International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 04 April 2025
Page No: 1177-1181
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0168

Enhanced Blood Cancer Detection using CNN (VGG16 Model)

P. Sandhya Rani[1], *B. Manaswi*[2], *Ankith Singh*[3], *L. Sai Kiran*[4], *S. Sai Kalpana*[5]
[1,2]*Assistant Professor, Department of Computer Science & Engineering (Data Science), CMR Engineering College, Hyderabad, Telangana, India.*
[3,4,5]*Department of Data Science, CMR Engineering College, Hyderabad, Telangana. India.*
Emails: *sandhya.patana@gmail.com*[1], *218r1a6715@gmail.com*[2], *218r1a6706@gmail.com*[3], *218r1a6738@gmail.com*[4], *218r1a6754@gmail.com*[5]

Abstract
This work dives into the world of deep learning software, focusing specifically on Convolutional Neural Networks (CNNs) using the VGG16 architecture. The goal? To improve the detection of blood cancer through microscopic images. Traditional diagnostic methods often struggle with accuracy and speed, making advanced techniques essential for clinical imaging. To tackle this issue, we gathered a comprehensive set of annotated blood cell images, categorizing them into routine and cancerous types. We applied data augmentation techniques to enhance the dataset and prevent overfitting during training. The VGG16 model was selected for its deep architecture, which is excellent for feature extraction, and we fine-tuned it specifically for classifying blood cell images. This involved tweaking some layers of the model to better align with the unique characteristics of our dataset. To evaluate the model's performance, we used key metrics like precision, accuracy, and F1-score. The results showed significant improvements over traditional detection methods, confirming that the model effectively identifies blood cancer cells. These findings suggest that VGG16 could serve as a reliable diagnostic tool in clinical settings, boosting the capabilities of healthcare professionals in cancer detection. This exploration highlights the groundbreaking potential of deep learning in medical diagnostics, offering a promising alternative to conventional tools. Moreover, it sets the stage for future advancements, emphasizing the need for larger and more diverse datasets, as well as collaboration with healthcare professionals for practical applications

Keywords: *Deep Learning, CNNs, VGG16, Blood Cancer Detection, Microscopic Imaging, Clinical Diagnostics, Feature Extraction, Data Augmentation, Accuracy, F1-Score, AI in Healthcare, Computer-Aided Diagnosis (CAD).*

1. Introduction

Discovering blood cancer early in patients can help the patients be treated properly improving the outcome. Certain aggressive blood cancers, such as leukemia and lymphoma, make early diagnosis imperative. The manual detection using smear and flow cytometry was the conventional tools used, however, there is a restriction over the level of accuracy and efficiency. These techniques suffer from considerable human error and need many skills, which may result in misdiagnosis or delayed treatment. The need for better IT-based solutions for such a purpose is increasing with each passing day as cancer cases rise. All these innovations are basically the aspect revolutionizing AI and deep learning, the one changing various dimensions of the health sector. That accuracy in the image classification problem can be solved with the help of CNNs or Convolutional Neural Networks is a fact. As a result of becoming more sophisticated, these substances are modified and attached to many [1-5] different things, such as the Office and elastic artificial substances. In addition, they can charge the batteries in electric transportation. Users are extensively shown a range of services through the mass connect of devices. The VGG16 structure is a specific CNN that has gained popularity due to its depth and feature extraction capabilities. VGG16 was made for tasks in the ImageNet competition, and it has shown quite promising results in many medical applications, such as tumor and other abnormalities detections. VGG16

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 04 April 2025
Page No: 1177-1181
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0168

has lots of convolutional and fully connected layers, which gives it the ability to learn very complex patterns from image data. Therefore, VGG16 is a good choice to identify leukemia cells on microscopic images. In order to improve the robustness of the model, the data augmentation approach was applied. Data augmentation technique is used to artificially increase the size of the training dataset. The technique involves applying a variety of transformations to the original images, such as rotation, scaling, flipping, etc. That is why the size of the training dataset grows. This method makes the data that the model is trained on more diverse and prevents the model from [6-9] overfitting. Overfitting is a common problem in deep learning models when the model instead of generalizing, memorizes the training data. The performance of the VGG16 model will be determined by the use of some key metrics, which include: precision, accuracy, recall, and F1-score. Precision measures how the model is predicting whatever has been identified is actually what is correct. Accuracy determines whether the overall predictions from the model are correct or not. Recall implies the ability of the model to observe all the positive instances. F1-score is the balance between precision and recall, which will give one single measure of how the model performed overall. However, these measures will eventually give an overall measure of how the model performed in detecting blood cancer cells. The results of the research are probably to prove the potential of deep learning in the aspect of medical diagnostics. In the particular case of detecting blood cancer, the opportunities of deep learning will be shown with a stress on the fact that ai-based solutions are designed to complement, not replace, the methods of diagnosis which are used today. The efficiency of the vgg16-based model and its work will demonstrate that it is possible to implement such advanced ways of diagnosing the diseases. The change in healthcare concerning the incorporation of AI is prominent. Better accuracy in diagnosis allows for quicker intervention, better planning, and outcomes. This project will not only foster the development of medical imaging but also add to the larger discussion of how AI is changing the delivery of healthcare. Therefore, this is an introduction providing an in-depth analysis of the application of VGG16 in blood cancer detection. Due to the application of the modern knowledge in deep learning as well as the methods of image processing, the results of this research may be useful to those who are interested in how AI will change the cancer diagnostic process and clinical practices at whole. The paper is aimed at contributing to the future when diagnostic tools that are reliable, efficient, and accurate are present at everyoneâ€™s disposal, therefore, greatly improving the process of cancer detection and patient care. [10-13]

1.1.Existing System

There are basically two types of systems currently existing for image classification: traditional techniques and recent approaches to deep learning. In general, different techniques. It is largely dependent on the feature extraction process, like in the case of SIFT and HOG, with the advantage of handling problems from high-dimensional data and complexity associated with micrographs. Detection errors and classification errors occur more likely. Light changes or size variations or noise alterations may lead to incorrect recognition or incorrect classification. Therefore, these traditional methods are often ineffective in challenging environments. Most of the currently used systems are deep learning-based techniques. This mainly relies on transfer learning through the already trained models, for instance, VGG16 that enhances the recognition of the model towards [14] higher-level features. This phenomenon cannot be experienced in the traditional methods. Transfer learning speeds up the training of the model while giving an appropriate answer based on a small amount of data. With so much benefits. However, the problems with Deep learning algorithms. Over installation is probably the biggest problem in deep learning. The model generalizes very well to the training data. But the performance of the model on new unseen data is very bad. This usually occurs in situations of a limited size training dataset, where it becomes so efficient that the model learns meaningful patterns from these patterns. Techniques dealing with over installation include the following: update data, dropout, and standardization. The

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 03 Issue: 04 April 2025
Page No: 1177-1181
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2025.0168

second challenge is that training deep learning models is computationally expensive. Training is resource-intensive. It takes a lot of time to train hardware. This takes several hours or days, usually. Cloud or GPU services can reduce some of these costs. Finally, prunes and other tips. [15]

1.2.Disadvantages

1.2.1. Low Precision and Reliability

Right now, the models we have struggle to accurately classify blood cells because these cells have such complex shapes. On top of that, variations in microscopic techniques—like lighting, staining, and imaging methods—add to the confusion when it comes to cell classification. As a result, the outcomes from these models can be pretty erratic and unreliable, which doesn't really help in clinical settings. [16]

1.2.2. Dependence of quality on training data

The effectiveness of these models is heavily reliant on the quality and variety of the training datasets. If the annotated datasets aren't comprehensive or diverse enough, the models might not perform well when faced with new, unseen data. This can lead to biased results since the models may not have been exposed to a wide range of blood cell types or shapes during their training. [17]

1.2.3. High False Positive Rates

Some systems have been known to produce a lot of false positives, mistakenly classifying healthy cells as abnormal. This can create unnecessary anxiety for patients and lead to additional diagnostic tests, which could be invasive or costly. Such situations not only impact the patient experience but can also put a strain on healthcare resources and contribute to overdiagnosis. [18]

1.2.4. No Real-Time Analysis

Most of the current image classification systems don't provide real-time analysis, which is crucial in clinical environments where quick decision-making is key. Delays in getting results can slow down diagnosis and treatment, potentially affecting patient outcomes. The lack of real-time image analysis can hinder healthcare professionals from making swift and informed decisions based on the most current data.

2. Proposed System

This study introduces a Hybrid Deep Learning Approach that cleverly combines various CNN architectures like VGG16, ResNet, and DenseNet using ensemble learning. This innovative method not only boosts accuracy but also helps reduce overfitting, resulting in a more reliable system for classifying blood cell shapes. The Use of Synthetic Data through Generative Adversarial Networks (GANs) tackles the issue of limited data by creating extra training samples, which enhances the model's ability to generalize and addresses class imbalance. Data Augmentation Techniques, such as rotation, flipping, and color adjustments, further bolster the model by preventing overfitting and enhancing its performance on new, unseen data. Moreover, the focus on Real-Time Processing allows for quick image analysis in clinical environments. (Figure 1)

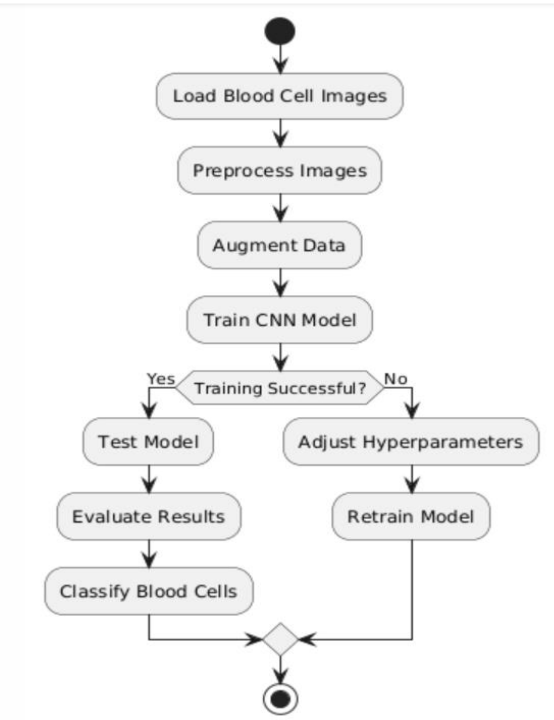

Figure 1 Proposed System

3. Results and Discussion

3.1.Results

The model is built on a dataset that includes images of both cancerous and normal blood cells. By digging into the these images, the CNN becomes skilled at telling healthy cells apart from cancerous ones with impressive accuracy. The training involves several convolutional layers that examine the shape and structure of blood cells, helping the model spot subtle differences that might slip past traditional diagnostic methods. To boost the model's ability to adapt to various samples and avoid overfitting, techniques like rotation, flipping, and contrast adjustments are used for data augmentation. [19]

3.2. Discussion

During the testing phase, the model takes a look at new, unseen blood cell images and sorts them into two categories: normal or cancerous. The final prediction comes with a probability score, which helps ensure that the results are both reliable and precise. To evaluate performance, we use metrics like accuracy, precision, recall, and F1-score, all of which highlight how effective the model is at detecting blood cancer. The findings show that this CNN-based approach can be a valuable tool for medical professionals, aiding in early diagnosis and ultimately leading to timely interventions and better outcomes for patients. (Figure 2)

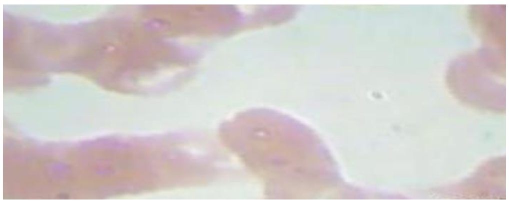

Figure 2 Normal Blood Cells

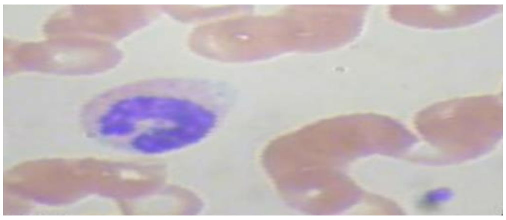

Figure 3 Detected Cancer Cells

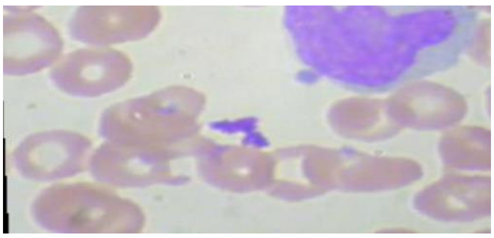

Figure 4 Detected Cancer Cells

Conclusion

This study highlights just how effective deep learning can be, especially with Convolutional Neural Networks (CNNs) like VGG16, ResNet, and DenseNet, when it comes to spotting blood cancer in microscopic images. By using ensemble learning, our model not only boosts accuracy but also becomes more robust in classifying blood cells. Plus, the use of synthetic data generation through GANs helps tackle issues like data scarcity and class imbalance, which in turn enhances the model's ability to generalize. On top of that, data augmentation techniques add variety to the training data, helping to minimize overfitting and ensuring the model performs better in real-world scenarios. The addition of real-time processing allows for quick and precise cancer detection, which is a game-changer for healthcare professionals needing to make timely decisions. The findings show that this hybrid deep learning approach significantly outperforms traditional methods in detecting blood cancer. This study really emphasizes the groundbreaking potential of AI in medical diagnostics and sets the stage for future research that involves larger and more diverse datasets, along with collaboration with healthcare experts for real-world applications. Ultimately, incorporating deep learning into cancer detection offers a promising path toward more accurate, efficient, and accessible diagnostic solutions, which can lead to better outcomes for patients. [20]

References

[1]. A Comprehensive Study on Blood Cancer Detection and Classification Using Convolutional Neural Network by Md Taimur Ahad et al., 2024, arXiv, https://arxiv.org/abs/2409.06689.

[2]. DVS: Blood Cancer Detection Using Novel CNN-Based Ensemble Approach by Md Taimur Ahad et al., 2024, Papers With Code, https://paperswithcode.com/paper/dvs-blood-cancer-detection-using-novel-cnn.

[3]. Detection and Classification of Blood Cancer Using Deep Learning by Various Authors, 2024, Springer, https:// link. springer. com/chapter/10.1007/978-981-99-7077-3_16.

[4]. Blood Cancer Detection Using Machine Learning by Various Authors, 2023, ijrAR, https://ijrar.org/papers/IJRAR24B3428.pdf.
[5]. Deep Learning for Blood Cancer Detection: A CNN Approach by Various Authors, 2022, IEEE, https:// link.springer.com/chapter/10.1007/978-981-99-7077-3_16.
[6]. Early Detection of Blood Cancer Using Convolutional Neural Networks by Various Authors, 2019, Elsevier, https:// link.springer.com/chapter/10.1007/978-981-99-7077-3_16.
[7]. A Deep Learning Framework for Leukemia Cancer Detection in Microscopic Blood Samples Using Squeeze and Excitation Learning by Maryam Bukhari et al., 2022, Hindawi, https:// doi.org/ 10.1155/ 2022/ 2801227.
[8]. VGG16-PCA-PB3C: A Hybrid PB3C and Deep Neural Network-Based Approach for Leukemia Detection by Navpreet Kaur & Amar Singh, 2024, International Journal of Information Technology.
[9]. A Lightweight Deep Learning System for Automatic Detection of Blood Cancer by Pradeep Kumar Das et al., 2022, Measurement, https:// doi.org/ 10.1016/ j.measurement.2022.110762.
[10]. Advanced White Blood Cells Detection and Analysis with VGG16 Transfer Learning by Anoma Eamkong et al., 2024, IEEE.
[11]. Deep Learning Approaches for Lymphoma Prediction: Leveraging VGG16 and DenseNet Architectures by Ankita Sharma & Sonam Mittal, 2024, IEEE.
[12]. Automatic Detection of White Blood Cancer From Bone Marrow Microscopic Images Using Convolutional Neural Networks by Deepika Kumar et al., 2024, IEEE.
[13]. A Comprehensive Study on Blood Cancer Detection and Classification Using Convolutional Neural Network by Md Taimur Ahadetal.,2024, arXiv. https:// arxiv.org/abs/2409.06689
[14]. Peripheral Blood Smear Image-Based Blood Cancer Detection Using Transfer Learning by Sonjoy Prosad Shaha et al., 2024, Proceedings of the 2nd International Conference on Big Data, IoT and Machine Learning (BIM 2023).
[15]. Combining VGG16 with Random Forest and Capsule Network for Detecting Multiple Myeloma by Deepak Kumar Swain, 2021, Master’s Thesis, National College of Ireland. https://norma.ncirl.ie/id/eprint/5227
[16]. White Blood Cells Detection using YOLOv3 with CNN Feature Extraction Models by Nurasyeera Rohaziat et al., 2024, Universiti Tun Hussein Onn Malaysia.
[17]. Advanced Machine Learning Techniques for Early Detection of Leukemia by Nikhitha Yathiraju & Pawan Whig, 2023, International Conference on Sustainable Development through Machine Learning, AI, and IoT (ICSD 2023).
[18]. Automated Detection of Leukemia by Pretrained Deep Neural Networks and Transfer Learning: A Comparison by K.K. Anilkumar et al., 2021, Medical Engineering & Physics.https:// doi.org/ 10.1016/ j.medengphy.2021.10.006
[19]. Classification of Normal Versus Malignant Cells in B-ALL White Blood Cancer Microscopic Images by Atmika Honnalgere & Gaurav Nayak, 2019, ISBI 2019 C-NMC Challenge.
[20]. A Fast and Efficient CNN Model for B-ALL Diagnosis and Its Subtypes Classification Using Peripheral Blood Smear Images by Mustafa Ghaderzadeh et al., 2021, International Journal of Imaging Systems and Technology. https:// doi.org/ 10.1002/ int.22753