورودی‌های دو مدل، بهره‌گیری از روش انتقال یادگیری باعث بهبود در تمام معیارها شده است بخصوص در معیار F1_score و Recall بیش از ۹٪ بهبود دارد.

برای ارزیابی تاثیر ماژول SE، این ماژول در انتهای هرکدام از لایه‌های مدل پایه اضافه شد و همان لایه هم مجدداً آموزش داده شد (مدل‌های ردیف ۳ تا ۶ جدول ۴). سپس ماژول SE در انتهای هرکدام از دو لایه اول و دوم، دو لایه سوم و چهارم، دو لایه دوم و سوم و نیز بصورت تصادفی دو لایه اول چهارم مدل پایه انتخاب و اضافه شدند و همان لایه‌ها هم مجدداً آموزش داده شدند (مدل‌های ردیف ۷ تا ۱۰ جدول ۴). همچنین ماژول SE در انتهای سه لایه اول (مدل ردیف ۹ جدول ۴) و هر چهار لایه (مدل ردیف ۱۰ جدول ۴) مدل پایه نیز اضافه شد و همان لایه‌ها هم مجدداً آموزش داده شدند. دو حالت دیگر نیز مورد ارزیابی قرار گرفت اول اضافه کردن ماژول SE به انتهای لایه‌های اول و دوم و آموزش مجدد لایه‌های سوم و چهارم (مدل ردیف ۱۱ جدول ۴) و دوم اضافه کردن ماژول SE به انتهای لایه‌های سوم و چهارم و آموزش مجدد لایه‌های اول و دوم (مدل ردیف ۱۲ جدول ۴) مدل برتر که با نام SEResNet-18 در این مقاله خوانده می‌شود.

ارزیابی مدل پایه ResNet-18 برای طبقه‌بندی تصاویر سیگنال‌های ECG در جدول ۲ گزارش شده است. کلاس NORM بالاترین میزان recall با مقدار ۰/۹۷۰۳ و Precision با مقدار ۰/۸۶۲۵ را نشان می‌دهد، که در مجموع منجر به F1_score برابر با ۰/۹۱۳۳ شده است. Specificity این کلاس نیز بسیار بالا و برابر با ۰/۹۸۳۵ است. کلاس CD با Precision برابر ۰/۸۰۹۸ و recall برابر ۰/۷۱۷۳ به F1_score معادل ۰/۷۶۰۸ رسیده و Specificity آن برابر با ۰/۹۷۸۸ است. در کلاس HYP، عملکرد مدل با Precision برابر ۰/۵۷۱ و recall برابر ۰/۲۸۵۷ منجر به F1_score پایین‌تر معادل ۰/۳۸۱۰ شده است، و Specificity این کلاس برابر با ۰/۹۹۲۴ است. در کلاس STTC، مدل به Precision برابر ۰/۷۱۱۳ و recall برابر ۰/۷۲۳۱ دست یافته که F1_score معادل ۰/۷۱۷۲ را به همراه داشته و Specificity آن برابر با ۰/۹۴۹۵ است. در کلاس MI، Precision و recall به ترتیب برابر با ۰/۵۳۱۲ و ۰/۶۱۴۰ بوده که منجر به F1_score برابر با ۰/۶۱۴۰ و Specificity برابر ۰/۹۶۳۴ شده است.

جدول ۳ نتایج ارزیابی مدل بهینه شده SEResNet-18 را برای طبقه‌بندی تصاویر سیگنال ECG نشان می‌دهد. کلاس NORM با بهترین عملکرد، دارای Precision برابر با ۰/۸۹۴۳، Recall برابر با ۰/۹۹۲۳ و F1 Score معادل ۰/۹۴۰۷ است. در حالی که Specificity برابر با ۰/۸۵۵۰ و AUROC برابر با ۰/۹۹۴۵ نیز نشان‌دهنده دقت بالای مدل در تشخیص این نوع سیگنال‌ها باشد. این مدل برای کلاس CD، در Precision برابر با ۰/۸۸۶۷ و Recall برابر با ۰/۷۲۲۸ به F1 Score برابر با ۰/۷۹۶۴ و Specificity برابر با ۰/۹۸۸۴ دست یافته است که نشان‌دهنده عملکرد خوب در این طبقه‌بندی می‌باشد. برای HYP، مدل عملکرد متوسطی دارد. Precision آن برابر با ۰/۷۱۴۳، Recall برابر با ۰/۵۳۵۷ و AUROC برابر با ۰/۹۵۲۴ است. در کلاس STTC و MI نیز مدل عملکرد نسبتاً خوبی نشان داده است، به‌ویژه در کلاس STTC که AUROC برابر با ۰/۹۶۱۷ است که نمودارهای AUROC در شکل ۷ هم قابل مشاهده هستند. برای اضافه کردن ماژول SE و نیز آموزش مجدد لایه‌های مدل پایه از پیش آموزش دیده ResNet-18 انتخاب‌های متعددی وجود دارد. برای بدست آوردن بهترین ترکیب، مدل‌های متعددی با قرار دادن ماژول



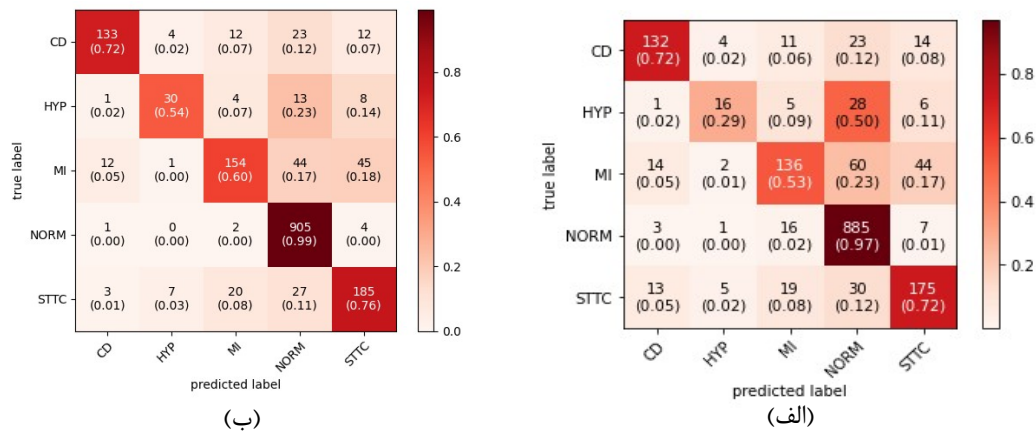
شکل ۷. نمودار AROC برای ۵ کلاس. نمودار (الف) مربوط به مدل ResNet-18 و نمودار (ب) مربوط به مدل SEResNet-18.

جدول ۳. ارزیابی مدل SEResNet-18 برای طبقه‌بندی تصاویر سیگنال‌های ECG.

SEResNet-18	Precision	Recall	F1 score	Specificity	AUROC
NORM	0.8943	0.9923	0.9407	0.8550	0.9945
CD	0.8867	0.7228	0.7964	0.9884	0.9763
HYP	0.7143	0.5357	0.6122	0.9925	0.9524
STTC	0.7283	0.7645	0.7460	0.9510	0.9617
MI	0.8021	0.6016	0.6875	0.9727	0.9494

جدول ۴. مقدار میانگین ارزیابی مدل ResNet-18 و ترکیبات مختلف ماژول SE روی تصاویر سیگنال‌های ECG.

Model	لايه‌های دارای SE لايه‌های Retrain	Precision	Recall	F1_score	Accuracy	Specificity	AUROC	زمان تست	زمان آموزش
۱	بدون SE ResNet-18 بدون پیش‌آموزش	0.6933	0.5507	0.5843	0.7770	0.9164	0.9297	27.08 s	4h54m44s
۲	بدون SE ResNet-18 ReTrain(1,2,3,4)	0.7363	0.6455	0.6777	0.8145	0.9386	0.9554	26.18 s	4h50m15s
۳	SE(1) ReTrain(1)	0.6828	0.5880	0.6177	0.7964	0.9312	0.9366	29.71 s	5h12m28s
۴	SE(2) ReTrain(2)	0.7561	0.7011	0.7214	0.8382	0.9498	0.9625	25.46 s	4h42m55s
۵	SE(3) ReTrain(3)	0.7164	0.6493	0.6703	0.7933	0.9321	0.9423	25.35 s	4h45m6s
۶	SE(4) ReTrain(4)	0.6946	0.6622	0.6765	0.7770	0.9290	0.9155	28.07 s	4h54m27s
۷	SE(1,2) ReTrain(1,2)	0.7551	0.7190	0.7317	0.8467	0.9542	0.9647	27.32 s	4h53m11s
۸	SE(3,4) ReTrain(3,4)	0.7116	0.6835	0.6941	0.7702	0.9309	0.9287	30.00 s	4h49m38s
۹	SE(1,4) ReTrain(1,4)	0.6499	0.5625	0.5802	0.7836	0.9226	0.9048	27.7 s	4h53m55s
۱۰	SE(2,3) ReTrain(2,3)	0.7058	0.6602	0.6810	0.8030	0.9382	0.9395	29.05s	4h52m60s
۱۱	SE(1,2,3) ReTrain(1,2,3)	0.8009	0.6991	0.7505	0.8352	0.9450	0.9682	26.57 s	4h54m9s
۱۲	SE(1,2,3,4) ReTrain(1,2,3,4)	0.7918	0.7128	0.7424	0.8213	0.9459	0.9540	27.19 s	4h52m15s
۱۳	SE(1,2) ReTrain(3,4)	0.7200	0.6832	0.6987	0.7909	0.9350	0.9349	30.00 s	4h49m38s
۱۴	SE(3,4) ReTrain(1,2) SEResNet-18	0.8051	0.7233	0.7565	0.8516	0.9515	0.9668	27.47s	4h54m12s



شکل ۸. ماتریس درهم ریختگی (الف) مربوط به مدل ResNet-18 و (ب) مربوط به مدل SEResNet-18.

طبقه‌بندی تصاویر سیگنال‌های ECG، نتایج مطلوبی کسب کرده است.

مضافاً برای ارزیابی بیشتر روش پیشنهادی روی دیتاست چاپمن [۳۳] اعمال شده و با سایر روش‌های موجود روی این دیتاست مقایسه شده است. این پایگاه داده دارای ۱۰۶۴۶ نمونه با نرخ نمونه‌برداری ۵۰۰ هرتز بوده و هر رکورد شامل ۱۰ ثانیه سیگنال ECG با ۱۲ لید است. این پایگاه داده شامل چهار نوع سیگنال است: برادی‌کاردی سینوسی (SB)، ریتم سینوسی (SR)، فیبریلاسیون دهلیزی (AFIB) و تاکی‌کاردی فوق بطنی عمومی (GSVT). این پایگاه داده به صورت تصادفی به سه مجموعه آموزشی، اعتبارسنجی و تست بصورت ۷۰ درصد آموزشی ۲۰ درصد تست و ۱۰ درصد اعتبارسنجی تقسیم بندی شده است. نتایج این ارزیابی در جدول ۶ آورده شده است. این مدل بالاترین کارایی را در Precision، Specificity و AUROC به ترتیب با مقادیر ۰/۹۵۶۴، ۰/۹۸۶۸ و ۰/۹۹۵۱ نسبت به سایر روش‌ها دارد. در روش پیشنهادی با مقدار Recall برابر ۰/۹۵۴۵ و F1 score برابر با ۰/۹۵۶۳ و Accuracy برابر با ۰/۹۶۰۹ نتایج قابل رقابت با برترین روش‌های موجود داشته است. در معیار recall روش ارائه شده توسط Zhang و همکاران [۲۴] با ۰/۹۵۷ بالاترین مقدار را داشته و در Accuracy و F1 score روش به ترتیب با مقادیر ۰/۹۷۰۹ و ۰/۹۶۶۴ بالاترین کارایی را داشته که تفاوت چندانی نسبت به روش ارائه شده ندارند و به خوبی قابلیت روش پیشنهادی را مشخص می‌سازد.

مدل نهایی SEResNet-18 (مدل ۱۲ جدول ۴) در معیارهای Precision با مقدار ۰/۸۰۵۱، Recall با مقدار ۰/۷۲۳۳، F1 Score با دقت برابر با ۰/۷۵۶۵ و Accuracy با مقدار برابر با ۰/۸۵۱۶ از سایر مدل‌های ارزیابی شده برتر می‌باشد. در معیارهای Specificity با مقدار ۰/۹۵۱۵ فقط از مدل ۷ کمتر و در معیار AUROC با مقدار برابر ۰/۹۶۶۸ فقط از مدل ۷ کمتر بوده و از سایر مدل‌ها بالاتر می‌باشد. این مقادیر نشان می‌دهند که افزودن ماژول SE به مدل ResNet-18 از پیش آموزش داده شده منجر به بهبود قابل توجهی در عملکرد مدل در طبقه‌بندی سیگنال‌های ECG شده است. برای ارزیابی بیشتر ماتریس درهم ریختگی مدل پایه با یادگیری انتقالی و مدل پیشنهادی در شکل ۸ نشان داده شده است.

۳_۳_ مقایسه مدل پیشنهادی با سایر مدل‌ها:

برای ارزیابی نهایی، مدل پیشنهادی SEResNet-18 با سایر مدل‌هایی که آنها نیز از مجموعه داده PTB_XL استفاده کرده‌اند در جدول ۵ مقایسه شده است. مدل SEResNet-18 در مقایسه با سایر مدل‌ها عملکرد بالاتری در تمام معیارهای کلیدی ارزیابی دارد که در جدول ۵ گزارش شده است. این مدل بالاترین Precision را با مقدار ۰/۸۰۵۱ و Recall را با مقدار ۰/۷۲۳۳ نشان داده و با F1 score برابر با ۰/۷۵۶۵ از مدل‌هایی مانند Zhang و همکاران [۱۸]، Liu، و همکاران [۱۴] و Zhang و همکاران [۲۴] پیشی گرفته است. Accuracy مدل برابر با ۰/۸۵۱۶ و Specificity آن ۰/۹۵۱۵ است که نشان‌دهنده توانایی این مدل در کاهش خطاهای تشخیص و تمایز دقیق کلاس‌ها می‌باشد. همچنین AUROC مدل با مقدار ۰/۹۶۶۸، بهترین عملکرد را در میان مدل‌های مقایسه شده ارائه کرده است. به طور خلاصه، مدل SEResNet-18 با دقت و اثربخشی بالا در

جدول ۵. مقایسه مدل SEResNet-18 با سایر روش‌ها روی دیتاست PTB_XL.

Models	Precision	Recall	F1 score	Accuracy	Specificity	AUROC
ST-REGE [18]	0.742	0.582	0.633	0.766	–	0.891
GCL + CE [21]	–	0.5715	–	0.7706	0.9199	0.8949
SSRL [14]	0.7128	0.5923	0.6169	–	–	0.9046
MT-MVT-Net [25]	–	–	0.7446	0.6546	–	0.9381
MT-DSVS-Net [25]	–	–	0.7352	0.6450	–	0.9313
MobileNet_Teacher[24]	–	–	0.689	0.801	–	0.874
MobileNet_Student[24]	–	–	0.661	0.781	–	0.882
Ensemble [22]	–	–	–	–	–	0.934
DConv-LSTM [23]	–	–	0.6883	0.6930	–	–
SEResNet-18	0.8051	0.7233	0.7565	0.8516	0.9515	0.9668

جدول ۶. مقایسه مدل SEResNet-18 با سایر روش‌ها روی دیتاست چاپمن.

Models	Precision	Recall	F1 score	Accuracy	Specificity	AUROC
MT-MVT-Net [25]	-	-	0.9664	0.9709	-	98.98
ST-REGE [18]	0.879	0.882	0.880	0.897	-	0.972
GCL + CE [21]	-	0.9447	-	0.9510	0.9840	0.9943
MobileNet[24]	-	0.957	0.945	0.962	-	0.990
SSRL [14]	0.9468	0.9393	0.9425	-	-	0.9943
SEResNet-18	0.9584	0.9545	0.9563	0.9609	0.9868	0.9951

۴- نتیجه‌گیری:

مراجع

- [1] H. M. Lynn, S. B. Pan, and P. Kim, "A Deep Bidirectional GRU Network Model for Biometric Electrocardiogram Classification Based on Recurrent Neural Networks," *IEEE Access*, vol. 7, pp. 145395–145405, 2019, doi: 10.1109/ACCESS.2019.2939947.
- [2] J. Gao, H. Zhang, P. Lu, and Z. Wang, "An Effective LSTM Recurrent Network to Detect Arrhythmia on Imbalanced ECG Dataset," *J. Healthc. Eng.*, vol. 2019, 2019, doi: 10.1155/2019/6320651.
- [3] V. Jahmunah, E. Y. K. Ng, R. S. Tan, S. L. Oh, and U. R. Acharya, "Explainable detection of myocardial infarction using deep learning models with Grad-CAM technique on ECG signals," *Comput. Biol. Med.*, vol. 146, no. February, 2022, doi: 10.1016/j.combiomed.2022.105550.
- [4] C. Han, S. Pan, W. Que, Z. Wang, Y. Zhai, and L. Shi, "Automated localization and severity period prediction of myocardial infarction with clinical interpretability based on deep learning and knowledge graph," *Expert Syst. Appl.*, vol. 209, no. February, p. 118398, 2022, doi: 10.1016/j.eswa.2022.118398.
- [5] A. A. Ahmed, W. Ali, T. A. A. Abdullah, and S. J. Malebary, "Classifying Cardiac Arrhythmia from ECG Signal Using 1D CNN Deep Learning Model," *Mathematics*, vol. 11, no. 3, pp. 1–16, 2023, doi:

در این پژوهش، مدل‌های ResNet-18 و SEResNet-18 برای طبقه‌بندی سیگنال‌های ECG چاپ‌شده استفاده شده است. نتایج نشان دادند که تنظیم بهینه شبکه مدل‌های ResNet-18 باعث پیشرفت دقت مدل در طبقه‌بندی بیماری‌های قلبی شده است. علاوه بر این اضافه کردن ماژول SE به معماری ResNet-18، هم باعث بهبود چشمگیر عملکرد مدل در تشخیص دقیق‌تر بیماری‌های قلبی شده است. در این مدل بهبود یافته، Precision به ۰/۷۵۶۵ و F1 Score به ۰/۷۲۳۳ و Recall به ۰/۸۰۵۱ یافت که در مقایسه با سایر مدل‌ها مقادیر بالاتری هستند. AUROC مدل با مقدار ۰/۹۶۶۸ به بالاترین سطح در میان مدل‌ها رسیده است و این نشان‌دهنده توانایی برتر مدل در طبقه‌بندی کلاس‌های مختلف سیگنال ECG است. با توجه به نتایج برجسته مدل SEResNet-18، این رویکرد می‌تواند در توسعه سیستم‌های خودکار تشخیص و تصمیم‌گیری پزشکی کاربرد زیادی داشته باشد و به کاهش خطاهای تشخیصی و افزایش سرعت و دقت تشخیص در مراکز درمانی کمک کند. استفاده از تصاویر ECG، این مدل را به یک ابزار عملی و کاربردی در حوزه پزشکی تبدیل کرده است. در آینده می‌توان به بررسی ترکیب معماری‌های عمیق‌تر مانند DenseNet و ResNet-50 به همراه ماژول‌های توجه پرداخت که ممکن است عملکرد مدل را ارتقاء دهد.

- deep neural network," *Complexity*, vol. 2021, 2021, doi: 10.1155/2021/5512243.
- [18] H. Zhang, W. Liu, S. Chang, H. Wang, J. He, and Q. Huang, "ST-ReGE: A Novel Spatial-Temporal Residual Graph Convolutional Network for CVD," *IEEE J. Biomed. Heal. Informatics*, vol. 28, no. 1, pp. 216–227, 2024, doi: 10.1109/JBHI.2023.3327025.
- [19] K. Fatema, S. Montaha, M. A. H. Rony, S. Azam, M. Z. Hasan, and M. Jonkman, "A Robust Framework Combining Image Processing and Deep Learning Hybrid Model to Classify Cardiovascular Diseases Using a Limited Number of Paper-Based Complex ECG Images," *Biomedicines*, vol. 10, no. 11, 2022, doi: 10.3390/biomedicines10112835.
- [20] L. Mhamdi, O. Dammak, F. Cottin, and I. Ben Dhaou, "Artificial Intelligence for Cardiac Diseases Diagnosis and Prediction Using ECG Images on Embedded Systems," *Biomedicines*, vol. 10, no. 8, pp. 1–16, 2022, doi: 10.3390/biomedicines10082013.
- [21] J. Shi et al., "Universal 12-lead ECG representation for signal denoising and cardiovascular disease detection by fusing generative and contrastive learning," *Biomed. Signal Process. Control*, vol. 94, no. April, p. 106253, 2024, doi: 10.1016/j.bspc.2024.106253.
- [22] N. Strodthoff, P. Wagner, T. Schaeffter, and W. Samek, "Deep Learning for ECG Analysis: Benchmarks and Insights from PTB-XL," *IEEE J. Biomed. Heal. Informatics*, vol. 25, no. 5, pp. 1519–1528, 2021, doi: 10.1109/JBHI.2020.3022989.
- [23] T. Dissanayake, T. Fernando, S. Denman, S. Sridharan, and C. Fookes, "DConv-LSTM-Net: A Novel Architecture for Single-and 12-Lead ECG Anomaly Detection," *IEEE Sens. J.*, vol. 23, no. 19, pp. 22763–22776, 2023, doi: 10.1109/JSEN.2023.3300752.
- [24] H. Zhang et al., "DDDg: A dual bi-directional knowledge distillation method with generative self-supervised pre-training and its hardware implementation on SoC for ECG," *Expert Syst. Appl.*, vol. 244, no. January 2023, p. 122969, 2024, doi: 10.1016/j.eswa.2023.122969.
- [25] Y. Qiang, X. Dong, X. Liu, and Y. Yang, "MT-MV-KDF: A novel Multi-Task Multi-View Knowledge Distillation Framework for myocardial infarction detection and localization," *Biomed. Signal Process. Control*, vol. 95, no. PA, p. 106382, 2024, doi: 10.1016/j.bspc.2024.106382.
- [26] A. Eleyan, F. Bayram, and G. Eleyan, "Spectrogram-Based Arrhythmia Classification Using Three-Channel Deep Learning Model with Feature Fusion," *Applied Sciences*, vol. 14, no. 21, 2024, doi: 10.3390/app14219936.
- [27] M. Ramkumar, M. Alagarsamy, A. Balakumar, et al. Ensemble classifier fostered detection of arrhythmia using ECG data. *Med Biol Eng Comput*, vol. 61, pp. 2453–2466, 2023, doi: 10.1007/s11517-023-02839-6
- [28] E.H. Houssein, M. Hassaballah, I.E. Ibrahim, D.S. 10.3390/math11030562.
- [6] J. Wang et al., "Automated ECG classification using a non-local convolutional block attention module," *Comput. Methods Programs Biomed.*, vol. 203, 2021, doi: 10.1016/j.cmpb.2021.106006.
- [7] M. F. Safdar, P. Pałka, R. M. Nowak, and A. Al Faresi, "A novel data augmentation approach for enhancement of ECG signal classification," *Biomed. Signal Process. Control*, vol. 86, no. PB, p. 105114, 2023, doi: 10.1016/j.bspc.2023.105114.
- [8] A. S. Sakr, P. Pławiak, R. Tadeusiewicz, J. Pławiak, M. Sakr, and M. Hammad, "ECG-COVID: An end-to-end deep model based on electrocardiogram for COVID-19 detection," *Inf. Sci. (Nij.)*, vol. 619, no. March 2020, pp. 324–339, 2023, doi: 10.1016/j.ins.2022.11.069.
- [9] B. M. Mathunjwa, Y. T. Lin, C. H. Lin, M. F. Abbod, and J. S. Shieh, "ECG arrhythmia classification by using a recurrence plot and convolutional neural network," *Biomed. Signal Process. Control*, vol. 64, no. February 2020, p. 102262, 2021, doi: 10.1016/j.bspc.2020.102262.
- [10] M. F. Safdar, R. M. Nowak, and P. Pałka, "A Denoising and Fourier Transformation-Based Spectrograms in ECG Classification Using Convolutional Neural Network," *Sensors*, vol. 22, no. 24, 2022, doi: 10.3390/s22249576.
- [11] Y. Li et al., "A deep learning approach to cardiovascular disease classification using empirical mode decomposition for ECG feature extraction," *Biomed. Signal Process. Control*, vol. 79, no. August 2022, 2023, doi: 10.1016/j.bspc.2022.104188.
- [12] A. Pal, R. Srivastva, and Y. N. Singh, "CardioNet: An Efficient ECG Arrhythmia Classification System Using Transfer Learning," *Big Data Res.*, vol. 26, p. 100271, 2021, doi: 10.1016/j.bdr.2021.100271.
- [13] H. Narotamo, M. Dias, R. Santos, A. V. Carreiro, H. Gamboa, and M. Silveira, "Deep learning for ECG classification: A comparative study of 1D and 2D representations and multimodal fusion approaches," *Biomed. Signal Process. Control*, vol. 93, no. December 2023, p. 106141, 2024, doi: 10.1016/j.bspc.2024.106141.
- [14] W. Liu, H. Zhang, S. Chang, H. Wang, J. He, and Q. Huang, "Learning Representations for Multilead Electrocardiograms From Morphology-Rhythm Contrast," *IEEE Trans. Instrum. Meas.*, vol. 73, pp. 1–15, 2024, doi: 10.1109/TIM.2024.3369152.
- [15] R. Brisk et al., "Deep learning to automatically interpret images of the electrocardiogram: Do we need the raw samples?," *J. Electrocardiol.*, vol. 57, no. xxxx, pp. S65–S69, 2019, doi: 10.1016/j.jelectrocard.2019.09.018.
- [16] T. Anwar and S. Zakir, "Effect of Image Augmentation on ECG Image Classification using Deep Learning," *2021 Int. Conf. Artif. Intell. ICAI 2021*, pp. 182–186, 2021, doi: 10.1109/ICAI52203.2021.9445258.
- [17] A. H. Khan, M. Hussain, and M. K. Malik, "Cardiac disorder classification by electrocardiogram sensing using



نسیم بیگزاده دانشجوی کارشناسی ارشد مهندسی کامپیوتر- هوش مصنوعی رباتیکز دانشگاه رازی می باشد. زمینه های پژوهشی مورد علاقه ایشان پردازش سیگنال ها و تصاویر پزشکی، یادگیری عمیق و شناسایی الگو می باشد.



عبدالحسین فتحی دانشیار گروه مهندسی کامپیوتر و فناوری اطلاعات دانشگاه رازی می باشد. زمینه های تحقیقاتی مورد علاقه ایشان پردازش تصویر، بینایی ماشین، شناسایی الگو، تحلیل داده های پزشکی، فشرده سازی اطلاعات و بیومتریک می باشد.

- Abdelminaam, Y.M. Wazery, "An automatic arrhythmia classification model based on improved Marine Predators Algorithm and Convolutions Neural Networks," *Expert Systems with Applications*, vol. 187, 2022, doi:10.1016/j.eswa.2021.115936.
- [29] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, vol. 2016-Decem, pp. 770-778, 2016, doi: 10.1109/CVPR.2016.90.
- [30] A. V. S. Abhisek, V. R. Gurralla, and L. Sahoo, "Resnet18 Model With Sequential Layer For Computing Accuracy On Image Classification Dataset," *Int. J. Creat. Res. Thoughts*, vol. 10, no. 5, pp. 176-181, 2022, [Online]. Available: www.ijcrt.org
- [31] J. Hu, "Squeeze-and-Excitation Networks," *Cvpr*, pp. 7132-7141, 2018, [Online]. Available: http://openaccess.thecvf.com/content_cvpr_2018/html/Hu_Squeeze-and-Excitation_Networks_CVPR_2018_paper.html
- [32] P. Wagner et al., "PTB-XL, a large publicly available electrocardiography dataset," *Sci. Data*, vol. 7, no. 1, pp. 1-15, 2020, doi: 10.1038/s41597-020-0495-6.
- [33] J. Zheng, J. Zhang, S. Danioko, H. Yao, H. Guo, and C. Rakovski, "A 12-lead electrocardiogram database for arrhythmia research covering more than 10,000 patients," *Sci. Data*, vol. 7, no. 1, pp. 1-8, 2020, doi: 10.1038/s41597-020-0386-x.