1220: VISUAL AND SENSORY DATA PROCESSING FOR REAL TIME INTELLIGENT SURVEILLANCE SYSTEM

Check for updates

WBCs-Net: type identification of white blood cells using convolutional neural network

Neeraj Baghel[1] ⬤ **· Upendra Verma[2] · Kapil Kumar Nagwanshi[3]**

Abstract

On monitoring an individual's health condition, White Blood Cells play a significant role. The opinion on blood-related disease requires the detection and description of the blood of a patient. Blood cell defects are responsible for numerous health conditions. The conventional technique of manually visualizing White Blood Cells under the microscope is a time-consuming, tedious process and its interpretation requires professionals. There are significant medical applications for an automated method for detecting and classifying blood cells and their subtypes. This work presents an automatic classification method with the help of machine learning for blood cell classification from blood sample medical images. The proposed method can identify and classify the function of each segmented White Blood Cells cell image as granular and non-granular White Blood Cells cell type. It further classifies granular into Eosinophil, Neutrophil and non-granular into Lymphocyte, Monocyte in various forms. Because of its high precision, the proposed framework includes a neural network model to detect white blood cell types. To improve the accuracy of multiple cells overlapping and increase the robustness, data augmentation techniques have been used in the proposed system. Which has improved the accuracy in binary and multi-classification of blood cell subtypes.

✉ Neeraj Baghel
nbaghel777@gmail.com

Upendra Verma
upendra.verma@nmims.edu

Kapil Kumar Nagwanshi
dr.kapil@ieee.org

[1] Dr. A.P.J. Abdul Kalam Technical University, Uttar pradesh, India

[2] SVKM'S NMIMS University, Maharashtra, India

[3] ASET, Amity University Rajasthan, Jaipur, India

⧠ Springer

1 Introduction

Blood cells consist mainly of platelets, white and red blood cells, commonly known as WBCs and RBCs respectively [31]. In the immune system of the human body, WBCs play an essential role. WBCs protect the body from extrinsic invaders and epidemic diseases like leucocytosis, leukopenia [6]. The WBCs can be classified mainly into two groups named Granular cells and Non-Granular cells [28]. Granular cells are the small particles that store enzymes that are released during the infection. It is further classified into three subtypes named Neutrophils, Eosinophil and Basophils. The Non-Granular cells are further classified into two subtypes named Monocyte and Lymphocyte. In the human body, the Monocyte helps in fighting bacteria, viruses, and other infections. Monocytes are the key elements of the immune response. Antibodies can be released by lymphocytes. In the bone marrow, they are developed and contained in the blood and lymph tissue. There is a whole branch of medicine known as haematology that studies the cause, prognosis, treatment and prevention of blood-related diseases. Haematologists usually use their expertise and traditional knowledge of haematology to classify WBCs and their subtypes [5]. This primitive data is consistently used by doctors to determine the seriousness of blood disease. Significant and essential requirements for medical diagnosis has thus become the study of the classification of white blood cells.

Figure 1 indicates the white blood cells composition. To resist foreign organisms, a particular defensive function is associated with each type of white blood cell. It is of two types granular and non-granular which is based on their shape. Further granular is classified into Eosinophil, Neutrophil, and Basophils. And non-granular is further classified into lymphocyte and monocyte. Neutrophils protect against bacteria and are one of the main white blood cells. The parasites are killed by Eosinophil and perform an allergy operation, although basophils are involved in different allergic reactions in the body. The monocyte destroys the body's damaged tissues by entering into the cell. To balance immunity mediated by cells, lymphocytes are very important and complicated cells. Lymphocytes are of two types (B-type and T-type) which are used for memorizing invasive micro-organisms and viruses, therefore the lymphocytes are entirely different from the other WBCs.

Similar to the traditional method on the classification of WBCs, automatic methods are required mainly on the classification of WBCs. In an automatic method some traditional image processing, machine learning and deep learning methods have been proposed. In traditional image processing, some features are extracted and then using mathematical relationships and controlled parameters a threshold function is defined. It is used to discriminate the features of different category which helps to make a better classifier. While in machine learning, an advanced feature like texture is extracted and the model is trained on those features. In deep learning, the automated feature engineering approach derives appearances, textures, patterns and different attributes. That play a primary role in image classification based on WBC characteristics. Image processing methods for classification, segmentation, extraction of features and recognition have been implemented. The model previously introduced by a haematology analyser [33] to identify the WBC images was time-consuming and manual. Often, it causes detection errors. Computer-aided image processing-based diagnostic machine learning algorithms such as decision tree [24], k-means [12], Recurrent Neural Network (RNN) [18], Convolutional Neural Network (CNN) [14] and support vector machine (SVM) [20]. Various types of white blood cells and their subtype classification are compared in Table 1.

(a) Classification of white blood cells

(b) Classification of Granular WBCs

(c) Classification of Non-Granular WBCs
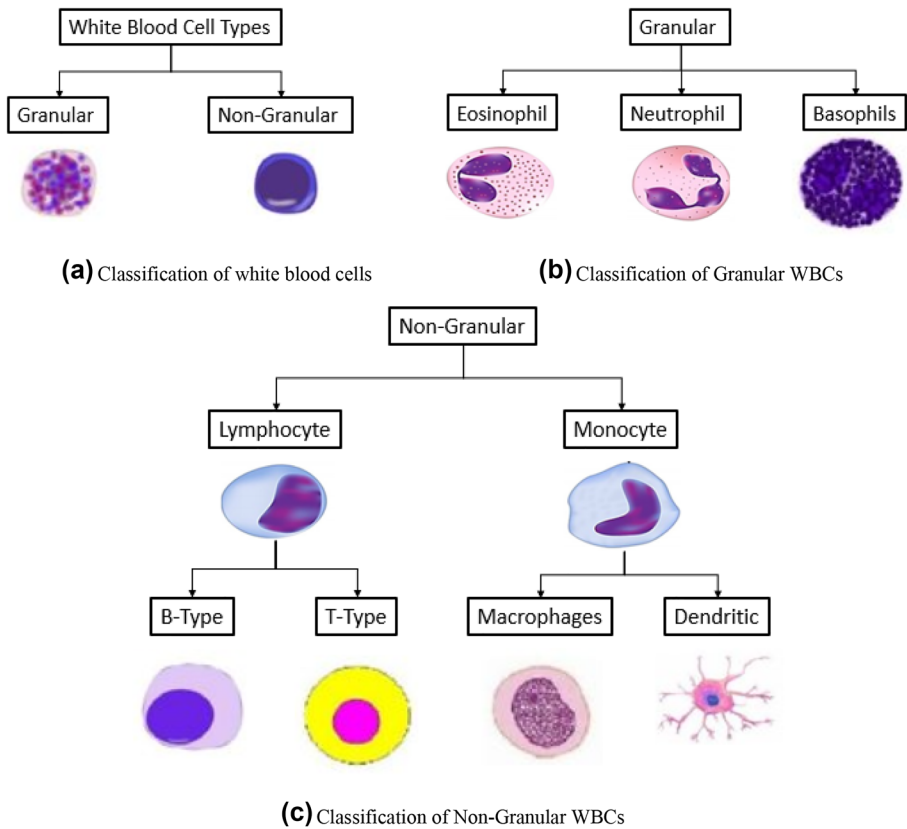
Fig. 1 White blood cell classification structure

Using the optical electron microscope on blood samples research has been introduced by Rosyadi et al. [23]. That has the potential to identify WBCs images.

Furthermore, the Ostu classification threshold approach is adopted by the authors with the KNN. Based on their study, it was concluded that discriminatory feature has developed after the implementation of clustering with k-means to identify WBCs. In [8] classification algorithms are used for the diagnosis of leukaemia where leukocyte segmentation is performed using CMYK colour space from two separate blood smears. The classification of support vector machines achieved an accuracy of 86.67%.

In Giemsasmear blood cell images using texture, histogram, and gradient region granulomatous characteristics with classifier as decision tree, Malihi et al. [17] suggested identification of malaria parasites. Pandey, C [4] Proposed Machine learning approach for image classification using SVM. Singh, A. [26] fused these features with ANN approach for classification. In [21] proposed CNN architecture for classification. In [9] another approach is used which have a classifier as naive Bayes for training the algorithm. It includes circularity, perimeter, area, eccentricity and semantic features. The Bayes classifier is used and achieved 80.88 per cent accuracy. In [25] the classification of cells with hyper-spectral imaging using a Gabor-wavelet and a CNN model is shown.

Table 1 Comparison of algorithms used for blood cell image classification

Techniques	Advantages	Description	Limitations
KNN [13]	Simple, semi-supervised learning approach	Classification of RBC and WBC in peripheral blood smear utilizing KNN	Lack of robustness, susceptible to noise
SVM [30]	SVM classifier is used to recognize WBC for future Leukaemia diagnosis	Performs a non-linear classification	Difficult to visualize and long training time for large datasets
Decision tree [8]	A model which anticipate the class by learning decision rules	Blood diseases detection using decision tree and blood cell pattern analysis	Complexity and Instability in applying regression
Naive Bayes [9]	It is Bayes Theorem, calculating the posterior probability	Automatic classification of leukocyte morphological features	Loss of accuracy. Classifier cannot modify dependencies
CNN [10]	CNN performs better in image	Classification and segmentation of Leukocyte using CNN	High computational cost, needs a million of training data
RNN [32]	It can model a sequence of data	R2U-Net for medical image processing	Can face underfitting as pixel values are distributed across the grid
CNN + RNN [15]	Developed framework highlights extracted features efficiently representing all the important details from the image	Combining CNN with RNN for blood cell image classification	Computational time is high, requires high-performance hardware

An automated machine learning (ML) approach is proposed in [1] for the identification of blood cell types as platelets, RBCs and WBCs. These types of blood cells are counted correctly by the algorithm provided. In [3] Automatic classification method for blood cells types using a transfer learning system has been suggested. The classification of acute lymphoblastic leukaemia based on WBCs segmentation using maximum feature selection with main component analysis is described in [27] with 94 per cent average accuracy.

In [29], random forest classifiers, SVM, multilayer perceptron (MLP) is compared with CNN classifier to classify WBCs image cells. In [19] for blood cell microscopic images, the automatic algorithm for classification is proposed. Discrete wavelet transformation is applied after image augmentation to obtain frequency information of the WBCs and ultimately CNN is used for classification. SVM and deep learning are used to establish irregular blood cell recognition techniques [2]. A CNN with the combination of RNN is applied in single-cell images with matrix transformation is proposed in [15] for the classification of blood cells. By integrating single image features of the cell nuclei, the CNN-RNN technique resulted in an accuracy of 90.87 per cent. CNN and the combination of CNN and RNN are mainly the subjects of the latest medical image classification research. The researchers have previously proposed different models of CNN capable of improving the accuracy of the classification, especially focusing on automatic feature learning techniques.

Blood cell defects are responsible for numerous health conditions. The conventional technique of manually visualizing White Blood Cells under the microscope is a time-consuming, tedious process and its interpretation requires professionals. This work presents an automatic classification method with the help of machine learning for blood cell classification from blood sample medical images. The proposed method can identify and classify the function of each segmented White Blood Cells cell image as granular and non-granular White Blood Cells cell type. It further classifies granular into Eosinophil, Neutrophil and non-granular into Lymphocyte, Monocyte in various forms. Which helps to identify leucocytosis, leukopenia of these sub cell types.

To classify the WBCs and achieve higher accuracy rates, these models can be further enhanced. The goal of this research is to enhance the accuracy of the classification with the extraction of discriminative features using automatic feature engineering techniques.

This work presents an automatic classification method with the help of machine learning for blood cell classification from blood samples. The proposed method can identify and classify the function of each segmented WBCs cell image as granular and non-granular WBCs cell type. It further classifies granular into Eosinophil, Neutrophil and non-granular into Lymphocyte, Monocyte in various forms. Because of its high precision, the proposed framework includes a neural network model to detect white blood cell types. To improve the accuracy of multiple cells overlap and increase the robustness, data augmentation techniques have been used in the proposed system for training binary classification of blood cell types and multi-classification of blood cell subtypes.

This paper is structured as follows: Sect. 2 describes the dataset used in this experiment Sect. 3 outlines the methods and methodology adopted to achieve the research objectives. In Sect. 4, along with the discussion, the experimental results are discussed. The last Sect. 5 ends with the conclusion and future work.

Fig. 2 Flow diagram for sample collection and dataset preparation

Table 2 Augmented WBCs dataset used for experiment

S. No	WBCs types images	WBCs sub-types images	Training images	Test images	Total
1	Granular	Eosinophil	2497	623	3120
		Neutrophil	2499	624	3123
2	Non-Granular	Lymphocyte	2483	623	3106
		Monocyte	2478	620	3098
Total			9957	2490	12447

2 Dataset

This dataset includes 12500 augmented blood cell images (JPEG format) with labels of the cell type (CSV format). This dataset contains approximately 3000 blood cell images for each 4 different sub-types of blood cell types. Granular and non-granular are the cell types. Eosinophil, neutrophil, monocyte, lymphocyte and are the four-cell sub-types. In particular, the dataset contains 2500 enhanced images and 4 additional (JPEG + CSV) sub-type labels. This dataset is a publically available standard dataset named as BCCD Dataset (https://www.kaggle.com/paultimothymooney/blood-cells) provided by github/Shenggan.

Figure 2 shows the flow diagram for sample collection and dataset preparation. In this process there are 5 steps. First step is to collect the blood sample from the subject. Second step is to take the blood sample to glass slide which is used to observe using electron microscope in step three. Step four is to acquire a good image for that sample with the help of experts in this filed. Last step is to prepare the dataset and the labelling of that sample for ground truth.

Table 2 describes the augmented dataset used for the experiment in this work. In this, there are major two types of WBCs as granular which contains 6243 images and non-granular 6204 images. Which is further sub-classified into 4 classes. In Eosinophil total of 3120 images are used (2497 images used for training and 623 images used for testing), Neutrophil total of 3123 images are used (2499 images used for training and 624 images used for testing), Lymphocyte total of 3106 images are used (2483 images used for training and 623 images used for testing) and Monocyte total of 3098 images are used (2478 images used for training and 620 images used for testing).

Figure 3 shows the various white blood cell images of different types. (a) Represents the Eosinophil type images. (b) Represents the Lymphocyte type images. (c) Represents the Monocyte type images. (d) Represents the Neutrophil type images.
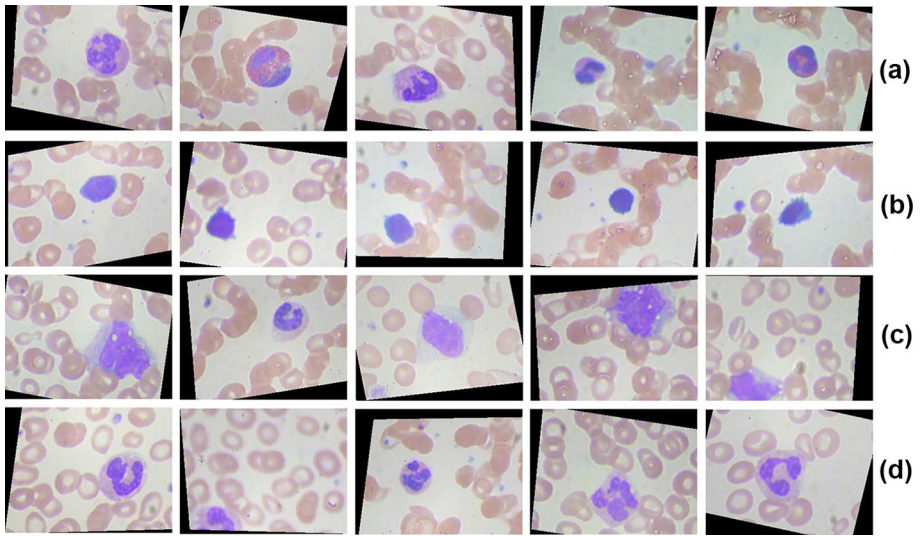
Fig. 3 Samples images of white blood cell images with augmentation

3 Proposed method

The main objective of the proposed work is to create an automatic classification model for white blood cell type detection. This section demonstrates the white blood cell image classification approach using the convolutional neural network method. The flowchart for the proposed algorithm is shown in Fig. 4. In this study, the image of blood cell samples is given as input then feature extraction is performed using a deep learning model which extracts deep features from these white blood cell images to achieve a higher rate of classification. This includes pre-processing and development of CNN architecture,
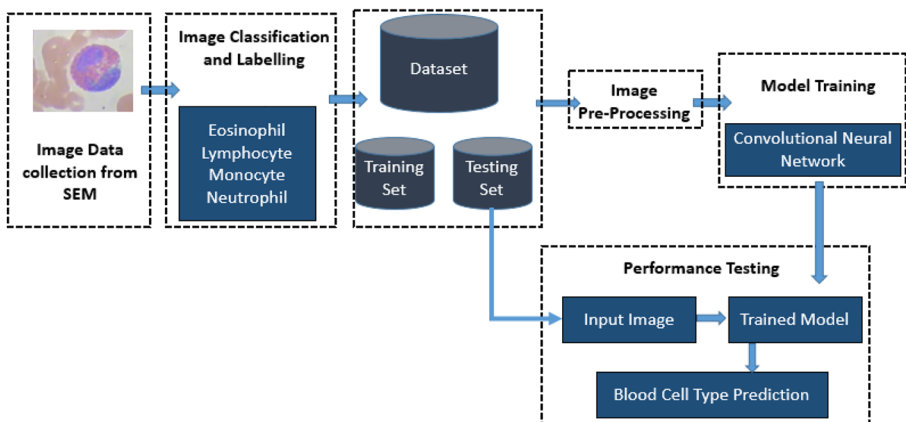

Fig. 4 Block diagram of the proposed framework for white blood cell type

accompanied by the discovery of influential features and discriminatory characteristics to differentiate these types of cells and used for training and testing.

The Flowchart of the proposed WBCs-Net model demonstrates the classification of WBC in the Fig. 4. The strategy is broken down into four parts: (a) collecting samples of blood from subjects or datasets (b) Image pre-processing and data augmentation is performed on image samples like cropping, rotation, scaling etc. (c) Proposed CNN architecture for white blood cell image multi-classification which can help for the identification and classification of Subtypes of blood cells have significant medical applications. The learning of blood cell type is carried out using model training and testing via the proposed CNN architecture. (d) Checked the results of the proposed CNN architecture for multi-classification using professional-generated ground reality.

3.1 Pre-processing on WBCs images

Pre-processing is the basic step in image processing. Dataset has a white blood cell images sample. Image pre-processing is done in this step to enhance accuracy and reduce processing time. In this step, rescaling is performed to convert image data from range 0–255 to 0–1 with the help of normalising the data by dividing the data by 255. The sheer range is used '0.2' and zoom range '0.2' is applied. Horizontal flip augmentation has also been used for better training.

3.2 CNN architecture

A deep learning model is suitably used in a variety of image processing applications. It is a neural network architecture of various kinds of layers. It is inspired by the human brain's biological visual cortex. Each CNN layer is responsible for performing the process of feature extraction with different operations and functions. Weight parameters are obtained by the CNN model by adjusting the initial weights by training on the dataset. Convolution neural networks is an architecture consist of a convolutional layer followed by a pooling and dense layer. This form of the model is useful for both types of input data such as images, audio. In this form, information is transferred from one layer to another layer. These features are processed with different hidden layers. Where the first few layers are processed the low-level features and the deeper hidden layer processed more high-level features are extracted. The CNN style, as shown in Fig. 5 is made up of a stack of different layers that relay data connected to each neuron from the input to the output level.

A detailed explanation and formulation of the various layers of the convolutional neural network are given in the following subsections.
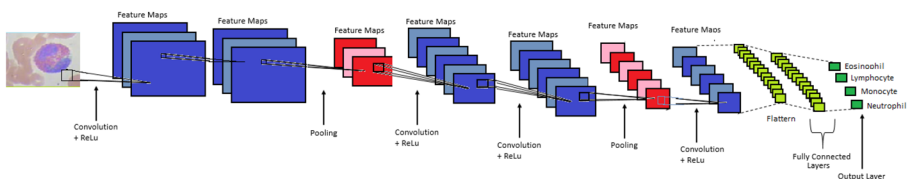
Fig. 5 Proposed CNN architecture for white blood cell type

3.2.1 The input layer

The input layer takes the raw image as an input to classify them into different types of WBC. These images consist of 3 dimensions such as height width and the number of colour channels. This sample is of size $64 \times 64 \times 3$.

3.2.2 The convolutional layer

The convolutional layer is the main feature extraction layer. The convolution layer highlighted the small changes in the images. This encourages the use of convolution neural networks in nano-scale biomedical images for pattern recognitions and identifies the optimal weights, activation function, and layer architecture. It takes input from the previous layer and extract important features and pass these feature to the next layer. This operation consists of three elements. Input layer feature, kernel, and convoluted features. The kernel slide over input WBC images from top-left to bottom right and extract the convolution features one by one. The output feature map is generated by the kernel which is multiplication over input data and inverted kernel. Also, the rectified linear unit (ReLU) activation function is used to activate/deactivate the neuron based on a function. It will provide irregularity in the neural network.

$$f(r) = \left\{ \begin{array}{l} 0, r < 0 \\ 1, r \geq 0 \end{array} \right\} \tag{1}$$

3.2.3 The pooling layer

The pooling layer is applied after this convolution and activation layer. It is used to down sample the feature map that reduces the computational and over fitting problems. Here max pooling is used with the pooling size of 2×2.

3.2.4 The fully connected layer

A fully connected layer gives desired output as a result of the network. The connection of this layer depends on two cases. If it is connected to the previous fully connected layer then a number of the parameter is defined by multiplication of parameter in previous layer and parameter in this fully connected layer. If it is connected to the convolutional layer then a number of the parameter is defined by multiplication of output feature map of the previous layer, size of kernel and number of a parameter in this fully connected layer. The output of different structures involving 2D-convolutional layers was studied while preparing the optimal CNN model for white blood cell image detection. The proposed structure, however, contains seven convolution layers that achieve optimum efficiency in the experiment. Figure 5 and Table 3 define the structure of the ideal CNN in detail.

3.3 Training and testing of the proposed model

CNN training is the main step in the neural network model, as the training approach has a significant influence on the subsequent performance of the entire system. First a separate

Table 3 The proposed CNN architecture for blood cell image multi-classification

S. No	Layers	Activation function	Output shape	Kernel- size	Filters	Strides	Weights
0	Input	–	(64, 64, 3)	–	–	–	–
1	Conv2D	ReLU	(62, 62, 16)	(3, 3)	16	1	448
2	Conv2D	ReLU	(62, 62, 32)	(3, 3)	32	1	4640
3	MaxPooling2D	–	(30, 30, 32)	(2, 2)	–	–	0
4	Conv2D	ReLU	(28, 28, 16)	(3, 3)	16	1	4624
5	Conv2D	ReLU	(28, 28, 32)	(3, 3)	32	1	4640
6	MaxPooling2D	–	(13, 13, 32)	(2, 2)	–	–	0
7	Conv2D	ReLU	(11, 11, 32)	(3, 3)	32	1	9248
8	Flatten	–	(3872)	–	–	–	0
9	Fully connected layer	ReLU	(128)	–	–	–	495744
10	Classification Layer	Softmax	(4)	–	–	–	516

Total Weights: 519860

collection of data was used for preparation, validation and testing using the proposed method. Second, using categorical cross-entropy loss, the training process is being optimized. We divide the entire training collection into so-called mini-batches due to memory constraints, and these mini-batches are used to train iterations of CNN. For a batch, 32 images are chosen randomly from the training data (without repetition) and trained for epochs. In a validation set, classification accuracy and fine-tune model parameters are used to evaluate the model, where prediction is made by slicing the test sample into overlapping TF-patches (1-frame hop), making a prediction for each TF-patch and finally selecting the prediction of the sample level as to the class with the highest mean output activation over all frames. Therefore, in 4 separate categories of white blood cell images with a training dataset, this CNN model is trained. These parameters of the model are stored and ported into low level computer systems. That low level system can be used for classification of WBCs images. The high level computers are used here for the training perspective. But this trained model don't require such type of high specification computers for classification.

For the multi-classification of white blood cell types, the hyperparameters used while training the proposed CNN architecture are shown in Table 4. These are the parameters

Table 4 Proposed hyper parameters for CNN

S. no	Hyperparameter	Value
1	Loss type	Categorical cross-entropy
2	Mini-batch size	32
3	Steps per epoch	800
4	Epoch	1000
5	Early stopping	25 Epoch with no improvement in val_loss
6	Model checkpoint	fas_mnist_1.h5
7	Padding	Valid
8	Optimization algorithm	Adam
9	multiprocessing	False

whose values are defined before the training process to control the learning process. Here 'categorical cross-entropy' is used as a loss type, mini-batch size 32 is used, maximum steps per epoch is 800, maximum epoch is 1000, early stopping is used as 25 epoch with no improvement. Model check pointing is used to store best epoch model and 'adam' optimization algorithm [11] is used. In the proposed model other parameters, strides, padding, processing, such as filters, kernel size, weights has been empirically determined based on several experiments.

The suggested model is trained using data from training, with validation data being used to fine-tuning the model parameters. The resulting model is tested using the test dataset to prevent overfitting and better splitting of the results. The trained deep neural network predicts whether the white blood cell images is one of the types (granular and non-granular) for binary classifications and (Eosinophil, Neutrophil, Lymphocyte, Monocyte) for multi-classification. These findings are linked to the principles of ground truth supplied by domain experts.

4 Experiment result and discussion

The convolution layers are applied to the entire white blood cell image dataset before training and validation processes are carried out by deep learning. Multiple cell images are considered instead of single-cell images because using the CNN technique, higher features are extracted and result in high precision. The hardware Using Keras, CNN is implemented in Python, with NVIDIA DGX-2 server, 2 petaFLOPS performance, 2×Intel Xenon Platinum CPUs, 16 NVIDIA Tesla V100 32 GB HBM2, 1.5 TB DDR4 RAM, 30 TB SSDs, 8×100 Gb/sec network speed.

Table 5 Describe the list of software and hardware used for implementation. The operating system used in the experiment is NVIDIA DGX Server Version 4.7.0 with customised NVIDIA Dockers is used in this experiment. NVIDIA customised NGC container is used is for experiment version "Nvidia:TensorFlow" where container tag is "20.02-tf1-py3". Python 3.7 was installed in this container, with TensorFlow GPU enable 1.15.2 version. The system has 16 NVIDIA tesla V100 GPUs, 81920 NVIDIA CUDA core, 10240 NVIDIA tensor core.

Table 6 demonstrates the binary classification confusion matrix based on types of WBCs by the proposed model on the test set. A total of 2453 images are classified correctly out of 2490. It shows the total accuracy the 98.51% on the test set.

Table 5 List of hardware and software used for experiments	S. No	Hardware and software	Version
	1	OS	NVIDIA DGX server version 4.7.0
	2	Docker	Nvidia-docker
	3	Python	3.7
	4	Keras	2.2.4
	5	Container	nvidia:tensorflow 20.02-tf1-py3
	6	Tensorflow_gpu	1.15.2
	7	GPU	16×NVIDIA Tesla V100
	8	CUDA Cores	81920 NVIDIA CUDA core
	9	Tensor Cores	10240 NVIDIA tensor core

Table 6 Confusion matrix of blood cells type identification

True label	Test dataset for binary classification	
	Predicted label	
	Granular	Non-Granular
Granular	1224	23
Non-Granular	14	1229

Table 7 demonstrates the multi-classification confusion matrix based on sub-types of WBCs by the proposed model on the test set which shows very little misclassification of a dataset. It shows the total accuracy the 97.7% on the test set for multi-classification.

Table 8 the results of the binary classification based on types of WBCs by the proposed model on the test set show the various other output parameters. A total of 98.51% accuracy has been achieved with 98.51% precision, 98.56% recall, 98.51% specificity and 97.50% F1 score on the binary classification.

Table 9 the results of the multi-classification based on sub-types of WBCs by the proposed model on the test set show the various other output parameters. A total of 98.91% accuracy has been achieved with 97.59% precision, 97.7% recall, 99.19% specificity and 97.63% F1 score on the multi-classification.

Figure 6a shows the plotted graph between the model training accuracy and testing accuracy. Where the blue colour shows the training parameter while the orange colour shows the testing parameter. Figure 6b shows the plotted graph between training loss and testing loss. Where the blue colour shows the training parameter while the orange colour shows the testing parameter.

Table 10 shows the comparison results of different models in 2D CNNs which were experimented on white blood cell images. This table shows that the model has 5 layers with maximum number of filter 32 has achieved the maximum accuracy rate among

**Table 7** Confusion matrix of blood cells sub-type identification

True label	Test dataset for multi-classification			
	Predicted label			
	Eosinophil	Neutrophil	Lymphocyte	Monocyte
Eosinophil	602	5	4	12
Neutrophil	8	609	4	3
Lymphocyte	3	5	608	7
Monocyte	4	2	3	611

Table 8 Various performance parameters and results for binary classification

Parameters	Equation	Granular (%)	Non-granular (%)	Average (%)
Validation accuracy	$\frac{\text{True Positive (TP)+True Negative (TN)}}{\text{Total Samples}}$	98.51	98.51	98.51
Precision	$\frac{\text{True Positive (TP)}}{\text{True Positive (TP)+False Positive (FP)}}$	98.86	98.16	98.51
Recall	$\frac{\text{True Positive (TP)}}{\text{True Positive (TP)+False Negative (FN)}}$	98.15	98.87	98.56
Specificity	$\frac{\text{True Negative (TN)}}{\text{True Negative (TN)+False Positive (FP)}}$	98.87	98.15	98.51
F1 score	$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision+Recall}}$	97.50	97.51	97.50

Table 9 Various performance parameters and results for multi-classification

Parameters	Equation	Eosinophil (%)	Lymphocyte (%)	Monocyte (%)	Neutrophil (%)	Average (%)
Validation accuracy	$\frac{\text{True Positive (TP)+True Negative (TN)}}{\text{Total Samples}}$	98.67	99.07	98.87	99.03	98.91
Precision	$\frac{\text{True Positive (TP)}}{\text{True Positive (TP)+False Positive (FP)}}$	97.56	98.22	96.52	98.06	97.59
Recall	$\frac{\text{True Positive (TP)}}{\text{True Positive (TP)+False Negative (FN)}}$	96.62	98.06	98.54	97.59	97.70
Specificity	$\frac{\text{True Negative (TN)}}{\text{True Negative (TN)+False Positive (FP)}}$	99.19	99.41	98.82	99.36	99.19
F1 score	$2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision+Recall}}$	97.08	98.13	97.51	97.82	97.63
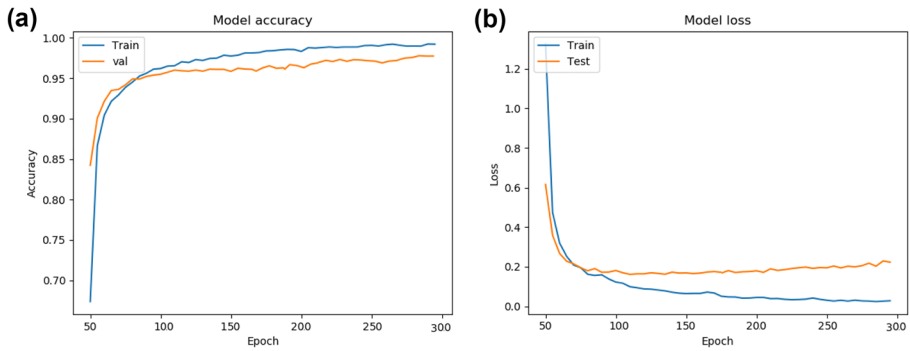
Fig. 6 Model accuracy and loss graph with training and testing samples

Table 10 Comparison results of different model in 2D CNNs

Different layers	Maximum number of filter	Accuracy rate (%)
3 Conv layers	13	0.92
3 Conv layers	16	0.95
4 Conv layers	13	0.92
4 Conv layers	16	0.96
5 Conv layers	16	0.88
5 Conv layers	32	0.97
6 Conv layers	16	0.94
6 Conv layers	32	0.93
7 Conv layers	16	0.94
7 Conv layers	32	0.87

these. Therefore this model is selected as the final model for this white blood cell type detection problem.

Table 11 shows the comparison of the performance of the proposed work on this subject with the current state of the art. The Livieris et al. [16] Achieved 93.29% accuracy KNN with is a semi supervised learning. Bani-Hani et al. [7] and Zhao et al. [10] Has used the Convolutional neural networks with an accuracy of 91.01 and 87.08% respectively. While Liang et al. [15] Has achieved 90.97% accuracy using fusion of CNN and RNN technique. The recent paper Patil et al. [22] has achieved 95.89% accuracy with CAA applied to CNN and RNN technique. The proposed method has achieved the 98.51% for binary and 97.70% for multi classification on WBCs. This method has resolved the issues of miss-classification of WBCs subtypes. This shows the proposed method has an advantage over the existing method while discriminating the white blood cell type.

Table 11 A comparative table of the proposed approach with the state-of-art

References	Classification type	Feature considered	Accuracy (%)
Livieris et al. [16]	Semi-supervised learning	KNN	93.29
Bani-Hani et al. [7]	GA-CNN	Convolutional neural networks	91.01
Zhao et al. [10]	Focuses on dimension using Gabor wavelet and hyper spectral microscopic blood images pixel by pixel	Convolutional neural networks	87.08
Liang et al. [15]	Single-cell image extracted and matrix transformation used	CNN + RNN	90.97
Patil et al. [22]	Overlapping the two distinct images, multiple cell nuclei, shapes extraction, and dimension reduction	CCA applied to (CNN + RNN)	95.89
Proposed method	CNN	Convolutional neural networks	WBCs type 98.51 WBCs subtypes 97.70

5 Conclusion

A completely convolutional neural network model based on deep learning is presented in this work to reduce the miss-classification of white blood cells. The data set augmentation approach is applied to improve the data set. There are convolutions and numerous other layers in the model. 98.51 per cent precision for the test set was reached by the proposed model for WBCs classification. 97.7 per cent precision for the test set was achieved by the proposed model for WBCs subtype classification. Through the use of a deep learning model on white blood cell images, such major improvements in performance indicate the superiority of the proposed system. Future studies will be aimed at recognizing various other kinds of cell types.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

1. Alam MM, Islam MT (2019) Machine learning approach of automatic identification and counting of blood cells. Healthc Technol Lett 6(4):103–108
2. Aliyu HA, Sudirman R, Abdul Razak MA, Abd Wahab MA. (2018) Red blood cell classification: deep learning architecture versus support vector machine. In: 2018 2nd international conference on biosignal analysis, processing and systems (ICBAPS), p 142–7
3. Andrea A, Santiago A, Anna M, Laura P, José R (2019) Recognition of peripheral blood cell images using convolutional neural networks. Comput Methods Programs Biomed 180:105020
4. Baghel N, Dutta MK, Burget R (2020) Automatic diagnosis of multiple cardiac diseases from PCG signals using convolutional neural network. Comput Methods Programs Biomed 197:105750
5. Bikhet SF, Darwish AM, Tolba HA, Shaheen SI (2000) Segmentation and classification of white blood cells 2259–61
6. Burton AG, Jandrey KE (2018) Leukocytosis and Leukopenia. Textbook of Small Animal Emergency Medicine, pp 405–412
7. Bani-Hani D, Khan N, Alsultan F, Karanjkar S, Nagaru N (2018) Classification of leucocytes using convolutional neural network optimized through genetic algorithm. In: Proceedings of the 7th annual world conference of the Society for Industrial and Systems Engineering, pp 1–7
8. Daqqa KASA, Maghari AYA, Sarraj WFMA (2017) Prediction and diagnosis of leukemia using classification algorithms. In: 2017 8th international conference on information technology (ICIT), pp 638–43
9. Gautam A, Singh P, Raman B, Bhadauria H (2016) Automatic classification of leukocytes using morphological features and naive Bayes classifier. In: 2016 IEEE region 10 conference (TENCON), pp 1023–7
10. Jianwei Z, Minshu Z, Zhenghua Z, Jianjun C, Cao F (2017) Automatic detection and classification of leukocytes using convolutional neural networks. Med Biol Eng Comput 55:1287–1301
11. Kingma D, Ba J (2014) Adam: a method for stochastic optimization. arXiv preprint. arXiv:1412.6980
12. Krishna K, Narasimha MM (1999) Genetic K-means algorithm. IEEE Trans Syst Man Cybern Part B 1999(29):433–439
13. Kulkarni SS, Hinge CS, Ambekar AG (2013) Classification of RBC and WBC in peripheral blood smear using KNN. Indian J Res
14. Lawrence S et al (1997) Face recognition: a convolutional neural-network approach. IEEE Trans Neural Netw 1997(8):98–113
15. Liang G, Hong H, Xie W, Zheng L (2018) Combining convolutional neural network with recursive neural network for blood cell image classification. IEEE Access 6:36188–36197

16. Livieris I, Pintelas E, Kanavos A, Pintelas P (2018) Identification of blood cell subtypes from images using an improved SSL algorithm. Biomed J Sci Tech Res 9:6923–6929
17. Malihi L, Ansari-Asl K, Behbahani A (2013) Malaria parasite detection in giemsastained blood cell images. In: 2013 8th Iranian conference on machine vision and image processing (MVIP), pp 360–5
18. Medsker LR, Jain LC (2001) Recurrent neural networks. Design Appl 5:64–67
19. Naz I, Muhammad N, Yasmin M, Sharif M, Shah JH, Fernandes SL (2019) Robust discrimination of leukocytes protuberant types for early diagnosis of leukemia. J Mech Med Biol 19:1950055
20. Noble WS (2006) What is a support vector machine? Nat Biotechnol 24(12):1565–1567
21. Pandey C, Baghel N, Dutta MK, Srivastava A, Choudhary N (2021) Machine learning approach for automatic diagnosis of Chlorosis in Vigna mungo leaves. Multimed Tools Appl 80(9):13407–13427
22. Patil AM, Patil MD, Birajdar GK (2020) White blood cells image classification using deep learning with canonical correlation analysis. IRBM
23. Rosyadi T, Arif Nopriadi A, Achmad B, Faridah (2016) Classification of leukocyte images using k-means clustering based on geometry features. In: 2016 6th international annual engineering seminar (InAES), pp 245–9
24. Safavian SR, David L (1991) (1991) A survey of decision tree classifier methodology. IEEE Trans Syst Man Cybern 21:660–674
25. Shiqi Y, Jia S, Chunyan X (2017) Convolutional neural networks for hyperspectral image classification. Neurocomputing 219:88–98
26. Singh A, Khan MA, Baghel N (2020) Face emotion identification by fusing neural network and texture features: facial expression. In: 2020 International Conference on Contemporary Computing and Applications (IC3A), pp 187–190
27. Sukhia KN, Ghafoor A, Riaz MM, Iltaf N (2019) Automated acute lymphoblastic leukaemia detection system using microscopic images. IET Image Process 13(13):2548–2553
28. Tai W-L, et al. (2011) Blood cell image classification based on hierarchical SVM. 2011 IEEE International Symposium on Multimedia. IEEE, 2011
29. Tatdow P, Siripen W, Kittiya K, Aniruth P (2019) Convolutional neural networks for recognition of lymphoblast cell images. Comput Intell Neurosci 2019:1–12
30. Ushizima DM, Lorena AC, de Carvalho ACPLF (2005) Support vector machines applied to white blood cell recognition. In: Fifth international conference on hybrid intelligent systems (HIS'05), pp 379–84
31. Young IT (1972) The classification of white blood cells. IEEE Trans Biomed Eng 19(4):291–298
32. Zahangir Alom Md, Hasan Mahmudul, Yakopcic Chris, Taha Tarek M, Asari Vijayan K (2018) Recurrent residual convolutional neural network based on U-Net (R2UNet) for medical image segmentation. arXiv:1802.06955
33. Zandecki M et al (2007) Spurious counts and spurious results on haematology analysers: a review. Part II: white blood cells, red blood cells, haemoglobin, red cell indices and reticulocytes. Inter J Lab Hematol 29:21–41