IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

LEUKEMIA BLOOD CELL IMAGE CLASSIFICATION USING CONVOLUTIONAL NEURAL NETWORK

Rikita Gohil, Dwiti Pandya

Computer Engineering Department

Vadodara Institute of Engineering, Vadodara, Gujarat, India

Abstract: Leukemia is a form of cancer that can be a fatal disease and treat it requires a correct and early diagnosis. Standard methods have transformed into automated computer tools for analyzing, diagnosing, and predicting symptoms. The purpose of this paper is to do a comparative analysis study on cancer detection that applies deep learning to address the topic of automated diagnosis. Leukemia is a type of blood cancer which occurs due to abnormal increase in WBCs (white blood cells) in bone marrow of human body. Leukemia can be classified as acute leukemia and chronic leukemia, in which acute leukemia grows very fast whereas chronic leukemia grows slowly.

Index Terms **- Leukemia Disease, CNN, machine learning, image classification, feature extraction.**

I. INTRODUCTION

One of the biggest risks to human health has been well established to be cancer [1], [2]. Cancer has been identified as one of the top causes of death in humans, second only to infectious, cardiovascular, and cerebrovascular disorders, according to epidemiologic data, posing a serious threat to human health [3]. One of the most prevalent aspects in the classification of cancer is thought to be gene expression. The origin of leukemia cancer is the bone marrow of human beings. Bone marrow generates the blood cells. Leukemia can be produced because of any problematic blood cells in bone marrow. As the human blood has three basic components i.e, Red Blood Cells, White Blood Cells and Platelets. Leukaemia usually smashed the white blood cells in the body.

Leukemia is categorized in two classes:

• Acute Leukemia: The defected white bloods cells stop performing their normal action and number increase gradually.

• Chronic Leukemia: Here the defected white blood cells perform their normal function also the number of cells increase slowly.

In our proposed work, a novel approach is presented in order to recognize the Leukemia cells using microscopic blood images.

The stained slides of the most frequent kinds of leukemia as illustrated in Fig. 1. Leukemia is divided into four categories, acute myeloid leukemia (AML) Fig. 1a, acute lymphoblastic leukemia (ALL) Fig. 1b, chronic myeloid leukemia (CML) Fig. 1c, and chronic lymphocytic leukemia (CLL) Fig. 1d, as well as a few less frequent kinds.

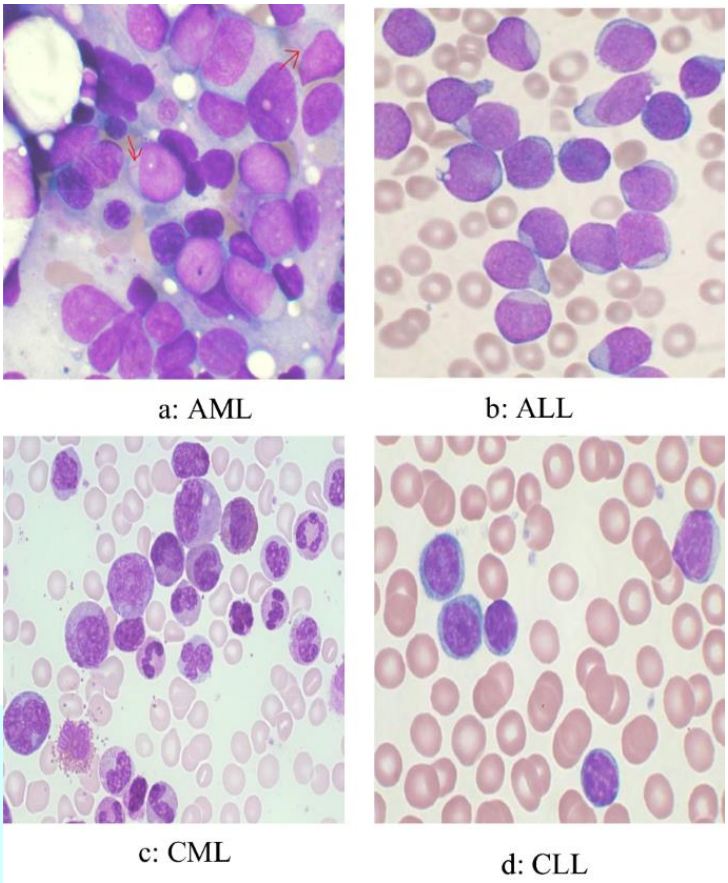

Fig.1 Types of leukemia

The purpose of our work is to develop a system that can automatically detect cancer from the blood cell images. This system uses a convolution network that inputs a blood cell images and outputs whether the cell is infected with cancer or not. The appearance of cancer in blood cell images is often vague, can overlap with other diagnoses, and can mimic many other benign abnormalities. These discrepancies cause considerable variability among medical personnel in the diagnosis of cancer. Automated detection of cancer from blood cell images at the level of expert medical personnel would not only have tremendous benefit in clinical settings, it would also be invaluable in delivery of health care to populations with inadequate access to diagnostic imaging specialists.

We develop a system which detects cancer from blood cell images at a level exceeding practicing medical personnel. This technology can improve healthcare delivery and increase access to medical imaging expertise in parts of the world where access to skilled medical personnel is limited.

Various techniques have been developed by researchers to detect leukemia. One of the most used technique is Convolution Neural Network (CNN) It is based on computer vision in recent years. The common algorithm for this approach consists of several rigid steps: image pre-processing, clustering, morphological filtering, segmentation, feature selection or extraction, classification, and evaluation.

II. RESEARCH METHODOLOGY

Existing methods for diagnosis

• Medical history and physical examination: The record of present symptoms, and problems a person has had in the past. The medical history of a person's family also helps in diagnose leukemia.

- Complete blood count (CBC): Blood is taken and checked under the microscope for the number of RBCs, WBCs and platelets.
- Bone marrow aspiration: Bone marrow is removed with the help of a needle from breastbone. The removed sample is observed under a microscope to look for abnormal cells.
- Cytogenetic analysis**:** Cytogenetic test takes blood or bone marrow to help identify individual chromosomes. It shows abnormalities in chromosomes, which help to diagnosis and identify the type of leukemia.
- Immunohistochemistry**:** Blood sample of cells are treated with special antibodies in immunohistochemistry. Under the microscope the change in color can be seen. It helps in determining the types of cells that are present.

Proposed Method

1. This technique proposed the segmentation method using color-based clustering to obtain nucleus region and cytoplasm area from stained blood smear images. SVM classifiers are applied with relevant features and gain satisfactory results.
2. have proposed an automatic detection of white blood cells (WBCs) from peripheral blood images and classification of five types of WBCs: eosinophil, basophil, neutrophil, monocyte, and lymphocyte. Eosinophil and basophil from other WBCs are first classified by SVM with a granularity feature. Other three types are then recognized using convolutional neural network to extract features, and random forest uses these features to classify those WBCs.

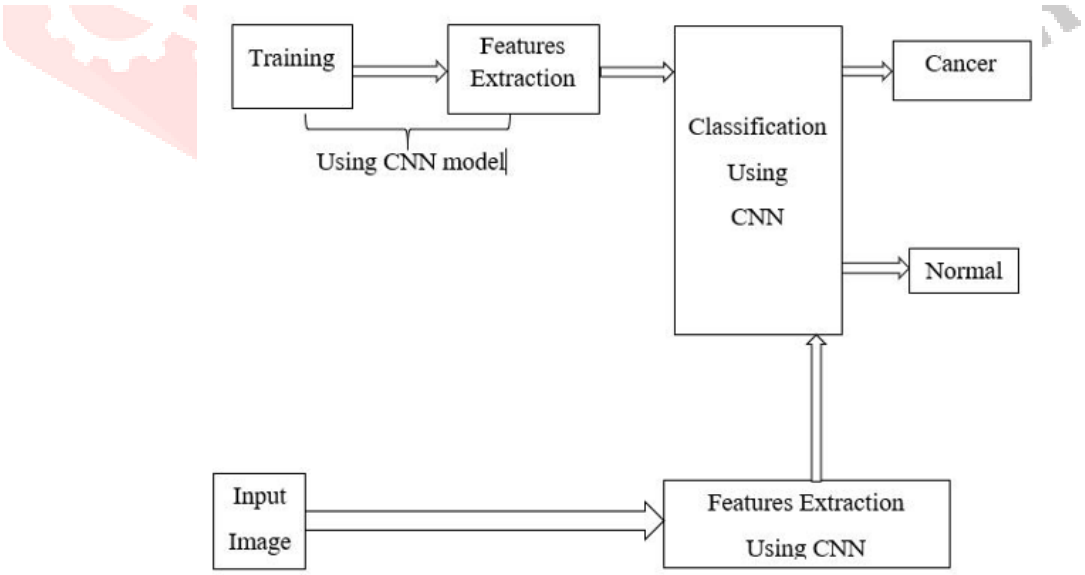

Fig. 2 System block diagram

Depending on the features extracted from the above step the classifier classifies a cell as normal cell or cancer affected cell i.e. blast cell. This is done by comparing some of the features like geometric, statistical, texture and size ratio from regions obtained in the segmentation process with standard feature. Then the results are analyzed to identify the types and subtypes of acute leukemia. automatically detect cancer in blood samples, neural networks are used. Due to the well-known method of the neural network as a good classifier for many practical applications, it is chosen as a classification tool. One of the crucial elements in creating a precise process model using CNNs is the training and validation processes. The training features set is used to train the CNN model, and the testing features set is used to check the accuracy of the trained utilising the feed-forward back propagation network. The dataset for the training and validation operations is divided into these two sections. Connection weights were continuously changed during the training phase until they reached the specified iteration number or the appropriate error. Using neural networks, cancer can be automatically detected.

III. RESULTS AND DISCUSSION

Training Data

Labelling data:

- **ALL** (Leukemia cells): **1**
- **HEM** (Healty cells): **0**

Model performance

In [16]:
```
ts_length = len(test_df)
test_batch_size = test_batch_size = max(sorted([ts_length // n for n in range(1, ts_length + 1) if ts_length%n == 0 and ts_lengt
test_steps = ts_length // test_batch_size

train_score = model.evaluate(train_gen, steps= test_steps, verbose= 1)
valid_score = model.evaluate(valid_gen, steps= test_steps, verbose= 1)
test_score = model.evaluate(test_gen, steps= test_steps, verbose= 1)

print("Train Loss: ", train_score[0])
print("Train Accuracy: ", train_score[1])
print('-' * 20)
print("Validation Loss: ", valid_score[0])
print("Validation Accuracy: ", valid_score[1])
print('-' * 20)
print("Test Loss: ", test_score[0])
print("Test Accuracy: ", test_score[1])
```

```
20/20 [==============================] - 3s 120ms/step - loss: 0.0860 - accuracy: 0.9962
20/20 [==============================] - 2s 101ms/step - loss: 0.1943 - accuracy: 0.9488
20/20 [==============================] - 22s 1s/step - loss: 0.1810 - accuracy: 0.9575
Train Loss:  0.08596161752939224
Train Accuracy:  0.9962499737739563
--------------------
Validation Loss:  0.19433359801769257
Validation Accuracy:  0.9487500190734863
--------------------
Test Loss:  0.18103870749473572
Test Accuracy:  0.9574999809265137
```

Confusion Matrics and Classification Report

In [18]:
```
g_dict = test_gen.class_indices
classes = list(g_dict.keys())

# Confusion matrix
cm = confusion_matrix(test_gen.classes, y_pred)
plot_confusion_matrix(cm= cm, classes= classes, title = 'Confusion Matrix')

# Classification report
print(classification_report(test_gen.classes, y_pred, target_names= classes))
```

```
Confusion Matrix, Without Normalization
[[1049   42]
 [  26  483]]
              precision    recall  f1-score   support

         all       0.98      0.96      0.97      1091
         hem       0.92      0.95      0.93       509

    accuracy                           0.96      1600
   macro avg       0.95      0.96      0.95      1600
weighted avg       0.96      0.96      0.96      1600
```

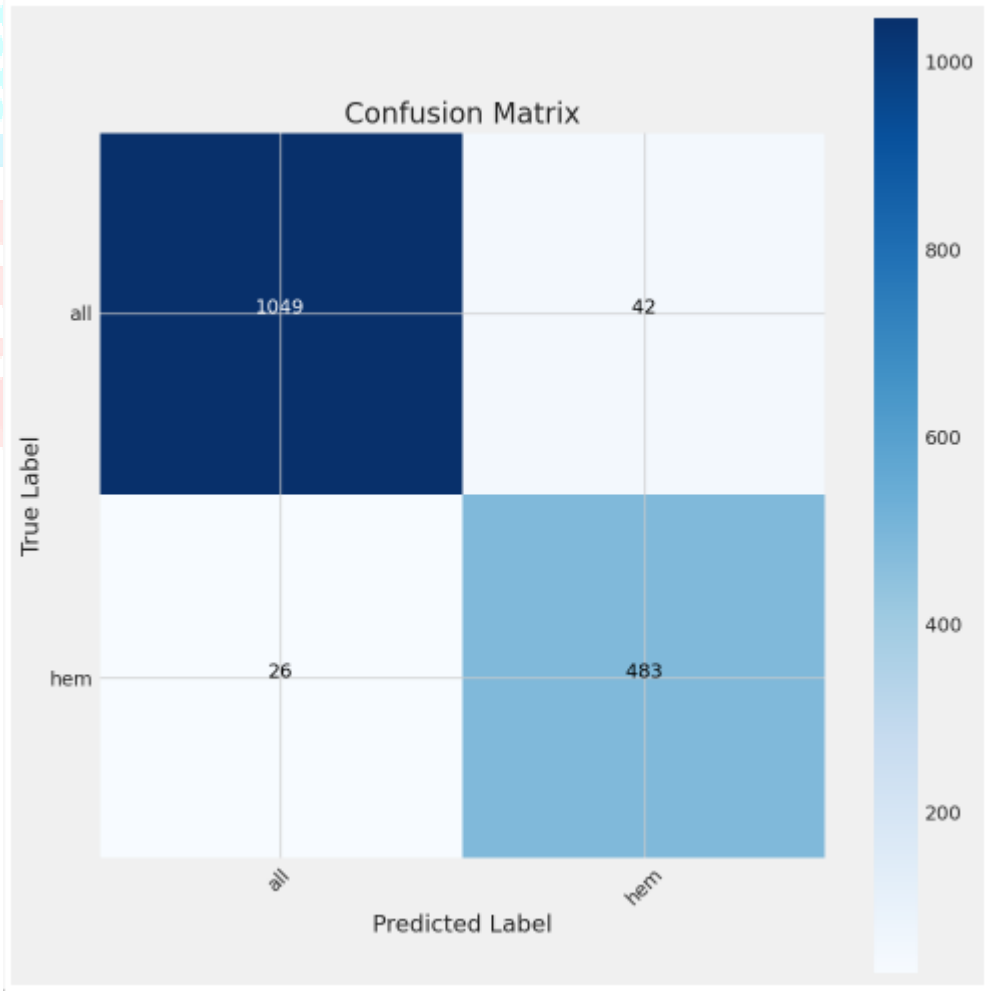

Fig. 3 Confusion Matrix

Fig.4 Loss and Accuracy plot of model

IV. **CONCLUSION**: A significant issue in the field of illness diagnostics is the early detection and diagnosis of leukaemia, that is, the accurate distinction of malignant leukocytes with minimal costs in the early stages of the disease. Flow cytometer equipment is few, and the methods used at laboratory diagnostic centres are laborious despite the high prevalence of leukaemia. Leukaemia can be treated more effectively if it is detected early. This work developed a new classification model for blood microscopic pictures that distinguishes between leukaemia-free and leukaemia-affected images. The general proposed method in this paper consists of three main steps which are: (i) Image_Preprocessing, (ii) Feature Extraction, and (iii) Classification. The proposed model give 96% accuracy.

REFERENCES

[1] F. V. Filipp, ‘‘Precision medicine driven by cancer systems biology,’’ Cancer Metastasis Rev., vol. 36, no. 1, pp. 91–108, Mar. 2017.

[2] T. C. Archer et al., ‘‘Systems approaches to cancer biology systems approaches to cancer biology,’’ Cancer Res., vol. 76, no. 23, pp. 6774–6777, 2016.

[3] T. Vos et al., ‘‘Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990–2015: A systematic analysis for the global burden of disease study 2015,’’ Lancet, vol. 388, no. 10053, pp. 1545–1602, 2016.

[4] (2020) Classification of Blasts inAcute Leukemia Blood samplesUsing k-NearestNeighbour— IEEEConference Publication.
J Denny, MM Rubeena, JK Denny, ―Cloud based Acute Lymphoblastic Leukemia Detection Using Deep Convolutional Neural Networks‖, 2020 Second International Conference on Inventive Research in Computing Applications (ICIRCA), IEEE (2020), pp. 530-536.

[5]Sachin Kumar ,Sumita Mishra, Pallavi Asthana, Pragya, ―Automated Detection of Acute Leukemia Using K-mean Clustering Algorithm, Advances in Computer and Computational Sciences. Advances in Intelligent Systems and Computing, Springer vol 554, (2018) pp: 655-670.