

Synthetic Photoplethysmogram (PPG) Generation Using Genetic Programming Based Generative Model

Fatemeh Ghasemi¹, Fardin Abdali Mohammadi^{1*}

¹ Department of Computer Engineering and Information Technology, Razi University, Kermanshah, Iran

Received: 16 December 2023, Revised: 01 October 2024, Accepted: 01 December 2024

Paper type: Research

Abstract

Today, advancements in Information and Communication Technology (ICT), particularly in healthcare and cardiac activity monitoring, have led to increased adoption of Photoplethysmogram (PPG) technology in smart devices and mobile phones. The development of generative models for producing artificial PPG signals faces challenges such as a lack of diversity and constraints in training data. This article employs a Genetic Programming (GP) based approach to introduce a generative model capable of producing PPG data with increased diversity and enhanced accuracy using an initial sample of PPG signals. In contrast to conventional regression, Genetic Programming automates the determination of the mathematical model's structure and compositions. The proposed approach, with a Mean Squared Error (MSE) of 0.0001, Root Mean Square Error (RMSE) of 0.01, and a correlation of 0.999, demonstrates superior performance due to appropriate optimization and acceptable accuracy in generating synthetic PPG data. It outperforms other methods in terms of efficiency and execution capability, especially in resource-constrained environments.

Keywords: Photoplethysmogram, Generative Model, Genetic Programming, Scalability, Mathematical Model.

* Corresponding Author's email: fardin.abdali@razi.ac.ir

تولید سیگنال فوتوپلتیسموگرام (PPG) مصنوعی با استفاده از مدل سازنده مبتنی بر برنامه‌نویسی ژنتیک

فاطمه قاسمی^۱، فردین ابدالی محمدی^{۲*}

^۱ دانشجوی کارشناسی ارشد، گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشکده فنی و مهندسی، دانشگاه رازی، کرمانشاه، ایران
^۲ دانشیار، گروه مهندسی کامپیوتر و فناوری اطلاعات، دانشکده فنی و مهندسی، دانشگاه رازی، کرمانشاه، ایران

تاریخ دریافت: ۱۴۰۲/۰۹/۲۵ تاریخ بازبینی: ۱۴۰۳/۰۷/۱۰ تاریخ پذیرش: ۱۴۰۳/۰۹/۱۱

نوع مقاله: پژوهشی

چکیده

امروزه تحولات فناوری اطلاعات و ارتباطات در حوزه سلامت، به‌ویژه در نظارت بر فعالیت‌های قلبی، به افزایش استفاده از تکنولوژی فوتوپلتیسموگرام (PPG: Photoplethysmogram) در دستگاه‌های هوشمند و تلفن‌های همراه منجر شده است. توسعه مدل‌های سازنده به جهت تولید سیگنال‌های مصنوعی PPG نیازمند حل چالش‌هایی مانند کمبود تنوع و محدودیت داده‌ها در آموزش مدل‌های یادگیری عمیق است. این مقاله، یک رویکرد مبتنی بر برنامه‌نویسی ژنتیک (GP: Genetic Programming) را به کار می‌گیرد تا مدل سازنده‌ای را ارائه دهد که با کمک یک نمونه اولیه از سیگنال PPG، قادر به تولید داده‌هایی با تنوع بیشتر و دقت افزوده شده باشد. برخلاف رگرسیون معمول، در برنامه‌نویسی ژنتیک ساختار و ترکیبات مدل ریاضی به صورت خودکار تعیین می‌گردد. رویکرد پیشنهادی، با داشتن اندازه خطای میانگین (MSE: Mean Squared Error) برابر با ۰,۰۰۰۱، ریشه میانگین مربعات خطا (RMSE: Root Mean Squared Error) به اندازه ۰,۰۱ و همبستگی ۰,۹۹۹ نشان می‌دهد که به دلیل بهینگی مناسب و دقت قابل قبول در تولید داده‌های PPG مصنوعی، نسبت به دیگر روش‌ها برتری دارد و از نظر کارایی و قابلیت اجرا در محیط‌های با منابع محدود نیز مؤثر عمل می‌کند.

کلیدواژه‌گان: فوتوپلتیسموگرام، مدل سازنده، برنامه‌نویسی ژنتیک، مقیاس‌پذیری، مدل ریاضی.

۱- مقدمه

در این بخش به روش‌های ارائه‌شده در جهت تولید سیگنال‌های PPG و ECG مختصراً اشاره خواهد شد. اغلب رویکردهای مدل‌سازی سیگنال‌های حیاتی عمدتاً بر اساس توابع گاوسی و شبکه‌های GAN می‌باشند. به‌طور مثال در روشی که در [۱۳] معرفی شده است، با استفاده از یک مدل پویا گاوسی مبتنی بر موج، مدل سازنده‌ای به منظور تولید داده‌های سیگنال الکتروکاردیوگرام ارائه شده است. در روشی دیگر با استفاده از آنالیز دنباله پالس‌های سیگنال PPG، سعی شده است تا با بررسی ویژگی‌های مهم دنباله، یک مدل سازنده سیگنال PPG ایجاد شود [۱۴]. در مدل‌های سازنده PPG کمینه بودن خطای نقاط مهم سیگنال PPG بسیار حائز اهمیت است. برای این کار وانگ و همکاران یک مدل سازنده PPG را با در نظر گرفتن مقدار خطای تخمین نقاط مهم سیگنال PPG با استفاده از ۴ یا ۵ تابع گاوسی را با دقت قابل قبولی ارائه دادند [۱۵]. لو و همکاران با هدف افزایش دقت در تولید سیگنال PPG مصنوعی روشی مبتنی بر لگاریتم نرمال و توابع گاوسی ارائه داده‌اند که استفاده از لگاریتم نرمال در این روش، سبب افزایش دقت مدل در ویژگی‌های طولانی سیگنال PPG شده است [۱۶] در ادامه تانگ و همکاران با استفاده از دو تابع گاوسی توانستند که سیگنال PPG را به‌صورت تک ضربانی تولید کنند که این روش در کنار دقت مناسب پارامترهای کمتری نیز دارد [۱۷].

شبکه‌های GAN نیز به‌عنوان یکی از مدل‌های یادگیری عمیق در جهت تولید داده‌های مصنوعی ارائه شده‌اند. این نوع از شبکه‌های عصبی در حوزه تولید سیگنال‌های پزشکی نیز بسیار کاربردی هستند [۱۸]. بروفی و همکاران توانستند با استفاده از شبکه DCGAN بر مبنای کانولوشن دوبعدی، سیگنال ECG و PPG تولید کنند. بدین‌صورت که این مدل ابتدا داده‌های ECG و PPG را به تصویر سیاه و سفید تبدیل می‌کند سپس این تصویر به‌عنوان ورودی مدل در نظر گرفته می‌شود. خروجی مدل سازنده ارائه‌شده یک تصویر سیاه و سفید است که در نهایت آن تصویر به یک سیگنال تبدیل می‌شود [۱۹]. در اغلب اوقات تولید داده در جهت افزایش دقت طبقه‌بند استفاده می‌شود. به‌طور مثال در روش [۲۰] با استفاده از CGAN، سیگنال PPG مصنوعی با این هدف تولید می‌شود. مدل انتشار نیز اخیراً به یکی از استانداردهای تولید داده‌های مصنوعی تبدیل شده است. استفاده از مدل انتشار در روش‌های [۲۱، ۲۲] در تولید سیگنال ECG نشان می‌دهند که این‌گونه از مدل‌ها پتانسیل بسیاری در زمینه تولید داده‌های مصنوعی مانند ECG و PPG خواهند داشت.

با این حال، مسئله حائز اهمیت آن است که ایجاد مدل‌های سازنده با استفاده از شبکه‌های عصبی، هرچند که دارای دقت خوبی می‌باشند، اما پیچیدگی بسیار بالای آن‌ها منجر به افزایش حجم مدل و

اخیراً استفاده از تکنولوژی فوتوپلتیسموگرام (PPG)، به‌عنوان یک روش کم‌هزینه و آسان، در تلفن‌های همراه و دستگاه‌های پوشیدنی هوشمند، مورد توجه قرار گرفته است [۱، ۲]. PPG یک روش غیرتهاجمی برای اندازه‌گیری تغییرات حجم خون در رگ انسان است که از طریق آن می‌توان تغییرات در نرخ ضربان قلب را به دست آورد. با استفاده از سیگنال PPG می‌توان بر فعالیت‌های قلبی از جمله ضربان قلب و فشارخون نظارت داشت و به همین علت سیگنال PPG در زمینه‌های ورزشی و نظارت بر سلامت انسان بسیار مورد استفاده قرار می‌گیرد [۳-۵]. امروزه با توجه به اهمیت سیگنال PPG، وجود داده‌های مربوط به این سیگنال به جهت آموزش مدل‌های یادگیری عمیق بسیار حائز اهمیت است. این‌گونه از مدل‌ها برای آموزش نیازمند مقادیر فراوان و متنوعی از داده‌های PPG می‌باشند. تاکنون مجموعه داده‌های PPG گوناگونی از جمله WristPPG [۶]، BioSec.Lab [۷]، PPG-BP [۸] و PPG-DaLiA [۹] ارائه شده‌اند. داده‌های PPG موجود در مجموعه WristPPG با فرکانس ۲۵۶ هرتز از ۹ شخص در بازه سنی ۲۲ تا ۳۲ در حالت‌های مختلف فیزیکی به‌دست آمده است. مجموعه BioSec.Lab شامل ۲۱۸ داده PPG از ۸۶ شخص در بازه سنی ۱۹ تا ۳۵ است که در حالت‌های مختلف با فرکانس ۱۰۰ هرتز نمونه‌برداری شده است. نمونه‌های PPG در PPG-BP شامل ۶۵۷ نمونه از ۲۱۹ شخص با فرکانس ۱ کیلوهرتز است و سن افراد شرکت‌کننده از ۲۰ تا ۸۹ متغیر است. در PPG-DaLiA، طیف وسیع‌تری از نمونه‌های PPG وجود دارد. این داده‌ها در طول ۳۶ ساعت از ۱۵ نفر در بازه سنی ۲۱ تا ۵۵ نمونه‌برداری شده است.

باتوجه به مجموعه داده‌های موجود PPG، برخلاف تنوع و فراوانی داده‌های الکتروکاردیوگرام (ECG)، به‌وضوح کمبود داده و عدم تنوع نمونه‌های سیگنال PPG مشهود است [۱۰]. این مسئله سبب ارائه مدل‌های سازنده در جهت تولید سیگنال‌های PPG به‌صورت مصنوعی شده است. از این‌رو، با استفاده از این مدل‌ها، داده‌ها به‌سادگی و با تنوع بسیاری تولید می‌شوند که این امر سبب پیشرفت در عملکرد مدل‌های یادگیری عمیق خواهد شد [۱۱، ۱۲]. با این حال ایجاد یک مدل سازنده برای تولید سیگنال PPG که جزو سیگنال‌های حیاتی محسوب می‌شود، چالش‌هایی دارد و نیازمند ارزیابی دقیق داده‌های تولیدشده از نظر صحت و کیفیت است.

در سال‌های گذشته، با توجه به توسعه حوزه پزشکی و دستگاه‌های پوشیدنی و نیاز به تولید داده‌های سری زمانی، ایده استفاده از مدل‌های سازنده به‌عنوان یکی از ابزارهای بسیار مؤثر و کارآمد در این زمینه مطرح شد. نکته حائز اهمیت آن است که خروجی مدل‌های سازنده سیگنال‌های حیاتی، نیازمند دقت بسیار بالا هستند.

$$F(x) = \begin{cases} f_1(x) = \exp_1 & 0 \leq x < m \\ f_2(x) = \exp_2 & m \leq x < 2m \\ \vdots & \vdots \\ f_n(x) = \exp_n & (n-1)m \leq x < n * m \end{cases} \quad (1)$$

در معادله ۱، بردار x ، نمونه‌های سیگنال PPG است. عبارت \exp بخشی از مدل سازنده سیگنال PPG است که توسط ترکیبی از عبارات ریاضی تولید می‌شود. در مجموع تابع تکه‌ای $F(x)$ شامل n زیر تابع می‌باشد و سیگنال PPG با اجرای تابع $F(x)$ تولید می‌شود. بهینگی این مدل توسط دو پارامتر تضمین می‌شود به شرح زیر:

- مقدار n کمینه است: مدل به نحوی پیاده‌سازی شده است که تعداد زیر توابع $F(x)$ کمترین تعداد ممکن خود را داشته باشد.
- طول تابع $F(x)$ کمینه است: اگر تعداد کاراکترهای عبارات، طول تابع $F(x)$ در نظر گرفته شود، طول تابع کمترین مقدار خود را خواهد داشت. $|f_i|$ طول کاراکترهای تابع f_i را نشان می‌دهد و مجموع تعداد کاراکترهای زیر توابع، طول تابع $F(x)$ می‌باشد که با $FTS(F(x))$ نمایش داده می‌شود.

$$FTS(F(x)) = \sum_{i=1}^n |f_i| \quad (2)$$

فلوچارت ایجاد یک مدل سازنده PPG مصنوعی با استفاده از برنامه‌نویسی ژنتیک، در شکل ۱ ارائه شده است. در ابتدا یک نمونه از سیگنال PPG انتخاب می‌شود. سپس به جهت افزایش دقت و کارکرد مدل سازنده، بر روی داده‌های PPG پیش‌پردازش‌های لازم انجام می‌شود. بخش مهم این روش ایجاد مدل سازنده با استفاده از برنامه‌نویسی ژنتیک است که مراحل آن در ادامه شرح داده می‌شود. به صورت جزئی‌تر، مراحل تولید سیگنال PPG مصنوعی با استفاده از برنامه‌نویسی ژنتیک در الگوریتم ۱ نمایش داده شده است. ورودی الگوریتم پیشنهادی، یک سیگنال PPG طبیعی از یک شخص است (خط ۱) و خروجی این الگوریتم، یک سیگنال PPG تولیدشده توسط مدل $F(x)$ می‌باشد (خط ۲). در ابتدا، تمام پیش‌پردازش‌های لازم بر سیگنال PPG اعمال می‌شود (خط ۳). همان‌طور که بحث شد، تابع F یک تابع تکه‌ای است که از n زیر تابع تشکیل شده است که در خط چهارم تعداد زیر توابع مشخص می‌شود (خط ۴). سپس سیگنال PPG به n قسمت مجزا تقسیم می‌شود (خط ۵). نقاط شروع و پایان نواحی سیگنال در دو مجموعه $StartingPoints$ و $EndingPoints$ قرار می‌گیرند (خطوط ۶ و ۷). الگوریتم GP به تعداد بخش‌های سیگنال و با بهره‌گیری از سیگنال اصلی اجرا می‌شود. نتیجه اجرای هر GP یک عبارت ریاضی است که در $ResultSet$ قرار می‌گیرد (خطوط ۹ تا ۱۱) سپس عبارات موجود در $ResultSet$ در تابع $F(x)$ ذخیره می‌شوند (خط ۱۳). نهایتاً تابع $F(x)$ قادر به ایجاد یک سیگنال مصنوعی PPG می‌باشد.

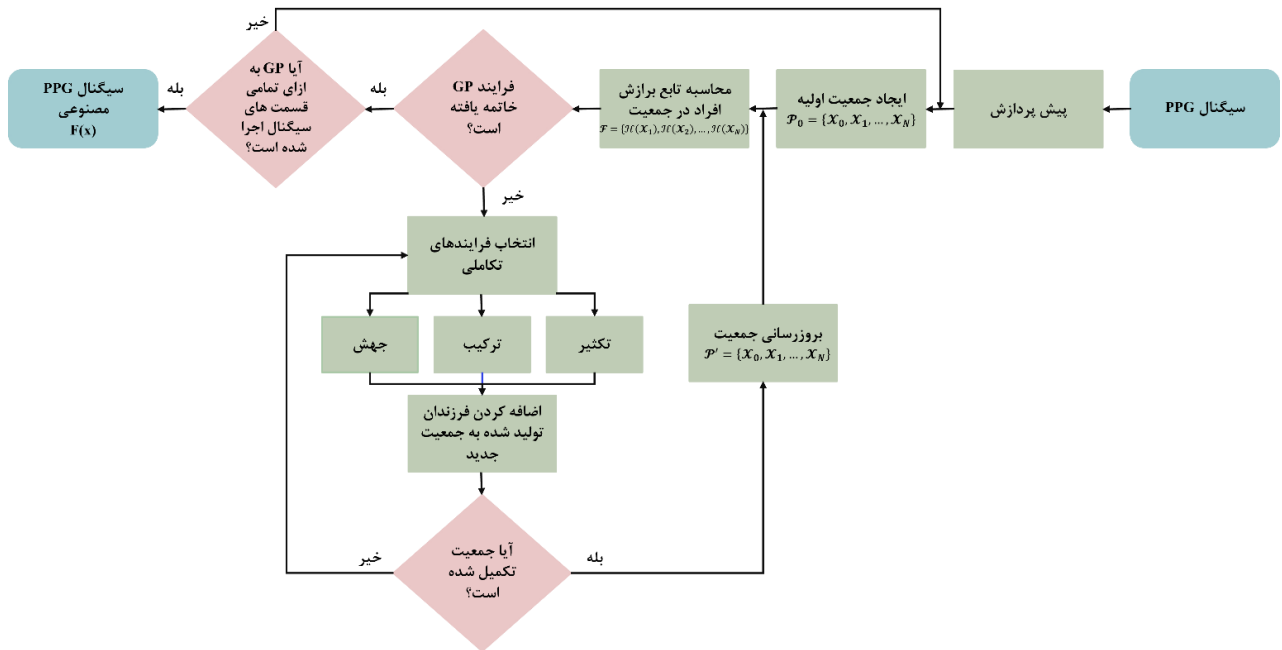
نیازمندی به منابع محاسباتی بیشتر خواهد شد. با توجه به گسترش استفاده از دستگاه‌هایی بر مبنای PPG [۲۳]، باید به این موضوع توجه داشت که استفاده از این‌گونه مدل‌ها در مواجهه با دستگاه‌هایی با منابع محدود به چالش کشیده می‌شود؛ بنابراین، امروزه نیازمند توسعه مدل‌های سازنده با دقت مناسب و درعین حال با پیچیدگی کمتر برای اجرا بر محیط‌هایی با منابع محدود هستیم.

در این پژوهش، تمرکز بر ارائه یک مدل سازنده مبتنی بر برنامه‌نویسی ژنتیک به منظور استخراج یک سیگنال PPG مصنوعی می‌باشد. برنامه‌نویسی ژنتیک که به‌عنوان مدل رگرسیونی نیز شناخته می‌شود [۲۴]، یک متد تکاملی بر مبنای درخت است که از طریق ایجاد راه‌حل‌های گوناگون و اعمال تکنیک‌های مختلف از جمله انتخاب، جهش و ترکیب تلاش می‌کند تا در طی نسل‌های مختلف و با بهره‌گیری از ترکیب عملگرهای حسابی (+, -, *, /) و عملگرهای هندسی (\sin , \cos) در بین فضای حالات موجود به بهترین جواب برسد [۲۵, ۲۶]. این روش به دلیل توانایی در یافتن راه‌حل‌های بهینه با استفاده از ترکیب عملیات ژنتیک در مسائل چالشی شناخته شده است [۲۷, ۲۸]. در برنامه‌نویسی ژنتیک، با تنظیم صحیح پارامترها، می‌توان عملکرد، دقت و میزان بهینگی مدل رگرسیونی را بهبود بخشید [۲۹]. این رویکرد از اهمیت ویژه‌ای در تولید مدل‌های پیشرفته و کاربردی در مسائل مختلف برخوردار است. در این روش، یک نمونه از سیگنال PPG به‌عنوان ورودی در نظر گرفته می‌شود. به‌منظور کاهش خطای مدل و بهبود عملکرد، پیش‌پردازش‌های لازم بر روی داده‌ها انجام می‌شود. سپس مدل GP با بهره‌گیری از نمونه اولیه PPG، در طی نسل‌های مختلف بهترین مدل را انتخاب می‌کند. در نهایت خروجی به‌صورت یک مدل ریاضی سازنده سیگنال PPG با میزان بهینگی و دقت قابل قبول ارائه می‌شود.

در بخش دوم این مقاله به تفصیل روش پیشنهادی مورد بحث قرار خواهد گرفت. سپس در بخش سوم، نتایج حاصل از اجرای این روش، ارزیابی پارامترها، دقت مدل و مقایسه آن با دیگر روش‌های موجود به طور جامع مورد بررسی و ارزیابی قرار خواهد گرفت و در نهایت بخش نتیجه‌گیری ارائه خواهد شد.

۲- مواد و روش‌ها

هدف از ارائه روش پیشنهادی، ایجاد یک مدل سازنده سیگنال PPG با دقت و بهینگی مطلوب است. به‌منظور دستیابی به این مدل، یک تابع تکه‌ای $F(x)$ تعریف می‌شود که از زیر توابع $\{f_1(x), f_2(x), \dots, f_n(x)\}$ تشکیل شده است. هر یک از زیر توابع توسط الگوریتم برنامه‌نویسی ژنتیک (GP) ایجاد می‌شود و قسمت مشخصی از یک سیگنال PPG را نمایش می‌دهد. فرمول تابع تکه‌ای $F(x)$ در معادله ۱ نمایش داده شده است.



شکل ۱. فلوچارت ایجاد یک مدل سازنده PPG با استفاده از برنامه‌نویسی ژنتیک

اصلی PPG اعمال می‌شود. سپس، به ازای هر یک از نواحی سیگنال، GP صورت مجزا اجرا خواهد شد.

۲-۲- ایجاد سیگنال PPG مصنوعی با استفاده از GP

برنامه‌نویسی ژنتیک، به جهت تولید عبارات با استفاده از رویکردهای تکاملی ایجاد شده است [۳۱]. در GP برنامه‌های کامپیوتری برای حل مسئله، به صورت تصادفی ایجاد می‌شوند. هر فرد در GP نشانگر یک برنامه است. این برنامه‌ها از ۳ عمل تکاملی انتخاب، جهش و ترکیب ایجاد می‌شوند و توسط تابع برازشی که تعریف شده است سنجیده می‌شوند. معمولاً از برنامه‌هایی با سطح سازگاری بیشتر با مسئله، برای ایجاد نسل‌های بعدی استفاده می‌شود. در نهایت بهترین راه‌حل تولید شده توسط GP، در حقیقت بهترین فرد سازگار در طی نسل‌های مختلف هست.

در روش پیشنهادی تمرکز بر ایجاد یک مدل سازنده PPG با استفاده از برنامه‌نویسی ژنتیک بر مبنای درخت است تا با ایجاد درخت‌های گوناگون، عبارت مطلوب یافت شود [۳۲، ۳۳]. در الگوریتم ۲، شبه کد برنامه‌نویسی ژنتیک با جزئیات بیشتری شرح داده شده است. ورودی این الگوریتم، یک ناحیه خاص از سیگنال PPG ($splited_data_j$) با نقاط شروع و پایان مشخص s_j و e_j می‌باشد (خط ۱). خروجی این الگوریتم نیز یک عبارت ریاضی (exp_j) است که بهترین پاسخ در کل نسل‌های GP می‌باشد و قابلیت ایجاد ناحیه-ای از سیگنال PPG را خواهد داشت (خط ۲).

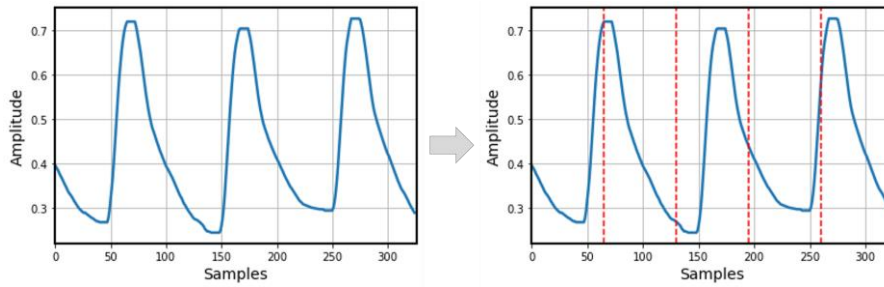
الگوریتم ۱. شبه کد روش پیشنهادی

1. Input: PPG Signal(data)
2. Output: Near Optimal Generated Model($F(x)$)
3. $data \leftarrow Preprocessing(data)$
4. $N \leftarrow Number\ Of\ Partitions$
5. $Splited_Data \leftarrow Split_Data(data, N)$
6. $StartingPoints = \{s_1, s_2, \dots, s_n\}$
7. $EndingPoints = \{e_1, e_2, \dots, e_n\}$
8. $ResultSet \leftarrow \emptyset$
9. for $j=1, \dots, N$ do.
10. $Exp = GP(Splited_Data_j, s_j, e_j)$
11. $ResultSet = ResultSet \cup Exp$
12. end for
13. $F(x) \leftarrow ResultSet$

۲-۱- پیش پردازش سیگنال PPG

پس از انتخاب سیگنال PPG، قسمتی از سیگنال که شامل سه پیک سیستولیک^۱ است، انتخاب شده و بر روی آن یک فیلتر میانه به جهت کاهش نویز اعمال می‌شود. فیلتر میانه یک فیلتر غیرخطی است که در حذف نویزهای سیگنال بسیار مؤثر عمل می‌کند [۳۰]. با تعیین طول پنجره، فیلتر، میانه اعداد در آن مجموعه را می‌یابد و عدد میانه را جایگزین کل مجموعه می‌کند. این فرایند سبب بهبود کیفیت سیگنال PPG برای استفاده در مراحل بعدی می‌شود. پس از اعمال فیلتر بر روی سیگنال PPG، سیگنال به n ناحیه مجزا تقسیم می‌شود (شکل ۲. Error! Reference source not found). این تقسیم‌بندی به جهت افزایش دقت مدل سازنده، بر روی سیگنال

^۱ systolic



شکل ۲. تقسیم‌بندی سیگنال اصلی PPG به نواحی مختلف، مبنای ایجاد یک مدل سازنده با دقت بیشتر در روش پیشنهادی است.

$$\mathcal{X}_{\text{selected}} = \arg \max_{\mathcal{X}_{i_j} \in T} \mathcal{H}(\mathcal{X}_{i_j}) \quad (4)$$

پس از انتخاب نمونه‌هایی با سطح سازگاری بیشتر در جمعیت فعلی، دو فرایند تکاملی ترکیب و جهش بر روی آن نمونه‌ها اعمال می‌شود (خط ۷). استفاده درست از عملیات تکاملی سبب بهبود سازگاری افراد در نسل‌های بعدی خواهد شد. نسل جدید ایجاد شده توسط فرایندهای تکاملی، جایگزین افراد در نسل قبلی می‌شود (خط ۸). پس از ارزیابی میزان سازگاری افراد در نسل جدید، سازگارترین فرد در نسل کنونی با بهترین فرد در نسل‌های گذشته مقایسه می‌شود و در صورت برتری، به‌عنوان بهترین پاسخ در نظر گرفته خواهد شد (خط ۹). تا زمانی که تمامی نسل‌ها ایجاد شوند، این مراحل تکرار خواهند شد و در نهایت بهترین فرد در میان نسل‌ها به‌عنوان پاسخ نهایی GP در نظر گرفته می‌شود (خط ۱۱).

الگوریتم ۲: الگوریتم برنامه‌نویسی ژنتیک (GP)

1. Input: splitted_data_j, s_j ∈ StartingPoints, e_j ∈ EndingPoints
2. Output: exp_j (mathematical expression)
3. P₀ ← Choose an initial random population of individuals
4. for i = 1 to Generation No do
5. H(X_i) ← Evaluate the fitness of the individuals in P_i
6. s ← Select the best individuals in P_i to be used by the genetic operators
7. U ← Create new individuals via crossover and mutation in s until the population is complete
8. P_{i+1} ← U
9. Update Best_individual
10. end for
11. exp_j ← Best_individual

۲-۳- بازنمایی خروجی توابع GP

مدل سازنده PPG با بهره‌گیری از توابع مختلفی که توسط GP ایجاد شده است، ساخته می‌شود. این توابع به‌صورت عبارت ۵

در GP، ابتدا یک جمعیت اولیه از افراد به‌صورت تصادفی ایجاد می‌شود که می‌توان آن را به‌صورت $\mathcal{P} = \{\mathcal{X}_1, \mathcal{X}_2, \dots, \mathcal{X}_i, \dots, \mathcal{X}_N\}, \forall i \in \{1, \dots, N\}$ نمایش داد. هر \mathcal{X}_i نشان‌دهنده یک فرد در جمعیت است. در این قسمت، N تعداد افراد در جمعیت را نشان می‌دهد که اغلب، این تعداد از طریق آزمایشات متعدد به دست می‌آید (خط ۳). سپس برای هر فرد در جمعیت تابع برازش محاسبه می‌گردد. این تابع به‌صورت $\mathcal{H}: \mathcal{X} \rightarrow \mathbb{R}$ نشان داده می‌شود که $\mathcal{H}(\mathcal{X}_i)$ نشان‌دهنده میزان سازگاری فرد \mathcal{X}_i در جمعیت \mathcal{P} با مسئله است. در این روش جهت ارزیابی میزان سازگاری هر یک از افراد با مسئله، از یک تابع ترکیبی به نام تخمین خطای منفی گرا (PLE) استفاده می‌شود (معادله ۳). در این تابع $e(\mathcal{g}_h^i)$ میزان خطای ژن \mathcal{g}_h^i فرد \mathcal{X}_i را در مقایسه با مسئله می‌یابد. به عبارت دیگر این خطا نشان‌دهنده میزان انحراف خروجی تولید شده توسط ژن از سیگنال هدف است که به‌عنوان یک شاخص برای سنجش عملکرد ژن در حل مسئله مورد نظر به کار می‌رود. $\Omega(\mathcal{g}_h^i)$ یک معیار کمی برای سنجش میزان پیچیدگی ژن \mathcal{g}_h^i است که می‌تواند به‌صورت تابعی از تعریف شود. تابع $\mathcal{T}(\mathcal{g}_h^i)$ تعداد گره‌های درخت در هر ژن \mathcal{g}_h^i در نمونه \mathcal{X}_i را محاسبه می‌کند که این معیار، اندازه و پیچیدگی ساختاری ژن را نشان می‌دهد. در نهایت متغیر L نشان‌دهنده مجموع کل تعداد کل گره‌های فرد \mathcal{X}_i می‌باشد.

$$\mathcal{H}(\mathcal{X}_i) = \frac{\sum_{h=1}^{N_{\text{model}}} [e(\mathcal{g}_h^i) + \Omega(\mathcal{g}_h^i)]}{\sum_{h=1}^{N_{\text{model}}} \mathcal{T}(\mathcal{g}_h^i)} = \frac{e(\mathcal{X}_i) + \Omega(\mathcal{X}_i)}{L} \quad (3)$$

پس از محاسبه تابع برازش به ازای تمام افراد جمعیت، افراد سازگارتر توسط بازای انتخاب با استفاده از روش تورنمنت انتخاب خواهند شد (خط ۶). در این فرایند، یک تورنمنت با اندازه k ایجاد می‌شود که به‌صورت $T = \{\mathcal{X}_{i_1}, \mathcal{X}_{i_2}, \dots, \mathcal{X}_{i_k}\}$ نشان داده می‌شود. \mathcal{X}_{i_j} نشان‌دهنده فرد انتخاب شده در تورنمنت است. در عبارت ۴ خواهیم دید که از میان افراد انتخاب شده، فردی که عملکرد بهتری دارد به‌عنوان فرد منتخب در نظر گرفته می‌شود. به این ترتیب، با استفاده از این روش، افراد با عملکرد بهتر احتمال بیشتری برای مشارکت در ایجاد نسل بعدی خواهند داشت.

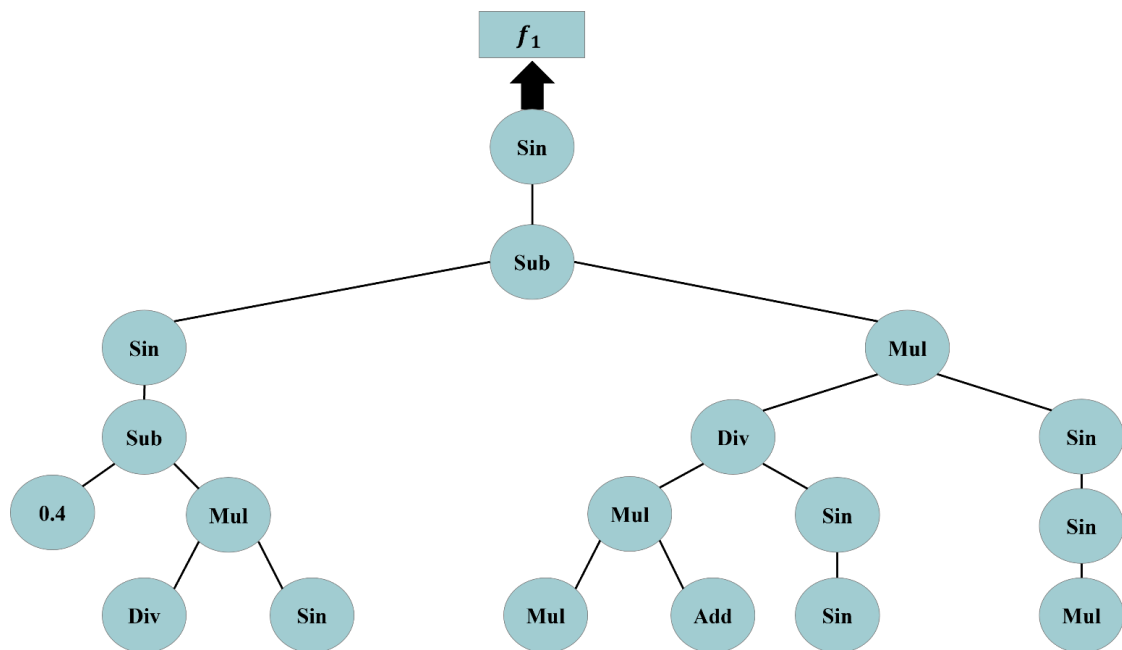
$$f_1 = \sin(\sin(0.4 - ((0.38 \times (x - \cos(\dots) \cos((-0.74 \times x) + 0.49 \dots \quad (۶)$$

در این روش، هر زیر تابع f_i بر اساس بخشی از سیگنال واقعی PPG که به صورت مجزا مشخص شده است، ایجاد می‌شود. فرایند GP تضمین می‌کند که هر تابع به صورت اختصاصی برای تولید بخش تعیین شده از سیگنال PPG آموزش داده می‌شود و با بخش‌های دیگر همپوشانی نخواهد داشت. هر یک از توابع به دست آمده به صورت مجزا اجرا می‌شوند و در نهایت خروجی آن‌ها جمع‌آوری شده و یک سیگنال مصنوعی PPG مشابه با سیگنال اصلی را تولید می‌کنند. در شکل ۴ فرایند تولید گام‌به‌گام سیگنال مصنوعی PPG نمایش داده شده است.

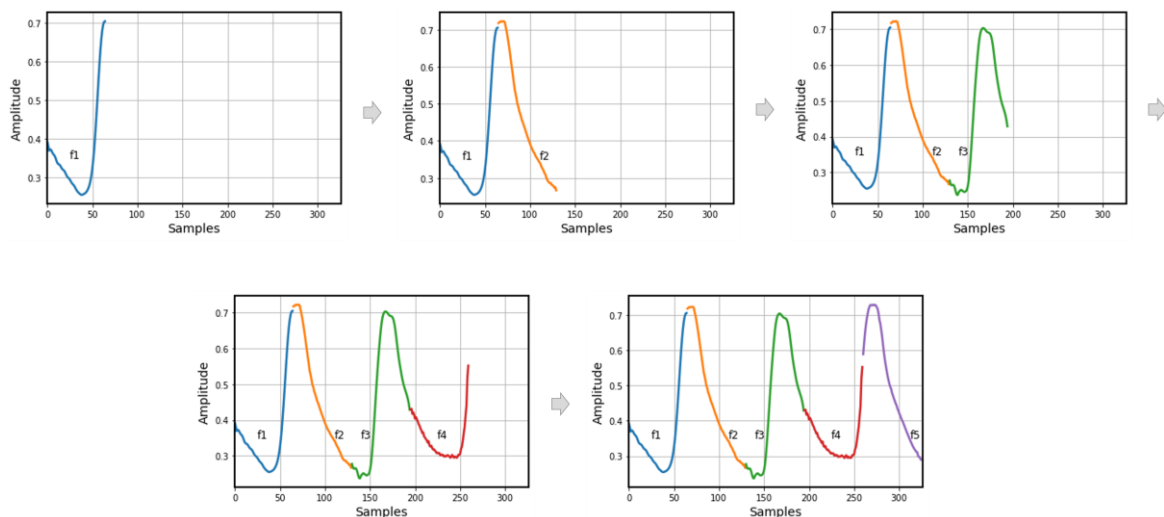
قابل بیان است که در آن تابع سازنده ناحیه i سیگنال PPG می‌باشد. هر یک از توابع f_i در حقیقت بهترین فرد در کل نسل‌های ایجاد شده توسط برنامه‌نویسی ژنتیک به ازای آن ناحیه می‌باشد. استفاده از این شیوه، منجر به ساخت یک سیگنال PPG با دقت و بهینگی بیشتر توسط GP خواهد شد.

$$Synthetic PPG = \{f_1, f_2, \dots, f_i, \dots, f_n\} \quad (۵)$$

به عنوان نمونه، عبارت f_1 ، بهترین فرد در اجرای GP به ازای بخش اول سیگنال PPG اصلی می‌باشد که ترکیبی از عملگرهای حسابی و هندسی است و به طور مختصر به صورت عبارت ۶ تعریف می‌شود. برای درک بیشتر از شیوه عملکرد توابع ایجاد شده توسط GP، یک نمایش از درخت GP توسط تابع f_1 در شکل ۳ نشان داده شده است.



شکل ۳. مثالی از درخت GP برای تابع f_1 که متشکل از عملگرهای (+, *, /) و (sin, cos) می‌باشد.



شکل ۴. ایجاد گام‌به‌گام یک سیگنال PPG مصنوعی با بهره‌گیری از توابع بهینه‌سازی شده تولید شده توسط برنامه‌نویسی ژنتیک

۳- بحث و نتایج

جدول ۱. تنظیمات اعمال شده بر پارامترهای GP در جهت ایجاد مدل

سازنده PPG	
پارامتر	مقدار
تعداد جمعیت (Population size)	۴۰۰
تعداد نسل‌ها (No. of Generation)	۱۷۰۰
توابع استفاده شده (Function set)	Add-Sub-Mul-Div-Cos-Sin
بیشترین عمق درخت‌ها (Max depth of trees)	۴
انتخاب (Selection)	تورنمنت با سایز ۳
جهش (Mutation)	۰,۲ (یکنواخت (Uniform))
ترکیب (Crossover)	۰,۷۸ (تک نقطه (one point))

جدول ۲. جزئیات مدل‌های ایجاد شده توسط GP به ازای

مجموعه $\{f_1, f_2, f_3, f_4, f_5\}$		
نام زیر تابع	زمان اجرای GP (ثانیه)	طول تابع
f_1	۳۲۹,۰۶	۱۴۱۳
f_2	۷۶۱,۱۷	۱۶۱۸
f_3	۶۳۵,۴۶	۱۶۸۵
f_4	۶۲۵,۶۲	۱۸۵۸
f_5	۸۴۱,۹۹	۱۹۶۹

در این بخش، مدل پیشنهادی، به‌طور کامل مورد بررسی و ارزیابی قرار گرفته است. در این روش به جهت ایجاد مدل سازنده، از قسمتی از سیگنال PPG که شامل سه پیک سیستولیک است استفاده می‌شود. داده PPG مورد استفاده در این روش از مجموعه داده MIMIC-III [۳۴] می‌باشد. این مجموعه داده شامل سیگنال‌های حیاتی فرد اعم از ECG و PPG از ۳۰۰۰۰ بیمار با نرخ نمونه‌برداری ۱۲۵ کیلوهرتز می‌باشد. به جهت افزایش دقت مدل، سیگنال به ۵ ناحیه مساوی حاوی ۶۵ نمونه تقسیم می‌شود. GP بر روی نواحی سیگنال با نقاط شروع ($startingPoint_j$) و پایان ($endingPoint_j$) اجرا می‌شود. همان‌طور که در بخش قبل گفته شد، خروجی هر فرایند GP یک عبارت ریاضی (exp) است که این عبارت ریاضی در حقیقت سازگارترین برنامه در طی نسل‌های مختلف GP در بین دیگر برنامه‌های ایجاد شده می‌باشد و قادر به ایجاد آن بخش از سیگنال PPG با کمترین خطا است. در جدول ۱ تنظیمات اولیه GP در جهت تحقق این مدل نمایش داده شده است که این مقادیر بر اساس آزمایشات متعدد به دست آمده است. پس از اجرای کامل GP به ازای تمامی نواحی سیگنال PPG، مدل‌ها به صورت $F = \{f_1, f_2, f_3, f_4, f_5\}$ قرار خواهند گرفت. در جدول ۲ جزئیات اجرای هر یک از توابع f شرح داده شده است. تمام پیاده‌سازی‌ها در سیستمی با مشخصات 11th Gen Intel Core i7 11800H، رم 32G انجام شده است. برای پیاده‌سازی برنامه‌نویسی ژنتیک، از پکیج DEAP^۱ در پایتون استفاده شده است [۳۵]. میزان دقت مدل سازنده ایجاد شده توسط برنامه‌نویسی ژنتیک توسط پارامترهای تابع میانگین مربعات خطا (MSE)، خطای جذر میانگین مربعات (RMSE) و ضریب همبستگی پیرسون برای کل بازه سیگنال مورد ارزیابی قرار گرفته شده است.

$$MSE = \frac{1}{n} \sum_{i=1}^n (F_r(i) - F_s(i))^2 \quad (7)$$

$$RMSE = \sqrt{MSE} \quad (8)$$

$$Corr(F_r, F_s) = \frac{\sum_{i=1}^n (F_{r_i} - \bar{F}_r)(F_{s_i} - \bar{F}_s)}{\sqrt{\sum_{i=1}^n (F_{r_i} - \bar{F}_r)^2} \sqrt{\sum_{i=1}^n (F_{s_i} - \bar{F}_s)^2}} \quad (9)$$

در فرایند ایجاد یک مدل سازنده PPG با بهینگی و دقت مناسب توسط برنامه‌نویسی ژنتیک، تنظیمات مختلفی بر روی پارامترهای کلیدی GP اعمال شده است تا به عملکرد قابل قبول توسط مدل سازنده ایجاد شده دست یافت. با آزمایش مقادیر مختلف احتمال وقوع فرآیند ترکیب، مقدار بهینه آن به‌عنوان ۰,۷۸ تعیین شده است (شکل ۵). سپس برای یافتن بهترین احتمال وقوع فرایند جهش، مقادیر مختلفی از ۰,۱ تا ۱ با گام ۰,۱ مورد بررسی قرار گرفته شده است. برای ارزیابی این آزمایشات از معیارهای MSE، RMSE و ضریب همبستگی استفاده شده است (شکل ۶).

تابع MSE یکی از پرکاربردترین توابع برای ارزیابی میزان دقت نتایج تولید شده در مدل‌های رگرسیونی است. کمینه‌بودن MSE و RMSE به معنای دقت بیشتر مدل در حل مسئله می‌باشد. در عبارت ۷ فرمول MSE نشان داده شده است. n طول سیگنال PPG می‌باشد. $F_r(n)$ و $F_s(n)$ به ترتیب سیگنال اصلی PPG و سیگنال تولید شده PPG است. همین‌طور با اندازه‌گیری جذر تابع MSE، مقدار RMSE به دست خواهد آمد (عبارت ۸). ضریب همبستگی پیرسون نیز یک معیار آماری است که میزان رابطه خطی بین دو متغیر را اندازه‌گیری می‌کند که در عبارت ۹ نمایش داده شده است [۳۶].

^۱Distributed Evolutionary Algorithm Package

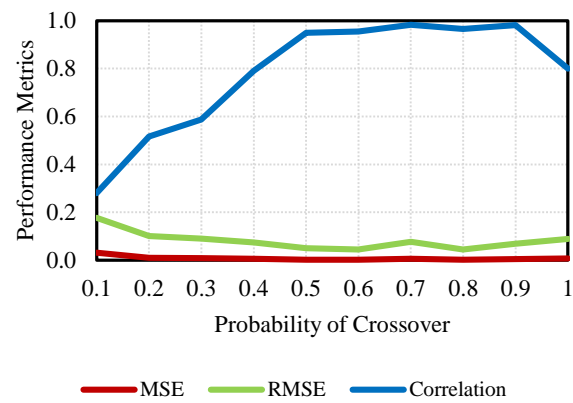
جدول ۳. تغییرات در تعداد نسل‌ها در برنامه‌نویسی ژنتیک: پس از گذشت ۱۷۰۰ نسل، بهترین نتیجه حاصل شده است.

Correlation	RMSE	MSE	تعداد نسل‌ها
۰,۶۱۸۱۰	۰,۰۹۱۶۵	۰,۰۰۸۴۰	۱۰۰
۰,۷۲۶۰۱	۰,۰۸۰۶۲	۰,۰۰۶۵۰	۲۰۰
۰,۸۱۲۹۶	۰,۰۷۱۴۱	۰,۰۰۵۱۰	۳۰۰
۰,۹۳۵۰۰	۰,۰۵۴۷۷	۰,۰۰۳۰۰	۴۰۰
۰,۹۳۸۹۲	۰,۰۵۳۰۰	۰,۰۰۲۸۰	۵۰۰
۰,۹۴۰۴۲	۰,۰۵۲۹۱	۰,۰۰۲۷۰	۶۰۰
۰,۹۲۷۲۱	۰,۰۵۵۶۷	۰,۰۰۳۱۰	۷۰۰
۰,۹۱۴۷۹	۰,۰۵۹۱۶	۰,۰۰۳۵۰	۸۰۰
۰,۹۳۷۴۹	۰,۰۵۴۸۶	۰,۰۰۳۰۱	۹۰۰
۰,۹۴۸۰۰	۰,۰۵۱۴۷	۰,۰۰۲۶۵	۱۰۰۰
۰,۹۵۷۸۷	۰,۰۵۰۴۹	۰,۰۰۲۵۵	۱۱۰۰
۰,۹۶۵۰۵	۰,۰۴۷۷۳	۰,۰۰۲۳۴	۱۲۰۰
۰,۹۸۰۰۰	۰,۰۴۴۷۲	۰,۰۰۲۰۰	۱۳۰۰
۰,۹۸۵۰۶	۰,۰۴۳۵۸	۰,۰۰۱۹۰	۱۴۰۰
۰,۹۶۷۳۴	۰,۰۳۰۱۶	۰,۰۰۰۹۱	۱۵۰۰
۰,۷۴۱۹۷	۰,۰۲۷۰۱	۰,۰۰۰۷۳	۱۶۰۰
۰,۹۹۹۹۸	۰,۰۱۰۰۰	۰,۰۰۰۱۰	۱۷۰۰
۰,۹۱۴۰۶	۰,۰۱۸۷۰	۰,۰۰۰۳۵	۱۸۰۰
۰,۹۶۸۰۱	۰,۰۱۵۸۱	۰,۰۰۰۲۵	۱۹۰۰
۰,۹۸۰۶۵	۰,۰۱۵۴۹	۰,۰۰۰۲۴	۲۰۰۰

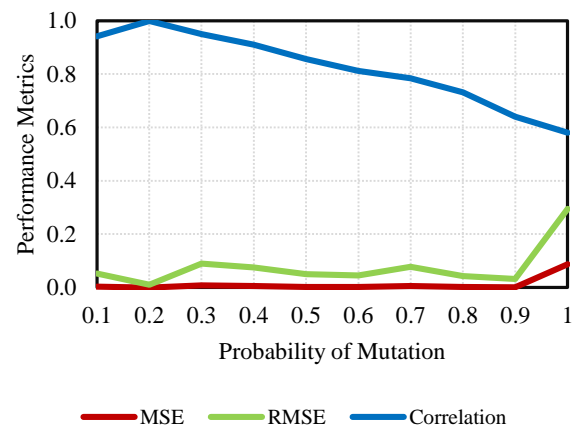
جدول ۴. تغییرات در تعداد افراد در جمعیت: تعداد ۴۰۰ فرد در جمعیت، بهترین نتایج را در میان دیگر تعدادها ارائه می‌دهد.

Correlation	RMSE	MSE	تعداد جمعیت
۰,۸۸۳۵۱	۰,۰۶۷۱۵	۰,۰۰۴۵۱	۱۰۰
۰,۹۳۵۵۵	۰,۰۵۸۷۳	۰,۰۰۳۴۵	۲۰۰
۰,۹۴۵۰۳	۰,۰۵۰۰۰	۰,۰۰۲۵۰	۳۰۰
۰,۹۹۹۹۸	۰,۰۱۰۰۰	۰,۰۰۰۱۰	۴۰۰
۰,۹۵۶۴۱	۰,۰۴۴۸۳	۰,۰۰۲۰۱	۵۰۰
۰,۹۳۶۶۸	۰,۰۵۴۸۶	۰,۰۰۳۰۱	۶۰۰
۰,۹۰۱۰۰	۰,۰۶۱۸۰	۰,۰۰۳۸۲	۷۰۰
۰,۸۹۱۲۲	۰,۰۶۷۹۹	۰,۰۰۴۶۲	۸۰۰
۰,۹۴۰۰۰	۰,۰۴۶۹۰	۰,۰۰۲۲۰	۹۰۰
۰,۹۱۸۸۸	۰,۰۵۶۹۹	۰,۰۰۳۲۴	۱۰۰۰

مدل سازنده PPG تولیدشده توسط برنامه‌نویسی ژنتیک در روش پیشنهادی در مدت‌زمان ۵۳ دقیقه ایجادشده است. با توجه به تعداد پارامترهای مدل در مجموع اندازه مدل سازنده تنها ۰,۰۶۵ مگابایت می‌باشد. مدل حاصل برخلاف مدل‌های سازنده ایجادشده توسط شبکه‌های عصبی به دلیل پیچیدگی و تعداد پارامترهای کمتر، برای دستگاه‌هایی با منابع محدود از جمله دستگاه‌های پوشیدنی بسیار مناسب است. همان‌طور که گفته شد هدف از ارائه این روش، ایجاد یک مدل سازنده سیگنال PPG مصنوعی با بهینگی مناسب می‌باشد که متعاقباً بایستی مدل دارای مقدار MSE و RMSE حداقل و ضریب همبستگی حداکثر باشد. یک سیگنال PPG مصنوعی ایده آل دارای MSE و RMSE صفر و ضریب همبستگی ۱ می‌باشد.



شکل ۵. تغییرات ارزیابی در پارامتر احتمال در فرایند تکاملی ترکیب در بازه ۰,۱ تا ۱,۰ نمودار نشان می‌دهد که مقدار احتمال فرایند ترکیب بهترین نتیجه را به دست آورده است.



شکل ۶. تغییرات ارزیابی در پارامتر احتمال در فرایند تکاملی جهش در بازه ۰,۱ تا ۱,۰ نمودار نشان می‌دهد که مقدار احتمال جهش ۰,۲ بهترین نتیجه را به دست آورده است.

در برنامه‌نویسی ژنتیک تعداد نسل‌ها و اندازه جمعیت نیز به‌عنوان پارامترهای مهم در نظر گرفته می‌شوند. در این روش GP با تعداد نسل‌های مختلف از ۱۰۰ تا ۲۰۰۰ اجراشده و نتایج ارزیابی در جدول ۳ ثبت‌گردیده است. همین‌طور در جدول ۴ نتایج تغییرات در تعداد افراد در جمعیت نشان داده‌شده است. تحلیل این آزمایشات این امکان را می‌دهد تا در بهترین تنظیمات برای GP اعمال شود.

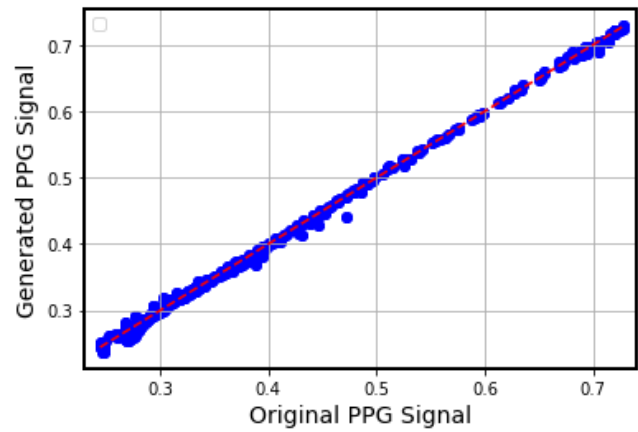
پس از تجزیه و تحلیل گام‌به‌گام نتایج بالا و انتخاب مناسب‌ترین تنظیمات، مدل نهایی سازنده سیگنال PPG مصنوعی به MSE ۰,۰۰۰۱، RMSE ۰,۰۱ و ضریب همبستگی ۰,۹۹۹ دست‌یافته است. در شکل ۷ میزان همبستگی $F_r(n)$ و $F_s(n)$ مشاهده می‌شود و در شکل ۸ میزان شباهت سیگنال تولیدشده PPG با سیگنال اصلی نشان داده‌شده است.

ترتیب، 0.001 ، 0.0316 و 0.988 است [۱۷]. این نتایج برای روش سالوشنکو و همکاران، 0.00018 ، 0.0136 و 0.930 می‌باشد [۳۷]. 0.00013 ، 0.0115 و 0.942 برای روش ارائه‌شده توسط لو و همکاران [۱۶] و 0.00025 ، 0.0159 و 0.926 نیز برای روش پیشنهادی وانگ و همکاران می‌باشد [۱۵].

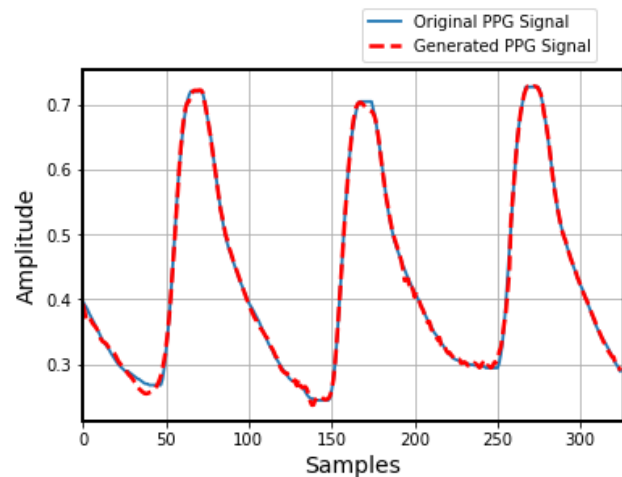
نتایج نشان می‌دهد که به دلیل آنکه مدل‌های ارائه‌شده در روش‌های موجود، با استفاده از صفات هندسی سیگنال و وابسته به دانش ریاضی افراد تخمین زده شده‌اند، تنها در یک نمونه خاص از سیگنال PPG بهتر عمل می‌کنند؛ اما با استفاده از برنامه‌نویسی ژنتیک، به صورت پویا امکان ایجاد سیگنال‌های PPG در تنوع بالا با ترکیب توابع ساده ریاضی و بدون هیچ‌گونه وابستگی امکان‌پذیر است [۳۸]. به نظر می‌رسد که به دلیل ویژگی‌های ذکر شده، سیگنال تولیدشده در روش پیشنهادی نسبت به سایر روش‌ها دارای کمترین MSE و RMSE و بیشترین مقدار ضریب همبستگی می‌باشد.

۴- جمع‌بندی

در این مقاله، با توجه به کمبود داده‌های سیگنال PPG و عدم تنوع آن‌ها در پایگاه‌های داده مرتبط، یک راهکار مبتنی بر مدل‌سازی سیگنال PPG ارائه شده است که با بهره‌گیری از برنامه‌نویسی ژنتیک، تعدادی مدل با استفاده از یک سیگنال PPG اصلی ساخته شود. در نهایت، اجرای این مدل‌ها منجر به تولید یک سیگنال PPG مصنوعی خواهد شد. استفاده از GP در این روش امکان تولید مدل‌هایی با دقت قابل قبول و بهینگی مناسب را فراهم می‌کند. مدل سازنده ارائه‌شده در این مقاله، قابلیت اجرا بر روی دستگاه‌هایی با منابع محاسباتی محدود را خواهد داشت. در این روش دو پارامتر مهم مقیاس‌پذیری و دقت، به طور ویژه‌ای مورد توجه قرار گرفته است. به جهت ارزیابی مدل ارائه‌شده، از معیارهای MSE، RMSE و ضریب همبستگی استفاده شده است. این روش نه تنها منجر به افزایش دقت مدل سازنده شده به دلیل استفاده از برنامه‌نویسی ژنتیک و ساختار درختی، حجم محاسباتی نیز بهبود یافته است. این پژوهش یک روش نوآورانه را برای تولید سیگنال PPG مشابه با سیگنال‌های واقعی ارائه می‌دهد که با بهینه‌ترین پارامترها، سیگنالی با دقت مطلوب تولید کند.



شکل ۷. میزان همبستگی سیگنال تولیدشده توسط GP با سیگنال اصلی PPG در این تصویر نشان داده شده است.



شکل ۸. موج سیگنال PPG مصنوعی و موج اصلی سیگنال که مدل بر اساس آن ایجاد شده است. در این تصویر میزان شباهت بالای سیگنال PPG تولیدشده توسط GP با سیگنال اصلی نشان داده شده است.

تاکنون روش مشابهی در جهت ایجاد مدل سازنده سیگنال PPG توسط برنامه‌نویسی ژنتیک ارائه نشده است لذا برای ایجاد دیدگاه بهتر، روش پیشنهادی با دیگر روش‌های مشابه رگرسیونی در جدول ۵، با استفاده از مجموعه داده‌های گوناگون مقایسه شده است. به عنوان نمونه در مجموعه داده MIMIC-III مقادیر MSE، RMSE و ضریب همبستگی در روش پیشنهادی توسط تانگ و همکاران به

جدول ۵. بررسی و ارزیابی سیگنال‌های PPG در مجموعه داده‌های گوناگون با استفاده از روش‌های رگرسیونی موجود.

Correlation	RMSE	MSE	روش	نویسنده	مجموعه داده
۰.۹۹۹۹۸	۰.۰۱۰۰۰	۰.۰۰۰۱۰	مبتنی بر GP	روش پیشنهادی	MIMIC III
۰.۹۸۸۳۱	۰.۰۳۱۶۰	۰.۰۰۱۰۰	مبتنی بر توابع گاوسی	تانگ و همکاران [۱۷]	
۰.۹۳۰۲۲	۰.۰۱۳۶۰	۰.۰۰۰۱۸	مبتنی بر توابع گاوسی و لگاریتم نرمال	سالوشنکو و همکاران [۳۷]	
۰.۹۴۲۰۰	۰.۰۱۱۵۰	۰.۰۰۰۱۳	مبتنی بر توابع گاوسی و لگاریتم نرمال	لو و همکاران [۱۶]	
۰.۹۲۶۰۰	۰.۰۱۵۹۰	۰.۰۰۰۲۵	مبتنی بر توابع گاوسی	وانگ و همکاران [۱۵]	
۰.۹۸۹۷۰	۰.۰۱۱۸۰	۰.۰۰۰۱۴	مبتنی بر GP	روش پیشنهادی	WristPPG
۰.۸۷۱۲۱	۰.۰۶۳۰۰	۰.۰۰۴۸۱	مبتنی بر توابع گاوسی	تانگ و همکاران [۱۷]	
۰.۷۹۳۳۳	۰.۰۷۱۴۰	۰.۰۰۵۱۱	مبتنی بر توابع گاوسی و لگاریتم نرمال	سالوشنکو و همکاران [۳۷]	
۰.۸۷۷۸۸	۰.۰۶۶۳۰	۰.۰۰۴۴۲	مبتنی بر توابع گاوسی و لگاریتم نرمال	لو و همکاران [۱۶]	
۰.۷۱۳۰۰	۰.۰۷۶۸۰	۰.۰۰۵۹۰	مبتنی بر توابع گاوسی	وانگ و همکاران [۱۵]	
۰.۹۵۲۰۰	۰.۰۱۴۴۰	۰.۰۰۰۲۱	مبتنی بر GP	روش پیشنهادی	BioSec.Lab
۰.۸۵۵۳۳	۰.۰۷۰۰۰	۰.۰۰۴۹۰	مبتنی بر توابع گاوسی	تانگ و همکاران [۱۷]	
۰.۸۶۱۲۶	۰.۰۶۸۹۲	۰.۰۰۴۷۵	مبتنی بر توابع گاوسی و لگاریتم نرمال	سالوشنکو و همکاران [۳۷]	
۰.۸۶۶۰۰	۰.۰۶۸۰۴	۰.۰۰۴۶۳	مبتنی بر توابع گاوسی و لگاریتم نرمال	لو و همکاران [۱۶]	
۰.۸۱۱۱۲	۰.۰۷۱۳۴	۰.۰۰۵۰۹	مبتنی بر توابع گاوسی	وانگ و همکاران [۱۵]	
۰.۹۷۴۳۲	۰.۰۴۰۰۰	۰.۰۰۰۱۶	مبتنی بر GP	روش پیشنهادی	PPG-BP
۰.۹۰۱۱۱	۰.۰۴۹۲۹	۰.۰۰۲۴۳	مبتنی بر توابع گاوسی	تانگ و همکاران [۱۷]	
۰.۸۵۲۲۵	۰.۰۷۰۰۰	۰.۰۰۴۹۱	مبتنی بر توابع گاوسی و لگاریتم نرمال	سالوشنکو و همکاران [۳۷]	
۰.۸۸۷۰۱	۰.۰۶۵۵۸	۰.۰۰۴۳۴	مبتنی بر توابع گاوسی و لگاریتم نرمال	لو و همکاران [۱۶]	
۰.۷۶۳۳۳	۰.۰۷۳۴۸	۰.۰۰۵۴۰	مبتنی بر توابع گاوسی	وانگ و همکاران [۱۵]	

- [3] B. Tarifi, A. Fainman, A. Pantanowitz, and D. M. Rubin, "A Machine Learning Approach to the Non-Invasive Estimation of Continuous Blood Pressure Using Photoplethysmography," *Applied Sciences*, vol. 13, no. 6, p. 3955, 2023. [Online]. Available: <https://www.mdpi.com/2076-3417/13/6/3955>.
- [4] M. A. Noor Faris Ali, "An efficient hybrid LSTM-ANN joint classification-regression model for PPG based blood pressure monitoring," *Biomedical Signal Processing and Control*, vol. 84, 2023, doi: <https://doi.org/10.1016/j.bspc.2023.104782>.
- [5] H. a. S. Asgharnezhad, Afshar and Bakhshayeshi, Ivan and Alizadehsani, Roohallah and Chamaani, Somayyeh and Alinejad-Rokny, Hamid, "Improving PPG Signal Classification with Machine Learning: The Power of a Second Opinion," 2023, pp. 1-5, doi: <https://doi.org/10.1109/DSP58604.2023.10167869>.
- [6] D. a. C. Jarchi, Alex, "Description of a Database Containing Wrist PPG Signals Recorded during Physical Exercise with Both Accelerometer and Gyroscope Measures of Motion," *Data*, vol. 2, 2016, doi: 10.3390/data2010001.
- [7] U. Yadav, S. N. Abbas, and D. Hatzinakos, "Evaluation of PPG Biometrics for Authentication in Different States," *2018 International Conference on Biometrics (ICB)*, pp. 277-282, 2017, doi: <https://doi.org/10.1109/ICB2018.2018.00049>.
- [8] Y. a. C. Liang, Zhencheng and Liu, Guiyong and Elgendi, Mohamed, "A new, short-recorded photoplethysmogram dataset for blood pressure monitoring in China," *Scientific Data*, vol. 5, p. 180020, 2018, doi: <https://doi.org/10.1038/sdata.2018.20>.
- [9] A. a. I. Reiss, Ina and Schmidt, Philip and Van Laerhoven, Kristof, "Deep PPG: Large-Scale Heart Rate Estimation with Convolutional Neural Networks," *Sensors*, vol. 19, p. 3079, 2019, doi: <https://doi.org/10.3390/s19143079>.
- [10] O. a. R. Mazumder, Dibyendu and Bhattacharya, Sakyajit and Sinha, Aniruddha and Pal, Arpan, "Synthetic PPG generation from haemodynamic model with baroreflex autoregulation: a Digital twin of cardiovascular system," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology*

در حال حاضر، استفاده از عملگرهای ساده در مدل‌های GP، از جمله +، -، *، /، sin و cos، محدودیت‌هایی در ایجاد مدل‌های پیچیده‌تر و بادقت بالا ایجاد می‌کند. در پژوهش‌های آتی، تلاش بر این است تا این دست از محدودیت‌ها را کاهش داده و مدل‌های پیچیده‌تر و مؤثرتری را ایجاد نمود. علی‌رغم ویژگی‌های مثبت برنامه‌نویسی ژنتیک در این روش، ضعف‌هایی در دقت مدل همچنان مشهود است؛ بنابراین، تلاش برای افزایش دقت این روش از اهمیت بسیاری برخوردار خواهد بود. در آینده، با استفاده از این روش می‌توان به شبیه‌سازی مدل‌های آریتمی دار نیز پرداخت. وجود انواع مدل‌های سازنده PPG منجر به ارتقای حجم و کیفیت پایگاه‌های داده مرتبط خواهد شد که همین مسئله می‌تواند اطلاعات بیشتری را در اختیار پژوهشگران قرار دهد و همچنین توانایی عملکرد مدل‌های یادگیری عمیق را افزایش دهد.

مراجع

- [1] M. Nardelli and R. Bailón, "Advances in Wearable Photoplethysmography Applications in Health Monitoring," (in eng), *Sensors (Basel)*, vol. 23, no. 16, Aug 10 2023, doi: <https://doi.org/10.3390/s23167064>.
- [2] K. B. Kim and H. J. Baek, "Photoplethysmography in Wearable Devices: A Comprehensive Review of Technological Advances, Current Challenges, and Future Directions," *Electronics*, vol. 12, no. 13, p. 2923, 2023, doi: <https://doi.org/10.3390/electronics12132923>.

- Computing*, vol. 60, pp. 447-469, 2017, doi: <https://doi.org/10.1016/j.asoc.2017.06.050>.
- [25] F. A.-M. Majid Sepahvand, "A novel multi-lead ECG personal recognition based on signals functional and structural dependencies using time-frequency representation and evolutionary morphological CNN," *Biomedical Signal Processing and Control*, vol. 68, 2021, doi: <https://doi.org/10.1016/j.bspc.2021.102766>.
- [26] m. amintoosi and H. Ezzati, "The aware genetic algorithm of the best member, applied to graph coloring and metric-dimension of the graph problems," *Journal of Information and Communication Technology*, vol. 12, no. 43, pp. 143-154, 2020. [Online]. Available: <http://rimag.ir/fa/Article/13601>.
- [27] S. Katoch, Chauhan, S.S. & Kumar, V, "A review on genetic algorithm: past, present, and future," *Multimed Tools Appl*, vol. 80, pp. 8091-8126 2021, doi: <https://doi.org/10.1007/s11042-020-10139-6>.
- [28] S. Bakhtiari and M. Khosroshahi, "A Multi-Objective Differential Evolutionary Algorithm-based Approach for Resource Allocation in Cloud Computing Environment," *Journal of Information and Communication Technology*, vol. 14, no. 51, pp. 113-128, 2022. [Online]. Available: <http://rimag.ir/fa/Article/39235>.
- [29] F. A.-M. Majid Sepahvand, "A novel representation in genetic programming for ensemble classification of human motions based on inertial signals," *Expert Systems with Applications*, vol. 185, 2021, doi: <https://doi.org/10.1016/j.eswa.2021.115624>.
- [30] L. a. J. Tan, Jean, *Image Processing Basics*. 2019.
- [31] m. nikseresht and M. Raji, "A Task Mapping and Scheduling Algorithm based on Genetic Algorithm for Embedded System Design," *Journal of Information and Communication Technology*, vol. 13, no. 47, pp. 186-197, 2021. [Online]. Available: <http://rimag.ir/fa/Article/16369>.
- [32] W. K. W. Pauline Wong, Filbert H. Juwono, Basil Andy Lease, Lenin Gopal, I.M. Chew, "Sensor abnormality detection in multistage compressor units: A "white box" approach using tree-based genetic programming," *e-Prime - Advances in Electrical Engineering, Electronics and Energy*, vol. 5, 2023, doi: <https://doi.org/10.1016/j.prime.2023.100209>.
- [33] M. P. Aghababa and a. b. baghbaderani, "Routing of Multipartite Computer Networks Using Ant Genetic Algorithm," *Journal of Information and Communication Technology*, vol. 9, no. 31, pp. 1-12, 2017. [Online]. Available: <http://rimag.ir/fa/Article/8199>.
- [34] A. E. Johnson *et al.*, "MIMIC-III, a freely accessible critical care database," (in eng), *Sci Data*, vol. 3, p. 160035, May 24 2016, doi: <https://doi.org/10.1038/sdata.2016.35>.
- [35] B. F. Jernej Zupančič, Matjaž Gams, "Genetic-programming-based multi-objective optimization of strategies for home energy-management systems," *Energy*, vol. 203, 2020, doi: <https://doi.org/10.1016/j.energy.2020.117769>.
- [36] A. Kianimajd *et al.*, "Comparison of different methods of measuring similarity in physiologic time series," *IFAC-PapersOnLine*, vol. 50, pp. 11005-11010, 2017, doi: <https://doi.org/10.1016/j.ifacol.2017.08.2479>.
- [37] A. a. P. Sološenko, Andrius and Marozas, Vaidotas and Sörmmo, Leif, "Photoplethysmogram Modeling During Paroxysmal Atrial Fibrillation," *Computers in Biology and Medicine*, vol. 81, 2017, doi: 10.22489/CinC.2017.049-011.
- [38] M. a. A.-M. Feli, Fardin, "12 lead Electrocardiography signals compression by a new genetic programming based mathematical modeling algorithm," *Biomedical Signal Processing and Control*, vol. 54, pp. 1-20, 2019, doi: 10.1016/j.bspc.2019.101596.
- Society (EMBC)*, 2019, pp. 5024-5029, doi: <https://doi.org/10.1109/EMBC.2019.8856691>.
- [11] T. W. C. Fanny, "Deep Learning for Imbalance Data Classification using Class Expert Generative Adversarial Network," *Procedia Computer Science*, vol. 135, pp. 60-67, 2018, doi: <https://doi.org/10.1016/j.procs.2018.08.150>.
- [12] J. M. Johnson, Khoshgoftaar, T.M, "Survey on deep learning with class imbalance," *Journal of Big Data*, vol. 6, 2019, doi: <https://doi.org/10.1186/s40537-019-0192-5>.
- [13] O. a. C. Sayadi, Gari, "Synthetic ECG Generation and Bayesian Filtering Using a Gaussian Wave-Based Dynamical Model," *Physiological measurement*, vol. 31, pp. 1309-29, 2010, doi: <https://doi.org/10.1088/0967-3334/31/10/002>.
- [14] D. Martin-Martinez, P. Casaseca-de-la-Higuera, M. Martin-Fernandez, and C. Alberola-Lopez, "Stochastic modeling of the PPG signal: a synthesis-by-analysis approach with applications," (in eng), *IEEE Trans Biomed Eng*, vol. 60, no. 9, pp. 2432-41, Sep 2013, doi: <https://doi.org/10.1109/TBME.2013.2257770>.
- [15] L. Wang, L. Xu, S. Feng, M. Q. Meng, and K. Wang, "Multi-Gaussian fitting for pulse waveform using Weighted Least Squares and multi-criteria decision making method," (in eng), *Comput Biol Med*, vol. 43, no. 11, pp. 1661-72, Nov 2013, doi: <https://doi.org/10.1016/j.compbiomed.2013.08.004>.
- [16] Y. Luo, W. Li, W. Rao, X. Fu, L. Yang, and Y. Zhang, "A New Modeling Method of Photoplethysmography Signal Based on Lognormal Basis," in *International Conference on Internet and Distributed Computing Systems*, 2016.
- [17] Q. Tang, Z. Chen, R. Ward, and M. Elgendi, "Synthetic photoplethysmogram generation using two Gaussian functions," (in eng), *Sci Rep*, vol. 10, no. 1, p. 13883, Aug 17 2020, doi: <https://doi.org/10.1038/s41598-020-69076-x>.
- [18] M. M. Alankrita Aggarwal, Gopi Battineni, "Generative adversarial network: An overview of theory and applications," *International Journal of Information Management Data Insights*, vol. 1, no. 1, 2021, doi: <https://doi.org/10.1016/j.ijimei.2020.100004>.
- [19] E. Brophy, Z. Wang, and T. E. Ward, "Quick and Easy Time Series Generation with Established Image-based GANs," *ArXiv*, vol. abs/1902.05624, 2019, doi: <https://doi.org/10.48550/arXiv.1902.05624>.
- [20] D. a. T. Kiyasseh, Girmaw Abebe and Le Nguyen Thanh, Nhan and Van, Tan and Thwaites, Louise and Zhu, Tingting and Clifton, David, "PlethAugment: GAN-Based PPG Augmentation for Medical Diagnosis in Low-Resource Settings," *IEEE Journal of Biomedical and Health Informatics*, vol. PP, pp. 1-1, 2020, doi: 10.1109/JBHI.2020.2979608.
- [21] J. M. L. Alcaraz and N. Strodthoff, "Diffusion-based conditional ECG generation with structured state space models," (in eng), *Comput Biol Med*, vol. 163, p. 107115, Sep 2023, doi: <https://doi.org/10.1016/j.compbiomed.2023.107115>.
- [22] M. A. Quiroz-Juárez, O. Jiménez-Ramírez, R. Vázquez-Medina, V. F. Breña-medina, J. L. Aragón, and R. A. Barrio, "Generation of ECG signals from a reaction-diffusion model spatially discretized," *Scientific Reports*, vol. 9, 2019, doi: <https://doi.org/10.1038/s41598-019-55448-5>.
- [23] Q. Zhang, D. Zhou, and X. Zeng, "Highly wearable cuff-less blood pressure and heart rate monitoring with single-arm electrocardiogram and photoplethysmogram signals," (in eng), *Biomed Eng Online*, vol. 16, no. 1, p. 23, Feb 6 2017, doi: <https://doi.org/10.1186/s12938-017-0317-z>.
- [24] M. M. E. Maryam Amir Haeri, Gianluigi Folino, "Statistical genetic programming for symbolic regression," *Applied Soft*