HYPERPARAMETER OPTIMIZATION IN CUSTOMIZED CONVOLUTIONAL NEURAL NETWORK FOR BLOOD CELLS CLASSIFICATION

GHAZALA HCINI¹, IMEN JDEY², ASHRAF HENI³, HELA LTIFI⁴

¹Faculty of Sciences and Technology of Sidi Bouzid, University of Kairouan, Kairouan, Tunisia

E-mail: ¹hcinighazala@fstsbz.u-kairouan.tn, ²imen.jdey@fstsbz.u-kairouan.tn, ³Ashraf.hani@fstsbz.u-kairouan.tn, ⁴hela.ltifi@fstsbz.u-kairouan.tn

ABSTRACT

Recently, various uses of supervised classification recognition algorithms for medical images are reported in literature. Specifically, in the current deep learning era, machine learning techniques are considered as the most important and used approach for automatic healthcare systems. In this context, many comparisons of supervised deep learning techniques, more precisely, the neural one, are proposed. The proposed approach provides a medical assistance based on relevant aspects of Machine -Learning methods applied for blood cells objects recognition while taking into consideration the property of uncertainty of this kind of image. The overview presented in this article examines the existing literature and the contributions already done in the field of intelligent healthcare systems for blood cell images classification. For this purpose, we summarize previous efforts made to define recognition process with supervised deep learning method, establishing a novel definition of personalized Machine- Learning with a major focus on the uncertainty input image. Departing from this definition, we propose and discuss the efficiency of Convolutional Neural Network for which the architecture is built and examined in detail. A Bayesian optimization of Convolutional Neural Network hyper parameters is also proposed. The main goal is to increase recognition rate while respecting time complexity. That is why an experimental comparison of Convolutional Neural Network with Support Vector Machine and K- nearest neighbor performance is discussed.

Keywords: Blood Cell Images, Machine Learning, Deep Learning, Convolutional Neural Network, Bayesian optimization

1. INTRODUCTION

Artificial intelligence is a vast computing field that allows machines to learn and gain knowledge from experience, adapt to data, and do human-like jobs intelligently, but without awareness, sensibility or mind. [1] [2] [3]. Deep learning, a subfield of it, contributes to the high quality of health-care services [4]. Its techniques are effective tools for augmenting classical machine learning and allowing computers to learn from data in order to create intelligent applications [5]. Because deep learning is achieving amazing results even at human level performance [6], it is currently accelerating the lively development of artificial intelligence in various fields and has had profound impacts on scientific fields such as computer vision and natural language processing [7] as well as medical image processing [8] [9] [10] [11] [12], even when the data set is unstructured and diverse [13]. We focus on intelligent health care systems in this paper by classifying blood cell images. Classifying cell types using statistical and moment invariant analysis is used to assess clinical state in order to entail the classification and identification of a blood sample from patients. In fact, we recommend that Convolutional Neural Network (CNN) algorithms be customized to deal with this specific image type, which is characterized by uncertainty and imprecision during the acquisition process. In order to provide context, we begin with a background study, followed by a discussion of artificial intelligence and related fields in the second section. We also go through the convolutional neural network architecture in detail. The final portion is to determine the deep learning approaches used to recognize blood cells images. In section 4, the proposed Convolutional Neural Networks technique is described by specifying the architecture performing by Bayesian optimization of convolutional neural network hyper parameters in order to improve recognition rate. Then, in Section 5, the experimental study. In section 6, an experimental comparison with similar methods like as support vector machine and K- nearest neighbor is discussed in the assessment optics of our proposed
system. Section 7 is where the conclusions are drawn.

2. BACKGROUND OF THE STUDY

White blood cells are important components of the human immune system. Blood cells are divided into three types: red blood cells (RBCs), which transport oxygen, white blood cells (WBCs), which are the face of the immune system, and platelets, which cause blood to clot in damaged tissues. One of the hurdles in medical science is identifying and quantifying the number of white blood cells. The quantity of red blood cells is the main reason. The WBCs account for about 1% of the total blood volume in a healthy adult. Because of the low proportion of WBCs in blood, identifying WBCs is difficult, making the task of defining WBC subtypes even more difficult [14]. The classification of White Blood Cells is critical in the detection of sickness in a person. Infections, allergies, anemia, leukemia, cancer, Acquired Immune Deficiency Syndrome (AIDS), and other disorders can all benefit from the classification [14]. Also the identification of red blood cells (RBCs) is a critical step in the early detection of blood-borne diseases including malaria and anemia, which must be followed by appropriate treatment [15].

To get around the manual gating limitations which is a technique, applied for the classification of blood cells images, that is expert-driven and introduces two major sources of operator bias. For starters, scatter plot gates are very subjective, and so might change significantly amongst operators. Automating the gating process with image processing technology combined with artificial intelligence and computer graphics systems to provide approaches that are more accurate, faster, and simple to use. Machine learning (ML) and Deep Learning (DL) could be feasible approaches to address these disadvantages which is the difficulty to duplicate, and that it takes a long time for large experiments [16]. The goal of machine learning, a part of artificial intelligence, is to combine the data we see with the experience that the computer gains in order to generate information that we can use. Supervised and unsupervised learning are the two broad categories of machine learning that can be used to solve problems involving classification and regression. Support Vector Machine (SVM), one of the most extensively used machine learning algorithms before neural networks became mature, is one of its methodologies [17] [18] [19] [20] [21].

And K- nearest neighbor (KNN). This technique is one of the most fundamental categorization and recognition techniques. That is why, in the last section, we employ it in a comparative study with our proposed approach to evaluating performance.

2.1 Support Vector Machine (SVM)

Before neural networks develop, the Support Vector Machine method is one of the most popular supervised learning algorithms [8]. The goal of this method is to identify the best separation between two classes of objects. Indeed, in order to optimize the distance between the hyperplane and the observations closest to the training sample, it chooses the hyperplane that divides the set of observations into two distinct classes (figure 1 and figure 2). Most of the time, the complexity of this strategy is polynomial. Indeed, it is mostly determined by the kernel function utilized [22], such as Gaussian, polynomial [23], and others, as inspired by [24] [25] [26].

Kernel Function:

\[ K(x_i, x_j) = \emptyset (x_i)^T (x_j) \] (1)

\[ f(x_i) = \sum_{n=1}^{N} (\alpha_n \gamma_n K(x_n, x_i) + b) \] (2)

Where, \( x_n \) represent support vector data, \( \alpha_n \) is Lagrange multiplier and \( \gamma_n \) is the label of membership class (+1, -1) with n=1, 2, 3, 4, ..., N.

- **Polynomial:**

\[ K(x_i, x_j) = (\gamma x_i^T x_j + r)^d, \gamma > 0 \] (3)

- **Gaussian:**

\[ K(x_i, x_j) = \exp (-\frac{||x_i - x_j||^2}{2\gamma^2}), \gamma > 0 \] (4)

- **Linear:**

\[ K(x_i, x_j) = x_i^T x_j \] (5)
2.2 K-Nearest Neighbor (KNN)

The k-nearest neighbor algorithm is a method for categorizing objects (figure 3). It is relatively simple classification model [30]. For this approach, the training procedure is limited to storing the vectors of different features as well as the labels of the training images. The unlabeled question point is simply assigned to the label of its nearest k neighbors throughout the classification procedure [3]. This is one of the most fundamental categorization and recognition methods [7].

What is the best value for K?

In actuality, there is no one-size-fits-all approach for selecting K; instead, it is dependent on the task at hand. If K is too large, the algorithm may incorrectly classify the new point because its nearest neighbors are located far away. Because of the noise in the training data, if k is too little, the algorithm is prone to overfitting the data. This will have an impact on the ability to generalize. The K size that minimizes the classification error is the best [32].

Distance function

The technique works by calculating the distance between the mathematical values of these points. It finds the chance of the points being similar to the test data by computing the distance between each data point and the test data. The distance function can be Euclidean. The shortest distance between two points, regardless of their dimensions, is known as the Euclidean distance. The Euclidean distance is the most used method for calculating the distance between two points.

\[
\text{dist}((x, y), (a, b)) = \sqrt{(x - a)^2 + (y - b)^2} \tag{6}
\]

To distinguish the difference between the two machine learning methods mentioned above Support Vector Machine (SVM) and k-nearest neighbor (KNN), the table below shows its pros and cons (table1).

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>High accuracy, even when data is not linearly separable.</td>
<td>- High memory requirements</td>
</tr>
<tr>
<td></td>
<td></td>
<td>- High level of complexity</td>
</tr>
</tbody>
</table>
2.3 Deep Learning

With the advancement of machine learning and the exponential growth of data [34], scientific research on Artificial Intelligence has shown significant limitations [35] associated with various machine learning techniques. This inspired data scientists to create the notion of deep learning [36] [37], a new machine learning paradigm [38] [39] based on artificial neural networks that consumes millions of data points [40] [41]. It has surpassed Machine-Learning algorithms in popularity (figure4) [42] [43].

Deep learning can be used to solve data analysis problems and produce high-quality summaries in order to make better decisions [45]. It has an impact on a variety of fields, with computer vision [46] [38] being one of the most notable. Object detection [38] [47] [48], face recognition [49], and other tasks like activity or event recognition [50] are examples of computer vision challenges. We are interested in the basic visual recognition application of object detection, which has had a lot of success in computer vision [51] [52]. When it comes to complicated problems like image classification, we find that deep learning really shines [53] [10] [54].

Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) are two common deep learning architectures for supervised learning [55] [56]. The hidden state of a recurrent neural network is one of its most notable features, as it aids it in remembering sequence information. Gated Recurrent Unit (GRU) and Long Short Term Memory (LSTM) are two common types of recurrent neural networks (LSTM). RNN saves all of the data that was previously computed. It employs the same parameter for all input and hidden layers, which lowers parameter complexity [57]. RNNs can handle time series and are utilized in a variety of applications such as stock price prediction, voice recognition, and video processing. The ability of this sort of network to preserve observations in order to extract information distinguishes it from others [56]. In the realm of computer vision, CNN is the most often utilized [58] [9]. It has had a good track record in practical applications [59]. Convolutional neural networks are widely used image classification models [60] [41] [61] [43] [62], and they have been found to perform well in this field [63] [58] [64] [65] [66]. The key advantage of CNN is that it has spread due to three factors; (1) it eliminates the need for manual feature extraction [9] because the key learns them automatically, (2) it is capable of being retrained to do new recognition tasks and gives excellent recognition results [67], and (3) you can also build on top of existing networks [68].

A CNN is a common deep learning architecture that learns relevant feature representations from image data [69] automatically [70]. It is not easy to come up with a powerful model to extract these attributes. As a result, we will need to go through the CNN architectural components in the next part.

2.4 Convolutional Neural Network Architecture

Convolutional Neural Networks get their name from the sort of hidden layers that make them up. Image and video recognition, recommendation systems, and natural language processing are just a few of the areas where they can be used [71]. CNN has produced excellent results in real applications [72]. The name Convolutional Neural Network declares that the network employs conversion as a mathematical process [53]. Convolutional Neural Networks [73] [9] [41] [74] have a variety of layers (figure 5).

Convolutional Neural Networks and Recurrent Neural Networks are two common deep learning architectures for supervised learning [55] [56]. The hidden state of a recurrent neural network is one of its most notable features, as it aids it in remembering sequence information. Gated Recurrent Unit (GRU) and Long Short Term Memory (LSTM) are two common types of recurrent neural networks (LSTM). RNN saves all of the data that was previously computed. It employs the same parameter for all input and hidden layers, which lowers parameter complexity [57]. RNNs can handle time series and are utilized in a variety of applications such as stock price prediction, voice recognition, and video processing. The ability of this sort of network to preserve observations in order to extract information distinguishes it from others [56]. In the realm of computer vision, CNN is the most often utilized [58] [9]. It has had a good track record in practical applications [59]. Convolutional neural networks are widely used image classification models [60] [41] [61] [43] [62], and they have been found to perform well in this field [63] [58] [64] [65] [66]. The key advantage of CNN is that it has spread due to three factors; (1) it eliminates the need for manual feature extraction [9] because the key learns them automatically, (2) it is capable of being retrained to do new recognition tasks and gives excellent recognition results [67], and (3) you can also build on top of existing networks [68].

A CNN is a common deep learning architecture that learns relevant feature representations from image data [69] automatically [70]. It is not easy to come up with a powerful model to extract these attributes. As a result, we will need to go through the CNN architectural components in the next part.

2.4 Convolutional Neural Network Architecture

Convolutional Neural Networks get their name from the sort of hidden layers that make them up. Image and video recognition, recommendation systems, and natural language processing are just a few of the areas where they can be used [71]. CNN has produced excellent results in real applications [72]. The name Convolutional Neural Network declares that the network employs conversion as a mathematical process [53]. Convolutional Neural Networks [73] [9] [41] [74] have a variety of layers (figure 5).

Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN) are two common deep learning architectures for supervised learning [55] [56]. The hidden state of a recurrent neural network is one of its most notable features, as it aids it in remembering sequence information. Gated Recurrent Unit (GRU) and Long Short Term Memory (LSTM) are two common types of recurrent neural networks (LSTM). RNN saves all of the data that was previously computed. It employs the same parameter for all input and hidden layers, which lowers parameter complexity [57]. RNNs can handle time series and are utilized in a variety of applications such as stock price prediction, voice recognition, and video processing. The ability of this sort of network to preserve observations in order to extract information distinguishes it from others [56]. In the realm of computer vision, CNN is the most often utilized [58] [9]. It has had a good track record in practical applications [59]. Convolutional neural networks are widely used image classification models [60] [41] [61] [43] [62], and they have been found to perform well in this field [63] [58] [64] [65] [66]. The key advantage of CNN is that it has spread due to three factors; (1) it eliminates the need for manual feature extraction [9] because the key learns them automatically, (2) it is capable of being retrained to do new recognition tasks and gives excellent recognition results [67], and (3) you can also build on top of existing networks [68].

A CNN is a common deep learning architecture that learns relevant feature representations from image data [69] automatically [70]. It is not easy to come up with a powerful model to extract these attributes. As a result, we will need to go through the CNN architectural components in the next part.
2.4.1 Convolution layer
This layer represents the key component of Convolutional Neural Networks [9]. The main purpose of this layer is to identify the existence of a set of features in the images received as input [76]. Filtering is carried out by convolution: its principle is to slide a window showing the feature on the image, and a product is calculated between the feature and each fragment of the image purged.

2.4.2 Pooling layer
This layer is generally installed between two convolution layers. The pooling operation is used to reduce the size of the images received as input, with all protection of important features. The processing is done by cutting the image into regular cells while keeping the maximum value of each cell.

2.4.3 Rectified Linear Unit (ReLU) layer
This layer acts as an activation function. It is a non-linear operation which replaces all negative values received with Zeros. The purpose of this operation is to introduce non-linearity in the model.

\[ \text{ReLU}(x) = \max(0,x) \]  
(7)

2.4.4 Fully-connected layer
It is the last layer of the Convolutional Neural Network. Its role is the input image classification. The fully-connected layer takes all neurons in the previous layer (Convolutional layer or pooling layer) and connects it to every single neuron it has.

3. RELATED WORKS

Researchers about blood cell image classification in recent years are mainly based on Deep learning techniques.

Shafique, S., & Tehsin, S. [77] used pretrained model AlexNet to automate the detection of acute lymphoblastic leukemia and the classification of its subtypes into four classes, applied on a public dataset which was split into two parts. Acute lymphoblastic leukemia IDB 1 was composed of 108 images, containing 59 images from healthy individuals and 49 images from leukemia patients. They used a data augmentation approach called image manipulation, in which they conducted image rotation and mirroring to improve the training data.

G. Liang & all [78] proposed an architecture that combined CNN and RNN in order to propose the CNN-RNN framework that consists of two different sections: the CNN section, which uses the Xception model, and uses the transfer learning method; the other section is the RNN, which uses the Bi-directional LSTM model applied to process a data set retrieved from the BCCD dataset (https://github.com/Shenggan/BCCDDataset), and publicly available dataset (https://www.kaggle.com/paultimothymooney/bloodcells/data). It contains 12444 augmented images of blood cells (JPEG), with 9957 and 2487 training and test images, and four additional subtype Eosinophil (2497 images), lymphocyte (2483 images), Monocyte (2478 images), and Neutrophil (2499 images). The entire network used RGB images of size 320*240*3 pixels.

Yildirim, M., Çinar, A. [79] proposed a model CNN that was coupled with AlexNet, Resnet50, Densenet201 and GoogleNet applied to Kaggle Dataset which consists of 4 different leukocyte cells (Eosinophil, Lymphocyte, Monocyte and Neutrophil) using Gaussian and median filters.

Andrea Acevedo & all [80] proposed a personalized CNN trained to classify eight classes of cells (Neutrophils, Eosinophils, Basophils, Lymphocytes, Monocytes, im-mature granulocytes (myelocytes, metamyelocytes and promyelocytes), Erythroblasts and Platelets) using a data set over 17000 cells images (88% Train, 12% Test) (RGB, 360*363 pixels) obtained from clinical practice. A confusion matrix was calculated to obtain performance parameters such as sensitivity, specificity and precision.

In [81], Sahlol & all improved Swarm optimization of deep features for efficient of white blood cells provided by Department of Information Technology. The input image is (224, 224, 3). They extracted the features using the VGGNet, then filtered these features using statistically enhanced Salp Swarm Algorithm (SESSA). This algorithm selected only 1K out of 25K features.

In [82], Loganathan, V. used the transfer learning approach to move the pre-trained weight parameters from the Image Net dataset to the CNN portion, and a custom loss function to make our network train and converge faster and with more accurate weight parameters. As a result, using segmentation and classification, this procedure is used to discover the irregularity of the cell. They ran numerous tests to assess the performance of the deep CNN in the special RBC classification scenarios, and they compared the findings. They used 434 raw microscopy images of 8 distinct SCD patients obtained from two separate hospitals in their trials to validate the resilience of their methods in dealing with varied imaging data. The classification results
showed that KNN had an accuracy of 88.04% while ANN had an accuracy of 54%.

Aliyu, Hajara Abdelkarim, & al. [83] used a framework that included image acquisition (were taken from 130 patients at Murtala Muhammad Specialist Hospital (MMSH) and Aminu Kano Teaching Hospital (AKTH), both in Nigeria, with authorized ethics numbers of MOH/off/797/T.1/849 and NHREC/21/08/2008/AKTH/EC/2307), image pre-processing (for further processing, it is necessary to remove undesired noise), automatic RBC blotch extraction (which was based on parameters like eccentricity, area, and bounding box), and the AlexNet model that had been trained.

Abou El-Seoud, S., Siala, M., & McKee, G. [84] proposed an approach based on convolutional neural network to detect and classify normal white blood cells. There are five layers in the proposed CNN structure. The first three layers are used to extract features, while the latter two levels are used to classify the features. The input image was [50x50x3], and the size of the receptive field (also known as filter) was 5x5. In each iteration, the stride is 1 and the filter is moved one pixel. On a dataset of 10,000 blood cell images, the proposed model was used.

The table 9 (Summary of Deep Learning based Blood Cells Diagnosis) displays the results of these different deep learning methods based on blood cells diagnosis.

As we notice not only the techniques, with the major core architectures of deep learning models Convolutional Neural Network(CNN), Long Short-Term Memory (LSTM) [39] Xception, Inceptionv3 [85] and the different pretrained models like ResNet50, AlexNet [86], DenseNet201, GoogleNet and Vgg-16 [87] [86] [4] [88] currently used by the authors have an effect on the results, but also the number of images for the processing, which affirms that deep learning has played an important role in Big Data analytical solutions [89] [90].

4. PROPOSED METHOD: CONVNET

The aim is to identify and characterize blood samples from patients. Several methods of detecting and classifying blood cells subtypes have important medical applications. CNN is a type of Artificial Neural Network (ANN) [34] [9]. The use of CNN for deep learning has spread for the main reason: the production of perfect recognition results [63]. It takes raw pixels as input and gives a result corresponding to the probabilities that the inputs belong to different classes. In the deep learning workflow using CNN, there are several repeated steps until the network reaches the desired level of accuracy such as (1) preprocess images, (2) set training options, (3) train network, and (3) testing to deploy the model.

The architecture of our model contains three Convolutional layers with two pooling layers installed between the Convolutional layers. It can have different combinations of initial layers but usually end with fully connected layer, a Softmax layer or a classification layer.

Hyper parameters related to our network structure

Important factors that affect the training time and the value of accuracy are notably, Network architecture, mini Batch Size and initial learning rate, also the choice of optimizer. Several optimizers are used in artificial intelligence. The best known are Stochastic Gradient Descent with momentum SGDM and Adaptive Moment Estimation ADAM [91] [92] [93], to minimize the loss. We trained the model and evaluated these optimization techniques in terms of convergence speed, accuracy and loss function. Also experimentation using certain activation functions, among which the most used are ReLU and tanh [93].

CNN Schematic (figure 6)

- The first Convolutional layer consists of 8 number of 4*4 sized filters
- The second Convolutional layer has 16 number of 2*2 filters
- The third Convolutional layer has 32 number of 6*6 filters
- Two max-pooling layers each of size are also being used in the network.
- The padding type used is "same", the size of the input matrix after each convolution does not change.
- An average pooling layer; its role is to sub-sample by dividing the input into rectangular grouping regions and calculating the mean values for each region.
- A FullyConnectedLayer represented the output layer.
Figure 6: ConvNet is a Convolutional Neural Network that is 15 layers deep. It reports zero errors.

5. EXPERIMENTAL STUDY

The used dataset for our study contains blood cells images (JPEG) with cells type labels. The cell types are Neutrophil, Monocyte, Lymphocyte, and Eosinophil (figure 7). The folder dataset used contains 12444 augmented images and 4 additional subtype labels (JPEG+CSV). The size of the images is [240*320], 3 channels (RGB). https://www.kaggle.com/paultimothymooney/blood-cells. We want to automatically classify each image based on what type of cells it contains.

To provide a quick overview, the following features and functions of WBC types and subtypes, as well as information on WBC are described [94]:

Neutrophils: are an important line of protection against germs and a part of the innate immune system.

Monocytes: activate osteoclasts, which are cells that can destroy bone. They are the most numerous WBCs.

Lymphocytes: are classified by their size and cytoplasmic granularity, and depending on their maturation stage, they might have a tiny or large nucleus.

Eosinophils: have the ability to release toxins from their granules, which can be used to destroy infections like parasites and worms. Their big granules make them easy to spot in stained smears. Eosinophil nuclei frequently have two lobes joined by a band of nuclear material.

5.1 CNN Architecture: ConvNet

This section provides the detailed description of the proposed architecture ConvNet using the blood cells dataset. Image preprocessing is the first stage, which includes image reading and image labeling. After the input images have been preprocessed, a convolutional neural network would be used to process the images (figure 8).

We separated the data into a training and test dataset. We analyzed the shapes of the training and test datasets. The training dataset was found to be consisting of 9957 images, 2497 Eosinophil, 2483 Lymphocyte, 2478 Monocyte and 2499 Neutrophil.
The test dataset had 2487 images, 623 Eosinophil, 620 Lymphocyte, 620 Monocyte and 624 Neutrophil. A study was made about optimization technique. We trained our model ConvNet and evaluated the optimization techniques Stochastic Gradient Descent with momentum SGDM and Adaptive Moment Estimation ADAM in terms of the value of accuracy. On the practical side, the Gradient Descent consists of evaluating the gradient of the cost function, then updating the parameters (the weights of a neural network) to reduce the error. The table below shows the significant impact of SGDM compared to ADAM.

Validation criteria:
We used accuracy to evaluate the proposed method's performance. The definition of this measure is as follows:

\[
\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{8}
\]

Where "TP" (true positives) refers to malignant samples that the classifier correctly labels, and "TN" (true negatives) refers to benign samples that the classifier correctly labels. False positives (FP) are malignant cells that have been mislabeled as benign, while false negatives (FN) are benign samples that have been mislabeled as malignant.

Table 2: Classification accuracy for both SGDM and ADAM

<table>
<thead>
<tr>
<th>Optimizer</th>
<th>Epochs</th>
<th>Max. iteration</th>
<th>Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>SGDM</td>
<td>10</td>
<td>770</td>
<td>70.45%</td>
</tr>
<tr>
<td>ADAM</td>
<td>10</td>
<td>770</td>
<td>62.48%</td>
</tr>
</tbody>
</table>

In deep networks, the choice of activation functions has a significant effect on training dynamics and task performance. The most successful and widely used activation function today is the linear rectified unit (ReLU). The table below shows the effect of the ReLU function compared to the tanh function.

Table 3: Classification accuracy for both ReLU and tanh

<table>
<thead>
<tr>
<th>Function</th>
<th>Epochs</th>
<th>Max. iteration</th>
<th>Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>ReLU</td>
<td>10</td>
<td>770</td>
<td>70.45%</td>
</tr>
<tr>
<td>tanh</td>
<td>10</td>
<td>770</td>
<td>53%</td>
</tr>
</tbody>
</table>

After choosing the optimizer and the activation function and configuring our architecture, we run ConvNet for a number of different epochs. According to the table below, we notice that each time we increase the number of iterations the value of the accuracy becomes better [95].

Table 4: Evaluation parameters: ConvNet

<table>
<thead>
<tr>
<th>Epoch</th>
<th>Max. iteration</th>
<th>Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>770</td>
<td>70.45%</td>
</tr>
<tr>
<td>30</td>
<td>2310</td>
<td>71.89%</td>
</tr>
<tr>
<td>40</td>
<td>3080</td>
<td>72.30%</td>
</tr>
<tr>
<td>50</td>
<td>3850</td>
<td>72.38%</td>
</tr>
<tr>
<td>100</td>
<td>7700</td>
<td>79.58%</td>
</tr>
</tbody>
</table>

5.2 Bayesian optimization

Bayesian optimization can have a great influence on model accuracy. To find optimal network hyper parameters and training options for CNN, we apply Bayesian optimization. This algorithm is well suited to optimize the hyper parameters of this classification model. After preparing the data, we must choose the variables to optimize. These variables are the network section depth. The role of this parameter is to control the depth of the network. Initial learning rate and Stochastic gradient descent momentum. With initialization of the weights and the bias.

An important hyper parameter that influences the dynamics of the model is the number of examples of the training dataset called the batch size. It controls the accuracy of the estimate of the error gradient. Also, to evaluate the results when we run ConvNet, we define metric function to test the prediction performance of our model. This function selects 300 test images. It makes it possible to evaluate the network formed on these images and to calculate the proportion of images that the network has misclassified.

The figures below are an explanation of the convolutional operation with stride 2*2.

Figure 9: displays the training metrics at each iteration (10 epochs)

Figure 10: displays the training metrics at each iteration (40 epochs).

Figure 11: shows the frequency of dataset items (2 epochs).
Figure 9. Training Progress specified as Plots value in training Options and start network training. This figure displays the training metrics at each iteration. Each iteration is an estimate of the gradient and an update of the network parameters (For 10 epochs).

Figure 10. Training Progress specified as Plots value in training Options and start network training. This figure displays the training metrics at each iteration. Each iteration is an estimate of the gradient and an update of the network parameters (For 40 epochs).

Figure 11. Histogram is a representation of statistical information showing the frequency of dataset items in successive intervals (4 cell types, for 2 epochs).

Figure 12: A pie chart using the dataset. Each slice of the pie chart represents a cell type (for 2 epochs).

Table 5: Evaluation Algorithm: Optimized-ConvNet before and after optimization

<table>
<thead>
<tr>
<th>Epoch</th>
<th>Partitioning (train, test)</th>
<th>Accuracy (before optimization)</th>
<th>Accuracy (after optimization)</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>75%, 25%</td>
<td>72.30%</td>
<td>77.93%</td>
</tr>
<tr>
<td>100</td>
<td>75%, 25%</td>
<td>79.58%</td>
<td>93.37%</td>
</tr>
</tbody>
</table>

6. DISCUSSION

This study examined the feasibility of using deep learning to classify blood cell images into 4 classes. The results of our study show that ConvNet, the proposed approach, can be trained to extract the features and classify the input. In order to increase recognition rate, a Bayesian optimization of this architecture hyper parameters is also proposed.

Comparison with other CNN architectures and related works:

In this subsection, the performance of the proposed approach is compared to other convolutional neural networks mentioned in the table of related works in terms of classification accuracy. For example, G. Liang & all [78], using the same dataset with 12444 images.

Table 6: Comparison with other CNN architectures and related works

<table>
<thead>
<tr>
<th>CNN models</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN-RNN Xception LSTM</td>
<td>90.79%</td>
</tr>
<tr>
<td>CNN-RNN Inception V3-LSTM</td>
<td>87.45%</td>
</tr>
<tr>
<td>CNN-RNN ResNet50-LSTM</td>
<td>89.38%</td>
</tr>
</tbody>
</table>
Comparison with machine learning methods:

Previous works were based on Machine-learning techniques employed for blood cells images classification. For example, in [96], tested 5 different machine learning algorithms for the classification decision tree, Naive Bayes, Random Forest, KNN and SVM. They were used for 5 different datasets. All these datasets are trained and tested at different rates for each machine learning algorithm: worst, best, mean, standard deviation, median and mod values are calculated. In this section, we will present the most used techniques in each group K-Nearest Neighbors KNN and Support Vector Machine SVM [96].

The dataset is trained and tested at different rates for each machine learning algorithm. The values are calculated for 40 epochs and 100 epochs.

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Partitioning (train, test)</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN</td>
<td>25%, 75%</td>
<td>74.90%</td>
</tr>
<tr>
<td>SVM</td>
<td>25%, 75%</td>
<td>77.19%</td>
</tr>
<tr>
<td>Optimized-ConvNet</td>
<td>75%, 25%</td>
<td>77.93%</td>
</tr>
</tbody>
</table>

Table 7: Comparison of recognition performance for 40 epochs

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Partitioning (train, test)</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>KNN</td>
<td>75%, 25%</td>
<td>81.61%</td>
</tr>
<tr>
<td>SVM</td>
<td>75%, 25%</td>
<td>82.76%</td>
</tr>
<tr>
<td>SVM</td>
<td>67%, 33%</td>
<td>84.48%</td>
</tr>
<tr>
<td>Optimized-ConvNet</td>
<td>75%, 25%</td>
<td>93.37%</td>
</tr>
</tbody>
</table>

Table 8: Comparison of recognition performance for 100 epochs

- **Advantages of the proposed method**
  The suggested model can learn the features of blood cell images on its own, eliminating the necessity for manual feature extraction (no cytoplasmic/nuclear segmentation is required). Moreover, we use CNN performing by Bayesian optimization to improve the accuracy and the quality of the output. The experiment confirmed the validity of our proposal. Also, we applied the model on public databases to take the route of a comparison with other methods and to decide which is the most effective.

- **Limitations of the proposed method**
  Despite its outstanding performance, our approach has a few flaws that could limit its practical use. In fact, the classification process is excessively slow. It can be resolved by increasing the performance of computer hardware or by optimizing the network's structure and reducing image size.

7. CONCLUSION

Due to its efficiency in diagnosing of blood-borne diseases including malaria and anemia at early stages, the convolutional neural network would be of great value to the medical diagnostic system used for detectors in blood cells. In this work, we proposed an architecture based on Convolutional neural network for a multi-class task for detection of four classes of blood cells. It can automatically extract and classify the deep features embedded in cell image patches. We conducted several experiments to evaluate the performance of our model by performing Bayesian optimization, we are able to achieve 93.37% accuracy. In comparison to other current methods, our proposed strategy outperformed them in terms of classification on the blood cell dataset. We hope that this model can be used to develop medical-aided diagnostic systems for blood related diseases in the future.

As perspective, model improvement is still needed in term of accuracy improvements. We are considering to test the proposed methodology on other datasets and more complex datasets and for the treatment of different diseases.
Table 9: Summary of Deep Learning based Blood Cells Diagnosis

<table>
<thead>
<tr>
<th>Year</th>
<th>Authors- Date</th>
<th>Number of images</th>
<th>Techniques</th>
<th>Performance (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2018</td>
<td>Shafique, S., &amp; Tehsin, S. [77]</td>
<td>108 images + Data augmentation</td>
<td>AlexNet</td>
<td>Sensitivity 96.74%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specificity 99.03%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accuracy 96.06%</td>
</tr>
<tr>
<td>2019</td>
<td>G. Liang &amp; all [78]</td>
<td>12444 images</td>
<td>CNN-RNN Xception LSTM</td>
<td>Accuracy 90.79%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CNN-RNN Inception V3-LSTM</td>
<td>Accuracy 87.45%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CNN-RNN ResNet50-LSTM</td>
<td>Accuracy 89.38%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CNN-RNN Xception-ResNet50-LSTM</td>
<td>Accuracy 88.58%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CNN Inception V3</td>
<td>Accuracy 84.08%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CNN-ResNet-50</td>
<td>Accuracy 87.62%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CNN Xception</td>
<td>Accuracy 88.70%</td>
</tr>
<tr>
<td>2019</td>
<td>Yildirim, M., Çinar, A. [79]</td>
<td>9400 images</td>
<td>CNN-AlexNet</td>
<td>Accuracy 82.16%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CNN-ResNet50</td>
<td>Accuracy 80.02%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CNNDensNet201</td>
<td>Accuracy 83.44%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CNN-GoogLeNet</td>
<td>Accuracy 75.21%</td>
</tr>
<tr>
<td>2019</td>
<td>Andrea Acevedo &amp; all [80]</td>
<td>17092 images</td>
<td>CNN-Vgg16</td>
<td>Accuracy 96%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CNN-InceptionV3</td>
<td>Accuracy 95%</td>
</tr>
<tr>
<td>2020</td>
<td>Sahlo &amp; all [81]</td>
<td>10661 images</td>
<td>CNN -VGGNet+ SESSA</td>
<td>Accuracy 87.9%</td>
</tr>
<tr>
<td>2019</td>
<td>Loganathan, V [82]</td>
<td>434 images</td>
<td>Transfer learning +KNN classifier +ANN classifier</td>
<td>Accuracy 88.04%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Accuracy 54%</td>
</tr>
<tr>
<td>2020</td>
<td>Aliyu, Hajara Abdelkarim, &amp; al. [83]</td>
<td>Over 9000 images</td>
<td>AlexNet</td>
<td>Accuracy 95.92%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sensitivity 77%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Specificity 98.82%</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Precision 90%</td>
</tr>
<tr>
<td>2020</td>
<td>Abou El-Seoud, S., Siala, M., &amp; McKee, G. [84]</td>
<td>10000 images</td>
<td>AlexNet</td>
<td>Accuracy 96.78%</td>
</tr>
</tbody>
</table>
REFERENCES


[20] A. Tharwat, "Parameter investigation of support vector machine classifier with


Online and Biomedical Engineering (iJOE), vol. 16, no. 15, 2020.


