

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

EDIA

International Journal of Innovative Research in SCIENCE | ENGINEERING | TECHNOLOGY

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

٢

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 11, Issue 5, May 2022

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.118

9940 572 462

6381 907 438

 \bigcirc

🛛 ijirset@gmail.com



| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|



Volume 11, Issue 5, May 2022

DOI:10.15680/IJIRSET.2022.1105242

Deep Learning in the Effective Prognosis of Leukemia

T.Selva Kumar^{*1}, Dr. J.Kasthuri^{*2}, Jerushaa Jane K H S³, BalaAnu Devi M⁴, Safiya Fathima M⁵,

Sangeetha K⁶

Assistant Professor, Department of Computer Science Engineering, PSNA College of Engineering and Technology,

Dindigul, TamilNadu, India^{*1}

Assistant Professor, Department of Zoology, The Standard Fireworks Rajaratnam College for Women, Sivakasi, Tamil

Nadu, India^{*2}

UG Students, Department of Computer Science Engineering, PSNA College of Engineering and Technology, Dindigul,

TamilNadu, India^{3,4,5,6}

ABSTRACT: Convolutional Neural Network Algorithm of Deep Learning has been adopted in the present investigation for better prediction and to provide automated analytic solution for the multi-class disease, blood leukemia by following the pipelined stages of pre-processing, segmentation, feature extraction, classification, leukemic stage detection and therapeutics associated with the classified leukemia. The dataset extracted from Kaggle.com has been adopted in the present programme. CNN eliminates the need of manual feature extraction.It makes useof softmax function to effectively to classify the images.CNN learn from the specific pattern within the image and recognize that pattern everywhere. The algorithm analyse the part of a image with 3*3 dimension and multiplies it with the filters to produce feature map. Rectified Linear Unit reduces non-linear features of the image such as transition of pixels, colours, borders etc. It improves the progression of colour of the input images and replaces the pixels with negative value 0."Max pooling", adopted in the present investigation extracts the maximum value of the area where the kernel convolvesand results in the feature map with the most prominent features.Median filtering algorithm, a non-linear approach, removes Gaussian, balanced and impulsive noise in the image. CNN architecture with several pre-processing techniques adopted in the prediction of leukemia achieved the accuracy of 98.80 %. Besides, the specific cancer type has also been diagnosed successfully.

KEYWORDS: Deep Learning, CNN Algorithm, Prediction, Leukemia, Accuracy

I. INTRODUCTION

Leukemia is a neoplastic disease characterized by an excessive proliferation and chaotic differentiation of immature white blood cells and their precursors. This condition warrants the marked increase in the circulating immature or abnormal white blood cells called "blasts". Multiplication of leukemic cell at the expense of normal haematopoietic cell lines causes decrement in the production of normal erythrocytes (Anaemia, weakness, fatigue and pallor), marrow failure, depressed blood cell count (cytopenia, granulocytopenia and thrombocytopenia) and death.Further, infiltration of leukemic cells into spleen, lymph nodes, skin, gingiva and central nervous system makes it as an aggressive cancer type. Based on the cells involved, it can be regarded as lymphocytic (lymphoblastic) leukemia (LL) and myeloid (myelogenous) leukemia (ML)In case of LL, the lymphocytes which could give rise to lymphatic tissue of immune system get impaired. Whereas in ML, the myeloid cells with the capability to evolve red blood cells (RBC), white blood cells (WBC) and platelet producing cells (PPC) are being affected. Besides, according to the rapidity of disease progression rate, leukemia can also be further classified as acute lymphocytic leukemia (ALL), chronic lymphocytic leukemia (CLL), acute myeloid leukemia (AML) and chronic myeloid leukemia (CML). Obviously, of all the four types, ALL seems to be the most predominant one in children. The remaining (CLL, AML. CML) types of leukemia are common among the adults[1].

よう ぶ IJIRSET | e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

Volume 11, Issue 5, May 2022

DOI:10.15680/IJIRSET.2022.1105242

	Lymphocytic Leukemia	Myelogenous Leukemia
Acute	Acute Lymphoblastic Leukemia (ALL)	Acute Myeloid Leukemia (AML)
Chronic	Chronic Lymphocytic Leukemia (CLL)	Chronic Myeloid Leukemia (CML)

Table 1: Four types of Leukemia

Until recently, the causatives of this deadly leukemia are being poorly defined. However, exposure to radiation, chromosomal abnormalities, chemical injuries and viral infections has been implicated. As far as the diagnostic methods are concerned, the microscopic image analysis system is being considered as the main procedure. Since this conventional methodology partly rely on the error-prone, time consuming manual examination and limited by the physical acuity of haematologist's, interventional radiology can also be adopted as an alternative technique in diagnosis of leukemia. However, radiological techniques, suffer from inheriting limitations of imaging modality, sensitivity and resolution of the radio images. Whereas the molecular techniques such as Cytogenetics, Long Distance Inverse Polymerase Chain Reaction (LDI-PCR) and Array - Based Comparative Genomic Hybridization (aCGH) on the other hand, require extensive work and incur high cost. Interestingly, different image-based prognoses are also being applied to diagnose leukemia by adopting Computed Tomography (CT), Magnetic Resonance Imaging (MRI) Scans, X-Rays, and Ultrasound [2].

However, the collections of these imaging modalities remain expensive and time-consuming. Besides these techniques do require the expertise in the fields of Pathology, Hematology and Oncology[3]. Moreover, according to a report published by WHO [4] the scanners of those images are still unavailable in under-developed and in the rural regions of the some developing countries. Consequently, with the view to minimize human intervention and to provide more accurate clinical information at affordable cost, the present investigation attempts to evaluate the automated detection systems for leukemia.

II. LITERATURE SURVEY

As early and accurate diagnosis of leukemia is believed to be successful in reducing the treatment cost, increasing the probability of remission and even the lives of patients [5], it has become imperative to adopt more accurate and rapid automated diagnostic tools which no longer require human intervention. Das *et al.*, [6] developed an algorithmic leukemia classification method to differentiate between various types of leukemia, an automated leukemia detection and classification procedure (ALL).Different algorithms in image processing are being generally employed for the implementation of the automated system. There are very effective and wide range of algorithms are being employed by the researchers in the detection and diagnosis of leukemia. The most popular algorithms include watershed transform, Zack algorithm, gray scale transformation, edge detection, Otsu's algorithm, filtering and thresholding [7] for pre-processing and segmentation purpose. Besides, machine learning and deep learning are proved to be exceptionally effective in the process of implementation.

Mohapatra *et al.*, [8] have proposed a fuzzy-based color segmentation method to segregate leukocytes from other blood components, followed by the nucleus shape and texture extraction as discriminative features. Fathi *et al.*, [9] have applied a Support Vector Machine (SVM) to detect leukemia in the blood cells. The k-means Clustering (KMC)-based segmentation [10] has been employed by Madhukar *et al.*, [11] to extract the leukocytes' nuclei using color-based clustering. Different types of features, such as shape (area, perimeter, compactness, solidity, eccentricity, elongation, form-factor), GLCM (energy, contrast, entropy, correlation), and fractal dimension were extracted from the segmented images[12]. Finally, they have applied the SVM classifier, utilizing K-fold, Hold-out and Leave-one-out cross-validation techniques. Mandal *et al.*, [13] created an image-based approach for cancer diagnosis by extracting critical information from the blood image data and training multiple classifiers. Recently, deep learning achieved many breakthroughs in different fields such as computer vision, natural language processing, and object recognition. Deep neural networks, such as convolutional neural networks (CNNs), can be used effectively to build computer-aided diagnostics.

With this background, the present investigation has been meticulously designed in such a way to make use of Convolutional Neural Network Algorithm of Deep Learning for better prediction and to provide automated analytic solution for the multi-class disease, blood leukemia. Thus the present system follows a pipelined stages of pre-processing, segmentation, feature extraction, classification, leukemic stage detection and therapeutics associated with the classified leukemia.

よう Mail JIRSET | e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

Volume 11, Issue 5, May 2022

DOI:10.15680/IJIRSET.2022.1105242

III. METHODOLOGY

DATASET :

The dataset shown in Figure 1 is extracted from Kaggle.com entitled as "Blood Cell Images". The dataset is also compiled by World Health Organization (WHO). It contains 12,500 augmented images of white blood cells in JPEG format. For each White Blood Cell types (Eosinophil, Lymphocyte, Monocyte, and Neutrophil) there are approximately 3,000 images clustered into 4 folders based on the cell type.

EOSINOPHIL												- 0 X
🕀 New - 🔏] ↑↓ Sort ∽	\equiv View \cdot								
$\leftarrow \rightarrow \sim \uparrow$	🗀 > This PC > De	esktop > BloodCell	> BloodCell > Ker	asCnn > Data > E	OSINOPHIL						~ C	
 ✓ Cuick access Desktop Documents ✓ Downloads ✓ Pictures ✓ Pictures ✓ Pictures ✓ MONOC/TE Review 2 Screenshots ✓ OneDrive - Perso ✓ Dosktop ✓ Documents ✓ Documents ✓ Downloads ✓ Music ✓ Pictures ✓ Videos ✓ Videos ✓ Network 	0.207 0.04480 0.04480 0.06628 0.09211 0.09211 0.09211 0.09211 0.09211 0.09211	0.651 0.4574 0.4574 0.9284 0.9284 0.9284 0.9284 0.9284 0.9284 0.9284	.0.797 .0.4747 .0.4747 .0.7353 .0.7353 .0.9508 .0.9508 .0.9508 .0.9508 .0.9508 .0.9508 .0.9508 .0.9508	0.991 0.5315 0.5315 0.7445 0.9795 0.9795 1.2863	.0_1169 .0_1169 .0_547; Dime .0_547; Dime .0_547; Dime .0_547; Dime .0_7721 .0	0,1414 ype: JPEG File mainons 320 x 240 8.62 KB 0,7803 1,127 1,3931 1,3931 1,7313	0.2142 0.5702 0.5702 0.7902 0.7902 1.293 1.293 1.3991 0.1391	0.2370 0.6660 0.6660 0.7989 1.464 0.14921 0.14921 0.14921	.0.3072 .0.3072 .0.6149 .0.9044 .0.9044 .0.9044 .0.9044 .1.752 .1.5002 .1.5002	4431 	1,009 items	
1,009 items					-						ENG	12:55
Mostly sunny				Q	🗏 👂 🖻	1 🖇 💽	🧧 🗷 💻	🧳 🖾	<u>.</u>		∧ ^{ENG}	(1) ■ 12:55 24:05-2022 ●

Fig 1: Dataset from Kaggle

CONVOLUTIONAL NEURAL NETWORK ALGORITHM :

CNN is a deep learning network architecture which learns directly from data. CNN eliminates the need of manual feature extraction, rather the features are directly learned by them. During training phase CNN passes each input image through the series of convolutional layers followed by pooling layers, fully connected layers and kernels. CNN then uses soft-max function to effectively classify the images.

WORKING OF CNN :

Convolution :

Based on the image resolution, the algorithm holds the pixel value of the input as h*w*d, where h is the image height, w image width and d holds the value three which refers to the dimension of the image. The CNN will learn from the specific pattern within the image and will recognize that pattern everywhere in the image. The algorithm analyse the part of the image with 3*3 dimensions and multiplies it with the filters to produce feature map. This process continues until the entire image is scanned.

Rectified Linear Unit (ReLU):

This phase reduces non-linear features of the image such as transition of pixels, colours, borders etc. It improves the progression of colour of the input images and replaces the pixels with negative value 0.

ReLUY={ 0, for Y<0 and 1, for Y0}

```
(1)
```

| e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|



Volume 11, Issue 5, May 2022

DOI:10.15680/IJIRSET.2022.1105242

Pooling :

Pooling reduces the image dimensionality thereby lowering the network weight to reduce computational complexity. It makes the algorithm capable of detecting patterns in the image irrespective of texture, position, angle etc. Thus pooling prevents over fitting of data. It can be classified as mean pooling, max pooling and sum pooling, The type of pooling adopted in the present investigation is "max pooling", which extracts the maximum value of the area where the kernel convolves. It results in the feature map with the most prominent features.

MEDIAN FILTERING ALGORITHM :

It is a non-linear approach to remove Gaussian, balanced and impulsive noise in the image. It works by moving through the image pixel by pixel and replaces each pixel value with the median value of the nearby pixel. Since median filters preserve edge properties of the image, they are widely used in digital image processing.

$$\sigma_{med}^2 = \frac{1}{4n f^2(\overline{n})} \approx \frac{\sigma_i^2}{n + \frac{\pi}{2} - 1} \cdot \frac{\pi}{2}.$$

(2)

(h)

WORKING OF MEDIAN FILTERING ALGORITHM :

• Pixel value of the image is stored in an array.

a.

- The neighbour pixel value for each pixel value is stored in an array called "Window".
- The median is calculated by first sorting all the pixel values from the window into numerical order, and then replacing the pixel being considered with the middle (median) pixel value.



Fig 2: Median Filtering Algorithm (a) Uploaded image with noise (b) Noise free Image after applying algorithm

GRAYSCALE CONVERSION:

Every image possess three colour channels ie., red, green, blue. Computing images with three colour channels symbolises huge data volume when compared to grayscaled image. Grayscaled image requires less space and is faster in complex computations. It makes use of only one component that is light. Three methods that can be used in effective grayscale conversion are : lightness method, average method and luminosity method. The method used in our analysis is luminosity method which is described as below :

Luminosity Grayscale = 0.3 * R + 0.59 * G + 0.11 * B (3)

Based on luminosity method, the contribution of blue to final value should decrease and the contribution of green to final value should increase.

e-ISSN: 2319-8753, p-ISSN: 2320-6710 www.ijirset.com | Impact Factor: 8.118



||Volume 11, Issue 5, May 2022||

| DOI:10.15680/IJIRSET.2022.1105242 |



(a)

(b)

Fig 3: Grayscale Conversion (a) Uploaded RGB Image (b) Image after grayscale conversion

ARCHITECTURE DIAGRAM :



Fig 4 : Architecture Diagram

IV. EXPERIMENTAL RESULTS

The predictive analysis for this work is implemented in Pycharm by inputting 12,500 images from Kaggle. Among 12,500 images, 8,750 (70%) images are used for training and remaining 3,750 (30%) for testing. Our proposed CNN architecture with several pre-processing techniques achieved the accuracy of 98.80 %. We found that after applying CNN algorithm the accuracy increases with epoch and the training and evaluation loss decreases with epoch which

よう は IJIRSET | e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

Volume 11, Issue 5, May 2022

| DOI:10.15680/IJIRSET.2022.1105242 |

makes our proposed model more robust. The proposed model successfully predicts the type of leukemia in a affected person with the high accuracy input provided.

Lpoen 55, io						
128/128 [=====]	- 74s	575ms/step	- los	5: 0.0624	- acc	: 0.9878
Epoch 36/40						
128/128 [=====]	- 85s	662ms/step	- los	5: 0.0716	- acc	: 0.9838
Epoch 37/40						
128/128 [=====]	- 80s	620ms/step	- los	5: 0.0611	- acc	: 0.9878
Epoch 38/40						
128/128 [=====]	- 74s	579ms/step	- los	s: 0.0937	- acc	: 0.9838
Epoch 39/40						
128/128 [=====]	- 73s	569ms/step	- los	5: 0.0614	- acc	: 0.9880

Fig 5: Accuracy Prediction-98.80%





Fig 6: Accuracy graph - Accuracy increases





Fig 8: Leukemia type Prediction

よう縁 IJIRSET | e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

Volume 11, Issue 5, May 2022

DOI:10.15680/IJIRSET.2022.1105242

V. CONCLUSION

The present investigation on automated prognosis of leukemia is likely to exhibit increased veracity rate with multiclass classification for prediction, vanquishes the improper shape detection of white blood cells. Besides, it is being reported with reduced False Positive Rate (FPR) or Fall-outs. The Median Filtering Algorithm is seemed to be effective in preventing the processing of noisy image inputs. Further, the feature extraction has performed automatically without the intervention of the researcher. CNN algorithm with convolution and pooling layers on the other hand, process the corresponding pixels and automatically extract only the relevant / significant features. Interestingly, the image augmentation with DL provide better coverage of data space and more precise results in terms of pattern recognition by overcoming the problem of over fitting of data. Further, CNN architecture with several pre-processing techniques adopted in the prediction of leukemia achieved the accuracy of 98.80 %. Besides, this system also diagnose the specific cancer type. Thus the authors propose this automated CNN-based detection framework with salient and discriminative features for a quick, safe and accurate early-stage diagnosis of leukemia. Thus this method is expected to play a key role in curing and saving patients' lives.

REFERENCES

- 1. T. T. P. Thanh, Caleb Vununu, Sukhrob Atoev, Suk-Hwan Lee, and Ki-Ryong Kwon International Journal of Computer Theory and Engineering, Vol. 10, No. 2, April 2018).
- 2. American Cancer Society,2018, *Acute lymphocytic leukemia detection and diagnosis*. <u>https://tinyurl.com/3dayesuy</u> [Accessed: 17 October, 2020].
- 3. Gehlot, S., Gupta, A., Gupta, R., 2020. Sdct-auxnetθ: Dct augmented stain deconvolutional cnn with auxiliary classifier for cancer diagnosis. Medical image analysis 61, 101661.
- 4. World Health Organization, . *Global Country Profiles on Burden of Cancer a to k*. <u>https://tinyurl.</u> <u>com/xbtybf7a</u> [Accessed: April, 2020
- Vogado, L.H.S.; Veras, R.D.M.S.; Andrade, A.R.; De Araujo, F.H.D.; e Silva, R.R.V.; Aires, K.R.T. Diagnosing leukemia in blood smear images using an ensemble of classifiers and pre-trained convolutional neural networks. In Proceedings of the 2017 IEEE 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), Niteroi, Brazil, 17–20 October 2017; pp. 367–373.
- Das PK, Jadoun P, Meher S. Detection and classification of acute lymphocytic leukemia. Proc. 2020 IEEE-HYDCON Int. Conf. Eng. 4th Ind. Revolution, HYDCON 2020; 2020. DOI: 10.1109/HYDCON48903.2020.9242745
- Bodzas A, Kodytek P and Zidek J (2020) Automated Detection of Acute Lymphoblastic Leukemia From Microscopic Images Based on Human Visual Perception.Front. Bioeng. Biotechnol. 8:1005. doi: 10.3389/fbioe.2020.01005
- Mohapatra, S., Samanta, S.S., Patra, D., Satpathi, S., 2011. Fuzzy based blood image segmentation for automated leukemia detection, in: 2011 International Conference on Devices and Communications (ICDeCom), IEEE. pp. 1– 5.56.
- 9. Fathi, E., Rezaee, M.J., Tavakkoli-Moghaddam, R., Alizadeh, A., Montazer, A., 2020. Design of an integrated model for diagnosis and classification of pediatric acute leukemia using machine learning. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine 234, 1051–1069.
- 10. Liu, Y., Long, F., 2019. Acute lymphoblastic leukemia cells image analysis with deep bagging ensemble learning, in: ISBI 2019 C-NMC Challenge: Classification in Cancer Cell Imaging. Springer, pp. 113–121.
- 11. Madhukar, M., Agaian, S., Chronopoulos, A.T., 2012. New decision support tool for acute lymphoblastic leukemia classification, in: Image Processing: Algorithms and Systems X; and Parallel Processing for Imaging Applications II, International Society for Optics and Photonics. p. 829518.
- 12. <u>Ondimu, S., Murase, H., 2008. Effect of probability-distance based markovian texture extraction on discrimination in biological imaging. Computers and Electronics in Agriculture 63, 2–12. 58</u>
- Mandal S, Daivajna V, Rajagopalan V., 2019, Machine learning based system for automatic detection of leukemia cancer cell. 2019 IEEE 16th India Counc. Int. Conf. INDICON 2019 - Symp. Proc. 2019;1–4.DOI: 10.1109/INDICON47234.2019.
- 14. M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting", in Proc. SIGGRAPH, pp. 417-424, 2000.
- 15. A. Criminisi, P. Perez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting.", IEEE Transactions on Image Processing, vol. 13, no.9, pp. 1200–1212, 2004.
- 16. Marcelo Bertalmio, Luminita Vese, Guillermo Sapiro, Stanley Osher, "Simultaneous Structure and Texture Image Inpainting", IEEE Transactions On Image Processing, vol. 12, No. 8, 2003.

よう総 IJIRSET | e-ISSN: 2319-8753, p-ISSN: 2320-6710| <u>www.ijirset.com</u> | Impact Factor: 8.118|

Volume 11, Issue 5, May 2022

DOI:10.15680/IJIRSET.2022.1105242

- 17. Yassin M. Y. Hasan and Lina J. Karam, "Morphological Text Extraction from Images", IEEE Transactions On Image Processing, vol. 9, No. 11, 2000
- 18. Eftychios A. Pnevmatikakis, Petros Maragos "An Inpainting System For Automatic Image Structure-Texture Restoration With Text Removal", IEEE trans. 978-1-4244-1764, 2008
- 19. S.Bhuvaneswari, T.S.Subashini, "Automatic Detection and Inpainting of Text Images", International Journal of Computer Applications (0975 8887) Volume 61– No.7, 2013
- 20. Aria Pezeshk and Richard L. Tutwiler, "Automatic Feature Extraction and Text Recognition from Scanned Topographic Maps", IEEE Transactions on geosciences and remote sensing, VOL. 49, NO. 12, 2011
- 21. Xiaoqing Liu and Jagath Samarabandu, "Multiscale Edge-Based Text Extraction From Complex Images", IEEE Trans., 1424403677, 2006
- 22. Nobuo Ezaki, Marius Bulacu Lambert, Schomaker, "Text Detection from Natural Scene Images: Towards a System for Visually Impaired Persons", Proc. of 17th Int. Conf. on Pattern Recognition (ICPR), IEEE Computer Society, pp. 683-686, vol. II, 2004
- Mr. Rajesh H. Davda1, Mr. Noor Mohammed, "Text Detection, Removal and Region Filling Using Image Inpainting", International Journal of Futuristic Science Engineering and Technology, vol. 1 Issue 2, ISSN 2320 – 4486, 2013
- 24. Uday Modha, Preeti Dave, " Image Inpainting-Automatic Detection and Removal of Text From Images", International Journal of Engineering Research and Applications (IJERA), ISSN: 2248-9622 Vol. 2, Issue 2, 2012
- 25. Muthukumar S, Dr.Krishnan .N, Pasupathi.P, Deepa. S, "Analysis of Image Inpainting Techniques with Exemplar, Poisson, Successive Elimination and 8 Pixel Neighborhood Methods", International Journal of Computer Applications (0975 – 8887), Volume 9, No.11, 2010





Impact Factor: 8.118





INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY





www.ijirset.com