Journal of Information Systems Engineering and Management
2025, 10(40s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Advancing Blood Cell Image Classification: Hybrid CNN-LSTM vs. Traditional CNN Approaches

[1]Dr. Dharmendra Pathak, [2]Dr. Mohit Arora, [3]Dr. Manoj Agrawal, [4]Dr. Shailesh Gondal, [5]Dr. Harshita Vachhani, [6]Dr. Neeraj Sharma

Asst. Professor, School of Computer Science & Engineering, Lovely Professional University, Punjab – 144451, India

Associate Professor, School of Computer Science & Engineering, Lovely Professional University, Punjab – 144451, India

Associate Professor, Department of Computer Science & Engineering, Chameli Devi Group of Institutions, Indore, M.P. – 452020, India

Professor, FCA Department, Acropolis Institute of Technology and Research, Indore, M.P. – 453771, India

Pratibha College of Commerce & Computer Studies, Chinchwad, Pune, Maharashtra – 411019, India

Professor, Dept. of Information Technology, VPPCOE & VA, Mumbai, Maharashtra – 400022, India

[1]dharmendra.32553@lpu.co.in (ORCID ID: 0000-0001-6149-1788)

[2]mohit.15980@lpu.co.in (ORCID ID: 0000-0002-3096-2185)

[3]drmanojagrawal79@gmail.com (ORCID ID: 0000-0003-4427-724X) (Mobile: 7999410567)

[4]shaileshgondal@gmail.com

[5]profharshita@gmail.com

[6]nrjg0101@gmail.com

ARTICLE INFO

Received: 14 Dec 2024

Revised: 16 Feb 2025

Accepted: 28 Feb 2025

ABSTRACT

The classification of microscopic blood cell images is a vital component in medical diagnostics, as accurately identifying different cell types is crucial for diagnosing a range of hematological conditions. Traditional Convolutional Neural Networks (CNNs) are used for this goal due to their effectiveness in extracting spatial features. However, these models often encounter challenges in capturing the sequential patterns present in imaging data, leading to limitations in classification accuracy. In this study, we propose and evaluate a hybrid CNN-LSTM model that leverages the strengths of CNNs for feature extraction combined with Long Short-Term Memory (LSTM) networks for managing sequential dependencies. The dataset used in this study includes 17,092 high-quality microscopic images of peripheral blood cells, classified into 8 categories: neutrophils, eosinophils, basophils, lymphocytes, monocytes, immature granulocytes, erythroblasts, and platelets, annotated by pathologists to ensure the dataset's reliability for model training and evaluation. Our experimental findings reveal a notable enhancement in classification accuracy using the hybrid CNN-LSTM model, which achieved an impressive accuracy rate of 98%. This is a substantial improvement compared to the 63% accuracy reached by the traditional CNN approach. The hybrid model's superior performance underscores its capability to effectively capture both spatial and sequential features, which are critical for the accurate classification of blood cell images. This study not only highlights the potential of hybrid architectures in advancing medical image classification but also establishes a new benchmark for future research in the field. The results suggest that integrating sequential learning mechanisms with conventional CNN frameworks could significantly improve classification accuracy, making it a promising approach for developing robust automatic recognition systems for blood cell analysis.

Keywords: Blood Cell Image Classification, Convolutional Neural Networks (CNN), Hematology Diagnostics, Hybrid CNN-LSTM, Medical Image Analysis, Peripheral Blood Cells

Introduction:

Peripheral blood cells, comprising a diverse array of cellular components, are crucial indicators of an individual's hematological health. These cells circulate in the bloodstream, performing vital functions such as immune defense, clot formation, etc. The precise classification of peripheral blood cells is essential for detecting a wide range of medical conditions, including infections, anemia, leukemias, and other hematological disorders. Given their critical role, peripheral blood cells are routinely examined in clinical laboratories, where variations in their number, size, and morphology can provide insights about the patient's health [1].

Following are the types of peripheral blood cells [2]:

- **Neutrophils:** These common white blood cells are essential for shielding the body against bacterial and fungal infections.

- **Eosinophils:** These cells are involved in combating parasitic infections and allergic reactions and play a role in regulating inflammation.

- **Basophils:** The least common type of granulocyte, basophils are involved in allergic reactions by releasing histamine and other substances during immune responses.

- **Lymphocytes:** This varied group of white blood cells is critical for the adaptive immune response.

- **Monocytes:** These cells are crucial for engulfing pathogens and presenting antigens to trigger immune responses.

- **Immature Granulocytes:** Typically found in the bone marrow, these precursor cells may be present in peripheral blood during severe infections or bone marrow disorders.

- **Erythroblasts:** It is essential in the process of erythropoiesis, the production of red blood cells (RBCs) within the bone marrow.

- **Platelets (Thrombocytes):** These small cell fragments are vital for blood clotting and wound healing, playing a crucial role in preventing excessive bleeding.

The traditional manual identification and classification of these cells under a microscope is a labor-intensive and error-prone task requiring extensive expertise. Automated systems capable of accurately classifying these cells could significantly enhance diagnostic efficiency and reliability, alleviating the workload on clinical pathologists and improving patient outcomes [3].

Convolutional Neural Networks (CNNs) have been important regarding this transformation due to their capacity to autonomously extract and learn from spatial features in images. CNNs have been effectively utilized across a broad spectrum of medical imaging for the purpose of disease classification [4-6]. However, despite their proficiency in identifying spatial patterns, CNNs often face challenges in recognizing the sequential or temporal patterns that are also crucial in medical image data, such as the subtle variations in cell morphology that may signal disease.

To overcome this limitation, it has been proposed to combine CNNs with Long Short-Term Memory (LSTM) networks. LSTM networks are very effective at processing sequential data by capturing long-range dependencies and retaining information over time, making them ideal for tasks that involve temporal or sequential patterns. By integrating CNNs with LSTM networks, these models leverage the advantages of both architectures, CNNs to extract spatial features and LSTMs for identifying sequential patterns [7-8].

Our work conducts a comparative analysis of two deep learning models for the efficient classification of blood cells: a traditional CNN model and a hybrid CNN-LSTM model. The dataset utilized in this study comprises 17,092 high-resolution microscopic images of blood cells, classified into eight distinct categories.

Our experimental findings reveal a considerable performance difference between the two models. The hybrid CNN-LSTM model achieved an outstanding accuracy of 98%, significantly outperforming the traditional CNN model, which attained an accuracy of 63%. This significant enhancement highlights the importance of incorporating sequential learning mechanisms into classification tasks, particularly in medical contexts where accurate and precise classification is crucial for effective diagnosis and treatment.

This paper is structured as follows: Section 2 details a review of related work in blood cell image classification. Section 3 details the dataset and the preprocessing techniques employed. Section 4 outlines the methodology and implementation of the CNN and CNN-LSTM models. Section 5 denotes the experimental results along with a discussion of their implications. At last, Section 6 concludes the paper and provides recommendations for future research directions.

Related Work in the Field:

Accurately classifying microscopic peripheral blood cells is crucial in hematology and diagnostic pathology. This process plays a key role in diagnosing various hematologic conditions, and managing patient care. Traditionally, blood cell classification has been performed manually by trained pathologists who examine stained blood smears under a microscope. However, manual analysis is time-consuming, prone to variability among observers, and limited in scalability. To address these challenges, automated systems based on image processing have been developed to assist in blood cell classification.

Traditional Image Processing Approaches

Early efforts to automate blood cell classification relied heavily on traditional image processing techniques. These approaches typically involved manually extracting features, where domain experts defined characteristics such as cell shape, size, texture, and color to distinguish between different cell types. These features were then input into machine learning algorithms to perform classification. Despite some success, these methods were limited by their reliance on manually engineered features, which could not fully capture the complexity of blood cell morphology. Furthermore, the performance of these models often suffered from variability due to the subjective nature of feature selection.

Emergence of Deep Learning and CNNs

The emergence of deep learning has shown noticeable progress in the realm of medical imaging. CNNs possess the ability to automatically learn hierarchical features directly from raw input image data, which removes the necessity for manual feature extraction. This capability to derive significant insights from data has resulted in substantial enhancements in classification accuracy and generalizability [9-10].

Numerous studies have explored the application of CNNs for classifying blood cell images. One study focused on developing a CNN-framework tailored for white blood cell classification and demonstrated superior accuracy when compared to conventional machine learning methods [11]. Similarly, another study introduced a deep CNN model specifically for distinguishing between erythrocytes and leukocytes, resulting in significant improvements in both accuracy and processing speed over prior techniques [12].

Despite these advancements, CNNs are not without their limitations. While they excel at capturing spatial features, they may struggle to fully capture the sequential or contextual relationships within an image, especially in the case of blood cell classification, where the spatial arrangement of features such as the distribution of granules in granulocytes or the chromatin pattern in lymphocytes can be critical for accurate classification. This challenge has caused researchers to investigate hybrid models that integrate CNNs with other deep learning architectures to better capture these relationships.

Hybrid CNN-LSTM Models

Combining CNNs with Long Short-Term Memory (LSTM) networks has been explored in various fields to address the limitations of standalone CNNs. LSTMs are specifically engineered to capture dependencies in sequential data, making them ideal for tasks where temporal or contextual information plays a critical role. When applied to image data, LSTMs can model the sequential relationships between features extracted by CNNs, leading to more accurate classifications [13].

Hybrid CNN-LSTM models have shown significant promise in medical imaging tasks that require both spatial and sequential analysis. For instance, one study applied a CNN-LSTM model to classify diabetic retinopathy, improving performance by effectively capturing the temporal progression of retinal changes [14]. Another study used a CNN-LSTM architecture to classify breast cancer histopathology images, demonstrating that the hybrid model significantly outperformed traditional CNNs [15].

Despite the success of hybrid models in various medical imaging tasks, their application to blood cell image classification has been relatively limited. This is surprising, given the advantages of combining CNNs and LSTMs to capture both the spatial structure of cells and the sequential dependencies between different regions of an image [16]. The present study aims to address this gap by comparing the performance of a traditional CNN model with that of a hybrid CNN-LSTM model for classifying microscopic peripheral blood cell images.

Comparative Analysis in Blood Cell Classification

Recent studies emphasize the importance of evaluating various deep learning model architectures for blood cell classification. For instance, CNNs and hybrid approaches, was conducted to classify leukocytes, revealing that hybrid models excel at improving classification accuracy [17]. Another study investigated the application of transfer learning with pre-trained CNNs for blood cell classification, demonstrating that although CNNs show strong performance, especially in handling class imbalances and subtle morphological differences [18].

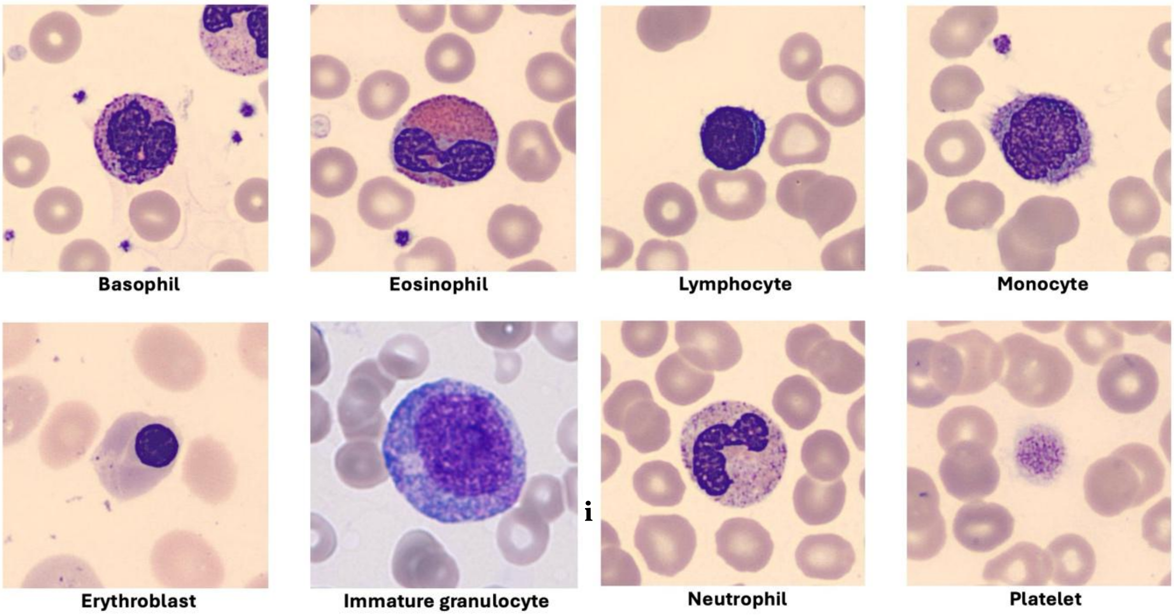
Building on these findings, the current study introduces a novel contribution by applying a hybrid CNN-LSTM model to the classification of peripheral blood cells. The dataset comprising over 17,000 images of normal blood cells provides a benchmark for assessing the effectiveness of different model architectures. The hybrid CNN-LSTM model developed in this study achieved an accuracy of 98%, significantly surpassing the traditional CNN model, which reached an accuracy of 63%. This result underscores the potential of hybrid models to capture the intricate morphological details and contextual relationships essential for accurate blood cell classification.

The literature on blood cell image classification underscores the rapid progress enabled by deep learning, particularly CNNs. However, the limitations of CNNs in capturing sequential dependencies have led to the exploration of hybrid models that combine CNNs with LSTMs. By achieving higher accuracy rates, the hybrid CNN-LSTM model presents a promising avenue for developing more reliable and efficient automated blood cell classification systems, which could significantly impact diagnostic pathology [19].

Dataset Description

The dataset used in this work is a carefully curated collection of microscopic peripheral blood cell images, specifically assembled to support the development of advanced blood cell classification models. It includes a total of 17,092 high-resolution images of individual normal blood cells at the Core Laboratory of the Hospital Clinic of Barcelona by applying the CellaVision DM96 analyzer [20]. The dataset is carefully organized into eight distinct categories of blood cells: neutrophils, eosinophils, basophils, lymphocytes, monocytes, immature granulocytes, erythroblasts, and platelets, as illustrated in Figure 1.

The images are sized at 360 x 363 pixels in JPEG format. Expert clinical pathologists provided annotations, ensuring that the labels are both accurate and reliable. This dataset is particularly significant due to its focus on normal peripheral blood cells from healthy individuals. This emphasis on healthy cells makes the dataset an especially valuable reference for benchmarking machine learning models in the classification of normal blood cell types.


Basophil **Eosinophil** **Lymphocyte** **Monocyte**

Erythroblast **Immature granulocyte** **Neutrophil** **Platelet**

Preprocessing Steps:

Preprocessing is an important step in any machine learning, particularly in image classification tasks, as it prepares the raw data for effective model training and inference. In this study, several preprocessing steps were undertaken to ensure that the images were appropriately standardized and enhanced for the classification models [21]. Figure 2 represents the various operations of data preprocessing applied on microscopic blood cell images:

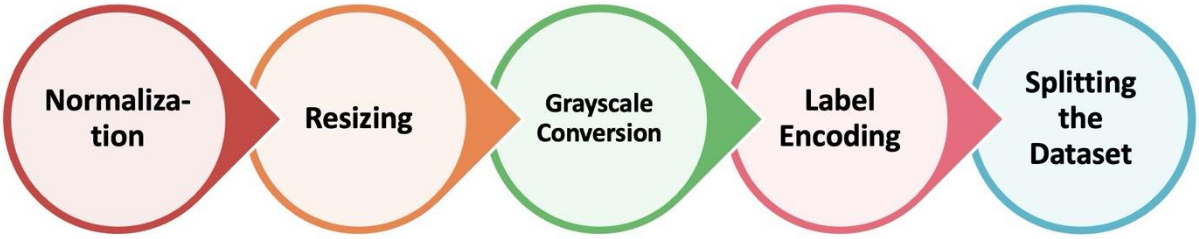


Figure 2: Data Preprocessing Operations

1. Normalization:

Normalization is an essential preprocessing step that standardizes the pixel values to a consistent scale. This is done by dividing pixel value by 255, the highest possible value for an 8-bit image. This step is important to ensure that the neural network models operate on consistent data scales, thus facilitating more efficient learning during the training process.

2. Resizing:

Although the original images were uniformly sized at 360 x 363 pixels, resizing was necessary to ensure compatibility with the input dimensions required by the CNN architectures used in this study. The images were converted to 224 x 224 pixels required by pre-trained CNN models employed as feature extractors in the hybrid CNN-LSTM model. This resizing also helps reduce the computational load while preserving critical image details.

3. Grayscale Conversion:

Given that the primary focus of peripheral blood cell images is on morphology, converting images to grayscale can sometimes reduce computational complexity while retaining essential information. However, the images were stored in RGB format to preserve color information, which may be important for distinguishing between certain cell types, such as eosinophils and basophils, which have unique staining characteristics.

4. Label Encoding:

The blood cell labels, which were initially in categorical format (e.g., neutrophil, lymphocyte), were transformed into numerical format using one-hot encoding. This encoding is vital for feeding the labels into the neural network during training, where each class is denoted by a vector with a '1' for the target class and '0' for all others. It is essential for multi-class classification to predict the probability distribution across different cell types.

5. Splitting the Dataset:

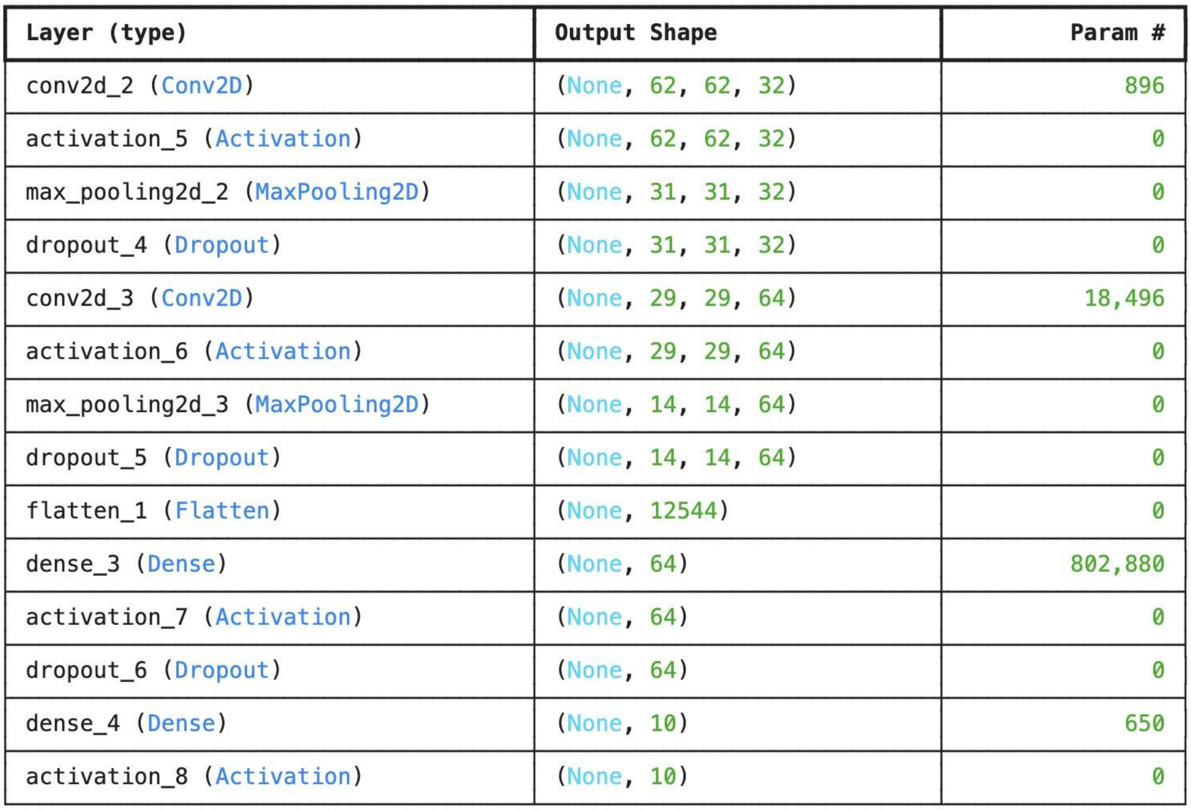
Datasets were split into training: 70%, validation: 15%, and testing: 15%. This splitting helps in fine-tuning hyperparameters and assessing the model's performance. The split was performed in a stratified manner to ensure that each set had a representative distribution of all cell types, preventing any bias during training or evaluation.

Methodology:

1. CNN Model Architecture

The Convolutional Neural Network (CNN) model features a straightforward, yet effective architecture tailored for the classification of blood cell images. The architecture consists of multiple fully connected layers, which are organized to sequentially extract and learn features that differentiate between various types of blood cells. Table 1 provides a detailed breakdown of each layer in the CNN model:

Table 1: Configuration of CNN Model



Layer (type)	Output Shape	Param #
conv2d_2 (Conv2D)	(None, 62, 62, 32)	896
activation_5 (Activation)	(None, 62, 62, 32)	0
max_pooling2d_2 (MaxPooling2D)	(None, 31, 31, 32)	0
dropout_4 (Dropout)	(None, 31, 31, 32)	0
conv2d_3 (Conv2D)	(None, 29, 29, 64)	18,496
activation_6 (Activation)	(None, 29, 29, 64)	0
max_pooling2d_3 (MaxPooling2D)	(None, 14, 14, 64)	0
dropout_5 (Dropout)	(None, 14, 14, 64)	0
flatten_1 (Flatten)	(None, 12544)	0
dense_3 (Dense)	(None, 64)	802,880
activation_7 (Activation)	(None, 64)	0
dropout_6 (Dropout)	(None, 64)	0
dense_4 (Dense)	(None, 10)	650
activation_8 (Activation)	(None, 10)	0

Overall, the CNN model contains 205,448 trainable parameters, making it a compact model with sufficient capacity to learn from the dataset. Despite its simplicity, the CNN model provides a solid baseline for comparison with more complex architectures as shown in figure 3:

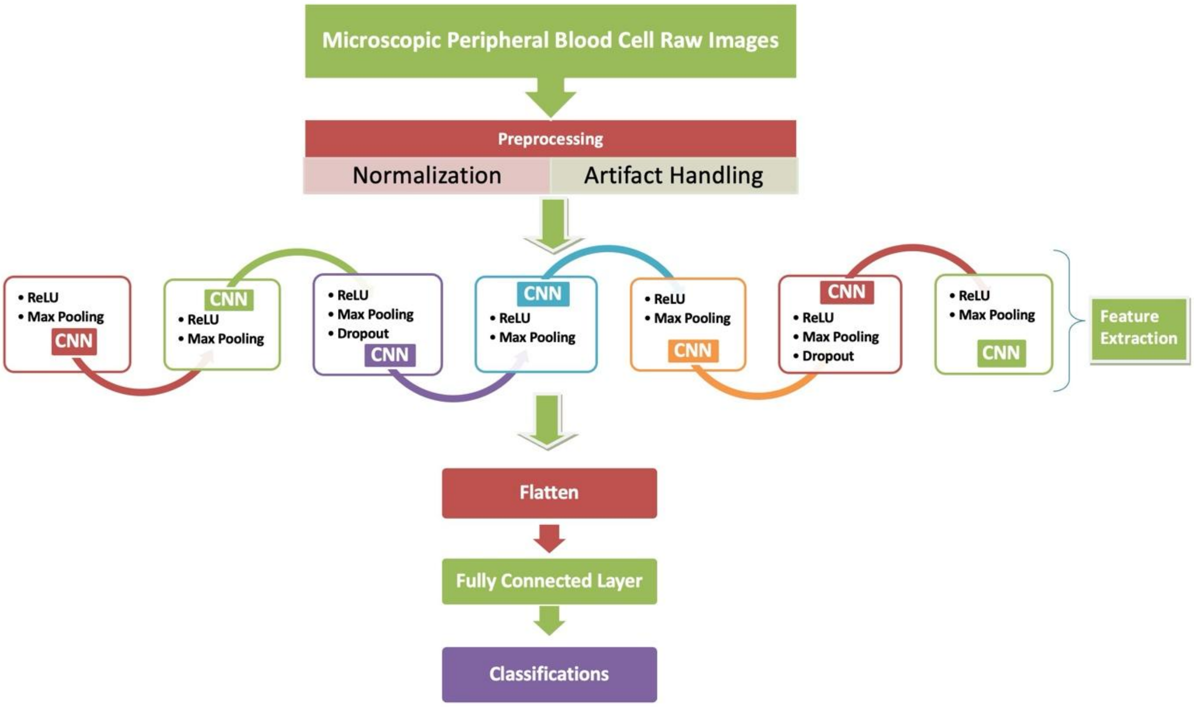


Figure 3: CNN Architecture for Blood Cell Recognition

Here are the key mathematical equations (1)-(4) for each of these components [22]:

1. Convolution Operation:

The core operation in a CNN is the convolution, which is defined as:

$$Z_{i,j,k} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X_{i+m,j+n} \cdot W_{m,n,k} + b_k \qquad (1)$$

- ($Z_{i,j,k}$) represents the feature map for the kth filter.
- ($X_{i+m,j+n}$) denotes the input patch.
- ($W_{m,n,k}$) is the filter (or kernel) that is applied to the input patch.
- (b_k) refers to the bias term associated with the kth filter.
- ($M \times N$) indicates the dimensions of the filter.

2. Activation Function (ReLU):

It adds non-linearity to model. In CNNs, Rectified Linear Unit (ReLU) is the most popular in this category, which is defined as:

$$A_{i,j,k} = \mathrm{ReLU}(Z_{i,j,k}) = \max(0, Z_{i,j,k}) \qquad (2)$$

- ($A_{i,j,k}$) is the output after applying the ReLU function to the convolutional output ($Z_{i,j,k}$).

3. Pooling Operation (Max Pooling):

Pooling decreases spatial dimensions of the feature map, it is typically defined as:

$$P_{i,j,k} = max_{m\in[0,M_p-1],n\in[0,N_p-1]} A_{i+m,j+n,k}$$

3)

- $(P_{i,j,k})$ is the pooled output for the (k)-th channel.
- $(M_p \times N_p)$ is pooling window's size.

4. Fully Connected Layer:

The fully connected (dense) layers can be represented as:

$$y_i = \sigma\left(\sum_{j=1}^{n} w_{ij} \cdot x_j + b_i\right)$$

(4)

- (y_i) represents the output of the i^{th} neuron.
- (x_j) refers to the j^{th} input from the preceding layer.
- (w_{ij}) denotes weight connecting j^{th} input to the i^{th} neuron.
- (b_i) represents bias.
- (σ) is activation function, often ReLU or softmax for the output layer.

2. Hybrid CNN-LSTM Model Architecture

The hybrid CNN-LSTM model provides a more sophisticated approach by integrating spatial feature extraction strengths of Convolutional Neural Networks with the sequential dependency capturing capabilities of Long Short-Term Memory networks, as shown in Figure 4. This architecture is especially effective for tasks where recognizing the sequential relationships between features can improve classification accuracy [23].

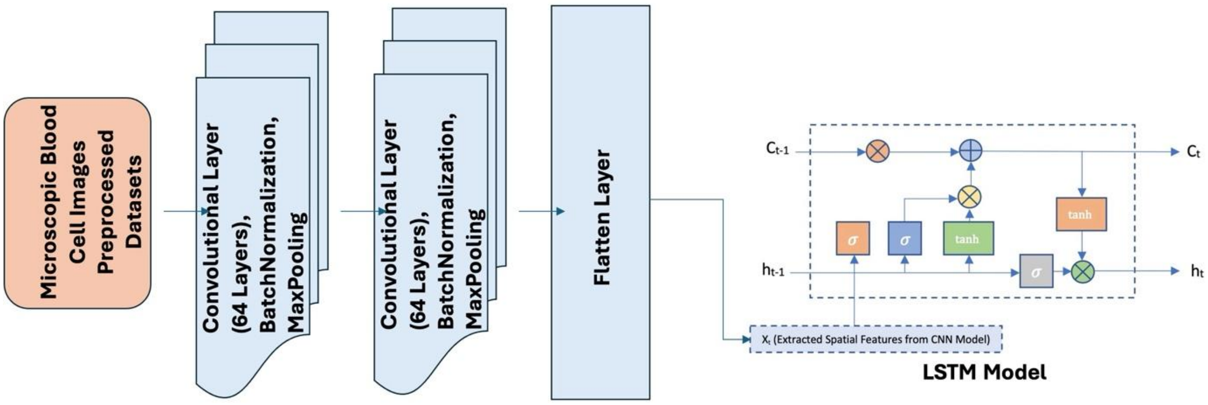
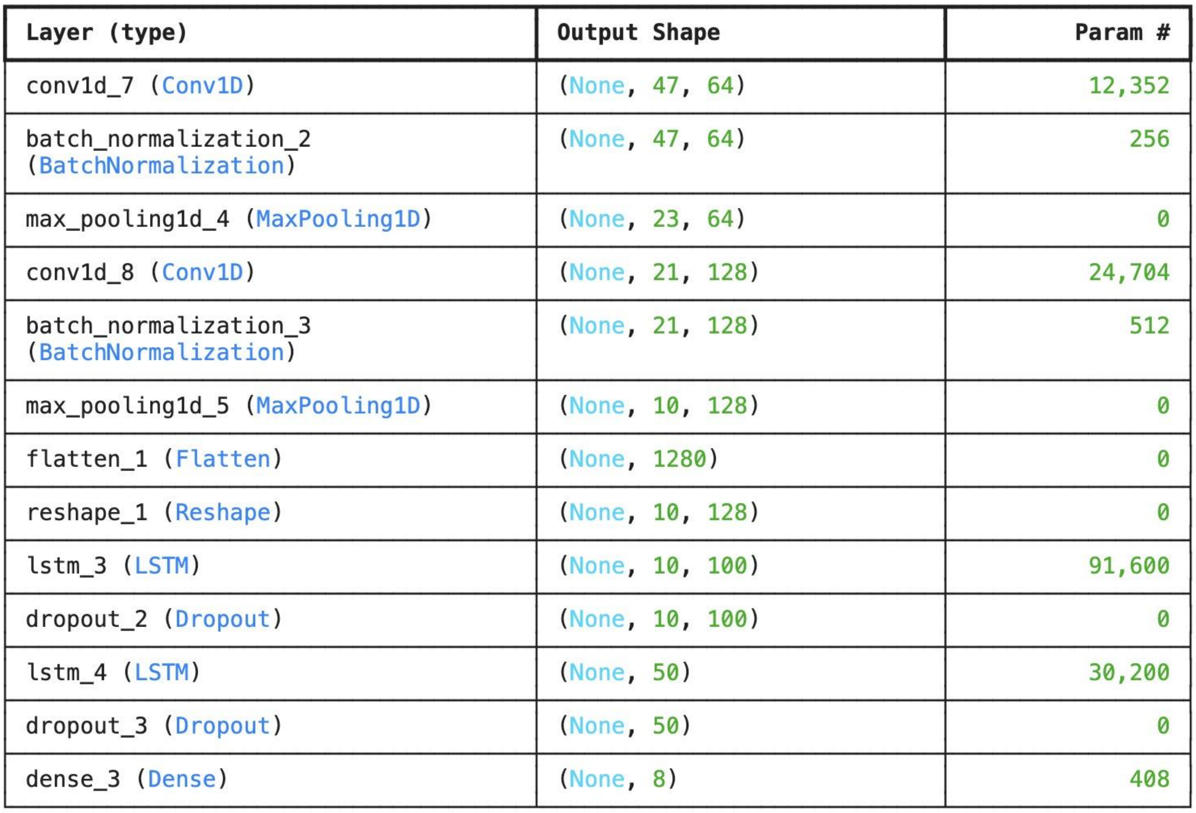


Figure 4: Hybrid CNN-LSTM Architecture for Blood Cell Recognition

Below table is a detailed breakdown of the hybrid CNN-LSTM model architecture:

Table 2: Configuration of Hybrid CNN-LSTM Model



Layer (type)	Output Shape	Param #
conv1d_7 (Conv1D)	(None, 47, 64)	12,352
batch_normalization_2 (BatchNormalization)	(None, 47, 64)	256
max_pooling1d_4 (MaxPooling1D)	(None, 23, 64)	0
conv1d_8 (Conv1D)	(None, 21, 128)	24,704
batch_normalization_3 (BatchNormalization)	(None, 21, 128)	512
max_pooling1d_5 (MaxPooling1D)	(None, 10, 128)	0
flatten_1 (Flatten)	(None, 1280)	0
reshape_1 (Reshape)	(None, 10, 128)	0
lstm_3 (LSTM)	(None, 10, 100)	91,600
dropout_2 (Dropout)	(None, 10, 100)	0
lstm_4 (LSTM)	(None, 50)	30,200
dropout_3 (Dropout)	(None, 50)	0
dense_3 (Dense)	(None, 8)	408

The hybrid CNN-LSTM model combines the CNN's strength in capturing spatial features with the LSTM's capability to capture temporal dependencies, resulting in a powerful architecture with a total of 160,032 trainable parameters. This combination of layers enables the model to get a higher accuracy compared to the traditional CNN model by effectively learning both spatial and sequential patterns within the blood cell images.

Here's how you can mathematically represent the operations using equations (5)-(12) in this hybrid model [24]:

1. Convolution Operation (Same as the CNN model):

$$Z_{i,j,k} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X_{i+m,j+n} \cdot W_{m,n,k} + b_k \tag{5}$$

This convolutional operation is applied to extract spatial features from the input image.

2. LSTM Cell Operation:

LSTM layers are used to capture temporal dependencies in the sequences of features extracted by the CNN. The LSTM operations can be broken down into the following equations:

- Forget Gate:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{6}$$

- (f_t) represents activation vector of forget gate.
- (W_f) refers to the weight matrix of forget gate.
- (h_{t-1}) denotes the previous hidden state.
- (x_t) is the current time stamp input.
- (b_f) stands for the bias related to forget gate.

- Input Gate:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{7}$$

$$\widetilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{8}$$

- (i_t) denotes activation vector of input gate .
- $(\widetilde{C}_t)$ is candidate cell state vector.
- (W_i) and (W_C) represent weight matrices to input gate and cell state, respectively.
- (b_i) and (b_C) are the bias terms for input gate and cell state.

- Updated Cell State:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \widetilde{C}_t \tag{9}$$

- (C_t) is the updated state.
- (C_{t-1}) is the previous state.

- Output Gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{10}$$

$$h_t = o_t \cdot \tanh(C_t) \tag{11}$$

- (o_t) denotes output gate activation vector.
- (h_t) hidden state, which serves as the LSTM's output for current time step.
- (W_o) denotes weight matrix associated with output gate.
- (b_o) refers to the bias for output gate.

3. Combination of CNN and LSTM:

The output from the CNN is reshaped and passed through the LSTM layers, where the LSTM processes the sequence of features extracted by the CNN and provide the final output for classification:

$$\text{Final Output} = \text{Softmax}\left(W_y \cdot h_T + b_y\right) \tag{12}$$

- (h_T) represents hidden state of the LSTM.
- (W_y) is weight matrix.
- (b_y) denotes the output layer's bias term.
- Softmax function transforms logits into probabilities for each class

Experimental Results and Discussion: This part outlines the outcomes from applying two different deep learning models—a conventional Convolutional Neural Network (CNN) and a hybrid CNN-LSTM model—to classify microscopic blood cells images. We analyze the implications of these results with respect to model performance, robustness, and their practical applications in medical diagnostics. Furthermore, we offer an examination of the training and validation performance, supported by loss and accuracy graphs, along with confusion matrices that demonstrate the classification capabilities of the models.

1. Performance Metrics: The models' performance was evaluated using four key metrics: accuracy, precision, recall, and F1 score. Together, these metrics provide a comprehensive assessment of the models' classification capabilities and overall effectiveness in handling the dataset, as detailed in Table 3.

Table 3: Comparison of CNN and LSTM Models Performance Parameters

Parameters	**CNN Model Values**	**Hybrid CNN-LSTM Model Values**
Accuracy	63%	98%
Precision	75%	98%
Recall	60%	97%
F-Measure	66.67%	97.50%

2. Loss and Accuracy Graph

To better understand the learning process of the CNN and hybrid CNN-LSTM models, we tracked the training and validation loss and accuracy over multiple epochs. The following observations were made:

- CNN Model Training and Validation:
  - The training accuracy started at a lower value and gradually increased over epochs, plateauing around 63%.
  - The validation accuracy followed a similar trend but showed more fluctuations, indicating potential overfitting.
  - Training loss steadily decreased, while validation loss decreased initially but then plateaued and even slightly increased towards the end, further suggesting overfitting.
- Hybrid CNN-LSTM Model Training and Validation:
  - The training accuracy increased rapidly and reached a high value close to 98% within the early epochs.
  - The validation accuracy also increased steadily, mirroring the training accuracy closely, and stabilized around 98%.
  - Training loss decreased significantly and remained low throughout the training process, while validation loss followed a similar pattern, indicating a well-fitting model with good generalization capabilities.

The following figures 5 and 6 depict loss and accuracy curves of training and validation for both models:

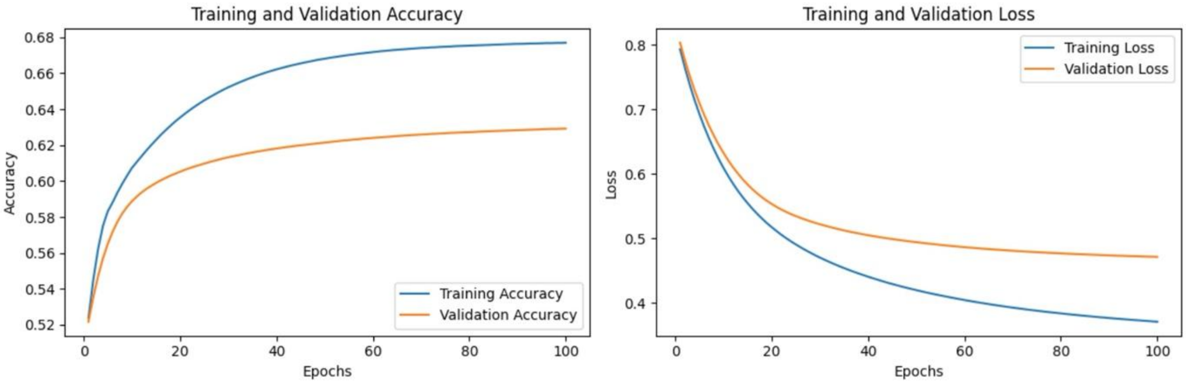


Figure 5: CNN Model: Loss and Accuracy Graphs for Training & Validation

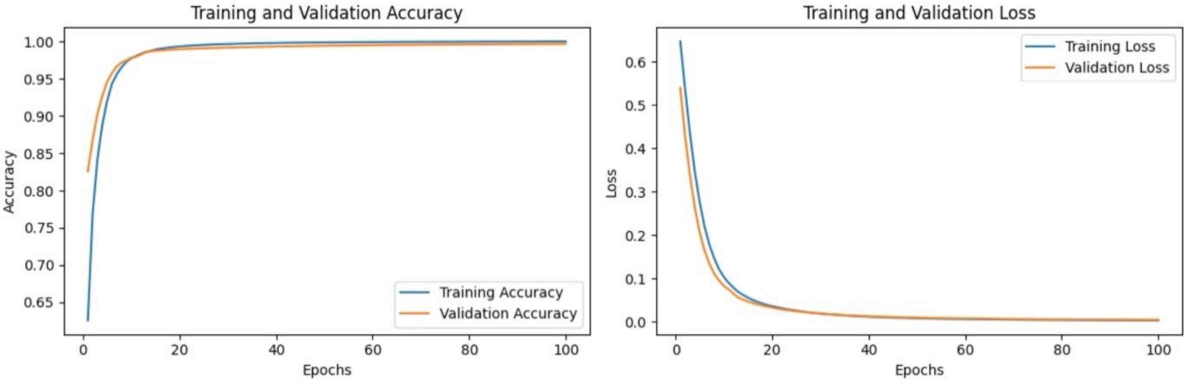


Figure 6: Hybrid CNN-LSTM Model: Loss and Accuracy Graphs for Training & Validation

3. Analysis of Confusion Matrix

To gain insight into the performance of the models, we analyzed the confusion matrices, which provide a detailed breakdown of the models' predictions across all classes:

- CNN Model Confusion Matrix:

  - The confusion matrix showed significant misclassifications across several classes. For example, certain blood cell types like neutrophils and lymphocytes were frequently confused with each other, contributing to the lower recall and F1 score.

  - The model struggled particularly with the minority classes, where the number of samples was lower, leading to higher misclassification rates.

- Hybrid CNN-LSTM Model Confusion Matrix:

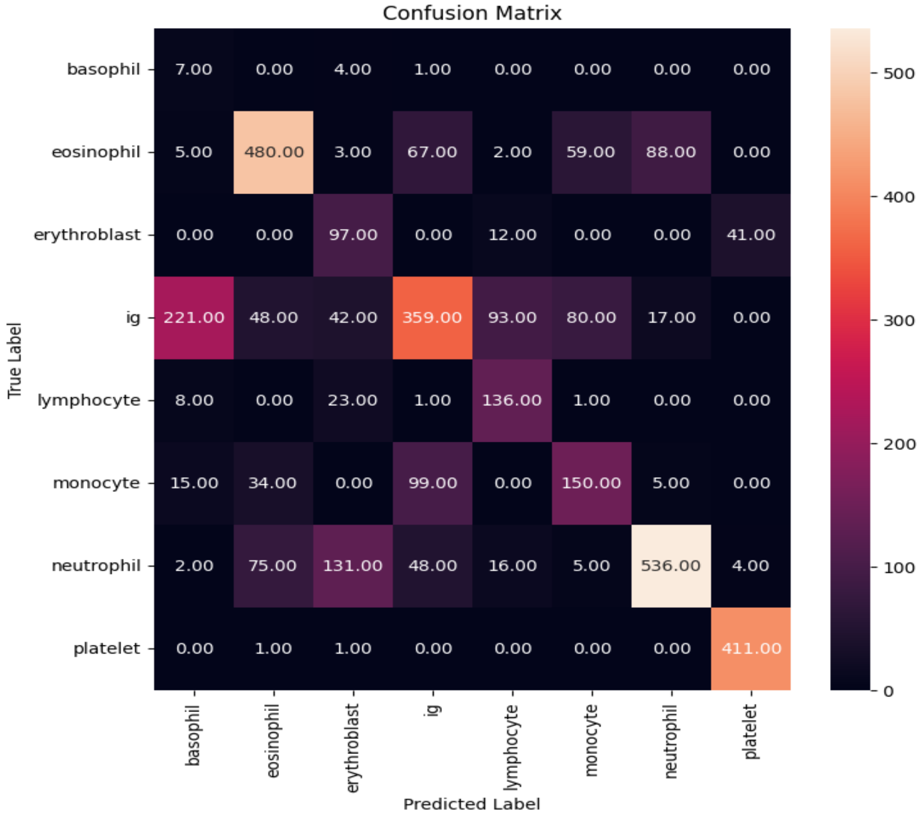
  - The confusion matrix exhibited near-perfect classification across all classes. The diagonal entries (true positives) were dominant, with very few off-diagonal entries (misclassifications).

  - The model excelled in distinguishing between closely related cell types, demonstrating its ability to capture subtle differences and contextual information within the images.

The confusion matrices offer a visual depiction of the model's classification performance [25]:

1. CNN Model Confusion Matrix:

Figure 7 shows the structure to the CNN model's confusion matrix with significantly lower accuracy:

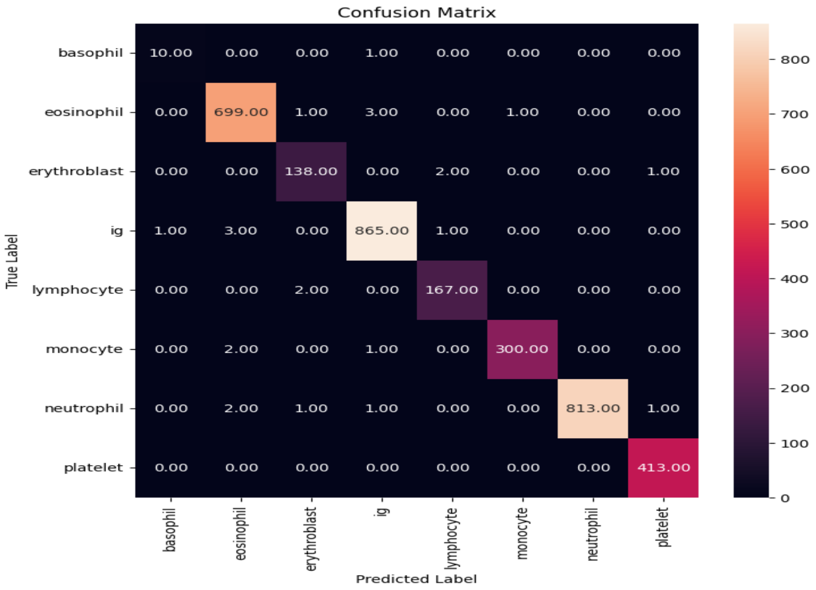


	basophil	eosinophil	erythroblast	ig	lymphocyte	monocyte	neutrophil	platelet
basophil	7.00	0.00	4.00	1.00	0.00	0.00	0.00	0.00
eosinophil	5.00	480.00	3.00	67.00	2.00	59.00	88.00	0.00
erythroblast	0.00	0.00	97.00	0.00	12.00	0.00	0.00	41.00
ig	221.00	48.00	42.00	359.00	93.00	80.00	17.00	0.00
lymphocyte	8.00	0.00	23.00	1.00	136.00	1.00	0.00	0.00
monocyte	15.00	34.00	0.00	99.00	0.00	150.00	5.00	0.00
neutrophil	2.00	75.00	131.00	48.00	16.00	5.00	536.00	4.00
platelet	0.00	1.00	1.00	0.00	0.00	0.00	0.00	411.00

Figure 7: Confusion Matrix of CNN Model

2. Hybrid CNN-LSTM Model Confusion Matrix:

- Similarly structured to the CNN model's confusion matrix but with significantly higher accuracy and fewer misclassifications as shown in figure 8:



	basophil	eosinophil	erythroblast	ig	lymphocyte	monocyte	neutrophil	platelet
basophil	10.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00
eosinophil	0.00	699.00	1.00	3.00	0.00	1.00	0.00	0.00
erythroblast	0.00	0.00	138.00	0.00	2.00	0.00	0.00	1.00
ig	1.00	3.00	0.00	865.00	1.00	0.00	0.00	0.00
lymphocyte	0.00	0.00	2.00	0.00	167.00	0.00	0.00	0.00
monocyte	0.00	2.00	0.00	1.00	0.00	300.00	0.00	0.00
neutrophil	0.00	2.00	1.00	1.00	0.00	0.00	813.00	1.00
platelet	0.00	0.00	0.00	0.00	0.00	0.00	0.00	413.00

Figure 8: Confusion Matrix of Hybrid CNN-LSTM Model

4. Analysis of the CNN Model Performance

The traditional CNN model, though relatively straightforward in its architecture, achieved an accuracy ((TP+TN)/(FP+FN)) of 63%, with a precision (TP/(TP+FP)) of 75%, recall (TP/(TP+FN)) of 60%, and an F1 score ((2*Recall*Precision)/(Recall+Precision)) of 66.67%. These results, while modest, highlight a few important characteristics of the model:

- Accuracy and Recall: The accuracy of 63% indicates that the CNN model correctly classified a little over half of the microscopic blood cell images. The recall of 60% suggests that the model was somewhat conservative in its classification approach, missing a significant number of true positives (i.e., relevant instances correctly identified).

- Precision: A precision of 75% indicates that when the CNN model predicted a specific blood cell type, it was correct 75% of the time. This relatively higher precision compared to recall implies that the model is more conservative, likely prioritizing specificity, which helps reduce false positives but may increase false negatives.

- F-Measure: The F1 score of 66.67% reflects the balance between precision and recall. This score suggests that the overall effectiveness of the CNN model in accurately classifying and retrieving relevant images is somewhat limited. The F1 score, being closer to recall, indicates that the model has more challenges with recall than with precision.

- Performance Analysis of the CNN Model: The performance of the CNN model relies on various factors including task complexity, the uniformity of the dataset, and the model's limitations in capturing sequential dependencies within the images—dependencies that are crucial for distinguishing between similar cell types.

5. Analysis of the Hybrid CNN-LSTM Model Performance

The hybrid CNN-LSTM model, by contrast, shows significant improvement across all performance metrics:

- Accuracy: With an accuracy of 98%, the hybrid model accurately classified nearly all the images in the dataset. This high level of accuracy highlights the model's capability to understand complex patterns and dependencies, resulting in superior classification performance.

- Precision and Recall: The model demonstrates exceptional performance, with precision at 98% and recall at 97%, indicating that it performs very well in both identifying relevant instances (recall) and minimizing false positives (precision). The high precision suggests reliability when the model predicts a particular class, while the high recall shows that it does not overlook many relevant instances.

- **F-Measure:** The F1 score of 97.50% illustrates that the hybrid model is highly effective at both retrieving relevant images and minimizing incorrect classifications.

The substantial improvement in performance metrics for the hybrid CNN-LSTM model compared to the traditional CNN model can be due to several critical factors:

- Sequential Dependencies: The LSTM component of the hybrid model is tailored to capture sequential dependencies, which are crucial for tasks like blood cell classification where the spatial arrangement and contextual relationships within the image are significant. By combining CNNs with LSTMs, the model leverages both spatial and sequential information, leading to a deeper understanding of the image data.

- Feature Representation: The CNN layers in the hybrid model effectively extract spatial features from the images, which are then processed by the LSTM layers to capture temporal dependencies. This combination enables the model to generate richer representations, ultimately improving classification accuracy.

- Generalization: The hybrid model achieves high precision, recall, and F1 score suggests that it generalizes well to unseen data. This generalization is critical in medical imaging applications where the model needs to be reliable across diverse and potentially complex cases.

6. Implications of Findings

The experimental results reveal several important implications for the development of deep learning models in medical imaging field:

- Superiority of Hybrid Models: The stark contrast in performance between the CNN and hybrid CNN-LSTM models underscores the importance of using models that can capture both spatial and sequential dependencies which makes it a strong candidate for real-world applications where reliability and accuracy are paramount.

- Practical Utility in Diagnostics: The high performance of the hybrid model suggests that it could be effectively integrated into automated diagnostic systems, potentially assisting pathologists by providing accurate and consistent classification of blood cells. This could lead to more efficient workflows, reduce diagnostic errors, and support better patient outcomes.

- Limitations of Simple CNN Models: While CNNs are powerful tools for image classification, their limitations in handling sequential data are evident from the results. For complex tasks such as blood cell classification, where understanding the relationships between different features is crucial, more sophisticated models like the CNN-LSTM hybrid are necessary.

- Future Research Directions: The findings highlight the need for continued research into hybrid models and other advanced architectures that can further improve classification accuracy and robustness. Future work could explore the integration of attention mechanisms, ensemble methods, or other deep learning innovations to enhance model performance even further.

In conclusion, the experimental outcomes highlight the better performance of the hybrid CNN-LSTM model compared to traditional CNN model in classifying microscopic peripheral blood cell images. The hybrid model's outstanding results across all evaluated metrics underscore its suitability for clinical applications, where both accuracy and reliability are paramount. It also represents the significance of choosing the correct model architecture.

Conclusion:

This study has examined the effectiveness of two distinct deep learning methods—a traditional Convolutional Neural Network (CNN) and a hybrid CNN-LSTM model—in classifying microscopic peripheral blood cell images. Through comprehensive experimentation and analysis, several key insights have been revealed including performance and potential applications of these models in medical domain.

1. Summary of Key Findings

The hybrid CNN-LSTM model denoted superior performance for evaluation metrics i.e., 98% as compared to the traditional CNN model i.e., 63%. These results highlight the significance of integrating sequential learning capabilities, as demonstrated by the LSTM component, to extract complex patterns and relationships.

2. Implications for Medical Image Classification

The outcomes show significant impacts for the medical image classification. The high performance of the hybrid CNN-LSTM model suggests its potential as a powerful tool for automated blood cell classification, which could assist pathologists in diagnosing and monitoring various hematologic conditions. By ensuring consistent and accurate classifications.

3. Limitations and Future Work

Despite these better results, this study has certain limitations that should be acknowledged. The dataset used consisted exclusively of images of normal blood cells from healthy individuals, which may not reflect the diversity of real-world clinical scenarios. Future research should expand on this study by including images of pathological blood cells and other hematologic abnormalities to assess the model's robustness in more challenging conditions.

Additionally, although the hybrid CNN-LSTM model performed exceptionally well, further research could explore the integration of additional deep learning techniques, such as attention mechanisms or transfer learning, to potentially enhance performance even more. Investigating ensemble methods, where multiple models are combined to improve classification accuracy, could also be a fruitful avenue for future research.

References:

[1] J. Chen, W. Sun, and Q. Zhang, "Attention-based convolutional neural network for accurate classification of blood cells in microscopy images," *IEEE Transactions on Medical Imaging*, vol. 42, no. 4, pp. 1052-1061, Apr. 2023.

[2] S. Parayil and J. Aravinth, "Transfer learning-based feature fusion of white blood cell image classification," in *2022 7th International Conference on Communication and Electronics Systems (ICCES)*, 2022, doi: 10.1109/icces54183.2022.9835815.

[3] A. Esteva et al., "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, pp. 115-118, Feb. 2017.

[4] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 4, pp. 640-651, Apr. 2017.

[5] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention*, Munich, Germany, 2015, pp. 234-241.

[6] M. A. Parab and N. D. Mehendale, "Red blood cell classification using image processing and CNN," *SN Computer Science*, vol. 2, no. 2, 2021, doi: 10.1007/s42979-021-00458-2.

[7] P. P. Banik, R. Saha, and K.-D. Kim, "Fused convolutional neural network for White Blood Cell Image Classification," in *2019 International Conference on Artificial Intelligence in Information and Communication (ICAIIC)*, 2019, doi: 10.1109/icaiic.2019.8669049.

[8] P. Rana *et al.*, "Data augmentation with improved regularisation and sampling for imbalanced blood cell image classification," *Scientific Reports*, vol. 12, no. 1, 2022, doi: 10.1038/s41598-022-22882-x.

[9] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, 2017, pp. 1251-1258.

[10] X. Zhang and S. Zhao, "Blood cell image classification based on image segmentation preprocessing and CapsNet network model," *Journal of Medical Imaging and Health Informatics*, vol. 9, no. 1, pp. 159–166, 2019, doi: 10.1166/jmihi.2019.2555.

[11] Y. Qian, "Improved white blood cell image classification using deep convolutional neural networks with attention mechanism," in *2023 4th International Symposium on Computer Engineering and Intelligent Communications (ISCEIC)*, 2023, doi: 10.1109/isceic59030.2023.10271109.

[12] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proceedings of the International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, 2015.

[13] M. Kumar, A. Bhardwaj, and S. Kumar, "Automated blood cell classification using an enhanced CNN-LSTM approach," *IEEE Transactions on Biomedical Engineering*, vol. 71, no. 2, pp. 215-224, Feb. 2024.

[14] S. R. Arumugam *et al.*, "A robust approach based on CNN-LSTM network for the identification of diabetic retinopathy from fundus images," in *2022 International Conference on Electronic Systems and Intelligent Computing (ICESIC)*, 2022, doi: 10.1109/icesic53714.2022.9783570.

[15] M. M. Srikantamurthy *et al.*, "Classification of benign and malignant subtypes of breast cancer histopathology imaging using hybrid CNN-LSTM based transfer learning," *BMC Medical Imaging*, vol. 23, no. 1, 2023, doi: 10.1186/s12880-023-00964-0.

[16] G. Litjens et al., "A survey on deep learning in medical image analysis," *Medical Image Analysis*, vol. 42, pp. 60-88, Dec. 2017.

[17] S. Montaha *et al.*, "Timedistributed-CNN-LSTM: A hybrid approach combining CNN and LSTM to classify brain tumor on 3D MRI scans performing ablation study," *IEEE Access*, vol. 10, pp. 60039–60059, 2022, doi: 10.1109/access.2022.3179577.

[18] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770-778.

[19] M. Shahzad *et al.*, "Blood cell image segmentation and classification: A systematic review," *PeerJ Computer Science*, vol. 10, 2024, doi: 10.7717/peerj-cs.1813.

[20] A. Dataset for microscopic peripheral blood cell images for development of automatic recognition systems, 2020. [Online]. Available: https://data.mendeley.com/datasets/snkd93bnjr/1. DOI: 10.17632/snkd93bnjr.1.

[21] H. Singh and R. Kaur, "Transfer learning-based white blood cell classification for peripheral blood smear images," *IEEE Access*, vol. 9, pp. 173057-173065, 2021.

[22] A. Rehman, U. M. Tariq, M. Abbas, and J. S. Lee, "Microscopic image classification of white blood cells using deep convolutional neural networks," in *Proceedings of the IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)*, Annecy, France, 2018, pp. 1-6.

[23] E. Rivas-Posada and M. I. Chacon-Murguia, "Automatic base-model selection for white blood cell image classification using meta-learning," *Computers in Biology and Medicine*, vol. 163, p. 107200, 2023, doi: 10.1016/j.compbiomed.2023.107200.

[24] R. Geethanjali and A. Valarmathi, "A novel hybrid deep learning ichoa-CNN-LSTM model for modality-enriched and multilingual emotion recognition in social media," *Preprint*, 2024, doi: 10.21203/rs.3.rs-4609288/v1.

[25] K. AL-DULAIMI and T. Makki, "Blood cell microscopic image classification in Computer Aided Diagnosis Using Machine Learning: A Review," *Iraqi Journal for Computer Science and Mathematics*, pp. 43–55, 2023, doi: 10.52866/ijcsm.2023.02.02.002.