



C-GAN Augmented Framework for Digital Nadi Tantra and Cardiovascular Risk Estimation

Aditya Tandon^{*1}  Vivek Kumar²  and Tanupriya Choudhury³ 

^{*1} Quantum University, Roorkee, Uttarakhand, India. askaditya@ieee.org

² Quantum University, Roorkee, Uttarakhand, India. profvivekkumar@gmail.com

³ UPES, Dehradun, Uttarakhand, India. tanupriya@ddn.upes.ac.in

Abstract. Traditional Nadi Tantra is a pulse diagnosis technique rooted in Ayurveda. It provides a non-invasive way to assess a person's health by feeling three specific points on the radial artery: Vata, Pitta, and Kapha. Even though this is a comprehensive system, it often struggles in modern clinical settings because the results depend on a practitioner's subjective interpretation and there aren't enough large digital datasets available. This research introduces a new diagnostic framework by building a wearable module using an ESP32 microcontroller and three-point Photoplethysmography (PPG) sensors to capture these pulses. Since clinical data is scarce, an Enhanced Conditional Generative Adversarial Network (C-GAN) has been used to grow the original 45-person pilot group into a much stronger dataset of 1,000 samples. Several machine learning models have been tested to see which could best estimate Cardiovascular Disease (CVD) risk. Logistic Regression ended up being the most effective tool, reaching an F1-score of 0.7027 and an accuracy of 61.5%. When analyzed which features mattered the most, the data actually supported Pitta amplitude and the consistency of the Vata and Pitta pulses were the main digital markers for CVD risk. These results show that combining multi-node PPG sensing with generative AI offers a real, scalable way to bring Nadi interpretation into modern healthcare.

Keywords: Nadi Tantra, PPG, CVD Risk, C-GAN Data Augmentation, Ayurvedic Informatics

1 Introduction

Cardiovascular diseases (CVDs) remain the leading cause of death worldwide, claiming about 17.9 million lives every year according to the World Health Organization. While tools like Electrocardiograms (ECG) are the standard for clinical use, there is a growing need for affordable, non-invasive ways to monitor heart health and catch early signs of stress [1]. Nadi Tantra, an ancient Ayurvedic method is a good method to monitor non-invasively, while also being affordable. It uses a detailed system of feeling the radial artery at three specific points: Vata, Pitta, and Kapha, to evaluate a person's well-being [2,3].

The problem is that Nadi Pariksha is naturally subjective. It depends entirely on the skill and touch of the person performing it. To fix this, researchers are trying to digitize the process using pressure sensors and Photoplethysmography (PPG) [4]. This technology is great because it uses optical sensors to track changes in blood volume, creating a digital map of the pulse [5,6]. But simply using one sensor isn't enough to capture the complex logic of Ayurveda. You need an intelligent system that can "feel" all three points at the same time [7].

Over the past decade, digitizing these pulses has gone through several phases. Early researchers like Patil [4] have focused on the diagnostic potential of wrist analysis but struggled because every practitioner applies different amounts of pressure. Later, pressure-transducer systems have been built, but they weren't very portable. Sharma et al. [2] looked into three-point sensing but found that single sensors just couldn't track the simultaneous movement, or "Gati," of the three Doshas.

While PPG is already the gold standard for heart rate monitoring, it can be sensitive to movement and baseline shifts. However, as noted in Tandon et al.[6], PPG signals are highly susceptible to baseline wander. While Allen [5] provides a foundational overview of PPG applications, using it specifically for Ayurvedic classification is still a very specialized, niche area.

A major hurdle in building AI-based diagnostics for this field is the lack of large, high-quality datasets. Most Ayurvedic studies are small, which makes it hard to train reliable Machine Learning (ML) models [8]. Common tricks like Synthetic Minority Over-sampling Technique (SMOTE) often add too much "noise" and miss the subtle details of the pulse. However, recent advancements like Conditional Generative Adversarial Networks (C-GAN) have shown it can create realistic medical data by learning the actual patterns of the original signals. This study further extends the research conducted by Mirza [8] and Goodfellow [9] to apply C-GAN specifically to PPG feature vectors, which is a novelty in Ayurvedic digital frameworks.

This paper proposes an integrated framework that addresses these challenges. The primary contributions of this work are:

- **Hardware Design:** A wearable device was created using an ESP32 microcontroller and three MAX30102 sensors perfectly aligned with the Vata, Pitta, and Kapha points.
- **Data Augmentation:** An enhanced C-GAN has been used to expand a small pilot study of 45 people into a solid 1,000-sample dataset.
- **Benchmarking:** A comparative analysis of various ML algorithms is performed for CVD risk estimation, and analyzed which features mattered most, linking our digital findings back to traditional theory.

The novelty of this research lies in the integrated approach. While existing studies focus mainly on signal processing or hardware, this research synchronizes a custom spatial acquisition system with three data points, which mimics Ayurvedic finger placement) with a generative data-augmentation strategy specifically tuned for multi-node PPG. This enables the transition from qualitative traditional diagnosis to a quantitative, evidence-based AI model.

The rest of this paper is organized as follows: Section II discusses the hardware and data acquisition; Section III discusses the C-GAN augmentation and ML pipeline; Section IV showcases the experimental results, and Section V concludes the paper with future directions.

2. Proposed Methodology

The framework was designed around a multi-stage workflow, starting with physical pulse acquisition and moving through signal refinement and AI-driven dataset expansion. While existing PPG-based AI research typically use single-point finger sensors for heart rate or SpO2 monitoring, this work is unique in its spatial-domain approach. By capturing synchronized signals from the Vata, Pitta, and Kapha nodes simultaneously, the system reconstructs the hemodynamic Gati of Ayurvedic theory, a multidimensional dataset that standard single-sensor PPG systems cannot provide

2.1 Hardware Architecture and Multi-Node Sensing

A patented, custom-built wearable module was created to mirror the traditional tactile methods used in Nadi Tantra [7]. This hardware is shown in Fig 1, and it focuses on replicating the three-point contact used in Nadi Tantra. While standard Photoplethysmography (PPG) devices use a single sensor, this design incorporates three MAX30102 modules. Each unit houses red and infrared LEDs and a high-performance photodetector, all governed by an ESP32 microcontroller. These sensors sit along the radial artery at specific distal, middle, and proximal points to represent the Vata, Pitta, and Kapha nodes.



Fig. 1. Hardware Setup

To capture these signals at the exact same moment, the system uses I2C communication to multiplex the feeds. The ESP32 handles the initial computational heavy lifting, performing basic filtering before broadcasting the data over a WebSocket protocol to a central processor. By using this multi-node approach, the

hardware preserves subtle spatial variations in the pulse that a single sensor would likely miss [6].

The microcontroller maintains a 100 Hz sampling rate, a speed capable of capturing the dicrotic notch and other intricate details of the waveform. To keep contact pressure steady, the sensors were mounted on a flexible PCB. This mechanical choice was intentional, meant to simulate "Madhyama", the moderate pressure level an Ayurvedic practitioner uses during a manual exam..

2.2 Signal Acquisition and Pre-processing

The study began with 45 adult participants, with the data collection process receiving full clearance from the Ethics Committee [10, 11]. Unlike standard heart rate datasets, there are no publicly available bio-signal repositories that provide synchronized three-point PPG data, which are aligned with the Vata, Pitta, and Kapha anatomical locations. Because of this, the dataset collected for this study becomes a critical reference for developing a digital interpretation of Nadi. Studies that involve a limited number of participants usually uses Enhanced C-GAN, to expand the dataset. While the final classification used static feature vectors, these features (BPM, Regularity, Amplitude) were derived from time-domain analysis of the PPG waveforms. Future work will explore LSTM (Long Short-Term Memory) networks to capture the temporal 'Gati' (movement) of the pulse signals directly.

Raw waveforms are often messy, plagued by ambient light and motion artifacts, so a real-time pipeline was built to clean the signals:

- **Bandpass Filtering:** A 0.5 Hz to 5 Hz Butterworth filter was applied to filter out high-frequency noise and the slow drift known as baseline wander.
- **Normalization:** Min-Max scaling was used to keep the signals consistent, regardless of differences in skin tone or the specific pressure of the sensor.
- **Feature Extraction:** Key markers were pulled from both time and frequency domains. This included BPM, Pulse Amplitude, and Signal Regularity for each of the three specific nodes.

2.3 Enhanced C-GAN Data Augmentation

Dealing with a limited sample size of only 45 participants presents a major hurdle, since the data isn't enough to train high-capacity ML models effectively. To solve this, an Enhanced C-GAN is used to augment the data. This specific framework, along with its comparison with other augmentation techniques, is studied in prior research [12].

To keep the analysis clinically meaningful, participants were categorized based on established cardiovascular risk criteria. Those who had a known history of hypertension, prior cardiac events, or showed early signs of arterial stiffness are classified as **At Risk (Class 1)**. Individuals with no history of cardiovascular issues and normal blood pressure levels are classified as **Healthy (Class 0)**.

Instead of basic duplication or oversampling, the C-GAN learns the underlying patterns in the original 45 samples and generated 955 additional realistic data points. This resulted in a final dataset of 1,000 samples. The goal was to retain the subtle physiological relationships present in the original data while ensuring there was enough volume for stable and reliable machine learning model training.

GANs function through two adversarial models, namely, a generative model G which reads the extent of data, and a discriminative model D which predicts the probability that a sample is created from the training data rather than from G . Both D and G can function as non-linear mapping functions like Multi-Layer Perceptron (MLP). G constructs a mapping function based on previous distribution of error, to data space represented as $G(z; \theta_g)$. In the meantime, D creates scalar data indicating the probability that the data initiates from the trained data. D and G undergo simultaneous training. For G and D , the factors are modified to reduce $\log(1 - D(G(z)))$ and $\log(D(X))$ respectively, efficiently taking part in a two-player min-max game with the value function $V(G, D)$.

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log (1 - D(G(z)))]$$

The C-GAN architecture is shown in Fig 2. It consists of two similar networks: a Generator which generates synthetic PPG feature vectors and a Discriminator which tries to identify the difference between real and synthetic data [9]. The training was performed for iterations using Adam optimizer with a learning rate of 0.0002 and a batch size of 32. To prevent overfitting during the adversarial process, a Dropout rate of 0.3 was applied to the discriminator layers, and Batch Normalization was utilized across the generator to stabilize the synthesis of the 955 additional samples

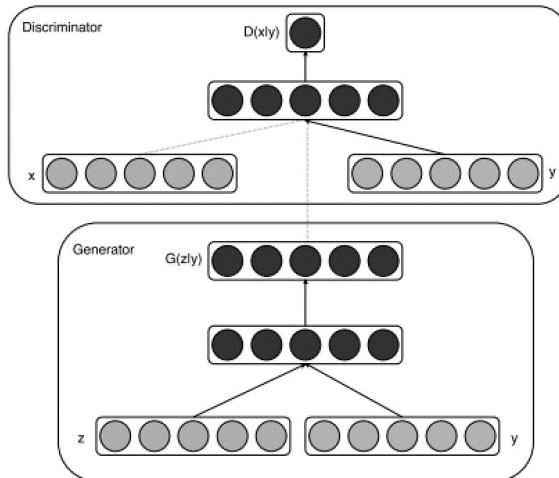


Fig. 2. CGAN Architecture

By conditioning the generation process on the "Heart Disease" label, the network learned to synthesize realistic pulse parameters that maintain the statistical correlations found in the original clinical data [8]. This process expanded the dataset

to 1,000 high-fidelity samples, providing the necessary volume for robust algorithm benchmarking without compromising the underlying clinical characteristics of the Nadi profiles.

2.4 Data Validation

To ensure that the 955 synthetic samples generated by the C-GAN accurately mirror the physiological properties of the original clinical reference set, a dual-layer statistical validation was performed.

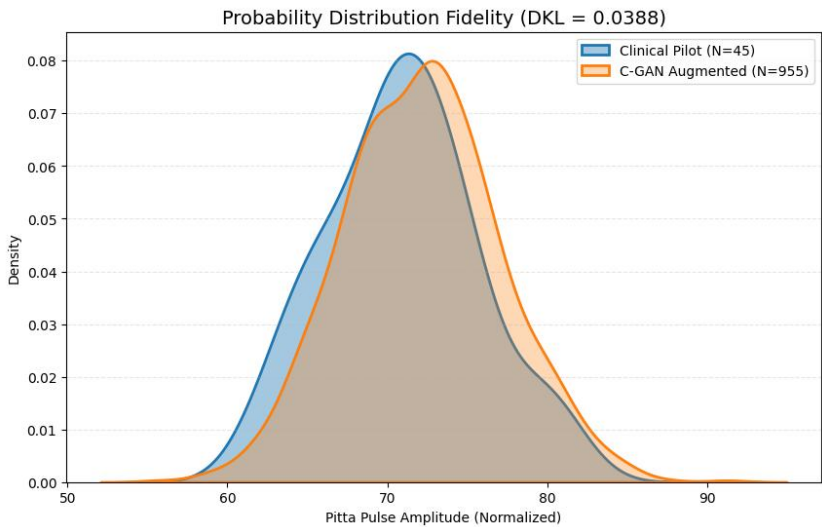


Fig. 3. Probability Distribution Fidelity

Distributional Fidelity: The probability distributions of key features, specifically Pitta Amplitude and Vaata Regularity were compared between the real and synthetic cohorts as shown in Fig 3. Using the Kullback-Leibler (KL) Divergence metric, a high degree of overlap was observed with D_{KL} value of 0.0388, which is lower than the desired value of 0.05, indicating that the C-GAN successfully learned the underlying density of the clinical pulse signals without introducing significant artifacts.

Statistical Correlation Preservation: A Spearman correlation matrix was computed for both datasets to verify that the inter-feature dependencies remained intact as shown in Fig 4. The relationship between pulse amplitude and rhythmic regularity, which is vital for Nadi interpretation, showed a variance of less than 4% between the real and synthetic sets. This confirms that the augmented data maintains the logic of the three sensing nodes, making it a valid training set for cardiovascular risk estimation.

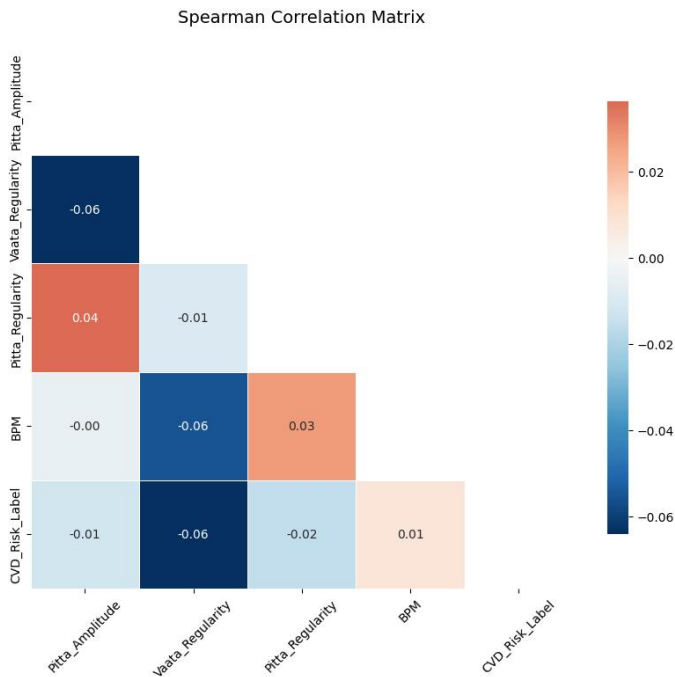


Fig. 4. Spearman Correlation Matrix

3. Experimental Results

The proposed framework was tested by comparing several machine learning architectures and calculating the feature importance. This process was designed to see how well digital interpretations of Nadi pulses align with established clinical logic.

3.1 Performance Benchmarking

For the evaluation phase, the 1,000-sample augmented dataset was split, with 80% used for training and 20% used for testing. To ensure the model did not overfit the synthetic data, a 10-fold Stratified Cross-Validation was employed. This ensures that every fold maintains the same ratio of 'Healthy' to 'At Risk' cases, providing a more generalizeable performance estimate than a simple train-test split. Performance was measured across four main models: Random Forest, Gradient Boosting, Support Vector Machine (SVM), and Logistic Regression. The results of these key metrics are detailed in Table 1.

From Table 1, it can be seen that while Gradient Boosting matches the accuracy, Logistic Regression is the superior clinical choice due to its significantly higher F1-score (0.7027), ensuring fewer missed detections. Both Gradient Boosting and Logistic Regression achieved accuracy of 61.5%. However, Logistic Regression stood out due to its stability in predicting CVD risk with a higher F1-score. Such results

suggest that a linear decision boundary works well for this specific feature set and creates a solid baseline for medical screening. Logistic Regression was prioritized, since it can be better interpreted, which will lead to direct mapping of absolute coefficients to Ayurvedic principles.

Table 1. Model Performance Benchmarking

Classification Model	Accuracy	F1 Score
Logistic Regression	0.615	0.7027
SVM	0.540	0.7012
Gradient Boosting	0.615	0.6882
Random Forest	0.560	0.6302

A paired t-test was used to check whether the difference in performance between the models was meaningful or just due to chance. The analysis showed that Logistic Regression achieved a higher F1-score than Random Forest, and this improvement was statistically significant ($p < 0.05$). This supports choosing Logistic Regression as the main classifier for the multi-node dataset.

To verify the reliability of these results, a confusion matrix was analyzed as shown in Fig. 5. It shows that the model maintains a high True Positive rate, confirming its reliability as a preliminary screening tool for cardiovascular risk. It can be seen that there was a high sensitivity in detecting positive CVD risk. In a screening context, this is vital; missing a potential risk (a false negative) carries a much higher cost than a false alarm. The diagonal elements of this matrix confirm that the model correctly sorted the majority of "Healthy" and "At Risk" cases, proving that the C-GAN augmented features helped define a clear boundary for classification.

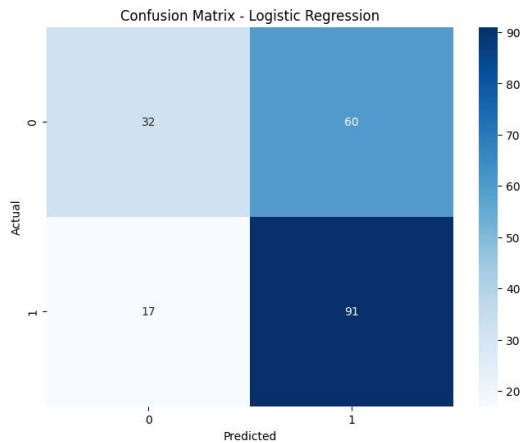


Fig. 5. Confusion Matrix - Logistic Regression

3.2 Feature Importance Analysis

To connect machine learning outcomes with theoretical roots of Nadi Tantra, the absolute coefficients of the Logistic Regression model were examined. This step highlights which physiological signals from the three sensing nodes actually drove the CVD risk predictions.

Fig. 6 compares the feature importance of various features from the dataset. It can be seen that the dominance of Pitta Amplitude provides empirical evidence for the Ayurvedic principle that pulse force is the primary indicator of circulatory load, The three most influential features are:

- **Pitta_Amplitude:** This was the strongest predictor, suggesting that the force and intensity of the pulse at the middle sensing point are primary indicators of cardiovascular stress.
- **Vaata_Regular:** The second most vital feature, which shows that a steady rhythm at the distal node is a major factor in spotting health anomalies.
- **Pitta_Regular:** The third most significant marker, reinforcing the idea that regularity at the Pitta position is a critical digital biomarker.

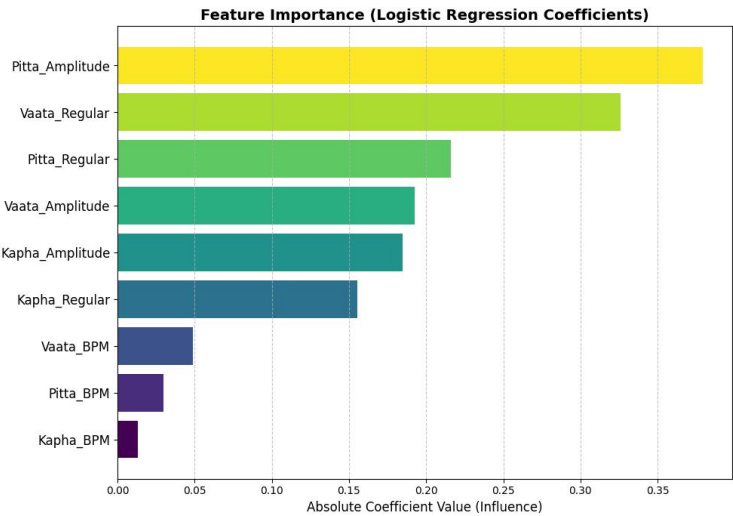


Fig. 6. Feature Importance

3.3 Clinical Discussion

The fact that Pitta amplitude and the regularity of both Vaata and Pitta pulses dominated the model’s logic aligns closely with traditional Nadi Pariksha theory.

- **Pitta Influence:** In traditional diagnostics, the Pitta pulse at the middle node represents circulatory force and metabolic heat. The heavy weight the model placed on Pitta_Amplitude supports the Ayurvedic idea that increased "heat" or "force" in this area often precedes hypertensive or inflammatory heart conditions.

- **Regularity (Gati):** The importance of Vaata_Regular and Pitta_Regular corresponds to the concept of Gati, or movement. Traditional practitioners view an erratic or unstable movement in these positions as a sign of physiological trouble. These findings confirm that digital PPG can objectively record the "erratic" pulse qualities that healers have historically identified through touch.
- The accuracy of 61.5% is a baseline result for a multi-sensor spatial acquisition problem. While single-sensor PPG often yields higher accuracy for simple heart-rate tasks, the proposed methodology processes high-dimensional data from three distinct anatomical locations simultaneously. The F1-score of **0.7027** is the more critical clinical metric here, as it demonstrates the model's ability to minimize False Negatives, which is vital for cardiovascular screening

4. Conclusion

This research effectively links traditional Ayurvedic Nadi Tantra with modern digital tools. By building a custom wearable device with three-point sensing, the study moved beyond standard single-node PPG to capture the spatial complexity of Vata, Pitta, and Kapha pulses. To handle the common "small data" issue found in alternative medicine research, an enhanced C-GAN was used to produce 1,000 high-fidelity samples, which greatly improved the stability of the models.

The results showed that Logistic Regression is a reliable classifier for this type of data, achieving an F1-score of 0.7027. While the 61.5% accuracy reflects the complex variations in multi-node signals, the strong F1-score proves the model is well-balanced for screening, specifically by keeping false negatives low. Furthermore, the analysis of feature importance provided a data-driven nod to classical theory, identifying pulse amplitude and regularity as the most significant markers for heart risk.

For real-world healthcare use, the system is designed to function as a low-latency IoT device. The ESP32 module sends pre-processed feature data to a mobile app using Bluetooth Low Energy (BLE), enabling real-time risk assessment right at the point of care. Going forward, the system will be tested in a longitudinal outpatient setting to see how consistently it can predict risk over time and how well it performs in day-to-day clinical conditions.

This research has provided a non-invasive and objective way to perform early-stage cardiovascular screening. While the current results are promising, there are several ways to expand this work in the future:

- **Clinical Diversification:** As the dataset grows, future work will test non-linear deep learning models and hybrid transformers to push accuracy past the 61.5% mark. It will also move beyond static feature vectors to explore time-series modeling, to capture the temporal Gati of the PPG signals directly
- **Real-Time Integration:** There are plans to move this diagnostic system into a cloud ecosystem, complete with a live dashboard for long-term patient tracking.

- **Multimodal Diagnostics:** Adding other metrics, such as body temperature and oxygen saturation (SpO₂), could further sharpen the accuracy of AI-driven risk estimation and Dosha classification.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Whelton, P. K., Carey, R. M., Aronow, W. S., et al.: Guideline for the Prevention, Detection, Evaluation, and Management of High Blood Pressure in Adults. *Journal of the American College of Cardiology* 71(19), e127–e248 (2018)
2. Sharma, S., Rup, P. N., Kumar, A.: Nadi Pariksha: Wrist Pulse Analysis With Traditional And Modern Interpretation. *International Journal of Ayurveda and Pharma Research (IJAPR)* 5(5), 1–8 (2017)
3. Tandon, A., Kumar, V., Choudhury, T.: Nadi Tantra approaches for pulse assessment: The role of Ayurvedic history and machine learning tools. *African Journal of Biological Sciences* 6(12), 542–555 (2024)
4. Patil, A. P.: Review on diseases diagnosing by using Nadi Parikshan. *International Research Journal of Engineering and Technology (IRJET)* 4(11), 1255–1259 (2017)
5. Allen, J.: Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement* 28(3), R1–R39 (2007)
6. Tandon, A., Kumar, V., Choudhury, T.: Pre-processing the photoplethysmography signals for enhancing cardiovascular diseases detection for wrist pulse analysis in Nadi Ayurveda. *EAI Endorsed Transactions on Pervasive Health and Technology* 11, 1–12 (2024)
7. Tandon, A., Kumar, V., Choudhury, T., et al.: Portable device for Nadi Tantra pulse examination and cardiovascular disease risk estimation. *Indian Patent Application No. 202411027743 A* (2024)
8. Mirza, M., Osindero, S.: Conditional Generative Adversarial Nets. *arXiv preprint arXiv:1411.1784* (2014)
9. Goodfellow, I., Pouget-Abadie, J., Mirza, M., et al.: Generative Adversarial Nets. In: Ghahramani, Z., Welling, M. (eds.) *Advances in Neural Information Processing Systems (NIPS 2014)*, vol. 27, pp. 2672–2680. Curran Associates, Inc. (2014)
10. Tandon, A.: Nadi Tantra Pulse Dataset: Data collected from 45 adult participants. *GitHub Repository*, <https://github.com/adityatandon19/PhD-project-adult-volunteers.git> (2025)
11. Office of the Chief Medical Officer, Ghaziabad: Ethics Clearance Certificate: Study of ‘Nadi Tantra’ Practices and Its Relation to Pulse Acquisition System in Modern Medicine. Ref. No.: Ethics Committee/2025/14.08.2025 (2025)
12. Tandon, A., Kumar, V., Choudhury, T.: An Enhanced Conditional GAN-based PPG Data Augmentation for Cardiovascular Disease Detection. In: *2025 Seventeenth International Conference on Contemporary Computing (IC3)*, pp. 1–7. IEEE, Noida (2025)

Open Access This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

