

BLOOD CELLS CLASSIFICATION USING DEEP LEARNING WITH CUSTOMIZED DATA AUGMENTATION AND EK-MEANS SEGMENTATION

ASHRAF HENI¹, IMEN JDEY², HELA LTIFI³

^{1,2,3}Research Groups in Intelligent Machines National School of Engineers (ENIS) Sfax, Tunisia

¹ashraf.heni@fstsbz.u-kairouan.tn, ²imen.jdey@fstsbz.u-kairouan.tn, ³hela.ltifi@fstsbz.u-kairouan.tn

ABSTRACT

White blood cells, also known as leukocytes, play an important role in the human body. By increasing immunity for fighting infectious diseases. The classification of white blood cells plays an important role in detecting disease in an individual. Classification can also help in the identification of diseases such as infections, allergies, anaemia, leukaemia, cancer, Acquired Immune Deficiency Syndrome (AIDS), etc., which are due to abnormalities of the immune system. Currently, there is a great deal of research being done in this area. In some cases, the classification of white cells is a medical emergency that requires rapid diagnosis. given the complexity of classifying them into subtypes, researchers have presented new techniques to help healthcare workers better facilitate this task. scientific reviews have shown the performance of classification systems using deep learning. We will use a deep convolutional learning technique (CNN) that can classify WBC images into its subtypes namely neutrophils, eosinophils, lymphocytes, and monocytes. The proposed approach will rely on data augmentation techniques for classification and the introduction of the EK-means algorithm for image segmentation which is a fusion between k-means and fuzzy c-means. we introduce a new technique for data augmentation called checkerboard image. this technique will be introduced in the context of further model learning by simulating the pixelization of images or hole images. By applying the emerging deep learning-based image classification technique "convolutional neural network" using Ek-means for segmentation and data augmentation for classification to a large database, we were able to classify WBC, using VGG19 model, into its subtypes with a validation accuracy of 96,24%. We also compare our proposed model with the reference models.

Key words: *White Blood Cells, EK-Means, VGG19, CNN, Data Augmentation*

1. INTRODUCTION

The image is perceived as an important source of information. Based on the elements of the production system, the image can be classified into two categories [1] : natural image (shadow or reflection) and artificial image that is produced with the help of an instrument. While this is relatively easy for the human brain, understanding an image remains a complex task for a machine. This challenge has given birth to many areas of research in the field of artificial intelligence such as computer vision, pattern recognition, etc. In any computer vision application, Segmentation is a significant step in the process of automatic image classification [2]. The segmentation step must be efficient because the quality of its result is the basis for other high-level operations such as the recognition and interpretation of extracted objects [3].

Medical image segmentation is of great interest to imaging researchers. However, until today, no

universal method for image segmentation exists [4]. Each proposed method is often effective for a given type of image, and for a given application. Researchers are constantly opening new horizons and exploring different techniques to create more and more efficient approaches. In the early 2010's, a technique is gaining momentum in the field of computer vision, namely deep learning neural networks. deep learning approaches have made a major re-emergence in the research community, which has led to their use in various projects requiring fine classification, such as medical imaging [5]. Deep learning is defined as the ability of an agent to learn to decide based on observations. In the biomedical context, the action of this agent results in additional information to help the doctor in his decision making. The management of a patient is affected at several stages, whether it be at the level of diagnosis, choice of treatment, follow-up over time, or even in the surgical intervention. The role of the agent in question is to classify biomedical images by

deep learning to discover clinically relevant pathology patterns. These classification operations remain the basis of computer-aided medical decision support tools.

It is in this context that our paper proposes a new method for image segmentation for object detection based on deep learning.

This paper is composed from three parts: the first section will represent object recognition and image segmentation methods. The second section will present the related research works and our method. And the last section will present the experiments and the results.

2. RESEARCH CONTEXT

Medical imaging is the process by which a physician can examine the inside of a patient's body without performing surgery. It can be used for clinical purposes to establish a diagnosis or for the treatment of pathologies, but also for scientific research studying the physiology of living beings.

The goal of medical imaging is to create an intelligible visual representation of medical information. This problem is more globally part of the scientific and technical image: the objective is to be able to represent in a relatively simple format of a large amount of information from a multitude of measurements acquired in a well-defined mode.

There are four types of medical imaging that rely on the use of X-rays, ultrasound, magnetic field or natural or artificial radioactivity. The objective is to create an intelligible visual representation of bones, tissues, and organs, today it offers an unobstructed view of these same organs in action and allows to visualize up to the cellular metabolism. In this article, we focus our research on the images of blood cells to approve the methods of image segmentation

Blood is composed of blood cells suspended in plasma. The whole is contained in the blood vessels. The total volume of blood of a human adult is 5 liters. The cells in suspension represent 45% of the total volume, which corresponds to the hematocrit. There are several types of cells:

- Red blood cells or erythrocytes.
- The white blood cells or leukocytes, divided into Polynuclear, monocytes and lymphocytes
- Platelets.

2.1. erythrocytes

Red blood cells are annucleated cells whose essential component is an oxygen-binding hemoprotein: hemoglobin (about 14.5g/100ml). The main role of these cells is to ensure the transport of oxygen and carbon dioxide between the alveoli pulmonary alveoli and the tissues.

2.2. White blood cells

These cells participate in the body's specific defenses. They are divided into two categories: granulocytes and lymphoid cells or agranulocytes. The term granulocyte is due to the presence of granulocyte-like granules, which helps to distinguish between them. In fact, these granules have different affinities towards neutral acidic or basic stains and give the cytoplasm different colors. Thus, granulocytes are distinguished between neutrophils, eosinophils (or acidophils) and basophils, while lymphoid cells are distinguished between lymphocytes and monocytes.

Each type of white blood cell is present in the blood in different proportions. we present a brief description of each cell type.

Lymphocytes

These are monocled cells, with a high nucleo/cytoplasmic ratio. Under light microscopy, they are small cells about 7 μ m in diameter with a nucleus occupying almost the entire cell. Their shape is regular and rounded. There is a small peripheral cytoplasmic fringe of purple appearance. The nucleus is spherical and dense.

Monocytes

Monocytes are the largest leukocytes, measuring 16 to 20 μ m. They have a large reniform or horseshoe-shaped nucleus, in some cases with two lobes. The cytoplasm is transparent but has a ground glass appearance

Neutrophils

They are cells about 12 μ m in diameter, the nucleus is usually three-lobed, but the number of lobes varies from 2 to 5 lobes and is an indication of the maturation of the cell. cytoplasm is transparent as its granules are tiny and have a slight pinkish tinge.

Eosinophils

The cytoplasm is filled with granules that take on a characteristic pink-orange colour.

The fixed and stained cells on a slide are analysed by a cytopathologist. This stage of analysis of a slide is called screening. It consists of a visual analysis under the microscope of all the cells present on a cytological slide. The purpose

of this step is to detect abnormal or suspicious cells to establish a reliable and valid diagnosis.

Computerised rescreening systems are designed to operate after manual screening. However, they can also be used in an earlier stage to operate. However, pathologists prefer computerised rescreening.

after presenting the context of the research, we present in the next part of this article the technological background and image analysis techniques and clustering literature

3. OVERVIEW OF DEEP RECOGNITION ARCHITECTURES FOR IMAGE ANALYSIS

3.1. Machine learning

Machine learning is a set of methods at the intersection of statistics, artificial intelligence, and computer science. It consists of applying algorithms to give computers the ability to learn based on the data presented to them [6]. It is used to detect objects in images and classify them [6]. A recent example of use is the classification of white cells according to their type [7].

According to [6], using machine learning is only necessary to use many images for the algorithm to determine the features needed to identify an object. Moreover, changing a single rule may require rewriting the entire system. An example of the failure of this hand-coded approach is detecting faces in images. Today, every smartphone can detect a face in an image using machine learning, but face detection was an unsolved problem until 2001.

Machine learning can be categorized into supervised and unsupervised.

Supervised learning: machine learning algorithms that implement pairwise learning of input and output data (inputs/outputs) are called "supervised" because there is human intervention and supervision [6]. Currently, supervised machine learning is the most widely used and often the most effective method for detecting an object in an image [8]. Among the most used supervised machine learning algorithms, we find: Decision trees, k-nearest neighbors, vector machines, logistic regression, Bayesian naive classification, and deep neural networks [6].

Unsupervised learning: in unsupervised learning, only the input data are known. The output data are not known [6]. This type of learning is used when the output data are not known or not well known.

3.2 artificial neural networks

An artificial neural network is an interconnected assembly of simple processing elements, units or nodes, whose functionality is broadly based on biological (animal or human) neurons. The information processing capacity of the network is stored in the interconnection strengths or weights (synaptic weights), obtained through a process of adaptation, or learning from a set of training patterns [9].

Neural networks learn inductively, i.e., by experience, by being confronted with input and output data.

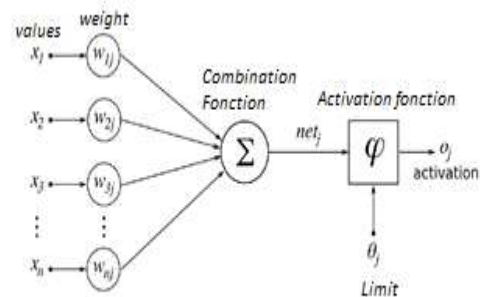


Figure 1 Schematic of an artificial neuron [9]

According to [10], "Typically, a neural network relies on many processors operating in parallel and organized into thirds. The first third receives raw information inputs, much like the human's optical nerves when processing visual signals. Then, each third receives the information output from the previous third. We find the same process in humans when neurons receive signals from neurons near the optic nerve. The last third, on the other hand, produces the outputs of the system."

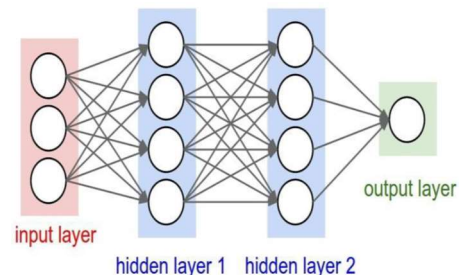


Figure 2 Example of an artificial neural network architecture

3.3 Deep learning

According to [11], Deep learning is a branch of machine learning. It is a novel approach to learning data representations that concentrates on gaining knowledge components of overly

greater representations. Deep networks of neurons are often renowned as 'multi-layer networks.'

According to [11], a deep learning system recognises the image of a person by combining the edges and extremities of the body in a hierarchical manner. Perhaps the day is not so far off when deep learning will be extended to applications that allow machines to think for themselves.

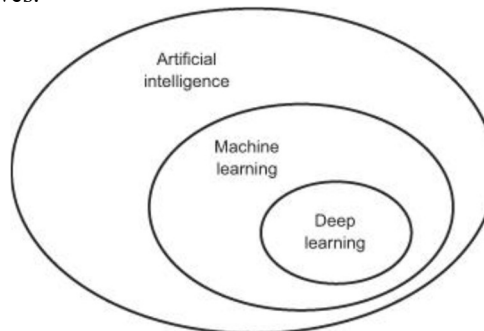


Figure 3 ; Illustration of deep learning versus artificial intelligence and machine learning [12]

Neural networks implementing the principle of deep learning are called deep neural networks (DNN).

In the following, we present the CNNs which are deep learning networks

3.4 CNN Architecture

Convolutional neural networks (CNNs) are an application of neural networks based on the application of filters inspired by the functioning of the part of the brain dealing with the visual field.

The CNN architecture was presented by [4] in their review on the progress of object recognition techniques. Convolutional neural networks have proven their contribution to image recognition by computer vision. It is a methodology like supervised learning methods. Its functional aspect is based on the reception of the image, the processing of the content and the generation of the classifier on it. A CNN is composed of a set of layers that are semantically and functionally interlinked through data transfer.

According to [13], convolutional neural networks (CNNs) emerged from the study of the visual cortex of the human and animal brain and have been used in image recognition and analysis since the 1980s. In recent years, thanks to the increase in computing power and the large amount of training data available, CNNs have been able to achieve superhuman performance on some complex visual tasks. They power image retrieval services, autonomous cars, automatic video, etc.

According to [11], CNNs, also known as "ConvNets", are very similar to regular neural networks. They still consist of neurons whose weights can be extracted from the data. Each neuron receives inputs and performs a scalar product. They always have a loss function on the last layer that is fully connected. They can use a non-linearity function. A regular neural network receives inputs as a single vector and passes through a series of hidden layers. Each hidden layer consists of a set of neurons, each neuron being fully connected to all neurons in the previous layer. Within a single layer, each neuron is completely independent and shares no connections. The last fully connected layer, also called output layer, contains scores for all classes (category of objects) to avoid image classification problems. In principle, there are three main layers in a simple convolution neural network (CNN).

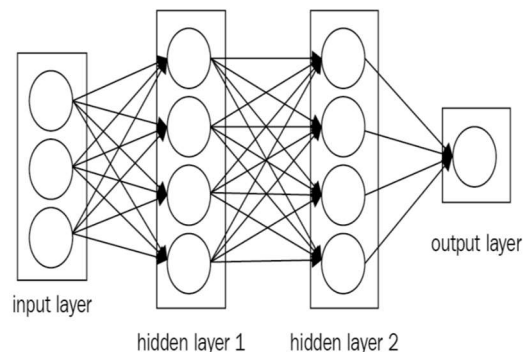


Figure 4 Simple three-level convolutional neural network [11]

One form of network that is consistently found is to repeat the previous scheme (above) by stacking the convolution and pooling layers (Subsampling) in succession until very small inputs are reached. This allows pre-processing of small amounts of information and therefore increases performance.

4 PROPOSED METHOD FOR IMAGE SEGMENTATION

4.1 Choice of approach for image analysis and object recognition

Among all the solutions studied, the one chosen is deep machine learning, more precisely convolution neural networks. It is with this technology that we find most recent solutions (after 2010) to the detection and classification of objects in images, and it is the solution proposed by all the recent studies.

4.2 Related works

The analysis and classification of medical images is a field that can help medical

experts to diagnose different types of diseases [14] [15] [16] [17]

The classification of blood cell images is the subject of much research. Researchers have tended to use deep learning as a technique for the segmentation of medical blood cells.

In the context of lymphoblastic leukemia detection, the pre-trained AlexNet model was used by [18] for the classification of leukaemia types into 4 subtypes. The model was applied on a dataset composed of 59 images for healthy individuals and 49 images for others with leukaemia. The authors used data augmentation techniques for the improvement of the training database by applying the image manipulation technique of rotation and mirroring.

The combination of the two architectures RNN and CNN was the proposal in [19]. Their approaches were based on two different modules. The first module is based on the CNN architecture with Xception using learning transfer. The second module is based on the RNN architecture using LSTM model was used to analyse a dataset extracted from the BCCD dataset, The dataset used contained 12444 augmented blood cell images. it contains 9957 and 2487 training and test data, respectively, as well as additional subtypes: 2497 Eosinophil image, 2483 Lymphocytes, 2478 Monocytes, and Neutrophil (2499 images). The complete array was made up of RGB pictures with dimensions of 320*240*3. The Resnet50, AlexNet, Densenet201 and GoogleNet models were coupled in the proposal in [20] as a model applied on the kaggle dataset. Gaussian filters were applied to the input images.

17 000 cell images (88% Train, 12% Test) (RGB, 360*363 pixels) from clinical practice were the dataset to be classified into cell classes.

In [21], authors applied a custom CNN architecture on these datasets by applying the confusion matrix to calculate the performance parameters of the proposed system (sensitivity and accuracy).

In [22], The authors propose a CNN architecture with a Bayesian optimization of the hyper parameters to increase the recognition rate while respecting the execution time factor. They proposed a comparison of the CNN with the support vector machine(SVM).

Their model's architecture consists of three Convolution layer having two pooling layers fitted between the Convolution layers levels. It has a variety of beginning layer configurations,

but it commonly ends with a completely linked layer, a Soft-max, or a classification layer.

4.3 Contribution

A blended method is proposed based on EK-means segmentation. EK-means is proposed in [23] and an adjusted CNN model, assisted by data augmentation. Our proposal is based on:

(a) normalizing the blood cell images.

(b) segmenting the blood cell image using EK-means.

(c) data improvement preceded by depths extracting features, using a refined CNN model, and classification of blood cells into the following categories: Neutrophils, Monocytes Lymphocytes, Eosinophils.

4.3.1. pre-processing: Normalizing blood cell images

The pre-processing phase is an obligatory step to carry out the following phases smoothly. This normalisation step influences the results by increasing the accuracy, given the observation of noise in the images that form the subject of the processing (such as Gaussian noise) [24]

In this step we try to normalise the intensities of the images. We use in the preprocessing step SimpleITK which implements a unified interface with the intensity-based registration framework [25]. The image size is set to (224*224)

4.3.2. segmentation of blood cells using EK-means clustering

Image segmentation consists of grouping together semantically related coherent regions of the image. It divides the pixels of the image into groups according to the highest similarity.

The K-means algorithm for automated data classifications is one of the most straightforward ones. The findings' dependence on the beginning values, however, is this method's fundamental drawback. Each random initialization of all the points given to such a cluster corresponds to a distinct solution, which in certain situations might be substantially different from the optimum alternative. [2]

Many variants have already been offered like k-medoids technique which employs group median value notwithstanding cluster means so that various distances other than the Euclidean distance may be employed to alleviate some restrictions in K-Means algorithms.

Cluster's initialization come up a problem of data analysis to classify objects, so the contribution in [23] is to replace this phase with a technique that fixes the clustering center from the onset.

As an outcome, the purpose of EK-means (developed algorithm as a variant of K-means) customisation is to adjust the randomized cluster initialization using the Fuzzy c-means concept based on the level of membership. In the classic K-means method, the following code snippet describes our goal of optimising the initialization task:

Algorithm

Begin

(1) cluster init:

$$V_i = [v_i^1, v_i^2, \dots, v_i^p], 1 \leq i \leq c$$

(2) init matrix $U^{(1)}$

(3)center of each cluster c:

$$V_i = \frac{\sum_{j=1}^n U_{ij}^m x_j}{\sum_{j=1}^n U_{ij}^m}$$

(4) Calculate $U^{(k+1)}$

DO (3) and (4), the algorithm stops if:

$$\|U^{k+1} - U^k\| \leq \epsilon; \epsilon \leq 10^{-4}$$

(5) After the clusters have indeed been fixed and the centre of each cluster has been determined, the K-means operation can proceed with the incarnation and termination steps.

End

4.3.3. Classification with VGG19 model

The model adopted in our proposal is VGG19. It is a model for transfer learning that increases the prediction accuracy and minimises the learning time. Previously, this model contained 19 layers of which 16 layers are convolutional and 3 layers are fully connected [26]. We have refined the weight of these layers.

A Max-pooling part the group of the initial two convolution layers in VGG19. The remaining 8 convolution layers are organised into groups with four convolution layers, preceded by one max-pooling layer with a fixed filter size of 3*3. The last three layers are dense layers with 4096 as well as 1000 features followed by Soft - max, as shown in the figure below.

The last FC-8 layer is the layer transmitted as a feature vector for the SoftMax function to predict the blood cell class.

We included the VGG19 in our proposal because the filter size in this model is minimal, i.e., 3X3, which was identical across all convolutional and with a step of 1, whereas in other CNN architectures the filter size is 11X11 or 7X7 given 4 or 5 strides. [3]

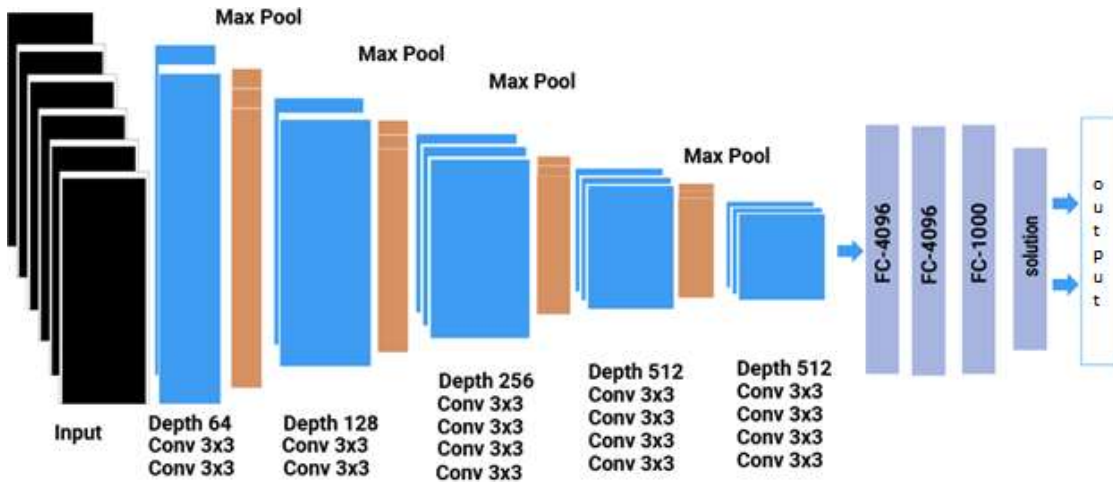


Figure 5 proposed VGG19 CNN model for blood cells segmentation

4.3.4. data augmentation

There are not many medical images that are labelled. This lack of availability of this type of image causes an anomaly in the learning phase which weakens the accuracy rate in object recognition. In deep learning, to have a good training of the model used, it is necessary to have a large quantity of labelled data. For this,

researchers use the technique of augmenting data in the learning phase to the prediction system [27] Data augmentation is done through cropping, filling, and horizontal flipping techniques to better shape the CNN.

For the pre-processing stage, we propose a new technique called "checkerboard image pixelation". This technique consists of going through the image matrix and forming two output images which, when merged, give the original image. Each output image is composed of pixels in coordinates (i,j) such that :

For each i , corresponding to the current row of the matrix, we loop over j (column of the matrix) with two increment steps.

The increment of j will itself be two steps. For the remaining pixels they will be drilled through the index $RGB=\#000000$

The pseudo-algorithm used is as follows:

```

Start Algorithm Checkerboard image
Load input image
Initialise  $i=0; j=0$ 
For each  $i$  in  $[0,240]$ 
For each  $j$  in  $[0,240]$ 
Keep pixel  $(i,j)$  and assign to imageOutput1 and assign a pixel  $\#000000$  to ImageOutput2
Keep pixel  $(i,j-1)$  and assign to imageoutput2 and assign pixel  $\#000000$  to Imageoutput1
Increment  $j+=2$ 
End loop  $j$ 
Increment  $i+=2$ 
End loop  $i$ 
End Algorithm
    
```

This new technique is used to triple the training images, while providing robustness to the system for recognising hole images with reference to an original image.

In the proposed method, 6 augmentation techniques are used, and 20 parameters are associated with these techniques to result in 20 augmented images relative to each original image.

5. EXPERIMENTATION AND RESULTS

5.1. Dataset

In our proposal, we use the dataset containing images of labelled blood cells. The dataset contains 12444 images in total. The original size of the images, before the intensity unification phase, is 240×320 , on 3 RGB channels. We would like to classify every image according to the type of cells it contains.

We have analysed the data and separate it into training data and test data. The processing data consisted of 9957 images, 2497 eosinophils, 2483 lymphocytes, 2478 monocytes and 2499 neutrophils. The testing data set included 2487 pictures, which included 620 Monocytes and same number for Lymphocytes, 623 Eosinophils, and 624 Neutrophils.

Table 1 Training Dataset

Blood cell classes	Basic classes	After data augmentation
eosinophils	2497	49940
lymphocytes	2483	49660
monocytes	2478	49560
neutrophils	2499	49980
Total	9957	199140

5.2. Proposed architecture

In this section, we describe the proposed framework for white blood cell dataset classification. For the classification of these images, the architectural proposal is based on the three steps mentioned in the previous contribution section. The sequence of processing units can be summarised as a dataset read, a pre-processing including image reading and labelling followed by the processing unit for image classification using the refined VGG19 model.

We have analysed the data and separate it into training data and test data. The processing data consisted of 9957 images, 2497 eosinophils, 2483 lymphocytes, 2478 monocytes and 2499 neutrophils. The testing data set included 2487 pictures, which included 620 Monocytes and same number for Lymphocytes, 623 Eosinophils, and 624 Neutrophils.

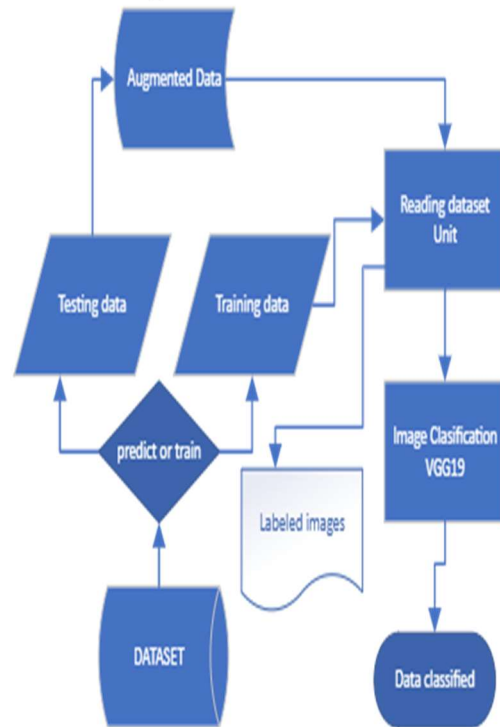


Figure 6 Processing Units Of The Proposed Architecture

5.3. Performance measurement:

We measure the performance of our proposal, compared to the work related to this research with the Accuracy metric. The formula is as follow :

$$Accuracy = \frac{TruePositive + TrueNegative}{TruePositive + TrueNegative + FalsePositive + FalseNegative}$$

Accuracy will provide a response to the question, "How accurate were the model's predictions?" True Positives and True Negatives are considered by Accuracy. We shall see later that not both are employed in some of the assessment indicators.

5.4. Analysis and comparison

Deep learning is the most widely used method in recently published blood cell classification work. A significant number of researchers have used CNN architectures with refinement for blood cell classification. The majority of these frameworks were 2D CNNs, but others were 3D CNNs [28]. However, with these deep learning architectures, one encounters the problem of the unavailability of sufficient data for the training phase. We have addressed this problem by using synthetic data augmentation.

We introduce through the following figure the progression of the accuracy during all the iterations as well as the loss curve of our proposal. The accuracy rate reaches 96.24%.

Here, we compare the performance of our proposal with the systems cited as related work by comparing the Accuracy rate.

Table 2 Comparison Of The Performance

CNN Model	Accuracy
RNN Xception LSTM	90.79%
RNN Inception V3-LSTM	87.45%
RNN ResNet 50 LSTM	89.38%
RNN Xception-ResNet 50 LSTM	88.58%
Inception 3	84.08%
ResNet 50	87.62%
Xception	88.70 %
Optimized ConvNet	93.37%
CNN-refined VGG 19	96.24%

In the recent literature, many approaches have been proposed to train CNNs end-to-end without splitting them into multiple patches. Unfortunately, medical imaging faces a major drawback that prevents these natural image segmentation methodologies from being easily transferred to medical image segmentation problems. Taking the example of blood cell segmentation, the region of interest is less than one percent of the overall image volume.

Below, the system is evaluated prior to and following data augmentation is compared. This comparison is done again with some related work.

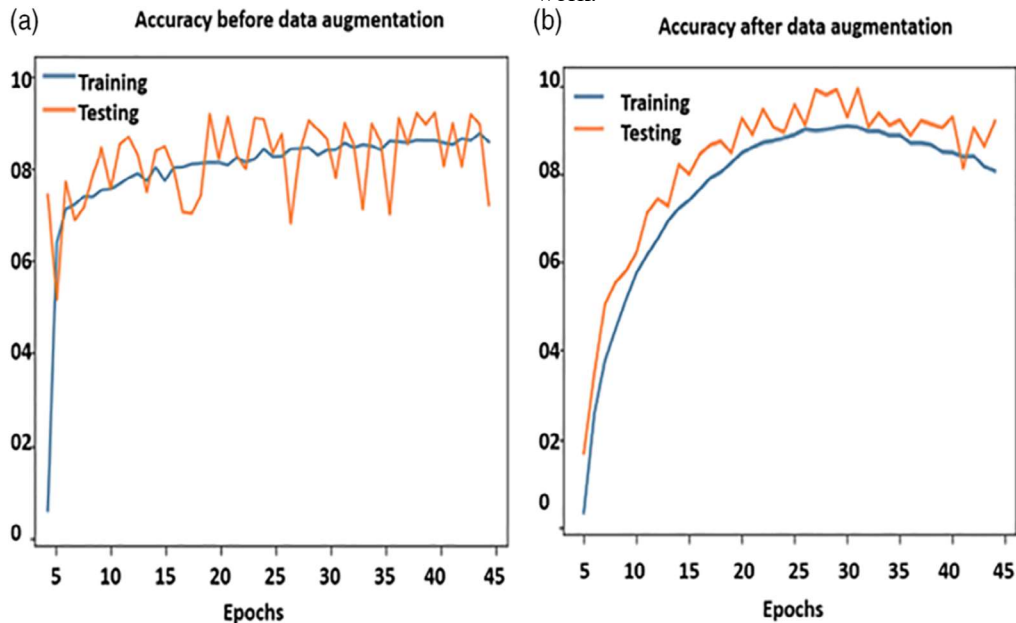


Figure 7 Comparison Of Accuracy Before And After The Data Augmentation

The accuracy rate has been 90.13% before data preprocessing as well as 96.24% after data augmentation with synthetic data. Finally, experimental results indicate that the proposed solution's overall accuracy has improved during the training and testing phases by employing various data augmentation techniques.

6. CONCLUSION

This paper introduces a system for segmentation and classification of blood cells using a CNN VGG19 model. To obtain better accuracy of the synthetic data of the classifier, the concept of augmentation is applied. The suggested technique implements a novel approach that has been termed as 'checkerboard image'.

The K-means clustering approach merged with fuzzy c-means is used to segment the blood cells. Finally, using Synthetic data enhancement techniques, a finely calibrated VGG-19 CNN model was trained to categorise the blood cells. A robust collection of studies is used to evaluate the proposed CNN-based approaches. Thus, the results demonstrate that the proposed method could help the medical experts to perform blood cell classification. The proposed method was contrasted to recent techniques available, and the results revealed that the suggested method has higher accuracy.

Representation learning algorithms are constantly evolving today and improving at a rapid pace. The opportunities in the biomedical field are therefore growing, provided that a database of sufficient size and representation is collected. Indeed, these methods of learning representations rely on the availability of a large amount of data, which must be availability of numerous data, which must be representative of the problem to be solved. The transfer of technology to the biomedical field therefore takes more time, as specialized medical expertise is needed both to gather observations for a database and to develop, but also to interpret the results obtained

It will be appropriate to give as perspectives to this work an improvement of the rate of precision of validation, because even if they are satisfactory at first, it is certain that they can be improved more by playing on the hyperparameters of the model as well as those of the algorithm and by seeking empirically the number of layers of convolutions and optimal decisional, we will be able to hardly but certainly exceed the threshold which this study reached.

REFERENCES:

- [1] F. Robert and A. Grandjean, *Philosophie de l'image*, m-Editer, 2010.
- [2] J. Chen, X. Qi, L. Chen, F. Chen and G. Cheng, "Quantum-inspired ant lion optimized hybrid k-means for cluster analysis and intrusion detection," *Knowledge-Based Systems*, vol. 203, 2020.
- [3] J. Bernal, K. Kushibar, D. S. Asfaw, S. Valverde, A. Oliver, R. Martí and X. Lladó, "Deep convolutional neural networks for brain image analysis on magnetic resonance imaging: a review," *Artificial intelligence in medicine*, vol. 95, pp. 64-81, 2019.
- [4] X. Wu, D. Sahoo and S. C. Hoi, "Recent advances in deep learning for object detection," *Neurocomputing*, no. 396, pp. 39-64, 2020.
- [5] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra and M. Riedmiller, *Playing Atari with Deep Reinforcement Learning*, 1312.5602 ed., 2013.
- [6] A. C. MÜLLER and S. GUIDO, *Introduction to Machine Learning with Python*, Etats-Unis: O'Reilly Media,, 2017.
- [7] T. T. Tang, J. A. Zawaski, K. N. Francis, A. A. Qutub and M. W. Gaber, "Image-based Classification of Tumor Type and Growth Rate using Machine Learning: a preclinical study," *Scientific Reports*, vol. 9, no. 1, 2019.
- [8] A. Géron, *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd Edition, O'Reilly Media, Inc., 2019.
- [9] C. Li, Y. Zhang and Y. Qu, "Object detection based on deep learning of small samples," in *2018 tenth international conference on advanced computational intelligence (ICACI)*. IEEE, 2018.
- [10] L. Bastien, "Réseau de neurones artificiels : qu'est-ce que c'est et à quoi ça sert ?," 2019. [Online]. Available: <https://www.lebigdata.fr/reseau-de-neurones-artificiels-definition>. [Accessed 12 07 2021].
- [11] M. Sewak, k. Md Rezaul and P. Pujari, *Practical Convolutional Neural Networks*, O'Reilly Media, 2018.

- [12] F. Chollet, Deep Learning with Python, O'Reilly media, 2017.
- [13] A. Geron, Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow, Paris: DUNOD, 2020.
- [14] H. Fahad, G. K. M. Usman, S. Tanzila, R. Amjad and I. Sajid, "Microscopic abnormality classification of cardiac murmurs using ANFIS and HMM," *microscopy research and technique*, vol. 81, no. 5, pp. 449-457, 2018.
- [15] A. Liaqat, M. A. Khan, M. Sharif, M. Mittal, T. Saba, K. S. Manic and F. N. H. Al Attar, "Gastric tract infections detection and classification from wireless capsule endoscopy using computer vision techniques: a review," *Current Medical Imaging*, 2020.
- [16] A. Mittal, D. Kumar, M. Mittal, T. Saba, I. Abunadi, A. Rehman and S. Roy, "Detecting pneumonia using convolutions and dynamic capsule routing for chest X-ray images," *Sensors*, 2020.
- [17] S. Perveen, M. Shahbaz, T. Saba, K. Keshavjee, A. Rehman and A. Guergachi, "Handling irregularly sampled longitudinal Data and prognostic modeling of diabetes using machine learning technique," *IEEE Access*, p. 21875–21885, 2020.
- [18] S. Shafique and S. Tehsin, "Acute lymphoblastic leukemia detection and classification of its subtypes using pretrained deep convolutional neural networks," *Technology in cancer research & treatment*, vol. 17, 2018.
- [19] G. Liang, H. Hong, W. Xie and L. Zheng, "Combining convolutional neural network with recursive neural network for blood cell image classification," *IEEE access*, pp. 36188-36197, 2018.
- [20] M. Yildirim and A. Çinar, "Classification of White Blood Cells by Deep Learning Methods for Diagnosing Disease," *Rev. d'Intelligence Artif.*, pp. 335-340, 2018.
- [21] A. Acevedo, S. Alférez, A. Merino, L. Puigví and J. Rodellar, "Recognition of peripheral blood cell images using convolutional neural networks," *Computer methods and programs in biomedicine*, p. 105020, 2018.
- [22] G. Hcini, I. Jdey, A. Heni and H. Ltifi, "Hyperparameter optimization in customized convolutional neural network for blood cells classification," *J. Theor. Appl. Inf. Technol.*, vol. 99, pp. 5425-5435, 2021.
- [23] A. Heni, I. Jdey and h. Ltifi, "k-means and fuzzy c-means fusion for object clustering," in *8th International Conference on Control, Decision and Information Technologies*, 2022.
- [24] A. Khan, A. Sohail, U. Zahoora and A. S. Qureshi, "A survey of the recent architectures of deep convolutional neural networks," *Artificial intelligence review*, vol. 53, no. 8, pp. 5455-5516, 2020.
- [25] K. Marstal, F. Berendsen, M. Staring and S. Klein, "SimpleElastix: A user-friendly, multi-lingual library for medical image registration," in *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 2016, pp. 134-142.
- [26] W. Yu, K. Yang, Y. Bai, T. Xiao, H. Yao and Y. Rui, "Visualizing and comparing AlexNet and VGG using deconvolutional layers," in *Proceedings of the 33rd International Conference on Machine Learning*, 2016.
- [27] T. Sadad, A. Rehman, A. Hussain, A. A. Abbasi and M. Q. Khan, "A Review on Multi-organ Cancer Detection Using Advanced Machine Learning Techniques," *Current medical imaging*, vol. 17, no. 6, pp. 686-694, 2021.
- [28] M. A. Khan, T. Akram, M. Sharif, T. Saba, K. Javed, I. U. Lali and A. Rehman, "Construction of saliency map and hybrid set of features for efficient segmentation and classification of skin lesion," *Microscopy research and technique*, vol. 82, no. 6, pp. 741-763, 2019.
- [29] R. Girshick, J. Donahue, T. Darrell and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," *Proceedings of the CVPR*, 2014.
- [30] R. B. Girshick, "From Rigid Templates to Grammars: Object Detection with Structured Models," Citeseer, 2012.
- [31] T.-Y. Lin, P. Goyal, R. Girshick, K. He and P. Dollár, "Focal loss for dense

- object detection," *Proceedings of the IEEE International Conference on Computer Vision*, p. 2980–2988., 2017.
- [32] Z. Yu, J. Yu and C. Xiang, "Beyond bilinear: generalized multimodal factorized high-order pooling for visual question answering," *IEEE Trans Neural Netw Learn Syst*, vol. 29, no. 12, p. 5947–5959, 2018.
- [33] J. Dai, Y. Li and K. He, "R-fcn: object detection via region-based fully convolutional networks," *Advances in neural information processing systems*, pp. 379–387, 2016.
- [34] J. Macqueen, "Some methods for classification and analysis of multivariate observations," In *5-th Berkeley Symposium on Mathematical*, p. 281–297, 1967.
- [35] P. Viola and M. J. Jones, "Robust real-time face detection," *Int. J. Comput. Vis.*, no. 57, p. 137–154, 2004.
- [36] V. Sharma and R. N. Mir, "A comprehensive and systematic look up into deep learning based object detection techniques: A review," *Computer Science Review*, vol. 38, p. 100131, 2020.
- [37] Q. Zhu, S. Avidan, M. Yeh and K. Cheng, "Fast human detection using a cascade of histograms of oriented gradients," in *CVPR*, 2006.
- [38] I. Kirati and Y. Tlili, "Colour image segmentation using spatial probabilistic clustering," *Int. J. Signal and Imaging Systems Engineering*, vol. 7, no. 3, 2014.
- [39] J. Yu, M. Tan and H. Zhang, "Hierarchical deep click feature prediction for fine-grained image recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 1-1, 2019.
- [40] Q. V. Le, "Building high-level features using large scale unsupervised learning," *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 8595-8598, 2013.
- [41] P. Arbelaez and L. Cohen, "Segmentation d'images couleur par partitions de voronoi," *Traitement du Signal*, vol. 21, no. 5, pp. 20-40, 2005.
- [42] R. Girshick, "Fast R-CNN," *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 1440-1448, 2015.
- [43] M. Mignotte, "A de-texturing and spatially constrained k-means approach for image segmentation.," *Pattern Recognition Letters*, vol. 32, no. 2, p. 359–367, 2011.
- [44] P. Felzenszwalb, D. McAllester and D. Ramanan, "A discriminatively trained, multiscale, deformable part model," *IEEE Conference on Computer Vision and Pattern Recognition, IEEE*, pp. 1-8, 2008.
- [45] M. Luo, Y.-F. Ma and H.-J. Zhang, "A spatial constrained k-means approach to image segmentation," in *Fourth IEEE Pacific-Rim Conference On Multimedia*, 2003.
- [46] S. Ben Chaabane, M. Sayadi, F. Fnaiech and E. Brassart, "Color image segmentation using automatic thresholding and the fuzzy c-means techniques," *Electrotechnical Conference, 2008. MELECON 2008. The 14th IEEE Mediterranean*, p. 857–861, 2008.
- [47] K. He, X. Zhang, S. Ren and J. Sun, "Deep residual learning for image recognition," 2016.
- [48] M. Mignotte, "Segmentation by fusion of histogram-based k-means clusters in different color spaces," *IEEE Transactions on Image Processing*, vol. 17, no. 5, pp. 780-787, 2008.
- [49] X. Ding, Y. Luo and Q. Li, "Prior knowledge-based deep learning method for indoor object recognition and application," *Syst Sci Control Eng*, vol. 6, no. 1, p. 249–257, 2018.
- [50] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Computer Society Conference on Computer Vision and Pattern recognition*, 2005.
- [51] I. Laptev, "Improvements of object detection using boosted histograms," in *BMVC*, 2006.
- [52] K. He, G. Gkioxari, P. Dollár and Ross Girshick, "Mask r-cnn," *Proceedings of the IEEE International Conference on Computer Vision*, p. 2961–2969, 2017.
- [53] G. Ghiasi, T.-Y. Lin, Le and QV, "NAS-FPN: learning scalable feature pyramid architecture for object detection,"

- Proceedings of *the IEEE conference on computer vision and pattern recognition*, p. 7036–7045, 2019.
- [54] K. S. Tan, W. H. Lim and N. A. M. Isa, "Novel initialization scheme for fuzzy c-means algorithm on color image segmentation," *Applied Soft Computing*, vol. 13, no. 4, p. 1832 – 1852, 2013.
- [55] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus and Y. LeCun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," *arXiv preprint arXiv:1312.6229*, 2013.
- [56] P. Dollar, C. Wojek and B. Schiele, Pedestrian detection: a benchmark, 2009.
- [57] P. Viola and M. Jones, "Robust real-time face detection," *Int J Comput Vis*, vol. 57, no. 2, pp. 137-154, 2004.
- [58] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy and A. Yuille, "Semantic image segmentation with deep convolutional nets and fully connected," CRFS, 2014.
- [59] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu and A. C. Berg, "Ssd: Single shot multibox detector," *European Conference on Computer Vision*, Springer, pp. 21-37, 2016.
- [60] J. Redmon and A. Farhadi, "YOLO9000: better, faster, stronger," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, p. 7263–7271, 2017.
- [61] X. Ding, Y. Luo and Q. Yu, "Indoor object recognition using pre-trained convolutional neural network," in *23rd international conference on automation and computing (ICAC) IEEE*, 2017.
- [62] S.-T. Bow, "Pattern Recognition and Image Preprocessing, number 77," *Electrical Engineering and Electronics*, 2015.
- [63] J. Arora and M. Tushir, "An Enhanced Spatial Intuitionistic Fuzzy C-means Clustering for Image Segmentation," *Procedia Computer Science*, vol. 167, pp. 646-655, 2020.
- [64] A. Husham, M. H. Alkawaz, T. Saba, A. Rehman and J. S. Alghamdi, "Automated nuclei segmentation of malignant using level sets," *Microscopy Research and Technique*, p. 993–997, 2016.
- [65] S. Iftikhar, K. Fatima, A. Rehman, A. S. Almazayad and T. Saba, "evolution based hybrid approach for heart diseases classification and associated risk factors identification," *Biomedical Research*, vol. 28, no. 8, p. 3451–3455, 2017.
- [66] S. L. Marie-Sainte, T. Saba, D. Alsaleh, A. Alotaibi and M. Bin, "an improved strategy for predicting diagnosis, survivability, and recurrence of breast cancer," *Journal of Computational and Theoretical Nanoscience*, p. 3705–3711, 2019.