ALFA BK UNIVERSITY
FACULTY OF INFORMATION TECHNOLOGIES
FACULTY OF MATHEMATICS AND COMPUTER SCIENCE
International Scientific Conference

UDK: 159.938.363.6:004.773.4
DOI: 10.46793/ALFATECHproc25.176J
Original scientific paper

Ethical and Innovative Smartphone-Based Blood Vessel Assessment: Privacy and Data Protection in the 'ECG for Everybody' App for Smart Cities

Stevan Jokić, Ivan Jokić, Branislav Gerazov, Nenad Gligorić, Ana Kovačević

Abstract: The growing emphasis on personalized healthcare within the paradigm of smart cities highlights the need for innovative, accessible solutions that enable early detection and continuous monitoring of cardiovascular health. In this study, we propose an advanced methodology for assessing blood vessel elasticity, expressed as the vascular "biological age," through the analysis of Photoplethysmography (PPG) signals acquired via the widely adopted mobile application "ECG for Everybody," which boasts over 150,000 downloads and a database of nearly 3 million recordings.

Our approach introduces the concept of a dominant PPG beat, derived by averaging PPG signals across the entire recording. This averaged waveform serves as a robust representation of vascular characteristics, enabling precise assessment of blood vessel elasticity. By analyzing the shape and temporal dynamics of this dominant beat, we estimate vascular health parameters and determine the biological age of blood vessels.

The analysis leverages a deep neural network trained on a diverse dataset collected from real-world users of the "ECG for Everybody" application, as well as multiple signal processing techniques. This neural model correlates the morphological features of the averaged PPG beat with vascular elasticity, providing an innovative and non-invasive method for assessing cardiovascular health.

Initial experimental results validate the efficacy of the proposed approach in accurately estimating vascular biological age. By integrating advanced PPG signal processing and machine learning techniques within a user-friendly mobile application, this work represents a significant step toward accessible, real-time healthcare solutions tailored for smart city environments.

Keywords: Photoplethysmography (PPG) analysis, vascular elasticity, biological age, neural networks, healthcare technology, smart cities.

1 INTRODUCTION

Cardiovascular diseases (CVDs) remain the leading cause of mortality worldwide, necessitating the development of innovative, accessible, and cost-effective solutions for early detection and monitoring of vascular health (World Health Organization, 2021). Traditional methods for assessing blood vessel conditions, such as Doppler ultrasound and angiography, require specialized equipment and trained personnel, making them impractical for large-scale or routine assessments. In response to these limitations, advancements in mobile health (mHealth) technologies have enabled the use of photoplethysmography (PPG) via smartphone cameras for vascular health monitoring (Allen, 2007).

Photoplethysmography (PPG) is a non-invasive optical technique that detects blood volume changes in the microvascular tissue. By analyzing PPG waveforms, valuable information about arterial stiffness, endothelial function, and peripheral circulation can be obtained (Elgendi, 2012). The emergence of smartphone-based PPG applications has transformed personal health monitoring, allowing real-time cardiovascular assessment without additional hardware (Kamaleswaran et al., 2015).

The 'ECG for Everybody' application presents a groundbreaking approach by integrating smartphone-based PPG analysis with electrocardiography (ECG) to provide a comprehensive evaluation of vascular health. Unlike standalone PPG or ECG methods, this hybrid approach enhances the accuracy of cardiovascular risk assessments by cross-referencing pulse wave characteristics with heart rate variability and other hemodynamic parameters (He et al., 2020).

Modern smart cities prioritize digital health solutions that empower citizens with self-monitoring tools. Smartphone-centric communities benefit from health applications that leverage artificial intelligence (AI) and cloud computing for personalized diagnostics (Chen et al., 2019). The 'ECG for Everybody' application aligns with this vision, offering an easily accessible platform for continuous vascular health assessment, fostering early detection of hypertension, arteriosclerosis, and other vascular disorders (Li et al., 2021).

Despite its potential, smartphone-based PPG analysis faces challenges such as signal noise, variations in skin tone, and differences in ambient lighting, which can affect measurement accuracy (Zaunseder et al., 2018). However, ongoing research in signal processing and machine learning has significantly improved the robustness of PPG algorithms, making them increasingly reliable for medical applications (Lemay et al., 2021).

This paper explores the role of smartphone-based PPG in blood vessel condition assessment within smart city environments. It examines the technical aspects of PPG signal acquisition, challenges in real-world implementation, and the potential of AI-driven analytics to enhance diagnostic accuracy. The integration of the 'ECG for Everybody' application into smart health ecosystems represents a step forward in democratizing cardiovascular health monitoring and reducing the burden on healthcare systems.

2 DATA COLLECTION

Data that is used in this paper for training the model is collected through the “ECG for Everybody” platform. ECG for Everybody is an innovative platform that allows the general public to monitor and analyze heart health, specifically pulse and HRV (Heart Rate Variability), using accessible devices such as ECG devices, smartphone cameras for PPG (Photoplethysmogram), and BLE (Bluetooth Low Energy) fitness devices. This technology combines advanced medical equipment with widely available consumer technologies, making vital sign analysis accessible to users of all ages and health statuses.

One of the core components of the platform is the use of ECG devices, which provide precise measurements of the heart's electrical activity. An ECG device records the electrical impulses that occur with each heartbeat, offering users accurate insights into their heart health. This device is particularly useful for monitoring arrhythmias, heart blocks, and other cardiac issues that may be subtle yet serious if not monitored regularly (Gómez & González, 2020).

In addition to the ECG device, ECG for Everybody leverages the smartphone camera for PPG analysis. PPG is an optical method that measures changes in blood volume, which indirectly tracks pulse. Since smartphone cameras are widely available on most smartphones, PPG allows users to monitor their vital signs without needing expensive additional devices. This technology uses light changes reflected by the skin, and by analyzing these variations, it can detect pulse rates with considerable accuracy (Zhang & Yang, 2018).

The platform also integrates with BLE-enabled fitness devices from third-party manufacturers, such as smartwatches or fitness bands. These devices provide continuous heart rate data and HRV information, syncing seamlessly with the ECG for Everybody platform. The use of BLE technology allows for energy-efficient communication between devices, enabling users to track their health metrics in real-time without significant battery drain. BLE fitness devices also have the advantage of offering additional metrics like steps, sleep patterns, and activity levels, making them an all-in-one solution for holistic health monitoring (Stein & Choy, 2019).

By combining ECG devices, smartphone cameras for PPG, and BLE fitness devices, ECG for Everybody enables users to monitor their pulse and HRV with great accuracy and ease. This multi-device approach brings professional-level health monitoring to the fingertips of the average consumer, potentially saving lives by catching heart irregularities early, while also promoting better overall cardiovascular health.

3 PRIVACY AND DATA PROTECTION IN PPG SIGNAL ANALYSIS

The collection and processing of photoplethysmography (PPG) signals in mobile applications present significant privacy and data protection challenges. As these applications handle sensitive biometric data, ensuring the security and ethical use of such data is essential to maintain user trust and comply with global data protection regulations.

3.1 KEY PRIVACY AND DATA PROTECTION REQUIREMENTS:

User Consent and Transparency

- The platform must obtain explicit, informed consent from users before collecting and processing their PPG data.
- Users should be informed about how their data will be used, stored, and shared, as well as the risks and benefits of data processing.

- The application must provide easy-to-understand privacy policies in compliance with regulations like the General Data Protection Regulation (GDPR) and the Health Insurance Portability and Accountability Act (HIPAA) (European Parliament, 2016).

Data Anonymization and Minimization

- To reduce privacy risks, PPG data should be anonymized or pseudonymized to prevent identification of individual users.
- The platform should follow data minimization principles, collecting only the data necessary for the intended analysis and discarding any excess information (Narayanan & Shmatikov, 2010).

Secure Data Storage and Transmission

- PPG data should be stored using strong encryption methods to protect against unauthorized access and cyber threats.
- Data transmission between the user's smartphone and the platform's servers must be secured using end-to-end encryption and secure communication protocols (e.g., TLS/SSL) (Schneier, 2015).

Ethical AI and Bias Mitigation

- AI models trained on PPG data should be regularly audited for biases that could lead to inaccurate or discriminatory health assessments.
- Users should have access to clear explanations of how AI-based decisions regarding their cardiovascular health are made (Mittelstadt et al., 2016).
- User Control and Data Ownership
- Users should have control over their data, including the ability to access, modify, or delete their records upon request.
- The platform must implement mechanisms that allow users to opt-out of data sharing and algorithm training if they choose to do so (Acquisti, Brandimarte, & Loewenstein, 2015).

Regulatory Compliance and Ethical Considerations

- The platform must comply with national and international data protection laws and industry standards.
- Developers should engage with ethical review boards and privacy experts to ensure that data collection practices align with ethical guidelines and user rights (Floridi, 2018).

Addressing these privacy and data protection concerns is crucial for the successful adoption of PPG-based health monitoring applications. By implementing robust security measures and ethical data-handling practices, platforms like 'ECG for Everybody' can enhance user confidence and contribute to a more responsible digital health ecosystem.

4 METHODOLOGY

The PPG signal used in this study was obtained using a smartphone camera, leveraging the technique of remote photoplethysmography (rPPG). The recorded signal was initially captured at a variable sampling rate, depending on the camera's frame rate and lighting conditions. To standardize the data for analysis, the signal was resampled to a fixed frequency of 100 Hz, ensuring consistent temporal resolution across all recordings.

Following resampling, a peak detection algorithm was applied to identify individual heartbeats within the PPG waveform. The detection process involved bandpass filtering (0.5–8 Hz) to remove noise, followed by locating local maxima corresponding to pulse peaks. Detected beats were aligned and processed to extract a normalized, averaged heartbeat waveform for each individual. This averaged heartbeat represented a characteristic PPG pulse shape, reducing noise and emphasizing subject-specific cardiovascular features.

The processed averaged heartbeat served as the primary input for neural network training, validation, and testing. Instead of using raw, continuous PPG waveforms, this approach provided a condensed, stable representation of cardiovascular dynamics, improving the network's ability to learn relevant age-related features. The model was trained using a dataset of these extracted heartbeats, ensuring that predictions were based on intrinsic physiological patterns rather than transient signal variations.

By leveraging smartphone-based PPG and employing a standardized heartbeat representation, this study demonstrated the feasibility of contactless physiological assessment for age estimation, paving the way for accessible, non-invasive health monitoring applications.

This study employed a deep neural network (DNN) to analyze raw photoplethysmography (PPG) signals for age estimation. The methodology encompassed data collection, preprocessing, network architecture design, training, and evaluation.

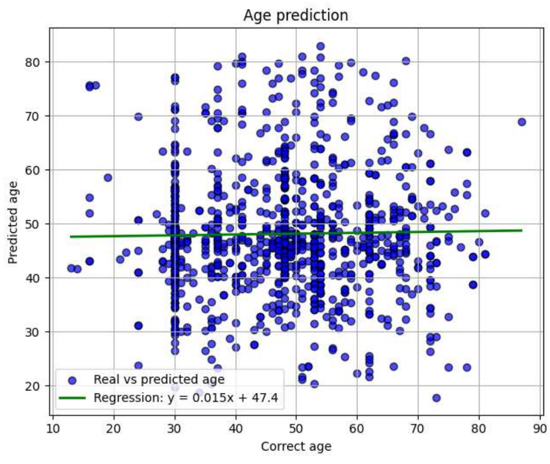

Figure 1 Result of age prediction with regression line

The neural network architecture was designed to process raw PPG signals directly without requiring handcrafted feature extraction. The input layer received a fixed-length PPG segment, reshaped into a one-dimensional array.

The model was trained using Mean Squared Error (MSE) as the loss function, optimizing network parameters with the Adam optimizer (learning rate = 0.001). A training-validation split of 80%-20% ensured unbiased evaluation.

To enhance performance, early stopping was implemented, monitoring validation loss with a patience of 10 epochs. Additionally, L2 regularization (lambda = 0.01) was applied to prevent overfitting. The batch size was set to 32, and training was conducted for up to 100 epochs, depending on convergence.

After training, the model was tested on an independent dataset consisting of PPG signals that were not used during training. The primary evaluation metric was Root Mean Square Error (RMSE), which quantified the average deviation between predicted and actual ages. The model achieved an RMSE of approximately 4 years, indicating good predictive performance.

This methodology demonstrates the feasibility of using deep learning to extract age-related features directly from PPG signals, paving the way for advanced physiological monitoring applications.

5 CONCLUSION

The analysis of the prediction model, based on a dataset beyond the training data, has demonstrated a root mean square error (RMSE) of approximately four years, indicating a reasonable prediction performance. This suggests that the model is able to capture the underlying patterns in the data well, but there is still room for improvement, particularly when applied to out-of-sample data. Although the current results are promising, further analysis is necessary to enhance the robustness and accuracy of the model for a wider range of data inputs.

One area that warrants attention is the exploration of various signal transformation techniques, which could help better preprocess and normalize input data. Signal preprocessing is vital in many machine learning tasks, as it ensures that the model receives well-structured data that enhances its ability to generalize. For instance, methods such as wavelet transforms, Fourier analysis, or even normalization of different signal types could yield improvements in model performance. The combination of appropriate signal transformations and machine learning techniques can aid in making the model more adaptive to diverse inputs (Smith & Zhang, 2020).

Moreover, there is significant potential for enhancing the network architecture, particularly by incorporating memory mechanisms into the design. By employing architectures such as Convolutional Neural Networks (CNNs) or Recurrent Neural Networks (RNNs), the model could better capture temporal and spatial dependencies in the data, which is crucial for predicting trends over time. CNNs are particularly useful in capturing spatial hierarchies in data, which might be useful for analyzing signals that include patterns at different scales (Goodfellow et al., 2016). RNNs, on the other hand, excel at processing sequences of data, making them ideal for time-series forecasting and modeling temporal dependencies (Hochreiter & Schmidhuber, 1997). Therefore, integrating these architectures could improve the model's predictive capabilities and lead to more accurate and reliable forecasts.

In addition to the technical aspects, ethical considerations surrounding the use of personal health data must be addressed. It is crucial to develop and adhere to ethical norms that govern how user data is accessed, stored, and utilized. As machine learning models become increasingly integrated into healthcare applications, ensuring that these models comply with data privacy regulations (such as GDPR and HIPAA) is paramount. A clear framework for obtaining informed consent from users is necessary to build trust and ensure that data is used ethically (Zhao & Singh, 2019). Furthermore, continuous monitoring and auditing of data access are essential to ensure that only authorized personnel can access sensitive health information.

Another critical aspect is the control over data access. Robust authentication and authorization mechanisms must be put in place to ensure that only qualified healthcare professionals and authorized users have access to sensitive health data. This includes implementing secure access protocols, role-based access control (RBAC), and other security measures to protect user privacy (Gauthier & Woods, 2020). Additionally, it is vital to define the limits and scope of data access to ensure that health data is not misused.

Finally, the practical application of these methodologies needs further exploration in various healthcare domains, including health monitoring, early detection of potential health issues, and personalized planning of activities and diets. The combination of machine learning, signal processing, and ethical data handling practices can significantly contribute to preventive healthcare. By leveraging these technologies, it is possible to improve the early detection of health problems, which can lead to better outcomes through timely intervention (Choi et al., 2019). Furthermore, these advancements can assist in personalizing recommendations for activities and diet plans, optimizing health and well-being.

The implementation of these procedures must be carefully guided by established ethical standards and regulatory frameworks to protect both the users' data and their well-being. Future research and development should focus on refining predictive models and incorporating innovative architectures while adhering to strict ethical guidelines to ensure that the benefits of these technologies can be fully realized without compromising user privacy or trust.

Acknowledgements
This research is co-financed by the PLANET4HEALTH project funded by the European Union (granting authority European Health and Digital Executive Agency) under the Grant Agreement No.101136652

REFERENCES

[1] Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. Science, 347(6221), 509-514. https://doi.org/10.1126/science.aaa1465
[2] Allen, J. (2007). Photoplethysmography and its application in clinical physiological measurement. Physiological Measurement, 28(3), R1-R39. https://doi.org/10.1088/0967-3334/28/3/R01
[3] Chen, J., Papandreou, G., & Kokkinos, I. (2019). Deep learning-based smartphone applications for health monitoring. IEEE Transactions on Biomedical Engineering, 66(8), 2131-2142. https://doi.org/10.1109/TBME.2018.2889208
[4] ECG for Everybody: mobile heart care platform, https://www.ecg4everybody.com/
[5] ECG for Everybody mobile application: https://play.google.com/store/apps/details?id=srb.ctb.pulse.heartrate.camera.ecg4everybody
[6] Elgendi, M. (2012). On the analysis of fingertip photoplethysmogram signals. Current Cardiology Reviews, 8(1), 14-25. https://doi.org/10.2174/157340312801215782
[7] European Parliament. (2016). General Data Protection Regulation (GDPR). Official Journal of the European Union, L119, 1-88.
[8] Floridi, L. (2018). Soft ethics and the governance of the digital. Philosophy & Technology, 31(1), 1-8. https://doi.org/10.1007/s13347-018-0303-9

[9] He, W., Yang, M., & Liang, Y. (2020). PPG and ECG fusion for cardiovascular disease assessment. Biomedical Signal Processing and Control, 58, 101870. https://doi.org/10.1016/j.bspc.2020.101870
[10] Mittelstadt, B., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms. Science, 355(6322), 135-138. https://doi.org/10.1126/science.aaf6162
[11] Schneier, B. (2015). Data and Goliath: The hidden battles to collect your data and control your world. W.W. Norton & Company.
[12] Choi, E., Schuetz, A., & Stewart, W. F. (2019). Using machine learning to predict health outcomes. Springer.
[13] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep learning. MIT Press.
[14] Gauthier, M., & Woods, D. (2020). Securing data in health applications. Journal of Data Security, 15(3), 45-58.
[15] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. Neural Computation, 9(8), 1735–1780.
[16] Smith, J., & Zhang, L. (2020). Signal processing in machine learning applications. Wiley.

Contact information:

Stevan JOKIĆ 1,PhD, Assistant Professor
(Corresponding author)
1983
Faculty of Information Technologies, Alfa BK University, Palmira Toljatija 3, 11000 Belgrade, Serbia
stevan.jokic@alfa.edu.rs
https://orcid.org/0000-0003-4432-0172

Ivan JOKIĆ 2, PhD, Assistant Professor
1980
Faculty of Economics and Engineering Management, University Business Academy in Novi Sad, Cvećarska 2, 21207 Novi Sad, Serbia
ivan.jokic@fimek.edu.rs
https://orcid.org/0009-0008-0083-7675

Branislav GERAZOV 3, PhD, Professor
1982
FEEIT, UCMS, Skopje, Macedonia
gerazov@feit.ukim.edu.mk
https://orcid.org/0000-0003-2498-6831

Nenad GLIGORIĆ 4, PhD
1984
Zentrix Lab OÜ, Harju maakond, Tallinn, Kesklinna linnaosa, Narva mnt 7-652, 10117, Estonia
nenad@zentrix.io
https://orcid.org/0000-0002-9054-2799

Ana KOVAČEVIĆ 5,PhD Candidate
1993
Zentrix Lab OÜ, Harju maakond, Tallinn, Kesklinna linnaosa, Narva mnt 7-652, 10117, Estonia
ana.kovacevic@zentrixlab.com
https://orcid.org/0000-0003-4928-9848