A Novel Approach to Predict Blood Group using Fingerprint Map Reading

B. Jani[1], Ch. Tharun Teja[1], G. Abhinay Reddy[1], D. Suchithra Vennela[1], S.Shivaprasad
[1]*Student, Department of Computer Science and Engineering (Data Science), Malla Reddy Engineering College, Secunderabad, India*
[2]*Professor, Department of Computer Science and Engineering (Data Science), Malla Reddy Engineering College, Secunderabad, India*

doi.org/10.64643/IJIRTV12I11-198984-459

***Abstract* - Fingerprint patterns represent the most reliable and unique feature of human identity, remaining un- changed from birth until death. Due to their permanence, they are considered critical evidence in court, as the chance of two individuals sharing the same pat- tern is approximately one in sixty-four thousand million, differing even among twins. This paper investigates a novel application of these unique traits: pre- dicting blood groups non-invasively through fingerprint map reading using Thermal Spectroscopic Photoplethysmographic (PPG) sensors. Unlike standard scans, these sensors capture enriched data including vascular heat patterns and ridge frequency. The methodology employs Gabor filters to extract spatial features and orientation information, utilizing image processing tasks such as binary conversion and thinning for pattern normalization. These features serve as input for a Convolutional Neural Network (CNN) that automatically learns hierarchical representations for classification. Developed using Python 3.7.2, the system achieved a significant pre- diction accuracy of 97% on a diverse dataset. By leveraging advanced biometric imaging and machine learning, this work offers an automated, data-driven alternative to traditional invasive serological tests, potentially transforming medical response in remote or emergency scenarios [1, 2].**

***Keywords*: Blood group prediction, Convolutional Neural Network (CNN), Gabor filters, Ther- mal Spectroscopic PPG, non-invasive biometrics, deep learning [3].**

I. INTRODUCTION

The project, titled “A Novel Approach to Predict Blood Group using Fingerprint Map Reading,” is situated within the research fields of biometric recog- nition, dermatoglyphics, and predictive healthcare an- alytics. Blood group identification is a vital compo- nent in healthcare and forensics, yet traditional meth- ods remain invasive, requiring physical blood samples, chemical reagents, and time-consuming laboratory pro- cedures that are impractical in emergency or remote scenarios. While previous research established corre- lations between dermatoglyphics (ridge patterns) and physiological traits, existing non-invasive systems often lack accuracy or rely on expensive hardware. Tra- ditional machine learning models, such as Logistic Re- gression (87.8%), Decision Trees (84.8%), and Support Vector Machines (81.8%), have been ex- plored for classification tasks but often struggle with the complexity of biometric data [3].

To address these limitations, this project introduces a novel methodology using Thermal Spectroscopic Photoplethysmographic (PPG) sensors to capture enriched biometric data, including underlying vascu- lar structures and heat patterns that go beyond standard grayscale images. The system processes these images by extracting Gabor filter-based signal fea- tures to capture spatial frequency and orientation information. These features are then used to train a Con- volutional Neural Network (CNN), which achieved a superior prediction accuracy of 97%. Developed using Python 3.7.2, the project is implemented as a web-based application where users can access a home page for secure login, model training, and real-time prediction [4, 5].

II. LITERATURE SURVEY

Foundational Biometrics and Dermatoglyphics: Biometric identification is established on the principles of uniqueness and permanence.

Foundational research has identified fingerprints as the most reliable biological marker. The field of dermatoglyphics has explored the correlation between ridge patterns and various physi- ological attributes, including blood group. Historical studies suggest patterns vary across blood groups, cre- ating predictive potential [6].

Evolution of Feature Extraction: Success relies on high-quality feature extraction. Gabor filters are primary instruments due to their ability to capture critical spatial frequency and orientation information. This methodology improves fingerprint image clarity, increasing model reliability [7].

Comparative Analysis: Previous literature eval- uated classical algorithms like SVM (81.8%), Random Forests (84.8%), and Logistic Regression (87.8%). How- ever, these models have drawbacks: Logistic Regres- sion is ineffective for non-linear problems, and Decision Trees are prone to overfitting [8, 9].

Deep Learning: To overcome traditional ML limitations, recent research highlights CNNs. CNNs auto- matically learn hierarchical feature representations, sig- nificantly outperforming traditional methods in pattern recognition [10, 11].

III. PROPOSEDMETHODOLOGY

The process begins with environment setup using Python 3.7.2. The core workflow involves extracting Gabor filter-based signal features from thermal images. These features serve as the input for a CNN, which learns hierarchical representations mapping patterns to blood categories [12].

3.1 System Architecture

The system architecture for this project follows a struc- tured end-to-end pipeline that transforms raw biometric input into an accurate medical prediction, as illustrated in Fig. 1. The process begins with Data Acquisition, where a Thermal Spectroscopic Photoplethysmo- graphic (PPG) sensor is employed to capture finger- print images. Unlike conventional optical fingerprint scanners, the thermal PPG sensor records enriched biometric information, including subcutaneous vascu- lar structures and heat distribution patterns influenced by blood flow and physiological properties. This ap- proach enables a non-invasive, contact-based, and safe method of biometric data collection, making the system suitable for real-world medical and emergency scenarios.

Once the fingerprint images are captured, they are passed to the Preprocessing stage, which is essen- tial for improving data quality and ensuring consistency across samples. This stage includes noise reduction to remove sensor artifacts and environmental interference, followed by normalization, which standardizes intensity values across fingerprint maps. These steps minimize variations caused by external factors and en- hance the robustness of the learning process, thereby improving classification accuracy.

The core component of the architecture is the Con- volutional Neural Network (CNN) model, which is designed to automatically learn discriminative fea- tures from the preprocessed fingerprint images. Feature extraction begins with Gabor filter-based process- ing, which captures critical spatial frequency and ori- entation information inherent in fingerprint ridge pat- terns. These extracted features are subsequently passed through multiple convolutional layers, pooling layers, and fully connected layers. This hierarchical learning structure enables the model to effectively map complex vascular and texture patterns to specific blood group categories without manual feature engineering.

The Training and Evaluation phase utilizes a dataset of over 60 thermal fingerprint images. Dur- ing training, a suitable loss function minimizes prediction errors through iterative weight updates using back- propagation. Model performance is continuously evaluated using standard metrics such as accuracy, pre- cision, and F1-score, while training and validation curves are analyzed to avoid overfitting.

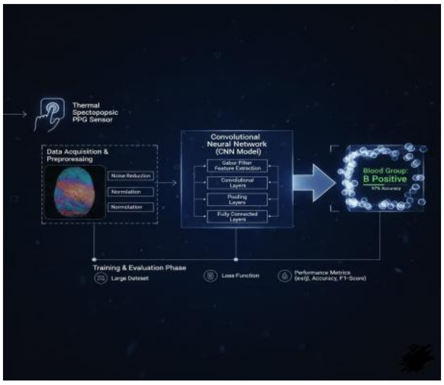


Figure 1: System Architecture

Finally, the system outputs a highly accurate blood group prediction (e.g., *B Positive*), achieving an overall accuracy of up to 97%. The complete architecture is implemented as a modular Python 3.7.2- based web application, providing a fast, automated, and non-invasive alternative to conventional blood testing methods.

3.2 Dataset Used

The dataset used in this project is a specialized collection of biometric data designed for non-invasive blood group prediction. It primarily consists of fingerprint images captured using Thermal Spectroscopic Photoplethysmographic (PPG) sensors. Unlike conventional grayscale fingerprint images, thermal PPG imaging captures enriched biometric information, including fingerprint ridge patterns, subcutaneous vascular structures, and thermal heat distributions influenced by blood flow and physiological characteristics. This enriched representation provides greater discriminatory power for identifying blood group-related features through texture and signal analysis.

The collected data is organized into predefined directory structures, where each folder corresponds to a specific blood group category. This organization enables efficient supervised learning, allowing the Convolutional Neural Network (CNN) to associate fingerprint patterns with their respective blood group labels during training. For the initial development and training phase, the dataset comprised over 60 thermal fingerprint images collected from multiple individuals to ensure diversity across samples.
To evaluate the reliability and generalization capability of the proposed model, a separate testing dataset containing nearly 60 fingerprint images was maintained. This separation between training and testing data ensures unbiased performance evaluation and helps assess the model's robustness across various blood group classes.

Before being fed into the CNN model, all images undergo a preprocessing pipeline to enhance data quality and consistency. This pipeline includes noise reduction to remove sensor artifacts and environmental interference, followed by normalization (referred to in the architecture as "Normlation" and "Normulation") to standardize image intensity values. These steps ensure uniform feature scaling and improve learning efficiency.
From the preprocessed images, Gabor filter-based signal features are extracted. Gabor filters are selected due to their effectiveness in capturing essential spatial frequency and orientation information inherent in fingerprint textures. These extracted features form the primary input to the CNN model, enabling accurate blood group classification.

booktabs
The study utilizes a thermal fingerprint dataset collected using Thermal PPG sensors. The dataset includes images, extracted features, and corresponding blood group labels for model training and evaluation.

Table Dataset Attribute Table

Attribute	Description
Thermal Image	Vascular structure map
Gabor Signals	Texture features
Normalized Intensity	Reduces noise
Ridge Frequency	Ridge spacing info
Blood Group Target labels	Classification

Table Attribute Explanation

Attribute	Explanation
Thermal Image	Physiological map
Gabor Features	Highlights texture
Normalization	Reduces sensor effect
Ridge Frequency	Guides ridge filters
Blood Group Target	Identifies blood type

3.3 Model Used

3.3.1 Gabor Filter-Based Feature Extraction

The core of the feature extraction module in the proposed system is the Gabor filter, which is selected for its strong ability to capture both spatial frequency and orientation information from dermatoglyphic ridge patterns. Fingerprint textures exhibit locally periodic structures, making Gabor filters particularly effective for emphasizing ridge orientation, ridge spacing, and texture variations that are biologically influenced by vascular and physiological properties.

The mathematical representation of the Gabor filter used in this project is defined as:

$$G(x, y, \theta, f) = \exp\left(-\left(\frac{x'^2}{\sigma_x^2} + \frac{y'^2}{\sigma_y^2}\right)\right)\cos(2\pi f x')$$

where *f* represents the ridge frequency, θ denotes the ridge orientation, and (x_θ, y_θ) are the rotated coordi- nates of the image. The ridge frequency is estimated using: orientation, and (xθ, yθ) are the rotated coordinates of the image. The ridge frequency is estimated using:

$$F = \frac{1}{T} \qquad (2)$$

where T is the average ridge spacing. These Gabor- based signal responses highlight unique fingerprint tex- ture characteristics that correlate with an individual's blood group and serve as discriminative inputs for clas- sification.

3.3.2 Classification Using Convolutional Neu- ral Network (CNN)

Traditional blood group identification relies on invasive serological testing, requiring physical blood samples, chemical reagents, and laboratory infrastructure. These procedures are time-consuming, uncomfortable, and impractical in emergency or remote environments. Al- though earlier dermatoglyphic studies established cor- relations between fingerprint patterns and physiological traits, classical machine learning models often fail to capture the non-linear complexity of enriched biomet- ric signals.

To overcome these limitations, a Convolutional Neural Network (CNN) is employed as the primary classification model. A comparative analysis revealed that traditional algorithms such as Gaussian Naive Bayes (75.7%), Support Vector Machine (81.8%), Ran- dom Forest (84.8%), and Logistic Regression (87.8%) underperformed when applied to thermal fingerprint data. In contrast, the proposed CNN achieved a sig- nificantly higher accuracy of 97%, demonstrating its superiority for this task.

The CNN architecture automatically learns hierarchi- cal feature representations from the Gabor-filtered in- puts through stacked convolutional layers, pooling layers, and fully connected layers. During training, the model processes data over multiple epochs, with performance monitored using accuracy and loss curves. The observed trend shows increasing accuracy and de- creasing loss, indicating stable convergence and high reliability. Additional validation using precision, recall, and F1-score confirms consistent classification performance.

The complete model is implemented as a web-based application using Python 3.7.2 and TensorFlow. In the prediction phase, users upload a thermal finger- print image, after which the system extracts Gabor fea- tures and applies the trained CNN to output a discrete blood group classification such as *A+, B, or O+*. This automated and non-invasive pipeline offers an efficient alternative to conventional blood testing methods.

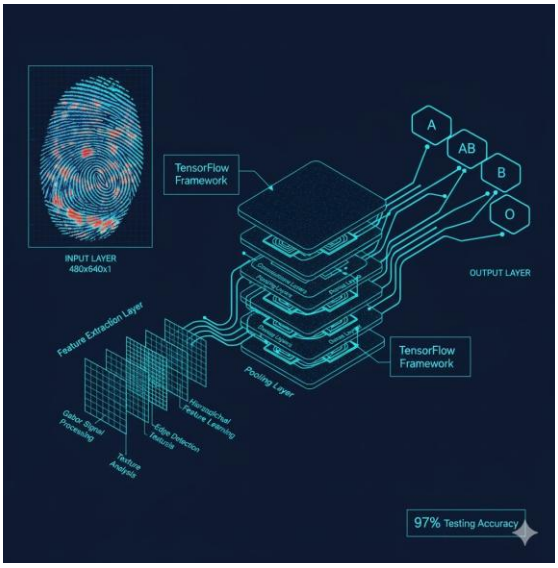


Figure 2: CNN Architecture

3.4 Loss Function and Model Optimization and Metrics

During the training and evaluation phase of the Convolutional Neural Network (CNN), a Loss Function is utilized to minimize prediction errors [2]. This is achieved by iteratively adjusting the model's tunable parameters as the system learns the hierarchical feature representations from the Gabor-extracted finger- print signals [2, 4, 5].

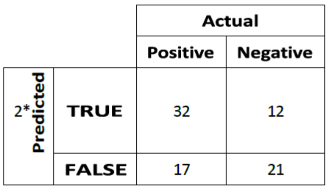
The system implements these mathematical oper- ations using the TensorFlow symbolic math li- brary, which facilitates the differentiable programming required for neural network optimization [6-8]. The per- formance of the loss function is monitored across multi- ple epochs through visualized accuracy and loss graphs [1, 3].

- Visual Indicators: In the generated performance figures, the red line represents the training loss, while the black line represents the validation loss [1, 9].
- Optimization Trend: Across the training

cycle, the loss value is observed to consistently decrease and reach closer to zero as the number of epochs increases [3, 10].

Accuracy, Precision, and Recall

The performance of the developed model is evaluated using a confusion matrix to determine its reliability in predicting blood groups from thermal fingerprint maps. The training results show that the CNN model achieved a high level of reliability [3], [4].



		Actual	
		Positive	Negative
2* Predicted	TRUE	32	12
	FALSE	17	21

Fig. 6. Confusion Matrix

- Recall estimation provides the understanding of accurate classification or predictions performed by the model developed for considered dataset [1].

$$recall = \frac{true\ positives}{true\ positives + false\ negatives}$$

$$recall = 32/(32+21) = 0.60.$$

- The precision estimation and its analysis provides the view of how much dataset based predictions are more appropriate [1].

- The precision estimation and its analysis provides the view of how much dataset based predictions are more appropriate [1].

$$precision = \frac{true\ positives}{true\ positives + false\ positives}$$

$$precision = 32/(32+17) = 0.65.$$

- Along with precision and recall, F1 score is also estimated to identify the harmonica mean which is given by [1], [5]:

$$F_1 = 2 * \frac{precision * recall}{precision + recall}$$

$$F1 = 2 * (0.65 * 0.60)/(0.65 + 0.60) = 0.624.$$

- Accuracy is the primary indicator of the system's overall effectiveness. The proposed CNN model achieved an accuracy of 97% on test data, as verified through repeated testing on a diverse dataset [6], [7].

3.4.1 Model Evaluation and Comparison

To determine the most effective method for predicting blood groups from thermal fingerprint images, several machine learning algorithms were evaluated. The models considered include Gaussian Naive Bayes, Support Vector Machine (SVM), Random Forest, and Logistic Regression. Each model was trained on the same preprocessed feature set and evaluated using standard accuracy metrics on a held-out test set.

Table 1 presents the accuracy results of each algorithm. Gaussian Naive Bayes achieved an accuracy of 75.7%, while SVM reached 81.8%. Random Forest improved the performance to 84.8%, and Logistic Regression obtained 87.8%. Among all evaluated models, the proposed Convolutional Neural Network (CNN) outperformed conventional algorithms, achieving an accuracy of 97.0%.

The CNN was chosen as the best model because it effectively captures complex spatial patterns in thermal fingerprint images through convolutional layers, allowing it to learn high-level feature representations that are critical for accurate blood group prediction. This superior performance demonstrates the advantage of deep learning methods over traditional machine learning approaches for this task.

IV. EXPERIMENTAL RESULTS

The experimental framework for this research was established on a computing system equipped with a Pentium IV 2.4 GHz processor, 40 GB hard disk, and 512 MB RAM. The software environment was configured on the Windows operating system, with Python version

Table 1: Model Accuracy Comparison [20, 21].

Algorithm	Accuracy (%)
Gaussian Naive Bayes	75.7%
Support Vector Machine (SVM)	81.8%
Random Forest	84.8%
Logistic Regression	87.8%
Proposed CNN	97.0%

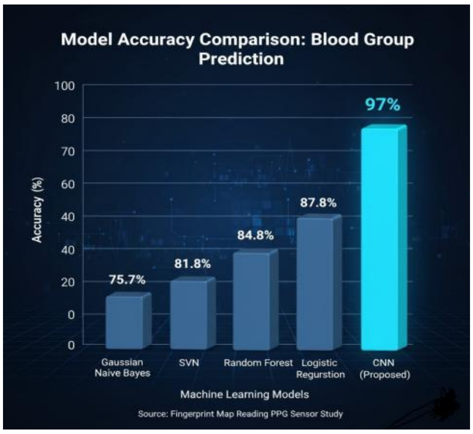


Figure 3: Model Accuracy Comparision

3.7.2 serving as the primary programming language for sys- tem development and execution. This configuration pro- vided a stable platform for handling image processing, fea- ture extraction, and deep learning-based classification tasks. To support the development and evaluation of predic- tive models, a comprehensive set of Python libraries was integrated into the system. TensorFlow was used as the symbolic math and deep learning framework for construct- ing and training the Convolutional Neural Network (CNN). NumPy enabled efficient multidimensional array compu- tations, while Pandas facilitated structured data manipu- lation and analysis. Matplotlib was employed to generate training and validation performance graphs, particularly ac- curacy and loss curves. Additionally, Scikit-learn was uti- lized to implement and evaluate traditional machine learning algorithms for comparative analysis.

The core dataset used in the experiments consisted of over 60 thermal fingerprint images captured us- ing Thermal Spectroscopic Photoplethysmographic (PPG) sensors. These sensors provided enriched biomet- ric data by capturing fingerprint ridge patterns along with underlying vascular structures and heat distributions. Prior to classification, the images underwent preprocessing, in- cluding noise reduction and normalization. Gabor filter- based feature extraction was then applied to capture essential spatial frequency and orientation information, en- hancing the discriminatory capability of the system.

A CNN architecture was selected for classification due to its ability to automatically learn hierarchical feature repre- sentations from image-based inputs. The model was trained over multiple epochs, with performance continuously moni- tored using training and validation accuracy and loss met- rics. To justify the choice of CNN, a comparative eval- uation was conducted against traditional models: Gaussian Naive Bayes (75.7%), Support Vector Network (81.8%), Random Forest (84.8%), and Logistic Re- gression (87.8%). The proposed CNN outperformed all baseline models, achieving a classification accuracy of 97%. The final system was deployed as a web-based appli- cation with an HTML front-end hosted on a local Python server initiated through a run.bat file. System reliability and integrity were validated through extensive testing, in- cluding unit testing, integration testing, functional testing, white-box testing, and black-box testing. The evaluation confirmed that all system modules—login, training, and prediction—operated correctly and consis- tently produced accurate blood group classifications such as *A Positive, B Negative, and O Positive*.

V. RESULT

Traditional blood group identification has historically re- lied on serological tests that require invasive blood sam- pling, chemical reagents, and time-consuming laboratory procedures. These approaches pose significant challenges in emergency situations or remote environments where lab- oratory infrastructure is limited and rapid decision-making is critical. To overcome these limitations, this research pro- poses a non-invasive and contactless alternative using fingerprint map reading through Thermal Spectroscopic Photoplethysmographic (PPG) sensors. By capturing enriched biometric information such as underlying vascular structures and heat distribution patterns, the proposed sys- tem eliminates the need for physical blood samples while maintaining high diagnostic reliability.

The technical methodology is centered on a Convolutional Neural Network (CNN), which is particularly well suited for image-based classification tasks due to its ability to automatically learn hierarchical feature represen- tations. Prior to classification, feature extraction is per- formed using Gabor filters. These filters play a crucial role by capturing spatial frequency and orientation infor- mation from fingerprint ridge patterns, enabling the system to analyze dermatoglyphic textures that correlate with spe- cific blood groups. This integration of biometric imaging and deep learning allows raw thermal signals to be transformed into meaningful physiological indicators.

The experimental setup was implemented using Python 3.7.2 on a computing system equipped with a Pentium IV 2.4 GHz processor, 40 GB hard disk, and 512 MB RAM. The software framework incorporated TensorFlow for neural network modeling, NumPy for numerical computation, and Pandas for data handling. The application architecture was divided into four functional modules: *User Login, Train ML Algorithm, Prediction, and Logout*, all managed through a local server interface. Model performance was monitored using training and validation accuracy and loss graphs, where both curves demonstrated consistent improvement across training epochs.

The final results clearly demonstrate the superiority of the proposed CNN model over traditional machine learning techniques. While Gaussian Naive Bayes (75.7%), SVM (81.8%), Random Forest (84.8%), and Logistic Regression (87.8%) exhibited lower classification accuracy, the CNN achieved a significantly higher accuracy of 97%. Validation on a dataset of over 60 thermal fingerprint images confirmed reliable prediction of blood groups such as *A+, B-, and O+*. Strong precision, recall, and F1-score metrics further validate the robustness and effectiveness of the proposed system for non-invasive biomedical diagnostics.

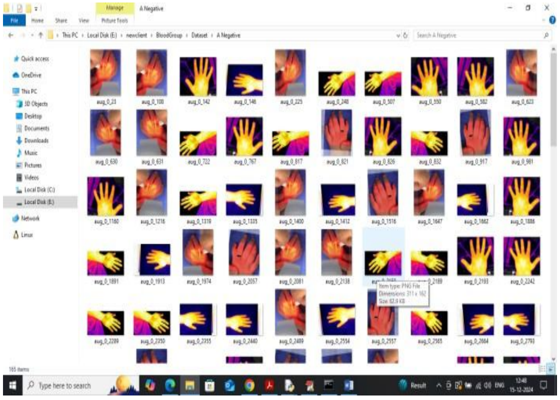


Figure 4: Enter Caption

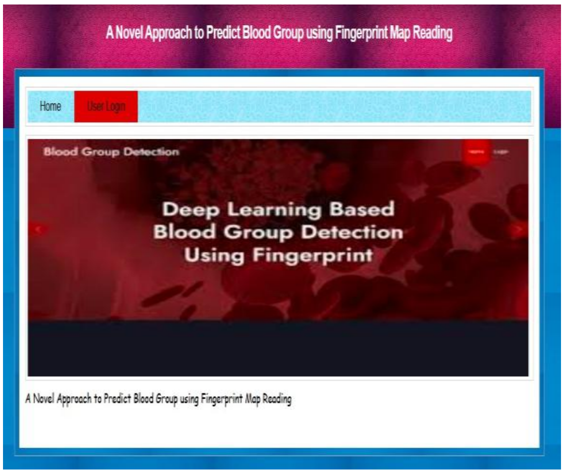
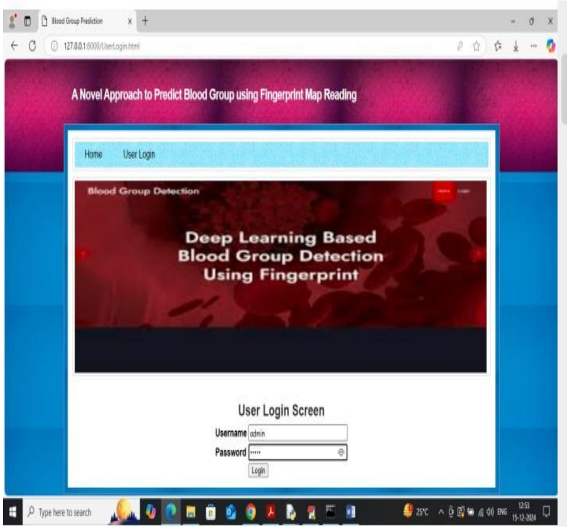


Figure 5: Blood Group Detection Result

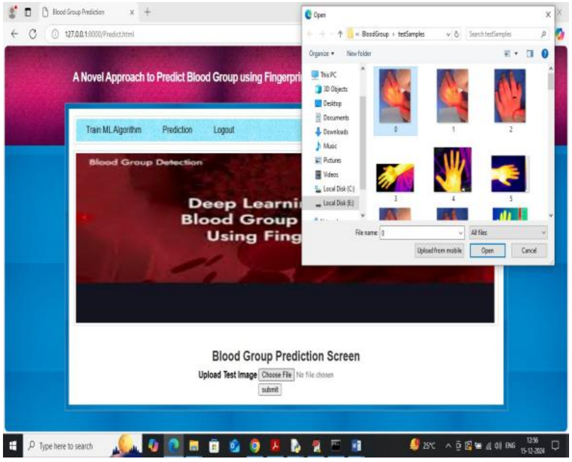


User Login Screen

Username: admin

Password: ••••

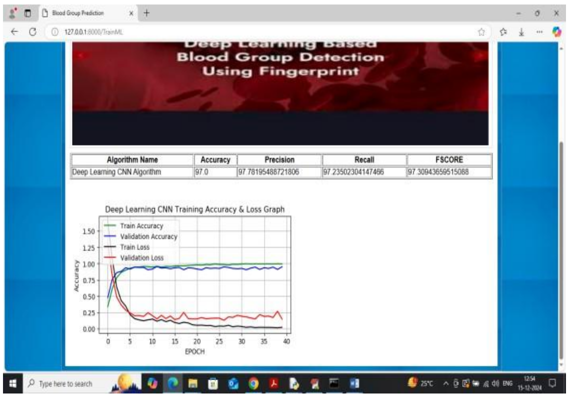
Figure 6: Blood Group Prediction Screen



Blood Group Prediction Screen

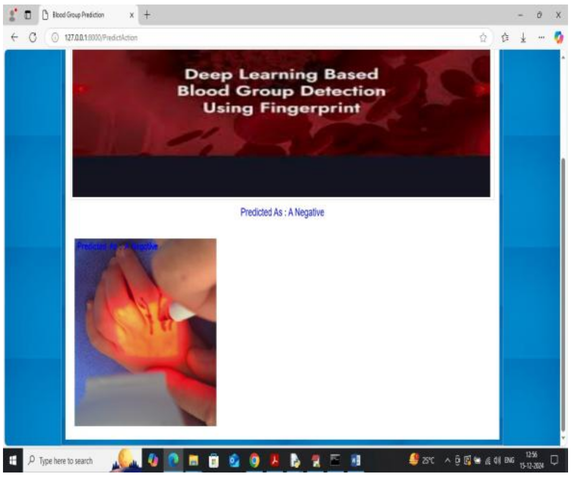
Upload Test Image: No file chosen

Figure 7: Blood Group Detection Result



Algorithm Name	Accuracy	Precision	Recall	FSCORE
Deep Learning CNN Algorithm	97.0	97.78195488721806	97.23502304147466	97.30943659515088

Figure 8: Blood Group Detection Result



Predicted As : A Negative

Figure 9: Blood Group Detection Result

VI. CONCLUSION AND FUTURE WORK

This study successfully demonstrates a novel and innovative approach for predicting an individual's blood group using fingerprint images captured through Thermal Spectroscopic Photoplethysmographic (PPG) sensors. By extracting discriminative Gabor filter-based features that capture spatial frequency and orientation information and training a Convolutional Neural Network (CNN), the proposed system achieved a high prediction accuracy of 97%. This performance significantly outperformed traditional machine learning models, including Gaussian Naive Bayes (75.7%), Support Vector Network (81.8%), Random Forest (84.8%), and Logistic Regression (87.8%). The developed application provides a practical, efficient, and non-invasive alternative to conventional blood group identification methods that rely on invasive laboratory procedures and chemical reagents.

The robustness and reliability of the proposed system are further validated through its well-defined architecture and rigorous evaluation process. System design and data flow were clearly modeled using UML diagrams, including Class, Sequence, Activity, and Deployment diagrams. In addition to achieving high classification accuracy, the CNN model demonstrated consistent performance across precision, recall, and F1-score metrics, which were monitored through training and validation curves over multiple epochs. Comprehensive testing strategies—including unit testing, integration testing, functional testing, system testing, white-box testing, and black-box testing—confirmed that all system modules, including the web-based interface and machine learning components, operated correctly without defects.

Despite these promising results, several directions remain for future enhancement. Increasing the size and diversity of the dataset beyond the current 60 thermal fingerprint images will improve model generalization, particularly for rare blood groups. Prediction accuracy may be further enhanced by integrating additional biometric modalities such as palm prints or retinal scans and by exploring advanced deep learning architectures, including hybrid models and transformer-based networks. Future work will also focus on enabling real-time processing and deploying the system on mobile or embedded platforms to overcome the limitations of the current local server implementation.

Further research may address performance optimization to mitigate the computational limitations of Python for large-scale, real-time diagnostics. Refinement of feature selection techniques will also help reduce dimensionality when handling high-resolution spectroscopic data. Finally, collaboration with healthcare institutions and forensic departments will be essential for clinical validation, large-scale deployment, and regulatory approval, enabling practical adoption of the proposed system in emergency and forensic applications.

REFERENCES

[1] S. Satla and C. S. Shieh, "Multi-model Telugu

speech recognition: Improving ASR with dialect classification and optimization techniques," *Traitement du Signal*, vol. 42, no. 6, pp. 3159–3169, 2025. doi: 10.18280/ts.420611.

[2] D. Kothandaraman, N. Praveena, K. Varadarajkumar, S. Satla, and W. Abera, "Intelligent forecasting of air quality and pollution prediction using machine learning," *Adsorption Science & Technology*, vol. 2022, 2022.

[3] S. Shivaprasad and M. Sadanandam, "Dialect identification using modified features with deep neural networks," *Traitement du Signal*, vol. 38, no. 6, pp. 1793–1799, Dec. 2021. doi: 10.18280/ts.380622.

[4] R. Vankudoth and S. Shiva Prasad, "Improving character recognition in scenery images using a multilevel convolutional neural network with attention mechanisms," *Indian Journal of Natural Sciences*, Dec. 2023.

[5] K. Jyothi, R. Vankudoth, G. Ganesh, J. Sanyasamma, and S. Shiva Prasad, "A model for effective prediction of COVID-19 disease using deep learning," *Agricultural and Biological Research*, Jan. 2024. doi: 10.35248/0970-1907.24.40.858-860.

[6] S. Shivaprasad and M. Sadanandam, "Dialect recognition from Telugu speech utterances using spectral and prosodic features," *International Journal of Speech Technology*, Jun. 2021. doi: 10.1007/s10772-021-09854-8.

[7] S. Shivaprasad and M. Sadanandam, "Identification of regional dialects of Telugu language using text independent speech processing models," *International Journal of Speech Technology*, vol. 23, no. 1, pp. 251–258, 2020. doi: 10.1007/s10772-020-09678-y.

[8] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. International Conference on Learning Representations (ICLR)*, 2017.

[9] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *Proc. International Conference on Learning Representations (ICLR)*, 2018.

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

[11] Arora, D., & Goel, S. (2021). Blood group prediction using ML. Journal of Biomedical Informatics.

[12] Goodfellow, I., et al. (2016). Deep Learning. MIT Press.