A.I Approach to Detect Blood Group Using Fingerprint

Abhishek K S, Ajith J Iyengar, Nithin A G, Shivamani S

School of Computer Science and Engineering, RV University, Bangalore, India
USN: 1RVU23CSE017, 1RVU23CSE034, 1RVU23CSE313, 1RVU23CSE430

***Abstract*— The use of Blood Group in the medical field is crucial as this is the first step for diagnostics and analysis of a human and it is required before any medical surgeries or operations. Its importance is also found in forensic science and the identification of biometrics as this is a unique and constant characteristic of the human body. Traditional methods of extracting Blood Group involves skilled professionals and laboratory space and this can be tedious and can be time consuming. Our project presents a non-invasive Blood Group detection system using fingerprint image processing through machine learning for more accurate results. It is portable and user friendly, which reduces the need for traditional methods of blood testing. The system also reduces the contamination of blood samples and can be applied to large scale. We have used statistical analysis and machine learning techniques like Convolutional Nueral Networks in our study to find patterns between fingerprints and blood groups. The machine learning technique takes the help of the fingerprint and its respective blood group dataset that has been provided to the system and identifies the blood group that has been asked by the patient through unique ridges present on the patient's finger and gives the blood group by the support of the trained data. The result indicates that there will be a step forward in integrating blood group detection through machine learning and helps in paving the way for future advancements in the medical field.**

***Index Terms*— Grayscale image of Fingerprint, Blood Group, Convolutional Neural Networks (CNN), Blood Group along with fingerprint for training dataset.**

I. Introduction

The Human Fingerprint is one of the most distinctive and unique characteristics feature due to the ridges, curves, loops present on the fingers and similar fingerprints in pairs cannot be found usually. Even though a human being grows in age, the unique fingerprint pattern remains constant until death. For detection of diseases and other infections and also for performing various surgeries, blood test is required for further analysis. These blood groups obtained can be indirectly linked to the person's fingerprint using pattern recognition by machine learning techniques. The traditional methods which are used in today's world are invasive as the person needs to be in contact with an injection and also time consuming, while this method is non invasive and a better alternative and can enhance medical diagnostics.

Understanding fingerprints is a prerequisite for knowing about this project. Fingerprints can be classified into four types mainly Loops, Whorls, arches and mixed types. Out of that, Loops are the most common type which comprises 65% and Whorls the second most common type with 25% and arches as well as mixed types account of the remaining percentage. These fingerprints are used to distinguish between the humans.

The Project suggests that fingerprint analysis for blood group detection improves efficiency in the medical field. This is done by the system through complicated features in fingerprints such as cores and ridges and then it corresponds them with the specific blood groups. It has applications in healthcare and forensics.

The process begins with the collection of fingerprints of the individual and this can be done by uploading the fingerprint image in the form of Grayscale to the system. The system will be having the blood group along with fingerprint dataset for training and once the image of the fingerprint is uploaded by the individual, the machine learning model analyses the uploaded fingerprint and tries to align with the tested dataset having millions of fingerprints along with their respective blood groups and later when the uploaded fingerprint corresponds to one of the fingerprints present on the dataset, the blood group respective to that fingerprint is displayed. The machine learning technique used for this process is Convolutional Neural Network. For pattern recognition applications this is the most used technique and a very important tool as CNNs are adapted at extracting features form complex images and are perfect for this analysis as fingerprint image will have complicated features. This innovative approach helps in enhancing medical diagnostics and this is the description of the project in short.

II. Problem Statement

The process of extracting blood form a human involves invasive methods such as drawing blood samples and conducting laboratory tests and this can be time consuming, expensive and require professionals. During this procedure, there can be spread of infections due to contamination of syringes that are used for extracting blood from a person. The usage of the same syringe on another person can be very dangerous as this can lead to the mixing of blood samples. With the advancement of medical technology, there needs to be a better alternative to the traditional methods. Here, this project plays an important role as it proposes a non-invasive machine learning based solution or technique to predict a person's blood group using fingerprint images. By applying extraction techniques like CNN which is used here, the system can identify patterns and can correspond to the specific blood group. The main goal or the objective of the project is to provide a fast, cost-effective and easily accessible tool or system for blood group identification that can support medical institutions and can resolve all the issues that were mentioned.

III. Related work

[1]. Research in Dermatoglyphics, which is a scientific study of fingerprints has shown possible correspondence in the ridge patterns as mentioned above like loops, whorls and arches, with the blood group types. Multiple statistical studies have identified patterns among specific blood groups, laying the idea for the use of fingerprints as predictors for traits like blood groups.

[2]. With the rise of deep learning methods like Convolutional Nueral Networks which is an image based classification. CNN extract the complex features form raw data and this makes it suitable for fingerprint image processing. There are also demonstrations that CNNs have the ability to distinguish between different fingerprint patterns as well as match fingerprint identities. This proves that CNNs are very ideal for this type of processing that is from grayscale imaging without manual processing.

[3]. A relevant research has been conducted and provided by Vijayakumar Patil, which was a method to predict blood groups using fingerprint readings. They used Linear Regression model to extract fingerprint features of blood types.

[4]. Another research named "Blood Group determination using Fingerprint" by the Author T.Nihar proposed that the dataset containing blood group along with fingerprints can be used to tell a new fingerprint of the individual by aligning the new fingerprint that was uploaded with the dataset and after it is aligned the blood group is displayed by the machine learning model.

IV. Methodology

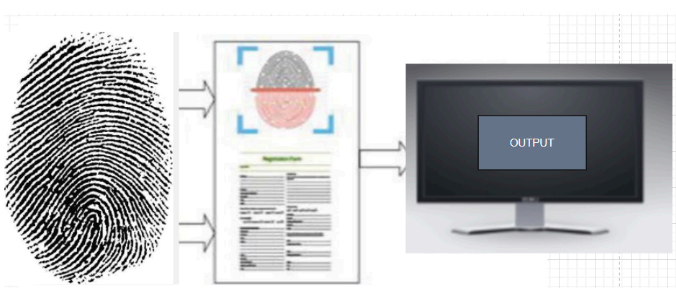

Fig. 1. Methodology diagram

The project proposes a non-invasive approach for blood detection by resolving the issues that were faced during the traditional methods blood extracting and then using it for detection. This approach follows a structured way of giving the output: Data collection, Preprocessing, model training through CNN, validation and prediction. First, the user would have to find the upload feature that is present in the platform and uploading of the image is required for processing. The user uploads a grayscale image of the fingerprint in this Upload feature. Now the processes that were mentioned above will take place. And then he will have to click on predict button, which will give the prediction of the blood group.

Data Collection: A dataset of images of fingerprints along with their respective blood groups were collected from various individuals. These collected images were later converted to grayscale image format to reduce complexity in the system. This dataset was distinguished into multiple blood group types or classes like A+, B+, AB+, O+, A-, B-, O-, AB-.

Data Preprocessing: Data preprocessing is required for the system to understand the uploaded image of the user in order to display the output in the form of blood group.

Image Resizing: The fingerprint images were resized to a fixed resolution so that consistency could be maintained overall in the dataset.
Augmentation: Inorder to increase the diversity in the data techniques like rotation and zooming are used also to decrease the overfitting.
Label Encoding: Blood group labels were encoded into numerical format for easier classification.

Model Architecture: The machine learning technique used is CNN or Convolutional Neural Network. A CNN is a special type of Deep Learning Model which is designed to process data with a grid like structure in the form of images. This is used in image classification tasks and also object identification.

The working of CNN can be explained in the following manner:

An image of the form grayscale image is uploaded to the system and this image is taken as the form of binary numbers that is 0 and 1 and these binary numbers are arranged in the form of a matrix. Then a filter or kernel which is a type of matrix slides over this image as the input in order to detect patterns like edges or specific shapes. These filters are responsible for identifying a particular feature and each filter gives an output and this output is nothing but a feature map. This process is called convolution.

After this process, the output goes through a non-linear activation function like the Rectified Linear Unit or the ReLU, which basically changes all negative values into zero. The network learns and enhances its ability to capture complex features. From this the data passes through Pooling layer where spatial dimensions of feature maps are reduced. Here most of the features are retained and reduces computational load.

These processes that is convolution and pooling are repeated several times to extract other complex features and this is called flattening and later the data is flattened into a one dimensional vector. The obtained vector is then fed into connected layers which act as a classifier and makes predictions. The final layer uses SoftMax activation function for multi-class distinguishing and classification and provides the output for each class.

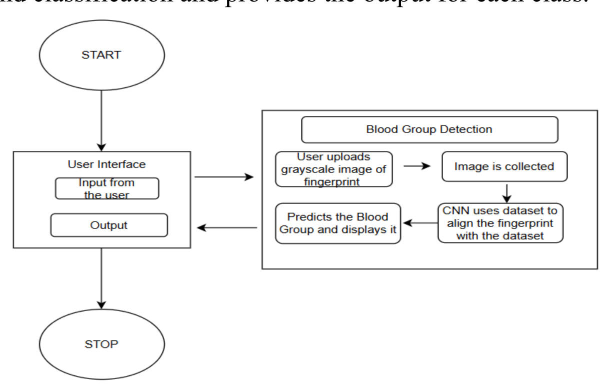

Fig. 2. Architecture of the study

Training the Model: The dataset was divided into 70% training and 15% validation 15% testing.The model is trained over multiple epochs and the most efficient epoch in this project came out to be epoch 25.

Validation and Evaluation: Monitoring model performance by training accuracy and loss was done. Validation accuracy was also used to evaluate so that the model performs well on raw or unseen data. Metrics like Precision, Recall and F1 score were also used for detailed performance analysis.

```
Accuracy: 0.8970
Precision (Macro): 0.8975
Precision (Micro): 0.8970
Precision (Weighted): 0.8980
Recall (Macro): 0.8995
Recall (Micro): 0.8970
Recall (Weighted): 0.8970
F1 Score (Macro): 0.8975
F1 Score (Micro): 0.8970
F1 Score (Weighted): 0.8964
```

Fig. 3. Performance Comparison Table

Full Classification Report:

	precision	recall	f1-score	support
A+	0.93	0.94	0.93	133
A-	0.89	0.84	0.86	138
AB+	0.86	0.95	0.90	144
AB-	0.91	0.93	0.92	162
B+	0.86	0.91	0.89	146
B-	0.91	0.95	0.93	131
O+	0.89	0.85	0.87	165
O-	0.93	0.82	0.87	165
accuracy			0.90	1184
macro avg	0.90	0.90	0.90	1184
weighted avg	0.90	0.90	0.90	1184

Fig. 4. Full Classification Report

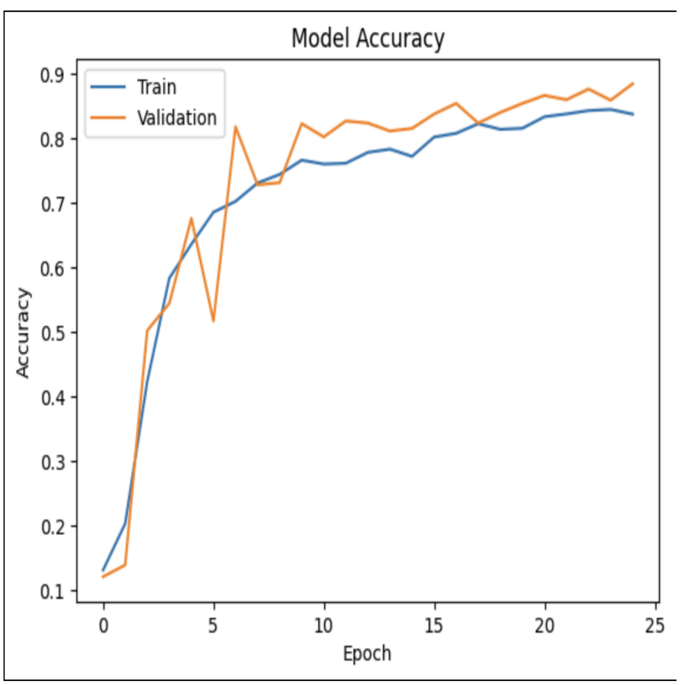

Fig. 5. Model Accuracy

Deployment: A user interface was built and designed to upload the fingerprint of the user and predict the blood group.

Tools and Technologies Used:

Programming language: Python for designing user interface and for machine learning model

Model Type: CNN

Libraries used: TensorFlow, NumPy, Matplotlib, Sklearn

Environment: Jupyter Notebook

Dataset: Kaggle

CNN Architectures:
We start initially from LeNet, AlexNet, VGG16 / VGG19, ResNet50 which are popular CNN architectures but are built for various purposes.

LeNet: Introduced by Yann LeCun in 1989, it is a type of CNN architecture which is usually referred to as LeNet-5. It was primarily designed for the purpose of reading handwritten digit present on cheques. It was also used to read zip codes. LeNet is best suited for studying the grayscale images used in this project and the images typically small in size. This makes it lightweight and perfect for low resolution image classification. LeNet consists of 7 layers including the convolutional layers.There are two convolutional layers, two average pooling layers and three fully connected layers that perform the task.

AlexNet:
AlexNet was introduced in 2012 which was marked a breakthrough in deep learning as it demonstrated the power of CNNs when it was trained with large datasets using high computational tools like GPUs. AlexNet was designed for colour images having a larger size and classification could be done into 1000 categories. It is more complex than LeNet as AlexNet comprises of eight layers which are five convolutional layers and three fully connected layers. This model showed that it was ideal for complex tasks and can be updated than traditional methods by a huge margin.

VGG16 / VGG19:
Proposed by the Visual Geometry Group at Oxford, VGG16 is known for its simple and uniform architecture using 3x3 convolutional layers stacked on top of each other. It has 16 weight layers and uses ReLU activation with max pooling. VGG16 is effective for transfer learning and is commonly used in fingerprint classification tasks due to its strong feature extraction capabilities. However, it requires more memory and computational power.

ResNet50:
Introduced by Microsoft in 2015, ResNet (Residual Network) uses skip connections or "residuals" to allow training of very deep networks without degradation. ResNet50 contains 50 layers and is highly effective in learning intricate fingerprint patterns. It solves the vanishing gradient problem and is ideal for large, complex datasets with many classes like blood group classification.

In this project, we used **LeNet** due to its simplicity and efficiency with grayscale fingerprint images, making it ideal for CPU use without requiring a GPU. While we also tested **AlexNet**, **VGG16**, and **ResNet50**, these more complex architectures were less suited for the task, either due to their computational demands or overfitting to the simpler dataset. LeNet offered the best balance between performance and resource efficiency.

The proposed project has certain limitations like the CNN trains on raw or unseen data but if any other blood group needs to be detected and if it is rarely found then there are no more training datasets for that blood group so the accuracy of the model decreases. For example, AB- is a rare blood group so for this blood type there are a smaller number of datasets as compared to other blood groups and this may lead to class imbalance. This can be resolved by oversampling this dataset. Oversampling is the method where the imbalanced dataset is duplicated so as to balance out the data.

V. Computational Efficiency

For this project, computational efficiency is determined according to trainable parameters and model structure. While FLOPs are generally employed in measuring arithmetic complexity in CNNs, in this case, we consider the parameter count and resource utilization of the model because of the grayscale fingerprint input and the moderate computational requirement.

The custom CNN model architecture implemented in this project includes:

- 5 convolutional layers
- 4 max pooling layers
- 1 flatten layer
- 2 dense layers
- Normal dropout layers to avoid overfitting

VI. Model Complexity

Model complexity is primarily reflected in the total number of trainable parameters and the hierarchical structure of the network. This project's architecture is moderately deep, making it well-suited for high-level feature extraction from grayscale fingerprint images. With nearly 3.7 million trainable parameters, the network is capable of learning intricate patterns without overfitting, thanks to the use of dropout regularization.

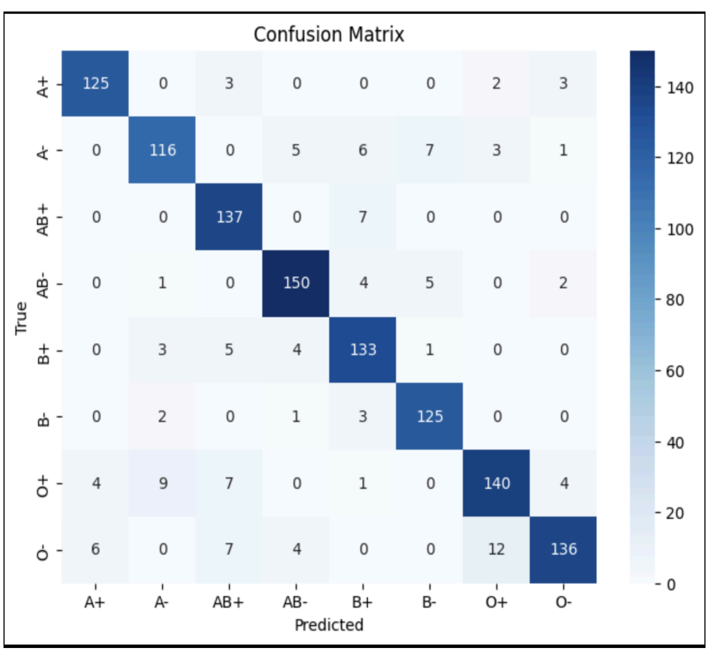

Fig. 6. Confusion Matrix

VII. Formulas

Precision = TP/(FP+TP)

Recall = TP/(FN+TP)

F1 Score = 2×(Precision×Recall)/Precision+Recall

Storage (in bytes) = Total Parameters×Precision

Trainable Params = hk×wk×cin×cout

VIII. Conclusion

The project on Blood Detection using Fingerprint provides an Artificial Intelligence approach by the integration with CNN model. This proposed model will be revolutionary system in blood group detection by scanning and uploading fingerprints. By analyzing the training dataset, the model offers a non-invasive and cost-effective method. This model can remove discomforts caused by the traditional methods making it suitable for medical applications.
The future scope of this is by improving and developing the model so as to get accurate results. Also, the model needs to be trained on larger and more diverse datasets including rare types. This model can be implemented with advanced machine learning models.

VI. Results

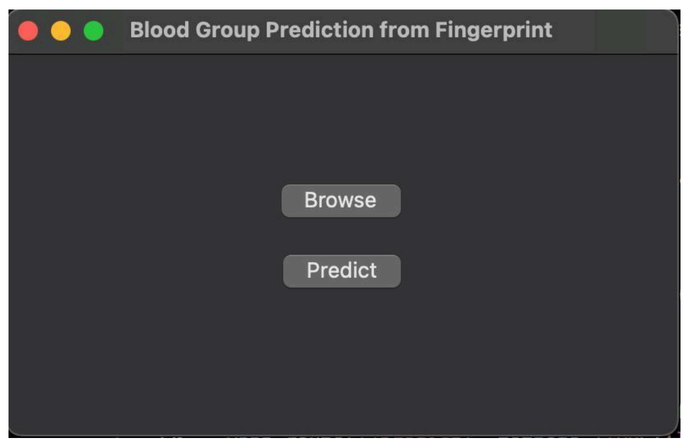

Fig. 7. User interface display

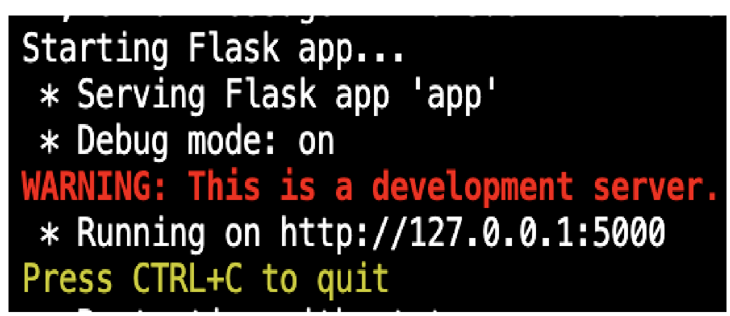

Fig. 8. Flask app starting

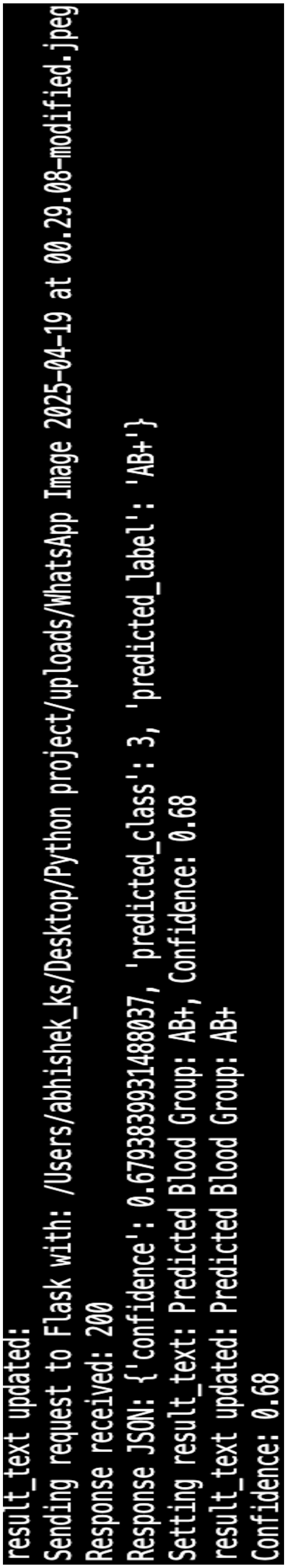

Fig. 7. Result display in terminal

VII. Acknowledgements

We would like to express our sincere gratitude to our esteemed professors, **Prof. Suman Majumder, Prof. Achala Hegde** and **Dr. Shabbeer Basha S H** from the School of Computer Science and Engineering at RV University. We would like to also thank R V University for providing this opportunity and support.

VIII. References

[1]. T. Nihar, K. Yeswanth, and K. Prabhakar, “Blood group determination using fingerprint,” *MATEC Web of Conferences*, vol. 392, p. 01069, 2024, doi: https://doi.org/10.1051/matecconf/202439201069.

[2]. “ADVANCED IMAGE PROCESSING FOR FINGERPRINT-BASED BLOOD GROUPING,” 1286: https://www.irjmets.com/uploadedfiles/paper//issue_7_july_2024/60089/final/fin_irjmets1720690577.pdf