IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Cancer Blood Disorder Detection Using Deep Convolutional Neural Networks: A Comprehensive Review

[1]Pulla Sujarani, [2]Dr. P. Sujatha
[1]Research Scholar, [2]Professor & Head,
[1]Department of Computer Science
[1]Vels Institute of Science, Technology and Advaned Studies (VISTAS), Chennai, India.

Abstract: One of the main health issues affecting people of all ages is cancer blood disorder. This paper's primary goal is to predict Cancer blood disorder in the early stage. This paved a way to propose a comparative study with previous studies based on Deep Convolutional Neural Networks. Furthermore, pre-processing techniques are used to predict these types of issues. In contrast to current DCNN models, several sophisticated models have recently been proposed that offer state-of-the-art performance for various computer vision and biomedical image analysis problems. In this work, we used these Advanced DCNN algorithms to solve issues related to cancer blood disorder that are assessed using several publically accessible microscopic blood image datasets. The outcomes show better performance for tasks, including classification, segmentation, and prediction, when compared to current machine learning and DCNN-based methods. Machine learning methods including Support Vector Machine, Naïve Bayes, Logistic Regression, K Mean Clustering Algorithm, and Decision Tree techniques are compared with the suggested methodology.

***Index Terms* – Deep Convolutional Neural Network, Cancer Blood Disorder, Classification, Machine Learning.**

I. Introduction

Cancer Blood disorder is one of the major issue in all types of age groups. A condition that affects the blood's capacity to function properly is known as a blood disorder. It results in the body's red blood cell concentration falling below normal, and unchecked white blood cell growth gives rise to cancer blood disorder[1]. Children can be harmed by certain blood cancer types. Children and adults may react and exhibit symptoms differently. An estimated 1 in 27 patients worldwide are diagnosed with a cancer blood disease each year. One of the main challenges in medical image processing is to capture the image without losing any crucial information.

A cancer blood disorder results from aberrant cells that have been abandoned and have the potential to spread to other regions of the body. At the moment, it is among the world's major causes of death. Cancer blood disorder impacts the production and function of blood cells in our bodies[2]. Bone marrow is the soft, sponge-like substance in the middle of your bones, and it is where most blood malignancies begin. The stem cells that develop into red blood cells, white blood cells, and platelets are produced by our bone marrow. Blood cells that are healthy control bleeding, combat infection, and transport oxygen throughout the body. When there is a disruption in our body's capacity to create blood cells, a cancer blood problem arises. Blood cancer can cause a variety of diseases because patients with the disease have more abnormal blood cells than normal blood cells[3]. The number of people with cancer blood disorders who have greater survival rates has grown due to advancements in therapy.

Image filtering is one of the most crucial stages in picture processing. It may be used for blur removal, edge detection, noise reduction, and more. Both linear and non-linear filters are used in filtering algorithms. If the input has a lot of noise but a low magnitude, a linear low-pass filter is employed .If the input has a lot of noise but a large magnitude, a non-linear filter is used. Linear filters are the most often used filters due to their speed and ease of usage [4]. Unlike non-linear filters, the linear filtering approach applies the algorithm to the image's input and surrounding pixels. It is possible to improve the image's pixel quality by using image filtering. Among the operations that change the pixel values of images are blurring and smoothing.

The goal of the deep learning subfield of machine learning is to replicate the neural connections found in the human brain in order to retrain artificial neural networks to perform complex tasks. When compared to conventional machine learning algorithms, deep learning techniques are essential for the results of image prediction or detection systems because they are superior at processing large datasets quickly and effectively. Algorithms for machine learning typically function well with small datasets but deep learning algorithms outperform with large datasets [5]. Deep learning algorithms require repeated training procedures during learning they can be time-consuming even though they enable a thorough evaluation of dataset features.

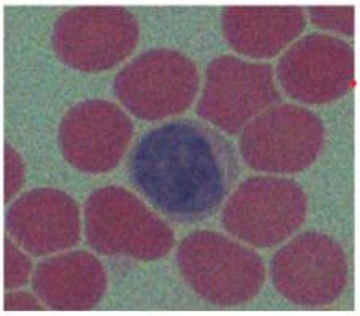

Figure 1.1 Microscopic Blood Image

II. Literature Survey

Based on earlier research, a comparison between cancer blood disorder prediction was created. The researchers employed a number of neural network methods and algorithms.

Nitish et al. [6] utilised a customised quadratic transformation-based Radon transform in conjunction with the Gaussian Filter approach for picture smoothing to analyse the skin cancer for hair detection and eradication. A pixel-wise interpolation technique was then applied for accurate feature extraction and optimum melanoma segmentation. The k-means clustering method is one of the clustering approaches used for picture segmentation. For training and testing, classification algorithms like CNN, Naïve Bayes, and decision trees are employed. The implementation tool is Matlab.

Boyraz et al. [7] suggested that median filtration was used to enhance the image's clarity. After filtering, the threshold technique was used to define the area. The Sobel edge detection technique was also used to evaluate the image. Finally, a tumor found after image processing was evaluated for cancerousness using machine learning. The proposed algorithms were implemented using the C programming language.

Yogeshwari [8] used a unique 2D AADF approach for image filtering. A method called Adaptive Mean Adjustment (AMA) has been developed to increase picture enhancement. To improve photos, methods like thresholding and clustering are used. The features were extracted using the GLCM approach. The last method for identifying illnesses is the DCNN classification architecture.

Pintelas et al. [9] suggested a novel framework for classifying images that extracts properties from the input picture by combining segmentation and clustering techniques. Several state-of-the-art black box strategies are compared with a novel hierarchy-based tree method for the best prediction. In the end, the recommended approach yielded the highest accuracy.

Keerthan et al. [10] proposed the GLCM in image processing for Skin cancer detection. A colour image can be converted to a greyscale image via greyscale translation. Every image processing method is applied to greyscale pictures. GLCM extracts the most common properties, including contrast, mean, energy, and homogeneity, from the grey level picture grid. It is advised to use a support vector machine classification method to determine if a patient's picture is malignant or not.

Yavuz and Eyupoglu [11] applied Principal component analysis is a novel technique and cascaded by median filtering. The categorisation was done using a generalised regression neural network model. The proposed approach has been developed and validated using the recently published Breast Cancer Coimbra Dataset (BCCD), which comprises 116 patients whose 9 clinical features were evaluated.

Thailambal [12] proposed novel methods for improving and filtering illness segmentation. For image thresholding, the Adaptive Otsu threshold approach was recommended, and for clustering, the Improved Fast Fuzzy C Means Clustering (IFFCMC) method was applied. Comparing the recommended approach to the existing processes, the latter yielded the best segmentation result.

Mehmood S. et al. [13] developed a variety of filtering algorithms, including median filter, bilateral filter, and gaussian vs. bilateral filters, and used a variety of filtering techniques and applications for image preprocessing. It alters the image's pixels to give it the desired shape using a variety of graphical editing methods, graphic design tools, and editing software.

Sallam N et al. [14] used a novel bio-inspired optimization-based filtering method called the bilateral filter to remove noise from medical pictures. The two swarm optimisation techniques, Modifiedfly (MF) and Dragonfly (DF), were employed to choose the parameters. Additionally, a suggested convolutional neural network classifier is used to recognise and determine if the denoised picture is normal or pathological. According on the testing findings, the suggested model outperforms the current filters and specific classifications.

Author	Images	Filter Method	Result
Meenu Sharma et al., [15]	256 MRI Scan Images	Hybrid Median Filter	PSNR of 36.56 and MSE of 1.05
Gizem Boyraz, et al., [16]	286 images	Median Filter	PSNR of 32.457 and the error rate of 1.05
Bhishman Desai et al., [17]	1500 Images	Adaptive Diffusion Filter	MSE of 0.892 and PSNR of 37.35
Mohamed Elhoseny, et al., [18]	500 MRI, CT Scan Images	Optimal Bilateral Filter	PSNR of 37.52 dB and an error rate of 1.23
Jihad Nader, et al., [19]	RGB color images	Salt & pepper noise	PSNR 39.7345 and MSE of 1.27

Table 1.1 Different Filtering Techniques

S. Shafique and S. Tehsin [20] In the field of MIA, DL accomplished palatable execution and somewhat simple to assemble a start-to-finish network utilizing CNN. TML models are prepared on physically separated highlights or learn highlights employing other straightforward ML procedures to perform different characterization tasks. Consequently, DL strategies stand out and propelled them to investigate DL's advantages for WBCs order. At present, DL has turned into a strong exploration device in AI, discourse examination, NLP, and clinical imaging.

S. Agaian et al. [21] proposed a simple method for automatically identifying and classifying AML in blood smears is demonstrated. The suggested framework was able to extract the lymphoblast cells from the sub images and entire images with 98% accuracy after testing 80 microscopic blood images.

R. D. Labati et al. [22] As indicated by a new survey on DL based MIA, DL calculations and especially convolutional networks, have turned into a decision for some for investigating clinical information. These techniques are especially reasonable to those areas where human-like intelligence is required to break down a lot of information. Furthermore, great knowledge is expected to extricate rich elements from a monstrous crude information volume.

R. Lipton et al. [131] set forward models that depended upon move learning and could separate among solid and sick BSI. First Model has adjusted pre-prepared CNN called Alex Net to get scored highlights, and different other notable classifiers like K-NN, LD, and SVM, DT were utilized for gathering. SVM classifiers have demonstrated to essentially be best classifier. Subsequent method has involved Alex Net for Double purposes, i.e., for include extraction as well as characterization. Pragmatic was hung on a data set involved 2820 images giving, and subsequent model accomplishes greatest exactness as compared to first model, as subsequent model acquired 100 percent order precision.

M. Hammad et al. [145] The suggested lightweight CNN model uses five learnable layers to make it easier to extract significant characteristics from chest X-ray (CXR) images. Using two larger CXR datasets, the suggested model is assessed. Less memory and fewer parameters are needed for this model. It is used as an additional tool to help radiologists diagnose conditions and is useful for detecting diseases.

R. Gad et al. [148] proposed a self-driven post-processing procedure after an automated segmentation approach for detecting cancerous cells. DenseNet and ResNet architectures are used to identify leukemia, It is seen that the diagnosing force of ResNet-34 and DenseNet-121 overrides generally past methodologies.

III. PROPOSED METHODOLOGY

The cancer blood disorder dataset was collected from Kaggle. Next, the pre-processing step and the filtration method are applied for microscopic blood images. It extracts and apply the model of Deep Convolutional Neural Networks using the selected data. Computers may learn from image samples by using deep Convolutional Neural Networks (DCNNs), which extract internal representations or features that support image grouping or categorization. The deep learning method is used to identify and classify cancer blood disorders. The CNN has two sections: training and testing. CNN's database is produced via deep feature extraction. After that, iterations are used to classify and identify the testing image based on the database. The following fig.1 shows the proposed architecture of blood disorder and cancer prediction.

Convolutional Neural Network (CNN) is a type of Deep Learning Neural Network architecture commonly used in Computer Vision. Computer vision is an area of artificial intelligence that allows a computer to comprehend and analyze visual information, such as images. Neural networks enable decision-making across various industries by enhancing productivity and streamlining processes. CNN performs effectively on image processing and recognition tasks. Convolution aims to identify a variety of characteristics in the input at different spatial locations. Convolutional layers are followed by fully connected layers for more complex reasoning based on the recovered characteristics and pooling to lower spatial dimensionality. CNN can learn hierarchical representations of information in images, which makes them highly effective at tasks like segmentation, object recognition, and picture categorization. The basic components of a convolutional neural network include the following:

Convolutional Layers: These layers utilize filters, often referred to as kernels and apply input images to convolutional processes. In order to recognize features like textures, edges, and complex patterns. Pixels' spatial relationships are preserved with the application of convolutional algorithms.

Pooling Layers: Pooling layers reduce the computational cost and the amount of network parameters by reduction the spatial dimensions of the input. A frequently used pooling method that selects the highest value from a group of adjacent pixels is called max pooling.

Fully Connected Layers: Based on the high-level characteristics that the preceding layers have learnt, these layers are in charge of formulating predictions. They link all of the neurons in one layer to all of the neurons in the next layer

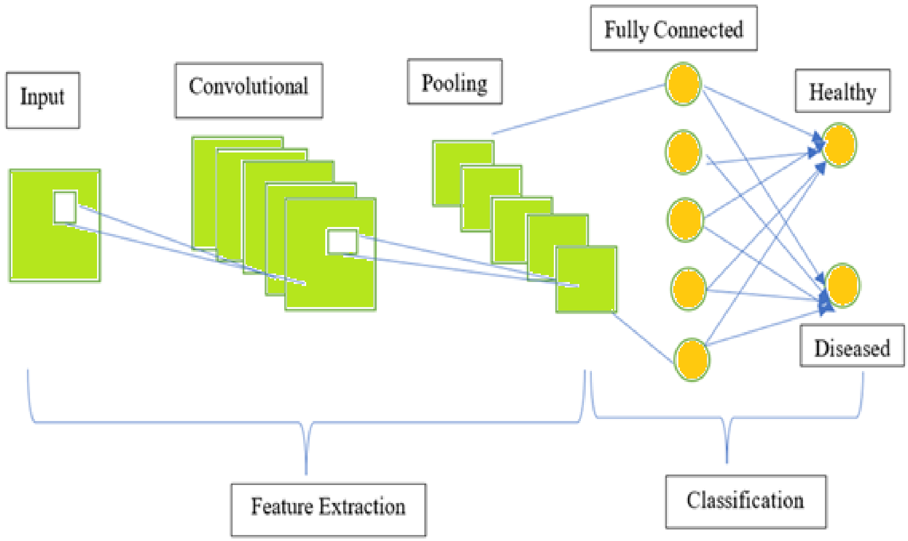

Figure 1. Architecture of Convolutional Neural Network

Naive Bayes

Text classification is one example of a classification application that frequently uses the Naive Bayes classifier. A member of the generative learning algorithm family, it reproduces the input distribution for a certain class or category. This method enables the algorithm to predict outcomes with speed and accuracy. Simple implementation, superior efficiency, handling of continuous data, and probabilistic forecasting capabilities are some of Naive Bayes' benefits. It manages both discrete and continuous data. It is unaffected by a superfluous feature.

Support Vector Machine

Regression and classification issues are the main applications for the SVM technique. The decision boundary for this approach is the hyperplane, which has to be located. When a group has several items, a decision process is required to divide them into different groups. It is necessary to use complex mathematical kernels to separate items into multiple classes if they cannot be divided linearly. SVM aims to identify the objects precisely using samples from the training data set. The following are some benefits of SVM: It can handle complicated functions, structured and semi-structured data, and both if the right kernel function can be identified. Overfitting is less likely using SVM's generalization strategy.

Logistic Regression

Logistic regression is a part of the supervised learning technique and is one of the most often used machine learning algorithms. To predict the result of a categorical dependent variable, logistic regression is utilized. The result must thus be deterministic or categorical. Rather of giving the exact numbers between 0 and 1, it gives the probability values that fall inside that range. True or False, Yes or No, 0 or 1, etc., are all possible. Logistic regression has several benefits, including processing economy, ease of regularization, and simplicity of application. Resizing the input's features is not necessary. This approach is often used to handle large-scale business-related issues.

Decision Tree

One of the best supervised learning methods for applications involving regression and classification is the decision tree. Every internal node defines a test on an attribute, and every leaf node has a class label. A test result is indicated by each branch. A tree structure that resembles a flowchart is the end consequence of this. It is created by continuously splitting the training data into subsets based on the attribute values until a stopping condition is satisfied, such as the tree's maximum depth or the bare minimum number of samples needed to split a node. The following advantages of decision trees include their suitability for classification and regression issues, their ease of interpretation, their ability to handle both quantitative and categorical variables, and their high performance due to the efficiency of the tree traversal method.

IV. Results and Methodology

The comparative analysis of several authors' research using various approaches and their accuracy is displayed in Table 2. Based on this performance, the accuracy of the suggested DCNN Network is higher than that of earlier research. A comparison of several methods for studying cancer blood disorders is presented in Table 2.2 below. The accuracy of DCNN is the greatest among the four machine-learning approaches.

Table 2. Comparison of different classification techniques

Method	Classification Algorithm	Overall Accuracy (%)
Proposed	**Deep Convolutional Neural Networks (DCNN)**	**96%**
Existing	**Naïve Bayes**	92.24%
	Support Vector Machine	91.64%
	Logistic Regression	.45%
	Decision Tree (DT	86.94%

V. CONCLUSION

Comparing with previous studies that have used neural networks is the primary goal. The researchers took a variety of factors into account when predicting cancer blood disorder. In this study, we present the best method for detecting cancer blood disorder, which is DCNN. Support Vector Machine, Naïve Bayes, Logistic Regression, K Mean Clustering Algorithm, and Decision Tree methods are among the machine learning algorithms that are compared to the suggested technique. The DCNN method is the most effective for prediction and has a quick training time. So, DCNN yields the best accuracy for Cancer blood disorder. The results showed that adding clinical data and image-based characteristics significantly improved the model's ability to differentiate between various cancer blood disorders.

References

[1] Y. Z. Tan, K.S Slim, F. F Ting, " Breast Cancer Detection using Convolutional Neural Networks for Mammogram imaging system", International conference on Robotics, Automation and Science (ICORAS), 2017.

[2] Kuo men, Jianrong dai, yexiong Li, " Automatic Segmentation of the clinical target volume and organs at risk in the planning CT for rectal Cance using Deep Dialated Convolutional Neural Networks", The international journal of Medical Physics research and Pracrice, Vol.44, 6377-6389, 2017.

[3] Shekoofeh Azizi, Sharareh Bayat, Pingkun Yan, Amir Tahmasebi, Jin Tae Kwak,Sheng Xu, Baris Turkbey Peter choyke, Peter pino, Bradford, "Deep recurrent neural networks for prostate cancer detection: Analysis of temporal enhanced ultrasound", IEEE transactions on Medical Imaging, Vol.37, 2018.

[4] Steven Walczak, Vic Velanovich. "Improving prognosis and reducing decision regret for pancreatic cancer treatment using artificial neural networks", ElSEVIER Decision Support Systems, Vol. 106, 110-118, 2018.

[5] Ashok kumar, Dwivedi, "Artificial Neural Network model for effective cancer classification using microarray gene expression data", Neural Computing and Applications, 29,1545-1554, 2018.

[6] Nitish, M., Pasupathieswaran, V., Mahalakshmi, B, "Skin cancer analysis using image processing techniques", Int. J. Adv. Res. Innov. Ideas Educ. (IJARIIE) 7(2), 1605–1611, 2021.

[7] Boyraz, G., Öztürk, F., Kırcı, P, " Cancer detection with an ımage processing application", Avrupa Bilim ve Teknoloji Dergis, 1442–1146 , 2021.

[8] Yogeshwari, M. and Thailambal, G. (2023) 'Automatic feature extraction and detection of plant leaf disease using GLCM features and convolutional neural networks', Materials Today: Proceedings, Vol. 81, pp.530–536, 2021.

[9] Pintelas, E., Liaskos, M., Livieris, I.E., Kotsiantis, S., Pintelas, A novel explainable image classification framework: case study on skin cancer and plant disease prediction. Neural Comput. Appl. 33(22), 15171–15189, 2021.

[10] Keerthan, N., Keerthi, S., Likhit, S., Samyama, M., Rao, A.V.: Skin cancer detection using ımage processing. J. Emerg. Technol. Innov. Res. 7(6), 1545–1548, 2020.

[11] Yavuz, E., Eyupoglu, C, " An effective approach for breast cancer diagnosis based on routine blood analysis feature", Med. Biol. Eng. Compu. 58, 1583–1601, 2020.

[12] G.Thailambal, M. Yogeshwari "Automatic segmentation of plant leaf disease using improved fast fuzzy C means clustering and adaptive otsu thresholding(IFFCM-AO) algorithm", European Journal of Molecular & Clinical Medicine(EJMCM), Volume 7, 2020.

[13] Mehmood, S., et al.., "Malignancy detection in lung and colon histopathology images using transfer learning with class selective image processing", IEEE Access 10, 25657–25668, 2022.

[14] Sallam, N.M, Saleh, A. I, Arafat Ali, H, Abdelsalam, "An efficient strategy for blood diseases detection based on grey wolf optimization as feature selection and machine learning techniques", Appl. Sci. 12(21), 1–23, 2022.

[15] Meenu Sharma, Parveen, R, "The application of image processing in liver cancer detection", International Journal of Advanced Computer Science and Applications (IJACSA), Vol. 12, No. 10, pp.448–457, 2021

[16] Gizem Boyaraz, Ozturk F, " An Image Processing Application for Cacner Detection", 1st International Conference on Applied Engineering and Natural Sciences ICAENS 2021, Konya, Turkey, 1 - 03, pp.11-12, 2021.

[17] Bhishman Desai, Utkarsh Kushwaha, Shivam Jha, "Image Filtering - Techniques, Algorithm and Applications" GIS Science Journal, Volume 7, Issue 11, 2020.

[18] Elhoseny M, Shankar k, "Optimal bilateral filter and convolutional neural network based denoising method of medical image measurements", Measurement 143, 125–135, 2019.

[19] Jihad Nader, Ziad A. A. Alqadi, Bilal Zahran, "Analysis of Color Image Filtering Methods", International Journal of Computer Applications, Volume 174, No.8, 2017.

[20] S. Agaian, M. Madhukar, and A. T. Chronopoulos, "Automated screening system for acute myelogenous leukemia detection in blood microscopic images," IEEE Systems journal, vol. 8, no. 3, pp. 995–1004, 2014.

[21] S. Shafique and S. Tehsin, "Acute lymphoblastic leukemia detection and classification of its subtypes using pre-trained deep convolutional neural networks," Technology in Cancer Research & Treatment, vol. 17, 2018.

[22] R. D. Labati, V. Piuri, and F. Scotti, "All-IDB: acute lymphoblastic leukemia image database for image processing," in Proceedings of 2011 18th IEEE International Conference on Image Processing, pp. 2045–2048, IEEE, Brussels, Belgium, Sept. 2011.
[23] R. Lipton, T. Schwedt, and B. Friedman, "GBD 2015 Disease and Injury Incidence and Prevalence Collaborators. Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries", 1990-2015. a systematic analysis for Global Burden of Disease Study 2015," Lancet, vol. 388, no. 10053, pp. 1545–1602, 2017.
[24] M. Hammad, M. H. Alkinani, B. Gupta, and A. A. Abd El-Latif, "Myocardial infarction detection based on deep neural network on imbalanced data," Multimedia Systems, pp. 1–13, 2021.
[25] R. Gad, M. Talha, A. A. A. El-Latif et al., "Iris recognition using multi-algorithmic approaches for cognitive internet of things (CIoT) framework," Future Generation Computer Systems, vol. 89, pp. 178–191, 2018.