*مقاله پژوهشی*

# ارائه یک روش برپایه‌ی شبکه‌ی عصبی حافظه‌ی کوتاه‌مدت ماندگار برای تشخیص ناهنجاری در داده‌های سری زمانی سیگنال قلبی

نصرت‌علی اشرفی‌پیامن و بهراد مظاهری

**چکیده: در این پژوهش، یک روش مبتنی بر یادگیری عمیق برای تشخیص ناهنجاری‌های سیگنال قلبی (ECG) ارائه شده است که از شبکه عصبی حافظه کوتاه‌مدت ماندگار (LSTM) بهره می‌برد. ابتدا سیگنال‌های ECG مجموعه‌داده MIT-BIH پیش‌پردازش شدند که شامل حذف نویز، نرمال‌سازی و قطعه‌بندی ضربان‌ها است. سپس ۹ ویژگی کلیدی شامل دامنه امواج P، Q، R، S و T و فواصل زمانی PQ_d، QR_d، RS_d و ST_d از هر ضربان استخراج شد. این ویژگی‌ها به‌عنوان ورودی شبکه LSTM استفاده شدند تا وابستگی‌های زمانی میان آن‌ها مدل‌سازی شود. نتایج نشان داد که مدل پیشنهادی قادر است ریتم‌های طبیعی و غیرطبیعی قلب را با صحت بالا از یکدیگر تفکیک کند. مدل پیشنهادی با استفاده از تنها ۹ ویژگی استخراج‌شده و معماری LSTM، موفق به شناسایی ۱۹ کلاس آریتمی با صحت ۹۷٫۳۳٪ شد؛ در حالی‌که بسیاری از روش‌های پیشین با استفاده از تعداد بیشتری ویژگی و مدل‌های پیچیده‌تر، عملکرد ضعیف‌تری گزارش کرده‌اند. نتایج به‌دست‌آمده نشان می‌دهد که روش پیشنهادی می‌تواند به‌عنوان ابزاری کارآمد برای تشخیص خودکار آریتمی و پایش هوشمند سلامت قلب مورد استفاده قرار گیرد.**

***کلیدواژه:*** ***تشخیص ناهنجاری، حافظه‌ی کوتاه مدت ماندگار، سیگنال قلبی، سری زمانی.***

## ۱- مقدمه

قلب یکی از حیاتی‌ترین اندام‌های بدن انسان است و هرگونه اختلال در عملکرد آن می‌تواند منجر به پیامدهای خطرناک و حتی مرگ شود. یکی از رایج‌ترین ابزارها برای بررسی وضعیت قلب، سیگنال قلبی[1] (ECG) است که فعالیت الکتریکی قلب را ثبت می‌کند. تحلیل این سیگنال‌ها به پزشکان کمک می‌کند تا مشکلاتی نظیر آریتمی یا بی‌نظمی در ضربان قلب را تشخیص دهند. با توجه به پیچیدگی این سیگنال‌ها و نیاز به تحلیل سریع و دقیق، استفاده از روش‌های هوش مصنوعی، به‌ویژه یادگیری عمیق، مورد توجه بسیاری از پژوهشگران قرار گرفته است.

بر اساس آخرین گزارش سازمان بهداشت جهانی[2] (WHO) در سال ۲۰۲۶، بیماری‌های قلبی-عروقی[3] (CVD) با اختلاف قابل‌توجهی، علت اصلی مرگ‌ومیر در جهان باقی مانده‌اند [۱] تا [۳]. بر اساس گزارش GBD 2023 که در سال ۲۰۲۶ به‌روزرسانی شده، تعداد مرگ‌های ناشی از این بیماری‌ها به ۲۰ میلیون نفر در سال رسیده است که حدود یک‌سوم از کل مرگ‌های جهانی را شامل می‌شود [۱] و [۲]. این آمار در حالی ارائه می‌شود که شیوع جهانی بیماری‌های قلبی-عروقی به ۶۲۶ میلیون مورد رسیده و تعداد افراد مبتلا نسبت به سال ۱۹۹۰ بیش از دو برابر شده است [۲]. پیش‌بینی‌های موجود برای سال ۲۰۵۰ نشان می‌دهد که سالانه ۳۵٫۶ میلیون نفر بر اثر این بیماری‌ها جان خود را از دست خواهند داد که لزوم اقدامات فوری پیشگیرانه و درمانی را بیش از پیش آشکار می‌سازد [۲]. بار اقتصادی این بیماری‌ها همچنان سنگین است و نابرابری‌های عمیقی در دسترسی به خدمات درمانی وجود دارد، به‌گونه‌ای که کشورهای با درآمد متوسط، مرگ‌ومیر ناشی از بیماری‌های قلبی را تقریباً دو برابر کشورهای پردرآمد تجربه می‌کنند [۴]. در میان انواع مختلف این بیماری‌ها، فیبریلاسیون دهلیزی[4] (AF) به عنوان شایع‌ترین آریتمی قلبی، با شیوع ۵۹ میلیون مورد در سال ۲۰۲۳، سهم قابل توجهی در آمار سکته‌های مغزی و قلبی دارد [۱]. با توجه به اینکه بخش عمده‌ای از بار این بیماری‌ها به عوامل خطر قابل تغییر مانند فشار خون بالا، چاقی و دیابت نسبت داده می‌شود [۲]، توسعه‌ی سیستم‌های هوشمند برای تشخیص به‌موقع و خودکار آریتمی‌هایی نظیر فیبریلاسیون دهلیزی، می‌تواند گامی حیاتی در کاهش بار عظیم مرگ‌ومیر ناشی از بیماری‌های قلبی باشد. این پژوهش با بهره‌گیری از روش‌های پیشرفته یادگیری عمیق، در راستای پاسخ به این نیاز حیاتی جهانی طراحی شده است. هدف این مقاله، طراحی و ارزیابی یک مدل برای تشخیص خودکار ناهنجاری در سیگنال قلبی است.

در ادامه، ساختار مقاله به این صورت است: در بخش ۲، پیشینه پژوهش و مهم‌ترین مطالعات مرتبط در حوزه تشخیص ناهنجاری‌های سیگنال ECG مرور می‌شود. در بخش ۳، مدل پیشنهادی شامل مراحل پیش‌پردازش، استخراج ویژگی، معماری شبکه LSTM و فرآیند آموزش مدل تشریح می‌شود. بخش ۴ به ارزیابی عملکرد مدل، تحلیل نتایج آزمایش‌ها و مقایسه با روش‌های پایه اختصاص دارد. در نهایت، بخش‌های ۵ و ۶ به‌ترتیب به نتیجه‌گیری، بیان محدودیت‌ها و ارائه پیشنهادهایی برای پژوهش‌های آینده می‌پردازند.

این مقاله در تاریخ ۱۰ دی ماه ۱۴۰۴ دریافت و در تاریخ ۱۳ تیر ماه ۱۴۰۵ بازنگری شد.

نصرت‌علی اشرفی‌پیامن (نویسنده مسئول)، دانشکده فنی و مهندسی، دانشگاه خوارزمی، تهران، ایران، (email: ashrafi@khu.ac.ir).

بهراد مظاهری، گروه برق و کامپیوتر، دانشگاه خوارزمی، تهران، ایران، (email: std_behradmazaheri@khu.ac.ir).

1. Electrocardiogram
2. World Health Organization
3. Cardiovascular Diseases
4. Atrial Fibrillation

## ۲- پیشینه پژوهش

در سال‌های اخیر، روش‌های متعددی برای تحلیل سیگنال ECG ارائه شده‌اند. الگوریتم‌های سنتی مانند Pan-Tompkins برای تشخیص نقاط R کاربرد دارند، اما در مواجهه با نویز یا تغییرات ناگهانی صحت پایینی دارند [۵]. در سیگنال قلبی، موج R اولین انحراف مثبت پس از موج P است و بخشی از کمپلکس QRS محسوب می‌شود [۶]. این موج نشان‌دهنده دپلاریزاسیون (فعال‌سازی الکتریکی) بطن‌ها از پایه به سمت نوک قلب است. همچنین برخی از روش‌ها برپایه‌ی یادگیری ماشین کلاسیک مانند ماشین بردار پشتیبان[۱] (SVM) [۷] یا جنگل تصادفی[۲] (RF) [۸] بوده‌اند. اما این روش‌ها معمولاً نیاز به استخراج دستی ویژگی دارند و قدرت درک وابستگی زمانی را ندارند.

در مطالعات اولیه، روش‌های یادگیری ماشین سنتی مانند ماشین بردار پشتیبان، شبکه‌های عصبی مصنوعی[۳] (ANN) طبقه‌بند بیز ساده[۴] (NB) و الگوریتم K نزدیک‌ترین همسایه[۵] (KNN) برای طبقه‌بندی اختلالات قلبی مورد استفاده قرار گرفته‌اند. با این حال، این الگوریتم‌ها برای دستیابی به عملکرد مطلوب نیازمند استخراج دستی ویژگی‌ها و داده‌های آموزشی گسترده هستند و در مقیاس‌های بزرگ یا داده‌های پیچیده عملکرد قابل‌قبولی ندارند [۹] تا [۱۱]. در مقابل، شبکه‌های عصبی بازگشتی[۶] (RNN) مانند شبکه عصبی حافظه‌ی کوتاه‌مدت ماندگار[۷] (LSTM) [۱۲] قابلیت یادگیری الگوهای زمانی را دارند و عملکرد بهتری در تحلیل داده‌های سری زمانی مانند سیگنال قلبی از خود نشان داده‌اند. پژوهش حاضر نیز بر پایه این مزیت طراحی شده است.

در مقایسه، LSTM با شبکه‌های عصبی پیشخور[۸] (FFNN)، شبکه‌های عصبی پیچشی[۹] (CNN)، واحدهای بازگشتی دروازه‌دار[۱۰] (GRU) و ترانسفورمرها مقایسه می‌شود. در حالی که FFNN فاقد قابلیت‌های حافظه مناسب برای داده‌های سری زمانی است، CNN در استخراج ویژگی‌های محلی برتری دارد اما در مدل‌سازی وابستگی‌های بلندمدت ناتوان است. GRU در مقایسه با LSTM ساده‌تر و سریع‌تر است اما برای کارهای پیچیده کمتر موثر است. ترانسفورمرها قدرتمند هستند اما ممکن است روی مجموعه داده‌های کوچکتر بیش‌برازش ایجاد کنند.

با ظهور یادگیری عمیق [۱۳]، رویکردهای جدیدی برپایه‌ی شبکه‌های عصبی پیچشی و شبکه‌های بازگشتی توسعه یافته‌اند. برای نمونه، وینگ و همکاران روش جدیدی برای طبقه‌بندی ضربان قلب ارائه کردند که شامل یک شبکه‌ی توجه و فرآیند آموزش دومرحله‌ای است. این روش صحت طبقه‌بندی را نسبت به روش‌های سنتی بهبود می‌بخشد [۱۴].

یک نمونه تشخیص ناهنجاری به کمک الگوریتم‌های یادگیری بدون ناظر مقاله‌ی «مدل یادگیری بدون نظارت برای تشخیص ناهنجاری در گراف‌های ناهمگون وزندار» است که توسط خزائی و اشرفی‌پیمن پیشنهاد گردید [۱۵]. این پژوهش به بررسی چالش‌های تشخیص ناهنجاری در گراف‌های ناهمگون می‌پردازد که ویژگی‌های متنوعی از جمله انواع مختلف همسایگی، ویژگی‌های گوناگون و انواع و تعداد مختلف پیوندها را دارند. نویسندگان با استفاده از مدل HetGNN و اعمال تغییراتی در توابع هزینه و پارامترها، ویژگی‌های ساختاری و محتوایی گراف را برای تشخیص ناهنجاری استخراج کرده و سپس از یک خودرمزگذار متغیر[۱۱] (VAE) برای شناسایی گره‌های ناهنجار بر اساس خطای بازسازی استفاده کرده‌اند. نتایج آزمایش‌ها بر روی مجموعه‌داده AMiner نشان‌دهنده عملکرد برتر مدل پیشنهادی نسبت به روش‌های پیشرفته موجود در گراف‌های ناهمگون است [۱۵].

ساوالیا و همکاران رویکردی ترکیبی با استفاده از الگوریتم‌های پرسپترون چند لایه[۱۲] CNN (MLP) پیشنهاد دادند که توانستند تا ۹ کلاس مختلف از ریتم‌های قلبی را با صحت‌های ۸۸٫۷٪ برای MLP و ۸۳٫۵٪ برای CNN طبقه‌بندی کنند. با این حال، این مدل به دلیل اجرای چند مرحله‌ای و بار محاسباتی بالا، کارایی محدودی در زمان اجرا دارد [۱۶].

پتروویچ و همکاران به بهینه‌سازی مدل‌های بازگشتی با استفاده از الگوریتم‌های فراابتکاری مانند بهینه‌سازی ازدحام ذرات[۱۳] (PSO)، الگوریتم ژنتیک[۱۴] (GA)، الگوریتم کرم‌شب‌تاب[۱۵] (FA)، الگوریتم بهینه‌سازی نهنگ[۱۶] (WOA) و الگوریتم سینوس کسینوس[۱۷] (SCA) پرداخته و با بهره‌گیری از لایه‌های توجه، صحت مدل را افزایش دادند. این مدل توانست نرخ خطا را به‌طور قابل‌توجهی کاهش دهد؛ اما به دلیل محدودیت‌های محاسباتی، از جمعیت‌های کوچک و دوره‌های آموزش محدود استفاده کرده و نیاز به بهینه‌سازی‌های بیشتر دارد [۱۷].

ثیل و همکاران مدلی بدون نظارت به نام TCN-AE ارائه دادند که بر پایه خودرمزگذارهای پیچشی زمانی طراحی شده بود. این مدل در یادگیری رفتار نرمال و تشخیص انحراف‌ها عملکرد خوبی نشان داد، اما در مواجهه با ناهنجاری‌های نادر و پیچیده دچار محدودیت بود [۱۸].

در رویکردی دیگر، لی و بولانگر با ترکیب ویژگی‌های دستی استخراج‌شده و طیف‌نگاری تبدیل فوریه کوتاه‌مدت[۱۸] (STFT) با شبکه‌های CNN، موفق به تشخیص ۱۶ نوع ناهنجاری ریتم قلب با صحت ۹۹.۷۹٪ شدند. هرچند این روش عملکرد بالایی داشت، اما وابستگی زیاد به ویژگی‌های دستی استخراج‌شده از محدودیت‌های اصلی آن محسوب می‌شود [۱۹].

همچنین، سابیک و همکاران (۲۰۲۰) کاربرد الگوریتم‌های یادگیری ماشین در شناسایی ناهنجاری در داده‌های ضربان قلب را بررسی کرده و نشان دادند که الگوریتم‌هایی مانند جنگل تصادفی و عامل پرت محلی[۱۹] (LOF) حتی در صورت نبود داده‌های برچسب‌دار، توانایی تشخیص ناهنجاری‌ها را دارند. با این حال، صحت این الگوریتم‌ها در بسیاری از موارد نسبت به روش‌های یادگیری عمیق کمتر است [۲۰].

1. Support Vector Machine
2. Random Forest
3. Artificial Neural Network
4. Naive Bayes
5. K Nearest Neighbor
6. Recurrent Neural Networks
7. Long Short Term Memory
8. Feed Foraward Neural Network
9. Convolutional Neural Network
10. Gated Recurrent Units
11. Variational AutoEncoder
12. Multi-Layer Perceptron
13. Particle Swarm Optimization
14. Genetic Algorithm
15. Firefly Algorithm
16. Whale Optimization Algorithm
17. Sine Cosine Algorithm
18. Short Term Fourier Transform
19. Local Outlier Factor

طبقه‌بندی و تشخیص ناهنجاری در سیگنال ECG با استفاده از مدل‌های یادگیری عمیق، موضوع تحقیقات گسترده‌ای در سال‌های اخیر بوده است. مطالعات متعددی با بهره‌گیری از منبع معتبر MIT-BIH آریتمی، رویکردهای مختلفی را برای حل این چالش مهم بالینی ارائه کرده‌اند.

در سال‌های اخیر، تحلیل خودکار سیگنال‌های قلبی (ECG) با استفاده از روش‌های یادگیری عمیق، توجه گسترده‌ای را در جامعه پژوهشی به خود جلب کرده است. در این بخش، مهم‌ترین دستاوردهای پژوهشی در این حوزه مورد بررسی قرار می‌گیرد.

## ۲-۱ روش‌های برپایه‌ی شبکه‌های بازگشتی

شبکه‌های عصبی بازگشتی به‌ویژه معماری‌های LSTM و Bi-LSTM به دلیل توانایی در مدل‌سازی وابستگی‌های زمانی، کاربرد گسترده‌ای در پردازش سیگنال‌های ECG یافته‌اند. ایرید و همکاران یک مدل Bi-LSTM برای طبقه‌بندی سیگنال‌های ECG برپایه‌ی مجموعه‌داده MIT-BIH ارائه دادند. نتایج پژوهش آنان نشان داد که استفاده از معماری دوطرفه موجب بهبود قابل‌توجهی در صحت تشخیص آریتمی‌ها می‌شود، چرا که این معماری قادر به درک همزمان زمینه زمانی گذشته و آینده سیگنال است. این سیستم به عنوان یک ابزار کمکی برای تشخیص پزشکان طراحی شده و صحت بالایی در شناسایی انواع آریتمی‌ها نشان داده است [۲۱].

در همین راستا، کارتیک و همکاران یک چارچوب ترکیبی شامل LSTM و مدل‌های مارکوف پنهان[1] (HMM) را برای طبقه‌بندی سیگنال سیگنال ECG و تشخیص ناهنجاری ارائه کردند. در این پژوهش، از LSTM برای استخراج ویژگی‌های پیچیده و وابستگی‌های زمانی بلندمدت و از HMM برای مدل‌سازی دنباله‌ای از حالت‌های پنهان مربوط به الگوهای ریتم قلبی استفاده شد. ادغام خروجی این دو مدل منجر به بهبود عملکرد در مقایسه با روش‌های برپایه‌ی تک‌مدل گردید [۲۲].

## ۲-۲ روش‌های مبتنی بر معماری‌های ترکیبی و ترنسفورمر

با پیشرفت‌های اخیر در حوزه یادگیری عمیق، استفاده از معماری‌های ترکیبی و ترنسفورمرها مورد توجه قرار گرفته است.

آیت بورخا و همکاران یک چارچوب ترکیبی متشکل از شبکه پراکندگی موجک بهینه‌شده و CNN برای طبقه‌بندی ضربان قلب از MIT-BIH ارائه دادند. در این روش، از شبکه پراکندگی موجک برای استخراج ویژگی‌های نامتغیر نسبت به جابجایی‌های زمانی و تغییرات مقیاس استفاده شد و سپس این ویژگی‌ها به یک شبکه CNN برای طبقه‌بندی تغذیه گردید. این رویکرد مقاومت بالایی در برابر نویز از خود نشان داد و عملکرد برتری در مقایسه با روش‌های سنتی داشت [۲۳].

الغیث در پژوهش پیشگامانه خود، مدل DeepECG-Net را به عنوان یک چارچوب ترکیبی برپایه‌ی ترنسفورمر برای تشخیص بلادرنگ ناهنجاری‌های ECG معرفی کرد. این مدل از ترکیب ترنسفورمر و CNN بهره می‌برد و با به‌کارگیری مکانیزم توجه، امکان تمرکز خودکار بر بخش‌های بحرانی سیگنال را فراهم می‌سازد. نتایج این پژوهش نشان داد که این معماری نه تنها صحت بسیار بالایی در تشخیص ناهنجاری‌ها دارد، بلکه سرعت استنتاج مناسبی برای کاربردهای بلادرنگ نیز ارائه می‌دهد [۲۴].

جدول ۱: توزیع کلاس‌های AAMI در مجموعه‌داده‌ی MIT-BIH

| کلاس AAMI | تعداد ضربان‌ها | توضیحات |
|---|---|---|
| N | حدود ۹۰,۰۰۰ | Normal and Bundle Branch Block Beats (شامل ضربان‌های نرمال N و انشعاب دسته‌ای His - L, R) |
| S | حدود ۳,۰۰۰ | Supraventricular Ectopic Beats (شامل ضربان‌های فوق بطنی نارس S و سایر ضربان‌های حاصل از کانون فوق بطنی A, J, a - ) |
| V | حدود ۷,۰۰۰ | Ventricular Ectopic Beats (شامل ضربان‌های بطنی نارس V و ضربان‌های بطنی escape – E) |
| F | حدود ۸۰۰ | Fusion Beats (شامل ضربان‌های Fusion – F ) |
| Q | حدود ۱۵,۰۰۰ | Unclassifiable Beats (شامل ضربان پاسیک P و ضربان فیوژن نرمال بطنی f و غیره) |

## ۲-۳ راهکارهای مقابله با عدم تعادل داده

مجموعه‌داده[2] MIT-BIH Arrhythmia یک مرجع استاندارد جهانی برای تحقیق در زمینه تشخیص آریتمی قلبی است که توسط بیمارستان عمومی ماساچوست (MGH) و دانشگاه هاروارد ایجاد شده است. جدول ۱، ۵ کلاس اصلی این مجموعه‌داده را نشان می‌دهد.

یکی از چالش‌های اساسی در این مجموعه‌داده، مشکل عدم تعادل داده‌ها است. ری و همکاران (در پاسخ به این چالش، معماری نوآورانه‌ای به نام GAN-SkipNet ارائه دادند. این معماری ترکیبی از شبکه‌ی مولد هماوردی[3] (GAN) و شبکه SkipNet است که به طور همزمان قادر به تولید نمونه‌های مصنوعی با کیفیت بالا از کلاس‌های اقلیت و استخراج ویژگی‌های مؤثر می‌باشد. نتایج این پژوهش نشان داد که GAN-SkipNet می‌تواند به طور مؤثری مشکل عدم تعادل داده را حل کرده و عملکرد مدل در کلاس‌های نادر را به طور قابل توجهی بهبود بخشد [۲۵].

## ۲-۴ مطالعات مقایسه‌ای و ارزیابی جامع

راجپوت و همکاران در یک مطالعه مقایسه‌ای سیستماتیک، عملکرد چندین معماری مختلف یادگیری عمیق شامل CNN، RNN و مدل‌های ترکیبی CNN-RNN را در تشخیص ناهنجاری‌های ECG مورد ارزیابی قرار دادند. یافته‌های این پژوهش نشان داد که مدل‌های ترکیبی CNN-RNN به دلیل بهره‌گیری همزمان از ویژگی‌های فضایی و زمانی، بهترین عملکرد را در تشخیص آریتمی‌ها دارند. همچنین این مطالعه نشان داد که LSTM در مدل‌سازی وابستگی‌های زمانی مؤثرتر بوده در حالی که CNN در استخراج ویژگی‌های محلی برتری دارد [۲۶].

## ۲-۵ خلأ پژوهشی و سهم پژوهش حاضر

با وجود دستاوردهای قابل توجه اشاره شده در پژوهش‌های پیشین، چندین خلأ پژوهشی قابل شناسایی است. اولاً، اکثر پژوهش‌های موجود

1. Hidden Markov Model

2. Dataset

3. Generative Adversarial Network

بر روی یک معماری خاص تمرکز داشته و مقایسه جامعی بین معماری‌های مختلف ارائه نداده‌اند. ثانیاً، مشکل عدم تعادل داده به طور کامل حل نشده و نیاز به راهکارهای نوآورانه‌تری احساس می‌شود. ثالثاً، قابلیت تفسیرپذیری مدل‌ها و امکان یکپارچه‌سازی آن‌ها در سیستم‌های تشخیصی واقعی کمتر مورد توجه قرار گرفته است.

دلایل اصلی انتخاب LSTM توانایی آن در یادگیری وابستگی‌های بلندمدت، مقاومت در برابر مسائل گرادیان ناپدید شونده، انعطاف‌پذیری با طول ورودی متغیر و تعادل خوب بین پیچیدگی و کارایی برای داده‌های پزشکی است. نتایج آزمایش‌ها روی داده‌های MIT-BIH نشان می‌دهد که LSTM در مقایسه با سایر معماری‌ها صحت بالاتری دارد [۲۷].

در مقایسه با روش‌های مذکور، مدل پیشنهادی این پژوهش برپایه‌ی LSTM طراحی شده است که به‌طور خاص برای داده‌های سری زمانی مانند سیگنال قلبی مناسب است. LSTM قادر به یادگیری وابستگی‌های طولانی‌مدت در داده‌ها و استخراج ویژگی‌های پنهان به‌صورت خودکار است. این ویژگی باعث می‌شود تا صحت مدل در تشخیص ناهنجاری‌ها افزایش‌یافته و نیاز به استخراج دستی ویژگی‌ها کاهش یابد. همچنین، برخلاف RNNهای کلاسیک، مدل LSTM در برابر مشکلاتی مانند ناپایداری گرادیان و گیرافتادن در کمینه‌های محلی مقاوم‌تر است.

در نتیجه، مرور ادبیات موجود نشان می‌دهد که اگرچه پیشرفت‌های قابل توجهی در حوزه تشخیص ناهنجاری‌های ECG حاصل شده است، اما مدل‌های برپایه‌ی LSTM به دلیل توانایی پردازش مؤثر داده‌های سری زمانی و یادگیری وابستگی‌های پیچیده، می‌توانند عملکرد برتری نسبت به سایر رویکردهای مطرح‌شده داشته باشند.

## ۳- مدل پیشنهادی

در این پژوهش، یک مدل یادگیری عمیق برای تشخیص ناهنجاری در سیگنال‌های قلبی طراحی شده است که در ادامه به تشریح معماری و ساختار ورودی و خروجی آن می‌پردازیم.

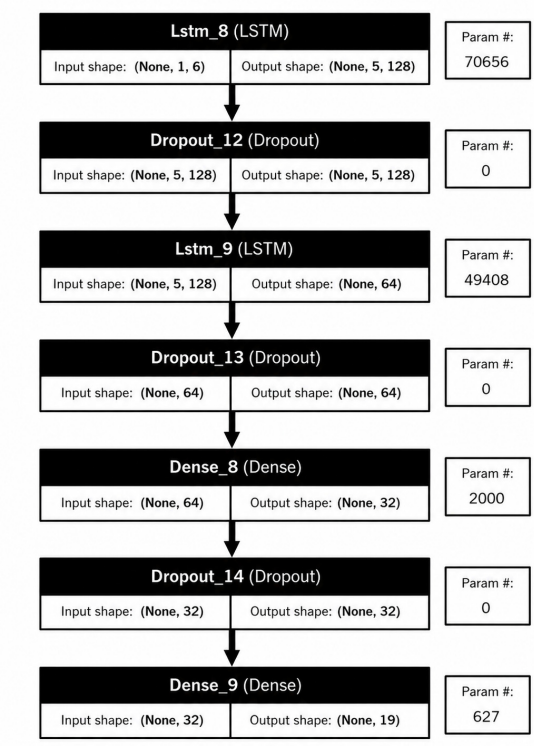
### ۳-۱ شرح لایه‌های مدل

مدل پیشنهادی با رویکردی هدفمند و کارآمد طراحی شده است که ترکیبی از قابلیت‌های مدل‌سازی زمانی LSTM و مکانیزم‌های جلوگیری از بیش‌برازش را در بر می‌گیرد. این معماری به صورت متوالی و لایه‌به‌لایه توسعه یافته است.

مدل پیشنهادی مطابق شکل ۱ از دو لایه LSTM، سه لایه Dropout و لایه تمام‌متصل[1] و یک لایه خروجی تشکیل شده است که به‌صورت متوالی برای استخراج ویژگی‌های زمانی و انجام طبقه‌بندی آریتمی‌های قلبی به کار گرفته شده‌اند. ساختار این مدل در ادامه تشریح می‌شود.

#### ۳-۱-۱ لایه اول LSTM

این لایه شامل ۱۲۸ واحد حافظه (سلول) است و به عنوان نخستین لایه استخراج ویژگی عمل می‌کند. ورودی آن به‌صورت توالی‌هایی با طول ۵ گام زمانی و ابعاد ویژگی مشخص وارد شبکه می‌شود. با استفاده از تنظیم return_sequences=True، خروجی تمامی گام‌های زمانی حفظ شده و به‌صورت یک توالی با ابعاد (۱۲۸×۵) به لایه بعدی منتقل می‌شود. این ویژگی امکان استخراج وابستگی‌های زمانی بلندمدت و حفظ اطلاعات




شکل ۱: .ساختار مدل پیشنهادی.

موجود در کل توالی را فراهم می‌کند. در طراحی معماری شبکه LSTM، لایه اول با ۱۲۸ واحد حافظه پیاده‌سازی شد. این پارامتر ظرفیت یادگیری مدل را برای شناسایی الگوهای پیچیده زمانی در سیگنال‌های ECG تعیین می‌کند. عدد ۱۲۸ در این لایه LSTM نشان‌دهنده تعداد واحدهای حافظه (سلول‌های LSTM) یا به عبارت دیگر، ابعاد فضای حالت پنهان[2] است. این پارامتر یکی از هایپرپارامترهای اصلی شبکه‌های بازگشتی محسوب می‌شود که پیچیدگی و ظرفیت یادگیری مدل را تعیین می‌کند. انتخاب ۱۲۸ واحد بر اساس توازن بین ظرفیت مدل و جلوگیری از بیش‌برازش صورت گرفته است، این مقدار فضای پارامتری کافی برای یادگیری وابستگی‌های بلندمدت در داده‌های الکتروکاردیوگرام فراهم می‌آورد، در حالی که از پیچیدگی محاسباتی بیش از حد که با واحدهای بیشتر (مانند ۲۵۶ یا ۵۱۲) همراه می‌شد، اجتناب می‌ورزد. این اندازه در مطالعات مشابه در حوزه پردازش سیگنال‌های زیستی به عنوان یک انتخاب بهینه گزارش شده است [۲۸].

#### ۳-۱-۲ لایه Dropout اول

به‌منظور کاهش بیش‌برازش، پس از لایه اول LSTM یک لایه Dropout قرار گرفته است. این لایه با غیرفعال کردن تصادفی بخشی از نورون‌ها در طول فرآیند آموزش، وابستگی بیش از حد شبکه به نورون‌های خاص را کاهش داده و موجب افزایش توانایی تعمیم مدل بر روی داده‌های جدید می‌شود. ابعاد خروجی این لایه تغییری نکرده و برابر با (۱۲۸×۵) باقی می‌ماند.

#### ۳-۱-۳ لایه دوم LSTM

لایه دوم شامل ۶۴ واحد LSTM است و خروجی توالی حاصل از لایه قبلی را دریافت می‌کند. در این لایه، از تنظیم ‘return_sequences=False’ استفاده شده است؛ بنابراین تنها خروجی

1. Dense

2. Hidden State Dimension

آخرین گام زمانی به لایه‌های بعدی منتقل می‌شود. این ساختار باعث می‌شود اطلاعات استخراج‌شده از کل توالی زمانی در قالب یک بردار ویژگی ۶۴ بعدی فشرده شده و برای مرحله طبقه‌بندی آماده شود.

### ۳-۱-۴ لایه Dropout دوم

پس از لایه دوم LSTM، یک لایه Dropout دیگر قرار گرفته است تا احتمال بیش‌برازش کاهش یابد. این لایه با حذف تصادفی تعدادی از نورون‌ها در طول آموزش، پایداری و تعمیم‌پذیری مدل را افزایش می‌دهد و خروجی آن همچنان دارای ۶۴ ویژگی است.

### ۳-۱-۵ لایه Dense میانی

این لایه شامل ۳۲ نورون بوده و نقش نگاشت ویژگی‌های استخراج‌شده توسط لایه‌های LSTM به فضای مناسب برای طبقه‌بندی را بر عهده دارد. در این لایه از تابع فعال‌سازی ReLU [۲۹] استفاده شده است که به دلیل سادگی، سرعت محاسباتی بالا، خاصیت تنکی[1] و کاهش مشکل محوشدگی گرادیان، یکی از پرکاربردترین توابع فعال‌سازی در شبکه‌های عصبی عمیق محسوب می‌شود. خروجی این لایه یک بردار ۳۲ بعدی است.

### ۳-۱-۶ لایه Dropout سوم

به منظور افزایش قابلیت تعمیم مدل و جلوگیری از وابستگی بیش از حد نورون‌ها، پس از لایه Dense یک لایه Dropout قرار داده شده است. این لایه بدون تغییر در ابعاد داده، از بیش‌برازش جلوگیری کرده و عملکرد مدل را بر روی داده‌های ناشناخته بهبود می‌بخشد.

### ۳-۱-۷ لایه خروجی Softmax

آخرین لایه مدل، یک لایه Dense با ۱۹ نورون است که تعداد نورون‌های آن برابر با تعداد کلاس‌های مورد نظر می‌باشد. در این لایه از تابع فعال‌سازی Softmax [۳۰] استفاده شده است. این تابع احتمال تعلق نمونه ورودی به هر یک از ۱۹ کلاس را محاسبه کرده و در نهایت کلاسی که دارای بیشترین احتمال است، به‌عنوان خروجی نهایی مدل انتخاب می‌شود. بدین ترتیب، شبکه قادر خواهد بود سیگنال ECG ورودی را در یکی از ۱۹ کلاس آریتمی موجود طبقه‌بندی کند که در جدول ۲ اشاره گردیده است.

همان طور که قبلاً اشاره شد برای پیاده‌سازی روش پیشنهادی، از مجموعه‌داده‌ی MIT-BIH استفاده شده که شامل نمونه‌های متنوعی از سیگنال‌های قلبی با برچسب‌های مشخص است [۳۱] و دو مرحله زیر روی آنها انجام می‌شود.

## ۳-۲ پیش‌پردازش

در این مرحله، پیش‌پردازش سیگنال الکتروکاردیوگرام با هدف حذف نویز و نرمال‌سازی دامنه صورت گرفت. ابتدا با استفاده از یک فیلتر باترورث[2] باندگذر از درجه سه با محدوده فرکانسی ۵/۰ تا ۴۵ هرتز، مؤلفه‌های فرکانسی نامرتبط شامل نویز پایه‌گرد[3] و نویز فرکانس بالا حذف گردید. به‌کارگیری تکنیک فیلترینگ دوطرفه از ایجاد اعوجاج فاز در سیگنال جلوگیری نمود. در گام بعدی، سیگنال فیلترشده با استفاده از نرمال‌سازی استاندارد آماری، به صورت کاسته شده از میانگین و

1. Sparsity
2. Butterworth
3. Low-Frequency Baseline Wander

جدول ۲: کلاس‌های مهم MIT-BIH

| ردیف | نماد کلاس | توضیح (انگلیسی) | توضیح (فارسی) |
|---|---|---|---|
| ۱ | N | Normal Sinus Rhythm | ریتم سینوسی طبیعی |
| ۲ | L | Left Bundle Branch Block Beat | ضربان با انشعاب دسته چپ |
| ۳ | R | Right Bundle Branch Block Beat | ضربان با انشعاب دسته راست |
| ۴ | B | Venticular Escape Beat | ضربان گریز بطنی |
| ۵ | A | Aberrated Atrial Premature Contraction | انقباض زودرس دهلیزی با اَبِرَکشن |
| ۶ | a | Atrial Premature Contraction | انقباض زودرس دهلیزی |
| ۷ | S | Supraventricular Premature Contraction | انقباض زودرس فوق بطنی |
| ۸ | V | Premature Ventricular Contraction | انقباض زودرس بطنی(PVC) |
| ۹ | F | Fusion of Ventricular and Normal Beat | ضربان فیوژن (بطنی و طبیعی) |
| ۱۰ | E | Ventricular Escape Beat | ضربان گریز بطنی |
| ۱۱ | j | Nodal (Junctional) Escape Beat | ضربان گریز اتصالی |
| ۱۲ | e | Atrial Escape Beat | ضربان گریز دهلیزی |
| ۱۳ | / | Fusion of Paced and Normal Beat | ضربان فیوژن (پیسمیکر و طبیعی) |
| ۱۴ | F | Atrial Premature Contraction | انقباض زودرس دهلیزی |
| ۱۵ | Q | Unclassifiable Beat | ضربان غیرقابل طبقه‌بندی |
| ۱۶ | ? | Beat not classified during learning | ضربان طبقه‌بندی نشده در فرآیند یادگیری |
| ۱۷ | ~ | Signal Quality Change | تغییر در کیفیت سیگنال |
| ۱۸ | \| | Isolated QRS-like Artifact | آرتیفکت |
| ۱۹ | J | Nodal (Junctinal) Premature Beat | ضربان زودرس گره‌ای (جانکشنال) |

تقسیم‌شده بر انحراف معیار درآمد تا یک توزیع آماری یکنواخت و بدون مقیاس داشته باشد. این رویکرد پیش‌پردازش، کیفیت سیگنال را برای استخراج ویژگی و ورود به مدل یادگیری عمیق به‌طور قابل توجهی بهبود بخشید.

برای تشخیص نواقص اصلی کمپلکس QRS و امواج مرتبط در هر ضربان، از یک الگوریتم برپایه‌ی پیک‌یابی استفاده شد. این روش با شناسایی نقطه اوج موج R (نشان‌دهنده دپولاریزاسیون بطن‌ها) آغاز می‌گردد که با جستجوی پیک‌های غالب در سیگنال با اعتبارسنجی برپایه‌ی ارتفاع و فاصله زمانی شناسایی می‌شود. مکان‌های موج‌های Q و S به ترتیب به عنوان حداقل‌های محلی در بازه‌های زمانی قبل و بعد از نقطه R تشخیص داده می‌شوند. متعاقباً، موج P (نمایانگر دپولاریزاسیون دهلیزها) با یافتن حداکثر محلی در بازه قبل از نقطه Q، و موج T (نمایانگر رپولاریزاسیون بطن‌ها) با یافتن حداکثر محلی در بازه پس از نقطه S، استخراج می‌گردند.

برای استخراج ویژگی‌های زمانی و مورفولوژیکی سیگنالECG، ابتدا نقاط شاخص موج‌های P، Q، R، S و T هر ضربان شناسایی شدند. در این روش، تشخیص نقطه R به‌عنوان مرجع اصلی انجام گرفت. بدین

جدول ۳: ویژگی‌های مورفولوژیکی و زمانی سیگنال.

| ویژگی | نوع | توضیح |
|---|---|---|
| P_amp | مورفولوژیکی | نمایانگر دپولاریزاسیون دهلیزها (دامنه موج P) |
| Q_amp | مورفولوژیکی | دامنه بخش Q از کمپلکس QRS |
| R_amp | مورفولوژیکی | (QRS) بخش اصلی و غالب کمپلکس QRS (دامنه موج R) |
| S_amp | مورفولوژیکی | بخش نزولی کمپلکس QRS (دامنه موج S) |
| T_amp | مورفولوژیکی | نمایانگر رپولاریزاسیون بطن‌ها (دامنه موج T) |
| PQ_d | زمانی | فاصله زمانی از شروع موج P تا شروع کمپلکس QRS |
| QR_d | زمانی | فاصله زمانی از موج Q تا موج R |
| RS_d | زمانی | فاصله زمانی از موج R تا موج S |
| ST_d | زمانی | فاصله زمانی از پایان کمپلکس QRS تا موج T |

منظور، با استفاده از تابع تشخیص قله و با در نظر گرفتن آستانه‌ای معادل ۶۰ درصد بیشینه دامنه ضربان و حداقل فاصله زمانی برابر با یک‌سوم فرکانس نمونه‌برداری، قله‌های برجسته شناسایی شدند و نخستین قله معتبر به‌عنوان نقطه R انتخاب گردید. برای جلوگیری از خطاهای ناشی از لبه‌های سیگنال، تنها ضربان‌هایی که نقطه R آن‌ها فاصله مناسبی از ابتدا و انتهای پنجره داشت، مورد پردازش قرار گرفتند.

پس از تعیین نقطه R، موج Q با جستجو در بازه‌ای شامل ۴۰ نمونه قبل از R استخراج شد. کمینه مقدار دامنه در این بازه به عنوان نقطه Q در نظر گرفته شد که متناظر با فرورفتگی ابتدایی کمپلکس QRS است. به‌طور مشابه، برای تعیین نقطه **S**، بازه‌ای شامل ۴۰ نمونه پس از R بررسی شد و کمینه دامنه موجود در این ناحیه به عنوان نقطه S انتخاب گردید. سپس جهت تعیین موج P، بازه‌ای تا ۸۰ نمونه قبل از نقطه Q مورد جستجو قرار گرفت و بیشینه دامنه موجود در این ناحیه به‌عنوان نقطه P در نظر گرفته شد که نمایانگر دپلاریزاسیون دهلیزی است. در ادامه، موج T با بررسی بازه‌ای از ۱۰ نمونه پس از S تا حداکثر ۱۲۰ نمونه بعد از آن استخراج شد و بیشینه دامنه در این محدوده به‌عنوان نقطه T انتخاب گردید که متناظر با مرحله رپلاریزاسیون بطنی است. در مواردی که به دلیل نویز یا شرایط غیرعادی، امکان یافتن نقاط مذکور وجود نداشت، مقادیر تقریبی مبتنی بر فاصله زمانی نسبی نسبت به نقطه R در نظر گرفته شد. این روش مبتنی بر جستجوی محلی پیرامون کمپلکس QRS، ضمن سادگی محاسباتی، امکان استخراج قابل اعتماد نقاط شاخص ECG و ویژگی‌های زمانی مرتبط را فراهم می‌کند.

## ۳-۲ استخراج ویژگی

در این پژوهش، داده‌های سری‌زمانی برای آموزش مدل LSTM با در نظرگیری ماهیت توالی‌ای سیگنال الکتروکاردیوگرام ایجاد گردید. برای هر بیمار، پنجره‌های متوالی با طول ثابت (time_steps) مطابق جدول ۳ از ۹ ویژگی مورفولوژیکی و زمانی استخراج‌شده شامل دامنه امواج P، Q، R، S، T و فواصل PQ، QR، RS، ST تشکیل شد.

انتخاب این ۹ ویژگی خاص شامل پنج ویژگی دامنه‌ای (P_amp, Q_amp, R_amp, S_amp, T_amp) و چهار ویژگی زمانی (PQ_d, QR_d, RS_d, ST_d) به‌دلیل توانایی آن‌ها در بازنمایی جنبه‌های حیاتی و مکمل الکتروفیزیولوژی قلب است. این ویژگی‌ها به‌طور جمعی سه عملکرد اصلی را ارائه می‌کنند:

– اول، آن‌ها ساختار اصلی چرخه قلبی را از دپولاریزاسیون دهلیزها (موج P) تا رپولاریزاسیون بطن‌ها (موج T) پوشش می‌دهند.
– دوم، این مجموعه امکان تشخیص انحرافات ظریف اما بحرانی را فراهم می‌سازد؛ برای مثال، تغییرات در دامنه موج R یا فاصله ST می‌تواند به ترتیب نشان‌دهنده هایپرترفی بطن چپ یا ایسکمی میوکارد باشد.
– سوم، این ویژگی‌ها از نظر محاسباتی کارآمد بوده و تمرکزی هدفمند بر روی پارامترهای کلیدی داشته که مستقیماً با پاتوفیزیولوژی آریتمی‌های شایع مرتبط هستند و در عین حال از پیچیدگی بیش از حد مدل که می‌تواند منجر به بیش برازش شود، جلوگیری می‌کنند.

در نتیجه، این ترکیب متعادل، پایه‌ای قوی و به لحاظ بالینی مرتبط برای آموزش مدل‌های تشخیص ناهنجاری فراهم می‌کند.

در ادامه، تأثیر حذف هر یک از این ۹ ویژگی بر عملکرد سیستم‌های طبقه‌بندی، با استناد به شواهد فیزیولوژیکی و یافته‌های پژوهش‌های پیشین، مورد بررسی علمی قرار می‌گیرد.

### ۳-۳-۱ تأثیر حذف دامنه امواج (P، Q، R، S، T)

دامنه امواج، شاخص مستقیمی از حجم و فعالیت الکتریکی حفرات قلبی هستند. حذف هر یک از این ویژگی‌ها، اطلاعات حیاتی مربوط به بخش خاصی از چرخه قلبی را از بین می‌برد.

- حذف دامنه موج P: موج P نشان‌دهنده دپلاریزاسیون دهلیزی است. دامنه و مورفولوژی آن برای تشخیص بزرگ‌شدگی دهلیزها حیاتی است. به‌عنوان مثال، افزایش دامنه موج P (P-pulmonale) نشان‌دهنده بزرگی دهلیز راست و افزایش پهنای آن (P-mitrale) نشان‌دهنده بزرگی دهلیز چپ است. حذف این ویژگی، توانایی تشخیص اختلالات دهلیزی نظیر فیبریلاسیون دهلیزی و بلوک‌های بین‌دهلیزی را به شدت کاهش می‌دهد.
- حذف دامنه امواج Q، R، S: کمپلکس QRS بازتاب‌دهنده دپلاریزاسیون بطنی است. دامنه این امواج، به‌ویژه موج R، در تشخیص هیپرتروفی بطنی (افزایش جرم عضلانی بطن‌ها) کاربرد دارد. به‌عنوان مثال، افزایش ولتاژ QRS نشانه هیپرتروفی بطن چپ است، در حالی که کاهش ولتاژ می‌تواند نشانه افیوژن پریکارد باشد. حذف دامنه این امواج، سیستم را در تشخیص افتراقی هیپرتروفی، انفارکتوس میوکارد (که با کاهش دامنه QRS همراه است) و بلوک‌های شاخه‌ای ناتوان می‌سازد.
- حذف دامنه موج T: موج T نشان‌دهنده رپلاریزاسیون بطنی است. تغییرات دامنه و جهت موج T (مانند معکوس‌شدگی یا افزایش شدید) یکی از نشانه‌های کلیدی ایسکمی میوکارد و تغییرات ST-T است. حذف این ویژگی، حساسیت سیستم را در تشخیص بیماری‌های عروق کرونر و اختلالات الکترولیتی به شدت کاهش می‌دهد.

### ۳-۳-۲ تأثیر حذف فواصل زمانی (PQ، QR، RS، ST)

فواصل زمانی، سرعت هدایت الکتریکی را در بخش‌های مختلف سیستم هدایت قلبی اندازه‌گیری می‌کنند. حذف آن‌ها به‌معنای نادیده‌گرفتن اختلالات هدایت[1] است.

- حذف فاصله PQ (PR): این فاصله مدت زمان هدایت تحریک از دهلیزها به بطن‌ها از طریق گره دهلیزی-بطنی (AV) را نشان

1. Conduction Disturbance

می‌دهد. طولانی‌شدن آن نشان‌دهنده بلوک درجه یک AV است و کوتاه‌شدن آن می‌تواند نشانه سندرم‌های پیش‌تحریک (مانند WPW) یا تاکی‌کاردی‌های ناشی از مسیر فرعی باشد. مطالعات نشان داده‌اند که فاصله‌ی PR یکی از ویژگی‌های مؤثر در تشخیص اختلالات هدایت است. حذف این ویژگی، سیستم را در تشخیص بلوک‌های قلبی ناتوان می‌سازد.

- حذف فاصله QR و RS (عرض QRS): این فواصل، بخشی از مدت زمان دپلاریزاسیون بطنی را نشان می‌دهند. پهن‌شدن کمپلکس QRS (افزایش فاصله QR/RS) از معیارهای اصلی تشخیص بلوک‌های شاخه‌ای (مانند LBBB و RBBB) و تاکی‌کاردی‌های بطنی است. پژوهش‌ها نشان می‌دهد که صحت تشخیص بلوک‌های شاخه‌ای وابستگی شدیدی به ویژگی‌های زمانی کمپلکس QRS دارد و حذف این ویژگی‌ها، طبقه‌بندی این اختلالات را با شکست مواجه می‌سازد .
- حذف فاصله ST: قطعه ST نشان‌دهنده فاصله زمانی بین پایان دپلاریزاسیون و آغاز رپلاریزاسیون بطنی است. انحراف قطعه ST از خط ایزوالکتریک (صعود یا نزول) یکی از حساسترین و اختصاصی‌ترین نشانه‌های ایسکمی میوکارد، انفارکتوس حاد و پریکاردیت است. بررسی‌ها نشان داده است که تمرکز صرف بر ویژگی‌های قطعه ST برای طبقه‌بندی برخی آریتمی‌ها کافی است. بنابراین، حذف این ویژگی به‌معنای از دست‌دادن مهم‌ترین شاخص تشخیصی بیماری‌های ایسکمیک قلبی است.

در این روش، هر نمونه آموزشی شامل یک توالی متشکل از ۵ مشاهده متوالی (گام‌های زمانی) است که اطلاعات مربوط به تکامل ویژگی‌های قلبی را در طول زمان حفظ می‌کند و برچسب نهایی هر توالی، بر اساس کلاس آخرین ضربان در آن پنجره اختصاص می‌یابد. این رویکرد، مدل LSTM را قادر می‌سازد تا نه تنها بر اساس ویژگی‌های یک ضربان منفرد، بلکه با در نظرگیری الگوی تاریخی چندین ضربان متوالی، تشخیص ناهنجاری را با صحت بالاتری انجام دهد.

## ۳-٤ آموزش مدل

برای آموزش این معماری، از «بهینه‌ساز Adam» [۳۲] با نرخ یادگیری پیش‌فرض و تابع زیان آنتروپی متقاطع طبقه‌بندی[1] استفاده شده است. از معیار صحت برای ارزیابی عملکرد مدل استفاده شده است.

در کدهای منتشرشده در گیت‌هاب[2] برای آموزش مدل، از تکنیک توقف زودهنگام[3] با نظارت بر مقدار تابع زیان در داده‌های اعتبارسنجی[4] استفاده شد. اگر عملکرد مدل در طی ۵ دوره متوالی بهبود نیابد، فرآیند آموزش متوقف و بهترین وزن‌های مدل بازیابی می‌شود. مدل نهایی با استفاده از ۸۰٪ داده‌های آموزشی و ۲۰٪ اعتبارسنجی، به مدت حداکثر ۲۵ دوره و با اندازه دسته ۳۲ آموزش داده شد. این تنظیمات با هدف جلوگیری از بیش‌برازش و بهبود تعمیم‌پذیری مدل انجام گرفته‌اند.

## ٤- ارزیابی

برای ارزیابی عملکرد مدل پیشنهادی، داده‌ها به دو بخش آموزش و آزمون تقسیم شدند. مدل بر روی داده‌های آموزشی آموزش دیده و سپس عملکرد آن بر روی داده‌های آزمون بررسی شد. معیارهای ارزیابی همان طور که در (۱) تا (۴) نشان‌داده شده است شامل صحت[5]، بازخوانی[6]، دقت[7] و F1-Score بوده‌اند [۳۳] تا [۳۷].

جدول ۴: نتایج الگوریتم‌های یادگیری ماشین با ۹ ویژگی.

| مدل | صحت (٪) | دقت (٪) | بازخوانی (٪) | F1-Score (٪) |
|---|---|---|---|---|
| AdaBoost | ۷۱٫۴۰ | ۵۴٫۱۶ | ۷۱٫۴۰ | ۶۱٫۴۴ |
| Naive Bayes | ۷۳٫۷۷ | ۷۴٫۲۳ | ۷۳٫۷۷ | ۷۳٫۲۵ |
| Logistic Regression | ۷۶٫۹۰ | ۷۱٫۰۴ | ۷۶٫۹۰ | ۷۱٫۷۵ |
| Decision Tree | ۹۴٫۱۵ | ۹۴٫۲۵ | ۹۴٫۱۵ | ۹۴٫۱۹ |
| SVM | ۹۶٫۲۶ | ۹۵٫۶۴ | ۹۶٫۲۶ | ۹۵٫۸۱ |
| Random Forest | ۹۶٫۶۲ | ۹۶٫۳۱ | ۹۶٫۶۲ | ۹۶٫۳۶ |
| MLP | ۹۶٫۵۴ | ۹۶٫۲۰ | ۹۶٫۵۴ | ۹۶٫۳۵ |
| KNN | ۹۶٫۶۸ | ۹۶٫۰۹ | ۹۶٫۶۸ | ۹۶٫۳۵ |
| LSTM پیشنهادی | ۹۷٫۳۳ | ۹۸٫۰۰ | ۹۸٫۰۰ | ۹۷٫۹۷ |

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (۱)$$

$$Recall = \frac{TP}{TP + FN} \quad (۲)$$

$$Precision = \frac{TP}{TP + FP} \quad (۳)$$

$$F_1 - Score = \frac{2 \times Precision \times Recall}{Precision \times Recall} \quad (۴)$$

که در آنها به ترتیب $TP$ و $TN$ مثبت و منفی صحیح و $FP$ و $FN$ به ترتیب مثبت و منفی کاذب هستند.

نتایج نشان می‌دهد که مدل توانسته است به صحت بالایی در تشخیص ناهنجاری‌ها دست یابد. همچنین تحلیل ماتریس درهم‌ریختگی[8] نشان می‌دهد که میزان خطای مدل در کلاس‌های مختلف پایین است. این ماتریس به صورت یک جدول مربعی نمایش داده می‌شود که در آن، نتایج پیش‌بینی مدل در مقابل نتایج واقعی قرار می‌گیرند.

بر اساس نتایج جدول ۴ مشاهده می‌شود که ویژگی‌های مورفولوژیکی و زمانی استخراج‌شده از نقاط P ، Q، R، S و T اطلاعات تشخیصی مناسبی را در اختیار تمامی الگوریتم‌ها قرار می‌دهند. الگوریتم‌های کلاسیک مانند AdaBoost، Naive Bayes، Logistic Regression و Decision Treeعملکرد قابل قبولی داشته‌اند، اما به دلیل ناتوانی در مدل‌سازی وابستگی‌های زمانی، صحت آن‌ها پایین‌تر است. روش‌های مبتنی بر یادگیری ماشین مانند KNN، Random Forest، MLP و SVM توانسته‌اند با استفاده از همین ۹ ویژگی به صحتی بیش از ۹۶٪ دست یابند.

برای ارزیابی عملکرد مدل پیشنهادی، از معماری مبدل بینایی[9]

1. Categorical Cross-Entropy Loss Function
2. https://github.com/BehradMazaheri/ECG
3. Early Stopping
4. val_loss
5. Accuracy
6. Recall (Sensitivity)
7. Precision
8. Confusion Matrix
9. Transformer Vision

جدول ۵: نتایج مقایسه با روش‌های پایه.

| مرجع | روش | مزایا | محدودیت |
|---|---|---|---|
| [۳۸] | ترنسفورمر چندشاخه سبک | استخراج ویژگی‌های پیچیده و وابستگی‌های بلندمدت | پیچیدگی محاسباتی بالا |
| [۳۹] | CardioSSL برپایه‌ی یادگیری خودنظارتی | کاهش نیاز به داده برچسب‌خورده | نیاز به پیش‌آموزش و منابع بیشتر |
| روش پیشنهادی | استخراج ویژگی P-QRS-T + LSTM | ساختار سبک و عملکرد رقابتی | افت عملکرد در کلاس‌های کم‌نمونه |

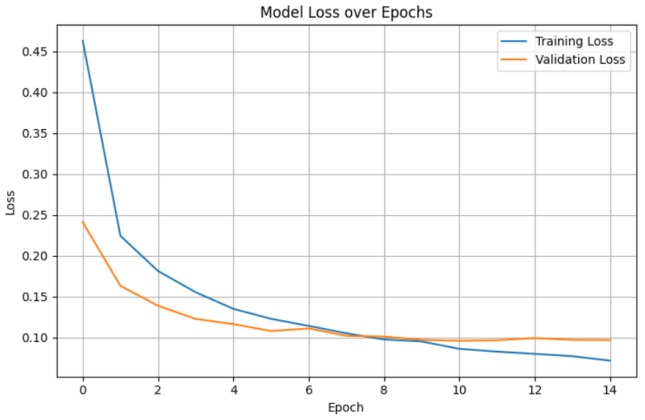
(ViT) به عنوان یک روش پایه مدرن و قدرتمند استفاده شد. در این پیاده‌سازی، سیگنال ECG یک بعدی ابتدا به یک بازنمایی[1] دوبعدی (مانند (مانند تصویر اسپکتروگرام یا ماتریس از پیش پردازش شده) تبدیل می‌شود. این تصویر سپس به دنباله‌ای از پچ‌ها تقسیم شده و پس از افزودن نشانش[2] موقعیتی، به عنوان ورودی به لایه‌های ترنسفورمر ارسال ارسال می‌گردد. مکانیزم توجه چند سر[3] (MHSA) در ViT این قابلیت را را فراهم می‌آورد تا مدل به صورت پویا به بخش‌های مرتبط‌تر سیگنال (مانند کمپلکس QRS یا بخش ST) برای تشخیص ناهنجاری تمرکز کند. انتخاب ViT به عنوان روش پایه به دلیل عملکرد برتر آن در وظایف بینایی کامپیوتر و همچنین قابلیت ذاتی آن در مدل‌سازی وابستگی‌های بلند-برد در داده‌ها است. این امر امکان مقایسه‌ای معنادار و رقابتی با روش‌های متعارف مانند LSTM و CNN را فراهم می‌سازد.

مقاله‌ی [۳۸] از یک «شبکه ترنسفورمر چندشاخه[4]» استفاده می‌کند. ترنسفورمرها که ابتدا برای پردازش زبان طبیعی ساخته شدند، به دلیل قابلیت فوق‌العاده در مدل‌سازی وابستگی‌های بلندمدت[5] در دنباله‌های داده، برای تحلیل سیگنال ECG ایده‌آل هستند.

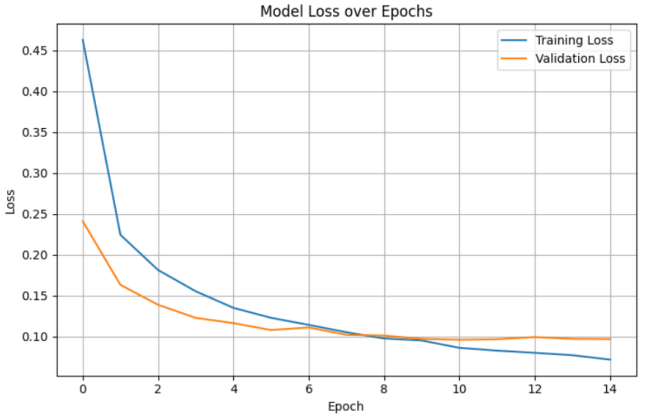
چالش بزرگ در حوزه پزشکی، کمبود داده‌های برچسب‌دار است. مقاله‌ی [۳۹] از «یادگیری خودنظارتی» استفاده می‌کند. در این پارادایم، مدل ابتدا روی حجم بسیار بزرگی از داده‌ی ECG بدون برچسب[6] پیش‌آموزش می‌بیند. در این فاز، مدل با حل کردن یک «مسئله جایگزین» (مانند بازسازی قطعات پنهان‌شده سیگنال یا تشخیص تغییرات زمانی) یاد می‌گیرد که یک بازنمایی غنی و عمومی از سیگنال قلب استخراج کند. سپس، این مدل از پیش آموزش دیده، با مقدار بسیار کمتری داده برچسب‌دار، برای کار اصلی تشخیص ناهنجاری، تنظیم دقیق[7] می‌شود. این روش به طور چشمگیری عملکرد را در شرایط کم‌دیتا کم‌دیتا بهبود می‌بخشد.

جدول ۵ مقایسه کیفی روش پیشنهادی با دو روش جدید را نشان می‌دهد. اگرچه روش‌های مبتنی بر ترنسفورمر و یادگیری خودنظارتی از قابلیت استخراج ویژگی بالایی برخوردارند، روش پیشنهادی با استفاده از ویژگی‌های مورفولوژیکی و شبکه LSTM، ساختاری ساده‌تر و عملکردی

1. Representaion
2. Embedding
3. Multi-Head Self-Attention
4. Multi-Branch Transformer
5. Long-Range Dependencies
6. Unlabeled
7. Fine-Tune




شکل ۲: نمودار صحت.




شکل ۳: نمودار تابع زیان.

رقابتی ارائه می‌دهد.

شکل ۲ نشان می‌دهد که با افزایش تعداد دوره‌های آموزشی، صحت مدل در داده‌های آموزش و اعتبارسنجی به‌صورت پیوسته افزایش‌یافته و در نهایت به بیش از ۹۴٪ رسیده است. این روند حاکی از یادگیری مؤثر مدل بدون نشانه‌ای از بیش‌برازش زودهنگام است. توضیحات حائز اهمیت شکل ۲ بدین شرح است:

۱. **تکنیک‌های قوی منظم‌سازی**، مانند نرخ بالای خروج (۰/۳ یا بالاتر) در طول آموزش، و منظم‌سازی L۲ که وزن‌های بزرگ را جریمه می‌کند، که می‌تواند صحت آموزش را کاهش دهد.

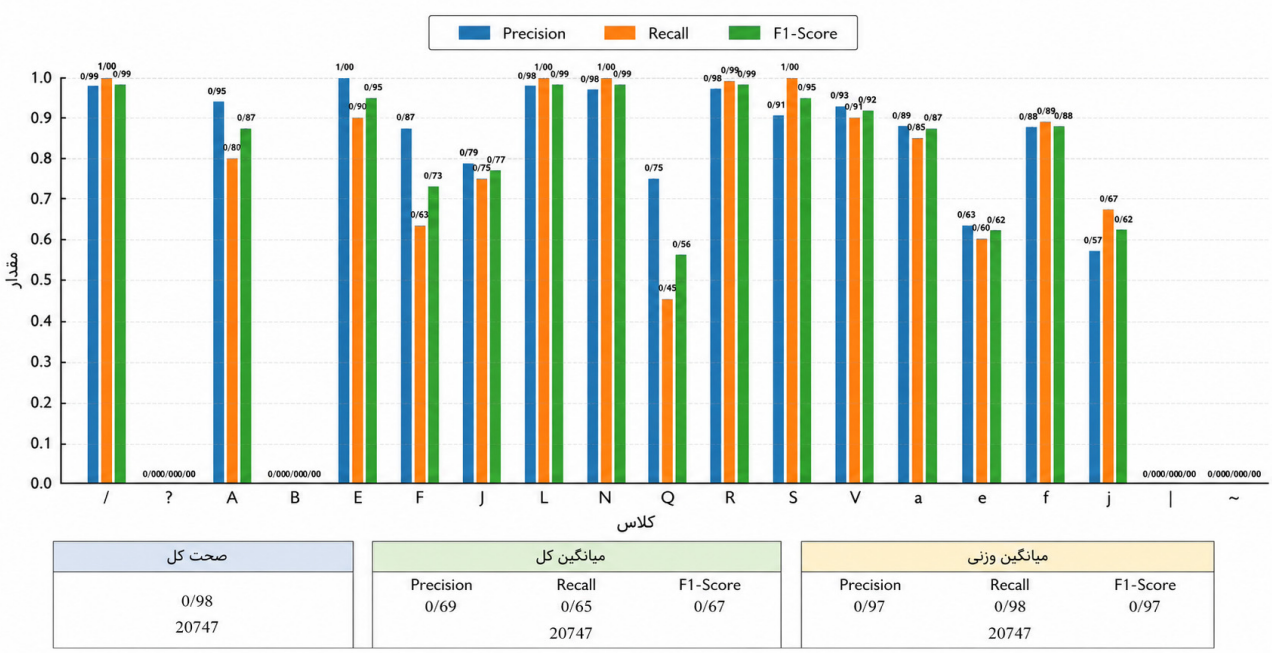
۲. **تفاوت در پیش‌پردازش داده‌ها**، که در آن غنی‌سازی داده فقط برای داده‌های آموزشی اعمال می‌شود و داده‌های اعتبارسنجی بدون تغییر باقی می‌مانند.

۳. **استفاده از دیگر روش‌های محاسبه**، استفاده از میانگین متحرک برای صحت آموزش و محاسبه آنی برای صحت اعتبارسنجی.

۴. **توزیع متغیر داده‌ها**، که در آن داده‌های اعتبارسنجی ممکن است در مقایسه با داده‌های آموزشی چالش‌برانگیزتر ساده‌تر باشند.

۵. **تأثیر توقف زودهنگام بر اساس از دست دادن اعتبارسنجی**، که ممکن است نتایج را قبل از آموزش کامل گزارش دهد.

مشاهدات نشان می‌دهد که صحت آموزش بهبود می‌یابد در حالی که صحت اعتبارسنجی پایدار می‌ماند. تجزیه و تحلیل نشان می‌دهد که اگر اختلاف صحت کمتر از ۲٪-۳٪ باشد و هر دو منحنی روند صعودی داشته باشند، این الگو طبیعی است. بهبودهای احتمالی شامل تنظیم پارامترهای فوق، بررسی توزیع داده‌ها و افزایش حجم داده‌های آموزشی است. به طور کلی، الگوی مشاهده شده نشان می‌دهد که بیش‌برازش وجود ندارد و فرآیند آموزش پایدار است و توانایی تعمیم خوبی را نشان می‌دهد.




| صحت کل | میانگین کل | | | میانگین وزنی | | |
|---|---|---|---|---|---|---|
| | Precision | Recall | F1-Score | Precision | Recall | F1-Score |
| 0/98 | 0/69 | 0/65 | 0/67 | 0/97 | 0/98 | 0/97 |
| 20747 | | 20747 | | | 20747 | |

شکل ۴: گزارش دسته‌بندی (با ۱۹ کلاس).

شکل ۳ نیز کاهش یکنواخت تابع زیان[1] را در هر دو مجموعه آموزش و اعتبارسنجی نشان می‌دهد که بیانگر پایداری فرآیند آموزش و بهبود تدریجی عملکرد مدل در کاهش خطای پیش‌بینی است.

این پدیده مشاهده شده که در آن ضرر آموزش از ضرر اعتبارسنجی بیشتر است، کاملاً طبیعی و ناشی از عوامل زیر می‌باشد:

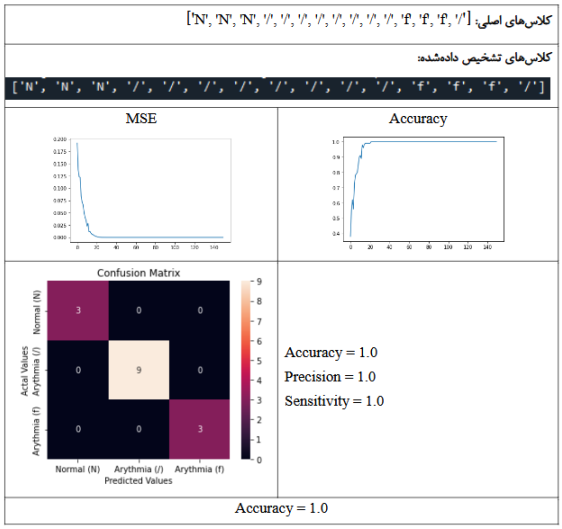
۱. **اعمال تکنیک‌های منظم‌سازی قوی** شامل Dropout و L2 regularization که تنها در فاز آموزش فعال هستند و عمداً مدل را با داده‌های «سخت‌تر» مواجه می‌کنند.

۲. **استفاده از داده افزایی**[2] که تنها بر روی داده‌های آموزشی اعمال شده و پیچیدگی مصنوعی به فرآیند آموزش اضافه می‌کند.

۳. **همگرایی پایدار هر دو منحنی** و کاهش یکنواخت آنها نشان‌دهنده سلامت فرآیند آموزش و عدم بیش‌برازش می‌باشد.

در این مطالعه، ما با چندین چالش عمده در طبقه‌بندی سیگنال‌های الکتروکاردیوگرام مواجه بودیم. نخست، مسئله «عدم تعادل شدید داده‌ها» بود. در مجموعه‌داده MIT-BIH توزیع کلاس‌ها به‌شدت نامتوازن است. بخش عمده ضربان‌ها به کلاس طبیعی (N) تعلق دارند و حدود ۸۲٫۸ درصد از کل نمونه‌ها را تشکیل می‌دهند. در مقابل، کلاس‌های آریتمی تعداد بسیار کمتری دارند؛ به‌طوری‌که کلاس S (ضربان‌های نابجای فوق‌بطنی) حدود ۲٫۵ درصد، کلاس V (ضربان‌های نابجای بطنی) حدود ۶٫۴ درصد، کلاس F (ضربان‌های فیوژن) حدود ۰٫۷ درصد و کلاس Q (ضربان‌های غیرقابل طبقه‌بندی) حدود ۷٫۶ درصد از کل داده‌ها را شامل می‌شوند. این عدم توازن شدید بین کلاس‌ها یکی از مهم‌ترین چالش‌های طراحی و ارزیابی مدل‌های یادگیری ماشین برای تشخیص آریتمی محسوب می‌شود، زیرا مدل‌ها تمایل دارند به سمت کلاس غالب N سوگیری پیدا کنند و عملکرد ضعیف‌تری روی کلاس‌های کم‌نمونه داشته باشند. دوم، «شباهت مورفولوژیکی قابل توجه بین کلاس‌های خاصی» مانند انشعاب دسته چپ (L) و انشعاب دسته راست (R) وجود داشت که تمایز این دو را برای مدل با دشواری روبرو می‌ساخت. سوم، «تغییرات

1. Loss
2. Data Augmentation




شکل ۵: یک نمونه ماتریس درهم‌ریختگی.

ذاتی بین بیماران مختلف» در ویژگی‌های سیگنال ECG مشاهده شد که ناشی از تفاوت‌های فیزیولوژیکی، شرایط آناتومیکی و موقعیت قرارگیری الکترودها بود و این امر تعمیم‌پذیری مدل را تحت تأثیر قرار می‌داد. این چالش‌ها در کنار یکدیگر، توسعه یک مدل طبقه‌بندی پایدار و دقیق را مستلزم به‌کارگیری راهبردهای ویژه در پیش‌پردازش داده‌ها و طراحی معماری مدل می‌ساخت.

مطابق شکل ۴ برای ارزیابی عملکرد مدل پیشنهادی در تشخیص ناهنجاری‌های ECG، از معیارهای دقت، بازخوانی، F1-Score و صحت کلی استفاده شد. کلاس N (ریتم نرمال) با ۱۴۵۵۶ نمونه بهترین عملکرد را داشته و مقادیر دقت، بازیابی و F۱-Score آن به‌ترتیب ۹۸٪، ۱۰۰٪ و ۹۹٪ بوده است. همچنین کلاس V با F1-Score برابر با ۹۲٪ و کلاس F با F۱-Score برابر با ۷۵٪، عملکرد مناسبی را نشان می‌دهند. در مقابل، برخی کلاس‌های دارای تعداد نمونه بسیار کم مانند B، E، J و Q و ~ عملکرد ضعیف‌تری داشته و در بعضی موارد مقدار F۱ -Score به صفر رسیده است که ناشی از عدم توازن شدید داده‌ها است. میانگین وزنی که سهم هر کلاس را متناسب با تعداد نمونه‌های آن در نظر می‌گیرد،

برای دقت، بازیابی و F1-Score به‌ترتیب برابر با ۹۷٪، ۹۸٪ و ۹۷٪ به‌دست آمده و بیانگر عملکرد مطلوب مدل در کل مجموعه داده است. در مقابل، میانگین کل که تمامی کلاس‌ها را بدون توجه به تعداد نمونه‌ها با وزن یکسان در نظر می‌گیرد، به‌ترتیب برابر با ۶۹٪، ۶۵٪ و ۶۷٪ بوده که نشان می‌دهد عملکرد مدل در کلاس‌های کم‌نمونه همچنان نیازمند بهبود است.

به منظور کاهش اثر این عدم تعادل داده، در مرحله پیش‌پردازش از راهبرد ترکیبی متعادل‌سازی داده‌ها شامل بیش‌نمونه‌گیری[1] و کم‌نمونه‌گیری[2] استفاده شد. در این روش، ابتدا تعداد نمونه‌های کلاس‌های اقلیت با ایجاد نمونه‌های جدید و متنوع افزایش یافت تا مدل بتواند ویژگی‌های این کلاس‌ها را به‌طور مؤثرتری فراگیرد. سپس، تعداد نمونه‌های کلاس‌های غالب به‌صورت کنترل‌شده کاهش داده شد تا از غلبه آن‌ها بر فرآیند یادگیری جلوگیری شود. استفاده همزمان از این دو راهکار، ضمن حفظ اطلاعات ارزشمند کلاس‌های پرتکرار، موجب ایجاد توزیع متعادل‌تر در داده‌های آموزشی، کاهش سوگیری مدل نسبت به کلاس‌های غالب و افزایش توانایی مدل در شناسایی آریتمی‌های کم‌تکرار شد. این متعادل‌سازی نقش مؤثری در بهبود قابلیت تعمیم مدل و افزایش دقت تشخیص کلاس‌های اقلیت ایفا کرده است، هرچند نتایج نشان می‌دهد تشخیص برخی کلاس‌های بسیار نادر همچنان به دلیل محدودیت ذاتی تعداد نمونه‌ها، نیازمند استفاده از روش‌های پیشرفته‌تر متعادل‌سازی یا تولید داده‌های مصنوعی در تحقیقات آینده است.

یک نمونه ماتریس درهم‌ریختگی در شکل ۵ نشان داده شده است که نتیجه اجرای آزمایش روی داده‌های نمونه ۱۰۱ منبع داده‌ای MIT-BIH بوده است. از ۱۴ نمونه، ۳ نمونه به‌درستی نرمال و ۹ نمونه به‌درستی رده / و ۳ نمونه به‌درستی رده f تشخیص داده شده‌اند.

در مجموع، نتایج نشان می‌دهد که مدل پیشنهادی در تشخیص ناهنجاری‌های رایج بسیار مؤثر عمل کرده و درعین‌حال، برای بهبود تشخیص ناهنجاری‌های نادرتر، نیاز به اتخاذ رویکردهایی نظیر افزایش نمونه‌های آموزشی، استفاده از تکنیک‌های پیشرفته‌تر داده‌افزایی یا روش‌های مقابله با عدم توازن کلاس می‌باشد.

## ۵- نتیجه‌گیری و کارهای آتی

در این مقاله، یک روش برپایه‌ی یادگیری عمیق برای تشخیص ناهنجاری در سیگنال‌های قلبی ارائه شده است. روش پیشنهادی با بهره‌گیری از شبکه LSTM و ویژگی‌های زمانی استخراج‌شده، قادر به تشخیص دقیق ریتم‌های نرمال و غیرنرمال بوده و می‌تواند در سامانه‌های مانیتورینگ بلادرنگ سلامت قلب مورد استفاده قرار گیرد.

روش پیشنهادی این پژوهش، با رویکردی ترکیبی و هدفمند، چندین نوآوری کلیدی را ارائه می‌دهد که بر «سادگی، تفسیرپذیری و کارایی عملی» متمرکز هستند. مهم‌ترین نوآوری، «استخراج تنها ۹ ویژگی مورفولوژیکی و زمانی معنادار بالینی» (مانند نقاط P، Q، R، S و T و فواصل مربوطه) است که به جای استفاده از صدها ویژگی پیچیده یا انتزاعی، مستقیماً با پاتوفیزیولوژی قلبی مرتبط هستند. این امر منجر به کاهش چشمگیر ابعاد داده، جلوگیری از مسئله‌ی ابعاد بالا، افزایش قابلیت تفسیر برای پزشکان و کاهش بار محاسباتی می‌شود.

معماری پیشنهادی، با «ترکیب بهینه دانش دامنه و یادگیری عمیق»، از یک لایه LSTM ساده با ساختار پنجره‌ای زمانی بهره می‌برد. این معماری امکان مدلسازی وابستگی‌های بلندمدت در سیگنال ECG را فراهم می‌آورد، در حالی که با داشتن تنها حدود ۷۶ هزار پارامتر، از پیچیدگی کمتری نسبت به شبکه‌های عمیق متداول برخوردار است. علیرغم این سادگی، روش به «صحت قابل توجه ۹۷/۳۳٪ در تشخیص همزمان ۱۹ کلاس آریتمی» دست یافته که با عملکرد مدل‌های بسیار پیچیده‌تر رقابت می‌کند. این موفقیت، حاکی از تعادل دقیق طراحی‌شده بین صحت و پیچیدگی است.

روش پیشنهادی دارای چندین مزیت مهم است که آن را برای کاربردهای بالینی مناسب می‌سازد. استفاده از تنها ۹ ویژگی مورفولوژیکی و زمانی، موجب کاهش بار محاسباتی و امکان پیاده‌سازی در سیستم‌های بلادرنگ و دستگاه‌های پوشیدنی شده است. بهره‌گیری از ویژگی‌های شناخته‌شده ECG، تفسیرپذیری و شفافیت تصمیم‌گیری مدل را افزایش می‌دهد. همچنین، با وجود معماری ساده شامل یک لایه LSTM و حدود ۷۶ هزار پارامتر، مدل توانسته به صحت ۹۷/۳۳٪ در طبقه‌بندی ۱۹ کلاس آریتمی دست یابد و عملکردی قابل رقابت با مدل‌های پیچیده‌تر ارائه کند. علاوه بر این، سادگی ساختار، خطر بیش‌برازش را کاهش داده و امکان یکپارچه‌سازی آسان با سامانه‌های پزشکی موجود را فراهم می‌کند. ترکیب دانش پزشکی با یادگیری عمیق، پایداری مناسب در برابر نویز و قابلیت بازتولید نتایج، از دیگر نقاط قوت این روش محسوب می‌شوند.

با وجود عملکرد مطلوب، روش پیشنهادی دارای برخی محدودیت‌ها است. عملکرد مدل به صحت مرحله پیش‌پردازش و استخراج نقاط P ، Q، R، S و T وابسته بوده و وجود نویز یا کیفیت پایین سیگنال می‌تواند باعث کاهش صحت طبقه‌بندی شود. همچنین، آموزش مدل بر روی مجموعه‌داده MIT-BIH ممکن است تعمیم‌پذیری آن را به سایر پایگاه‌های داده محدود کند. استفاده از پنجره‌های ثابت، توانایی مدل را در تشخیص آریتمی‌های طولانی‌مدت یا بسیار نامنظم کاهش می‌دهد و تمرکز بر ویژگی‌های زمانی-مورفولوژیکی ممکن است برخی اطلاعات مهم موجود در سایر حوزه‌ها را نادیده بگیرد. علاوه بر این، افزایش تعداد کلاس‌ها می‌تواند نیازمند معماری پیچیده‌تر و ویژگی‌های بیشتری باشد و با وجود تفسیرپذیری مناسب، سازوکار داخلی LSTM همچنان به‌طور کامل شفاف نیست.

از جنبه عملی و کاربردی، سیستم طراحی‌شده «متعادل از نظر کارایی-صحت» است و با زمان استنتاج کوتاه و مصرف حافظه پایین، پتانسیل استقرار در سیستم‌های بلادرنگ و سخت‌افزارهای کم‌مصرف (مانند دستگاه‌های پوشیدنی) را داراست. همچنین، با «رویکرد تفسیرپذیر برپایه‌ی ویژگی‌های بالینی شناخته‌شده»، شفافیت و قابلیت اعتماد لازم برای محیط‌های پزشکی را فراهم می‌سازد. پشتیبانی از استراتژی‌های آموزشی کارا (مانند توقف زودهنگام)، مدیریت هوشمندانه عدم تعادل داده‌ها، و ارائه کد متن‌باز و کاملاً مستند، از دیگر نقاط قوت این روش هستند که تکثیرپذیری و توسعه‌پذیری آن را تضمین می‌کنند. در مجموع، این پژوهش مسیری نوین را با تأکید بر «ادغام هوشمندانه سادگی محاسباتی، صحت بالا و تفسیرپذیری بالینی» در حوزه تشخیص آریتمی قلبی ارائه می‌دهد.

در راستای توسعه و تکمیل این پژوهش، مسیرهای متعددی برای کارهای آتی پیشنهاد می‌شود. اولویت نخست، توسعه یک «معماری ترکیبی پیشرفته» (مانند CNN-LSTM) خواهد بود تا امکان استخراج خودکار ویژگی‌های مکانی و مدل‌سازی وابستگی‌های زمانی به صورت

1. Oversampling
2. Undersampling

همزمان فراهم گردد. برای مقابله با چالش عدم تعادل داده‌ها، به کارگیری تکنیک‌های «افزایش داده هوشمند» و روش‌های نمونه‌برداری پیشرفته مانند SMOTE برای کلاس‌های اقلیت حیاتی است. همچنین، انتقال مدل به محیط‌های عملیاتی از طریق توسعه «سیستم‌های تشخیص بلادرنگ و سبک‌وزن» (مانند پیاده‌سازی بر پایه MobileNet) برای اجرا روی دستگاه‌های همراه و سخت‌افزارهای تعبیه شده (مانند FPGA) در دستور کار قرار دارد. افزودن قابلیت «تفسیرپذیری[1]» با بهره‌گیری از روش‌هایی مانند Grad-CAM برای شفاف‌سازی منطق تصمیم‌گیری مدل و نیز ادغام داده‌های چندوجهی (مانند سن، جنسیت و سیگنال‌های فیزیولوژیک دیگر) به منظور افزایش صحت تشخیصی، از دیگر اهداف کلیدی محسوب می‌شوند.

در کارهای آینده، می‌توان از یادگیری فدرال[2] برای حفظ حریم خصوصی داده‌ها و از داده‌های چندکاناله برای افزایش بیشتر صحت مدل استفاده کرد. یادگیری فدرال یک روش یادگیری ماشین است که یک الگوریتم را در چندین دستگاه لبه غیرمتمرکز یا سرورهای نگهدارنده نمونه داده‌های محلی، بدون مبادله آنها آموزش می‌دهد. توسعه این چارچوب برای تشخیص سایر ناهنجاری‌های پزشکی نیز پیشنهاد می‌شود.

## مراجع


[1] British Heart Foundation Health Analytics Team, *Global Cardiovascular Disease Factsheet*, British Heart Foundation, Jan. 2026.

[2] U. Naeem, *et al.*, "Global burden of cardiovascular disease: What should be expected in the next 25 years?" *Current Cardiology Reports*, vol. 28, no. 1, Article ID: 41826546, Mar. 2026.

[3] S. Buzby "Heart disease remains leading cause of death, despite declining overall mortality," *Healio*, 21 Jan. 2026.

[4] S. Antipolis "Avoidable inequalities remain in cardiovascular disease burden and care," *European Society of Cardiology*, 19 May 2026

[5] N. Khan and M. N. Imtiaz, *Pan-Tompkins++: A Robust Approach to Detect R-peaks in ECG Signals*, arXiv preprint, arXiv:2211.03171, Nov. 2022.

[6] J. Hampton and D. Adlam, *The ECG Made Easy E-Book*, 10th ed. Amsterdam, The Netherlands: Elsevier Health Sciences, 2024.

[7] A. Darmawahyuni, *et al.,* "An improved electrocardiogram arrhythmia classification performance with feature optimization," *BMC Medical Informatics and Decision Making*, vol. 24, no. 1, Article ID: 412, 2024.

[8] T. Chen, H. Zhang, and J. Wang, "ECG arrhythmia classification using random forest with improved feature selection," *Biomedical Signal Processing and Control*, vol. 82, Article ID: 104537, 2023.

[9] C. Zhang, *et al.*, "A low-power ECG processor ASIC based on an artificial neural network for arrhythmia detection," *Applied Sciences*, vol. 13, no. 17, Article ID: 9591, 2023.

[10] S. L. Kailan, W. H. M. Kurdi, A. H. Najim, and M. N. Kadhim, "Efficient ECG classification based on machine learning and feature selection algorithm for IoT-5G enabled health monitoring systems," *International Journal of Intelligent Engineering & Systems*, vol. 18, no. 1, pp. 1187-1199, 2025.

[11] M. Hassaballah, Y. M. Wazery, I. E. Ibrahim, and A. Farag, "ECG heartbeat classification using machine learning and metaheuristic optimization for smart healthcare systems," *Bioengineering*, vol. 10, no. 4, Article ID: 429, 2023.

[12] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 1997.

[13] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning*, Cambridge, MA, USA: MIT Press, 2016.

[14] Y. N. Wang, G. Zhou, and C. Yang, "Interpatient heartbeat classification using modified residual attention network with two-phase training and assistant decision," *IEEE Trans. on Instrumentation and Measurement*, vol. 72, pp. 1-15, 2022.

[15] M. Khazaei and N. Ashrafi-Payaman, "An unsupervised anomaly detection model for weighted heterogeneous graph," *Journal of Artificial Intelligence and Data Mining*, vol. 11, no. 2, pp. 237-245, 2023.

[16] S. Savalia and V. Emamian, "Cardiac arrhythmia classification by multi-layer perceptron and convolution neural networks," *Bioengineering*, vol. 5, no. 2, Article ID: 35, 2018.

[17] A. Petrovic, *et al.*, "Anomaly detection in electrocardiogram signals using metaheuristic optimized time-series classification with attention incorporated models," *International Journal of Hybrid Intelligent Systems*, vol. 20, no. 2, pp. 159-183, 2024.

[18] M. Thill, W. Konen, H. Wang, and T. Bäck, "Temporal convolutional autoencoder for unsupervised anomaly detection in time series," *Applied Soft Computing*, vol. 112, Article ID: 107751, 2021.

[19] H. Li and P. Boulanger, "Structural anomalies detection from electrocardiogram (ECG) with spectrogram and handcrafted features," *Sensors*, vol. 22, no. 7, Article ID: 2467, 2022.

[20] E. Šabić, D. Keeley, B. Henderson, and S. Nannemann, "Healthcare and anomaly detection: Using machine learning to predict anomalies in heart rate data," *AI & Society*, vol. 36, no. 1, pp. 149-158, 2021.

[21] S. M. H. Irid, D. Moussaoui, M. Hadjila, and O. Azzoug, "Classification of ECG signals based on MIT-BIH dataset using Bi-LSTM model for assisting cardiologists diagnosis," *Traitement du Signal*, vol. 41, no. 6, pp. 3245-3253, Dec. 2024.

[22] G. Karthik, T. Muppala, V. Chowdary, and S. Jayan, "ECG signal classification and anomaly detection using LSTM and HMM models," in *Proc. 11th Int. Conf. on Advances in Computing and Communications*, 6 pp., Kochi, India, 6-8 Nov. 2024.

[23] M. E. Ait Bourkha, A. Hatim, D. Nasir, and S. El Beid, "Optimized wavelet scattering network and CNN for ECG heartbeat classification from MIT-BIH arrhythmia database," *International Journal of Advanced Computer Science and Applications*, vol. 16, no. 2, Article ID: 116, 2025.

[24] M. Alghieth, "DeepECG-Net: A hybrid transformer-based deep learning model for real-time ECG anomaly detection," *Scientific Reports*, vol. 15, no. 1, Article ID: 20714, 2025.

[25] H. M. Rai, J. Yoo, and S. Dashkevych, "GAN-SkipNet: A solution for data imbalance in cardiac arrhythmia detection using electrocardiogram signals from a benchmark dataset," *Mathematics*, vol. 12, no. 17, Article ID: 2693, 2024.

[26] A. Rajput, N. R. Patel, R. S. Bhati, A. P. Singh, and H. Fatima, "Anomaly detection in ECG using deep learning," in *Proc. 11th Int. Conf. on Computing for Sustainable Global Development*, pp. 76-82, New Delhi, India, 28 Feb.-1 Mar. 2024.

[27] S. Hiriyannaiah, S. G. M., K. M. H. M., and K. G. Srinivasa, "A comparative study and analysis of LSTM deep neural networks for heartbeats classification," *Health and Technology*, vol. 11, no. 3, pp. 663-671, 2021.

[28] S. Sahoo, M. Dash, S. Behera, and S. Sabut, "Machine learning approach to detect cardiac arrhythmias in ECG signals: A survey," *Innovation and Research in BioMedical engineering*, vol. 41, pp. 185-194, Aug. 2020.

[29] V. Nair and G. E. Hinton, "Rectified linear units improve restricted Boltzmann machines," in *Proc. 27th Int. Conf. Machine Learning*, pp. 807-814, 21-24 Jun. 2010.

[30] J. S. Bridle, "Training stochastic model recognition algorithms as networks can lead to maximum mutual information estimation of parameters," *Advances in Neural Information Processing Systems*, vol. 2, pp. 211-217, Denver, CO, USA, 27-30 Nov. 1989.

[31] G. B. Moody and R. G. Mark, *MIT-BIH Arrhythmia Database*, PhysioNet, 2005. [Online]. Available: https://physionet.org/content/mitdb/1.0.0

[32] D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, arXiv preprint, arXiv:1412.6980, 2014.

[33] D. M. Powers, *Evaluation: From Precision, Recall and F-measure to ROC, Informedness, Mmarkedness and Correlation*, arXiv preprint, arXiv:2010.16061, 2020.

[34] Y. Sasaki, *The Truth of the F-Measure*, *Teach Tutor Materials*, 5 pp., 26 Oct. 2007. Available: https://people.cs.pitt.edu/~litman/courses/cs1671s20/F-measure-YS-26Oct07.pdf [Accessed: Oct. 25, 2026].

[35] J. Brabec, T. Komárek, V. Franc, and L. Machlica, "On model evaluation under non-constant class imbalance," in *Proc. Int. Conf.*


1. Explainable AI
2. Federated Learning

*on Computational Science*, pp. 74-87, Amsterdam, The Netherlands, 3-5 Jun. 2020.

[36] W. Siblini, J. Fréry, L. He-Guelton, F. Oblé, and Y. Q. Wang, "Master your metrics with calibration," in *Proc. 18th Int. Symp. on Advances in Intelligent Data Analysis*, pp. 457-469, Konstanz, Germany, 27-29 Apr. 2020

[37] D. Chicco and G. Jurman, "The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation," *BMC Genomics*, vol. 21, no. 1, Article ID: 6, 2020.

[38] C. Qiu, H. Li, C. Qi, and B. Li, "Enhancing ECG classification with continuous wavelet transform and multi-branch transformer," *Heliyon*, vol. 10, no. 5, ArticleID: e26147, Mar. 2024.

[39] J. Park, K. Lee, N. Park, S. C. You, and J. Ko, "Self-Attention LSTM-FCN model for arrhythmia classification and uncertainty assessment," *Artificial Intelligence in Medicine*, vol. 142, ArticleID: 102570, Aug. 2023.

**نصرت‌علی اشرفی‌پیامن** مدرک کارشناسی مهندسی نرم‌افزار را از دانشگاه خوارزمی، مدرک کارشناسی ارشد علوم کامپیوتر را از دانشگاه صنعتی شریف و مدرک دکتری در رشته مهندسی نرم‌افزار را دانشگاه علم و صنعت ایران دریافت کرد. او در حال حاضر استادیار دانشکده فنی و مهندسی در دانشگاه خوارزمی است. علایق تحقیقاتی ایشان شامل یادگیری ماشین، داده‌کاوی، گراف‌کاوی و آسیب‌پذیری نرم‌افزار است.

**بهراد مظاهری** در سال ۱۳۷۸ در تهران، ایران متولد شد. او مدرک کارشناسی مهندسی نرم‌افزار را در سال ۱۴۰۲ از دانشگاه خوارزمی (KHU) دریافت کرد. ایشان در حال حاضر دانشجوی کارشناسی ارشد مهندسی کامپیوتر – هوش مصنوعی و رباتیکز دانشگاه خوارزمی ایران است. علایق تحقیقاتی فعلی او شامل داده‌کاوی، داده‌های حجیم و علم داده است.