Indonesian Journal of Electrical Engineering and Informatics (IJEEI)
Vol. 11, No. 3, September 2023, pp. 778~790
ISSN: 2089-3272, DOI: 10.52549/ijeei.v11i3.4855

Acute Lymphoblastic Leukemia Blood Cells Prediction Using Deep Learning & Transfer Learning Technique

Omkar Subhash Ghongade[1], S Kiran Sai Reddy[2], Yaswanth Chowdary Gavini[3], Srilatha Tokala[1],Murali Krishna Enduri[4]
[1,2,3,4]Algorithms and Complexity Theory Lab, Department of Computer Science and Engineering, SRM University-AP, Amaravati, India

Article Info

Article history:

Received Jun 9, 2023
Revised Sep 3, 2023
Accepted Sep 22, 2023

Keywords:

Leukemia
Deep Learning
Transfer Learning
Neural Network
Convolutional Neural Network

ABSTRACT

White blood cells called lymphocytes are the target of the blood malignancy known as acute lymphoblastic leukemia (ALL). In the domain of medical image analysis, deep learning and transfer learning methods have recently showcased significant promise, particularly in tasks such as identifying and categorizing various types of cancer. Using microscopic pictures, we suggest a deep learning and transfer learning-based method in this research work for predicting ALL blood cells. We use a pre-trained convolutional neural network (CNN) model to extract pertinent features from the microscopic images of blood cells during the feature extraction step. To accurately categorize the blood cells into leukemia and non- leukemia classes, a classification model is built using a transfer learning technique employing the collected features. We use a publicly accessible collection of microscopic blood cell pictures, which contains samples from both leukemia and non-leukemia, to assess the suggested method. Our experimental findings show that the suggested method successfully predicts ALL blood cells with high accuracy. The method enhances early ALL detection and diagnosis, which may result in better patient treatment outcomes. Future research will concentrate on larger and more varied datasets and investigate the viability of integrating it into clinical processes for real-time ALL prediction.

Corresponding Author:

Srilatha Tokala,
Department of Computer Science and Engineering,
SRM University-AP,
Neerukonda, Amaravati, Andhra Pradesh, India.
Email: srilatha_tokala@srmap.edu.in

1. INTRODUCTION

Blood cell malignancies are among the most prevalent and lethal types of cancer, posing a complex and challenging disease that affects millions of people worldwide [1]. Timely and accurate diagnosis of these conditions is crucial for effective treatment and improved patient outcomes. In a single drop of blood, there are over two hundred million red blood cells and nearly 500,000 white blood cells. To maintain these quantities, our bodies must continuously produce fresh blood cells while replacing old ones. The human bone marrow, composed of stem cells responsible for generating all blood cells, serves as the site for this continuous process of blood cell production. Understanding the intricacies of blood cell malignancies requires a comprehensive analysis of various blood cell types and their characteristics. Each stem cell within the bone marrow possesses the remarkable ability to divide repeatedly, generating copies that differentiate into distinct blood cell types. Abnormalities or mutations in the genetic material of these stem cells can lead to the development of blood cell malignancies, such as leukemia, lymphoma, or myeloma.

Early detection and precise diagnosis of blood cell malignancies are critical for implementing appropriate treatment strategies. The identification of specific blood cell types and their abnormalities can provide valuable

insights into the progression and severity of the disease. Consequently, advanced diagnostic techniques and tools are necessary to accurately classify and differentiate between different blood cell types, enabling healthcare professionals to make informed decisions regarding treatment options. In recent years, deep learning models have demonstrated remarkable capabilities in various medical applications, including cancer diagnosis. These models leverage the power of artificial neural networks to analyze complex data patterns and make accurate predictions. By training deep learning models on large datasets of blood cell images, it becomes possible to develop automated systems for the identification and classification of blood cell malignancies. These systems have the potential to enhance the efficiency and accuracy of diagnosis, enabling healthcare providers to intervene at an early stage and initiate appropriate treatment protocols.

Leukemia [2], lymphoma, and myeloma are three of the most common forms of blood cell malignancies, and cancer of the blood is one of the major causes of mortality in the world. The excessive production of aberrant white blood cells results in leukemia, a kind of blood and bones-marrow cancer. Leukemia comes in four different forms. There are four different kinds of leukemia. Children have a greater probability to be diagnosed with acute lymphoblastic leukemia (ALL) [3], but adults are more likely to get acute myelogenous leukemia (AML), which is not hereditary. Both are generated by an excess production of WBC's. The two different types of leukemia are Chronic lymphocytic leukemia and Chronic myelogenous Leukemia [4]. Chronic lymphocytic leukemia is a type of leukemia that is most frequent in adults over 60 and results from the body producing an excessive number of inefficient white blood cells. Chronic myelogenous Leukemia [5] develops in stages and is most prevalent in men. It is brought on by the body's overproduction of white blood cells. Leukemia symptoms include frequent infections, pain in the joints or bones, fever, and sweats at night. The cancer type lymphoma is another one that commonly causes solid tissue mass in the form of lymphadenopathy but can also manifest as an extranodal mass, such as in the brain. The two primary types oflymphoma include non-Hodgkin and Hodgkin lymphoma. There are numerous subtypes of lymphoma [6], which are categorized according to the lymphocytes afflicted, the location and spread of the malignancy, and other considerations. Here is a summary of the non-Hodgkin lymphoma, subgroups of Hodgkin lymphoma. Lymphoma symptoms can include enlarged lymph nodes itching and unexplained weight loss. Another form of cancer called myeloma causes healthy tissues to suffer long-term harm when myeloma cells spread through the circulation and gather in the bone marrow. The disease is known as multiple myeloma because, in most portion instances, cellular myeloma migrates into spaces in all the main bones in the body, causing numerous tiny lesions.

There are three different varieties of myelomas namely [7]

- Multiple Myeloma: It makes up roughly 90% of all cases and is the most prevalent variety. It can beidentified by the bone marrow's accumulation of abnormal plasma cells.
- Smoldering Myeloma: Although the patient does not exhibit any signs of the illness, this type ofmyeloma is distinguished by the presence of abnormal plasma cells within the marrow of the bone.
- Solitary Plasmacytoma: This uncommon form of myeloma grows outside of the bone marrow as asingle tumor made up of aberrant plasma cells.

Myeloma symptoms can include bone discomfort and fractures, anemia, kidney issues, and nerve issues.

Uncontrolled proliferation of aberrant bone marrow and blood cells characterizes the complexand diverse group of disorders known as blood cell malignancies. The three main subtypes of blood cells that can be impacted by these tumors are RBC's, WBC's, and platelets [8]. The prognosis for cancer depends on the form and level of the disease, blood cell cancer symptoms might vary, but they frequently include weakness, exhaustion, fever, weight loss, and swollen lymph nodes. Physical examinations, blood testing, bone marrow biopsies, and microscopic inspection of blood cell samples by a qualified pathologist are currently used to diagnose blood cell malignancy. These techniques have some drawbacks, –such as the fact that theytake a lot of time and are subjective and reliant on the pathologist's experience and knowledge. These techniques can also be ineffective in detecting early-stage blood cell malignancies, which can delay diagnosisand reduce the effectiveness of treatment. For these diseases to be effectively treated and for patient outcomesto be improved, early and precise identification is essential. Convolutional Neural Networks (CNNs) have emerged as a promising tool in the field of image processing and medical diagnostics, holding immense potential for improving cancer detection rates. Numerous studies have demonstrated the ability of CNNs to detect blood cell malignancies [9] with high levels of accuracy, often surpassing conventional diagnostic techniques. The application of CNNs in medical diagnostics has paved the way for more precise and efficient detection of cancer, enabling earlier interventions and improved patient outcomes.

One of the notable advantages of CNNs is their ability to predict how cancer cells will respond to different therapies based on their biological characteristics. By analyzing the intricate patterns and features present in

blood cell images, CNNs can provide insights into the behavior and treatment response of cancer cells. This personalized approach to treatment has the potential to revolutionize cancer therapy by tailoring interventions to the specific needs of individual patients, thereby maximizing treatment efficacy and minimizing adverse effects. In this study, we propose a CNN-based model for the detection of blood cell cancer. Our model takes advantage of the latest advancements in deep learning and image analysis to develop a reliable, scalable, and interpretable framework. By training the model on large and diverse datasets of blood cell images, we aim to create a robust system capable of accurately identifying and classifying different types of blood cell malignancies. To evaluate the performance of our model, we will conduct experiments on a publicly available dataset of blood cell images. We compared the results obtained by our CNN-based model with those obtained using conventional diagnostic approaches. Through this comparison, we seek to demonstrate the effectiveness and potentiality of deep learning techniques in blood cell cancer detection. Furthermore, by providing a proof-of-concept for the application of deep learning in this domain, we hope to inspire further research and advancements in the field. This research aims to harness the power of CNNs and deep learning toenhance the detection and diagnosis of blood cell malignancies. By leveraging the capabilities of these advanced technologies, we strive to contribute to the development more accurately using efficient diagnostic tools for early cancer detection, which ultimately leads to improved patient outcomes and a greater understanding of these complex diseases.

1.1 RESEARCH CHALLENGES

Due to its efficacy, Convolutional Neural Networks (CNNs) have become crucial in the identification and classification of images; yet, various research hurdles still exist. These issues include data accessibility and quality, which are characterized by time-consuming annotation, as well as data imbalance, which results in biassed models that favor majority classes. With small datasets in particular, overfitting is still a problem, prompting research into methods like data augmentation and regularization. The "black-box" nature of CNNs necessitates initiatives to improve model interpretability, and adversarial attacks necessitate novel robustness measures. Domain adaption strategies are required for generalization to unknown data distributions. Furthermore, the high computing demands of CNNs create efficiency problems, which stimulate research into compression and quantization. Integrating knowledge from various sources will boost performance through multi-modal and cross-modal learning.

In order to differentiate minute subcategories within classes, fine-grained categorization is used. Low-latency processing is necessary for real-time picture categorization, which is essential for applications like autonomous driving. The importance of ethical issues, such as bias reduction, as well as privacy-preserving methods for sensitive image data, has increased. There is a need for creative techniques to expand CNNs to low-resource environments. The dynamic nature of CNN-based image categorization is highlighted by these complex problems, which spur on-going research to improve precision, equity, efficacy, and interpretability in a variety of contexts.

1.2 RELATED WORK

The classification of white blood cells (WBCs) has undergone a significant paradigm change with the introduction of machine learning, which divides them into two distinct groups: leukemia-affected cells and unharmed cells. Deep learning approaches have taken on a crucial role in this changing environment, quickly advancing the boundaries of this field's study and consistently producing outstanding results. The quality and richness of the dataset used, the wise selection of pertinent features, the optimization of classifier models, and the computational speed with which the classification procedure is carried out all play a crucial role in the effectiveness of this classification process. The significance of recent contributions that have resonated throughout this discipline will be thoroughly explained in this section. [10]

In the past, meticulous attention has been paid to the time-honored practice of gathering and preserving enormous repositories of image datasets, then subjecting them to a meticulous regimen of methodical training and thorough review. This thorough procedure, which has been specifically designed to account for different leukemia subtypes [7], has been particularly noticeable in the field of acute lymphoblastic leukemia (ALL) diagnosis. These innovative projects have been zealously propelled by an uncompromising effort to improve the accuracy and dependability of ALL detecting systems, supported by a firm dedication to offering cutting-edge solutions. The seamless integration of real-time image processing capabilities into these earlier investigations has also been a key aspect, ushering in a world where relevant healthcare professionals are quickly alerted to the detection of ALL subtypes. This real-time intervention represents a significant advancement in the pursuit of prompt and precise diagnoses, supporting a dedication to proactive healthcare. Through the application of deep learning architectures, most notably as shown by the famous AlexNet, the intricate nuances of preprocessing and the extraction of important characteristics from microscopic leukemia pictures have been carefully addressed in this context.

We have made an effort to reproduce and broaden this powerful approach as part of our own study, extracting an astonishing collection of 1000 important attributes from each image. The Support Vector Machine (SVM) has risen to the top of the classification model hierarchy, routinely outperforming its rivals. This consistent finding is in line with a long list of other research where SVMs have proven their unrivaled effectiveness in a wide range of ALL subtype recognition tasks. As a result, our research adds to and strengthens the vast body of information that already permeates the field of acute lymphoblastic leukemia diagnosis, making it a lasting contribution. By working together, we fan the flames of advancement and advance the persistent development of this crucial aspect of medical diagnostics.

2. RESEARCH METHOD

In this paper, we have used various deep learning [11] and transfer learning techniques like CNN,VGG16-Net and Inception-Net [12]. Transfer learning offers the advantage of reducing the training duration for a neural network model while potentially yielding a decrease in generalization error. These models are widely favored for transfer learning due to their exceptional performance and their introduction of specific architectural innovations. VGG introduced consistent and repeating structures, Google Net introduced inception modules, and ResNet introduced residual modules. In this research, we used inception and VGG models. The model VGG16 was created by the Visual Graphics Group (VGG) at Oxford University and InceptionV3, which is the third iteration of the inception architecture, originated from the development of the Google Net model.

Transfer learning offers a compelling advantage by drastically cutting the time required to train neural network models and possibly lowering generalization error. Pre-trained models that have already discovered meaningful representations from huge datasets are used to do this. These pre-trained models provide a great foundation from which neural networks can learn about low-level features, patterns, and correlations when they are applied to a specific task. Since the model starts with a set of properly initialized parameters, convergence during fine-tuning is accelerated. In situations where computational resources are scarce or where quick model building is necessary, this reduction in training time is very beneficial.

Furthermore, when knowledge is transferred from the source domain to the target domain during transfer learning, generalization error is frequently reduced. During its initial training, the pre-trained model gathers reliable and general features that can be modified to fit the requirements of the target task. Better generalization performance is obtained as a result of the fine-tuning of the model's parameters to match the target data distribution. Transfer learning improves the model's capacity to make precise predictions on unobserved data by utilizing the feature extraction capabilities of the pre-trained model and tailoring them to the specifics of the target task, ultimately leading to better overall performance and higher training efficiency for neural networks.

2.1. Convolutional Neural Network (CNN):

Research challenges related to image identification and classification often rely on convolutional neural networks (CNNs) due to their effectiveness in image processing tasks [13]. CNNs are composed of multiple layers, each serving a specific purpose within the image-processing pipeline. One of the key characteristics of CNNs is their ability to handle high-resolution images with multiple channels, such as RGB channels. The input layer of a CNN typically takes high-resolutionimage, which contains rich spatial and color information. This allows the CNN to capture intricate details andpatterns present in the image. The subsequent layers of CNN, including convolutional layers, pooling layers, and fully connected layers, work together to extract and learn features at various levels of abstraction. The convolutional layers employ filters to convolve across the input image, detecting local patterns and spatialrelationships. The pooling layers down sample the feature maps, reducing their spatial dimensions while preserving the important features. Finally, the fully connected layers aggregate the learned features and make predictions based on the extracted information.

However, the utilization of CNNs for picture identification and classification also have certain challenges.One of the challenges is the need for a large amount of labeled training data. CNNs are data-hungry models and require a substantial number of labeled images to learn the complex representations effectively.

Additionally, CNNs can be computationally intensive, especially when dealing with high-resolution images and large datasets. Training and evaluating CNNs may require significant computational resources and time. Moreover, designing an optimal architecture for a CNN is another challenge. Determining the number of layers,their sizes, and the overall network structure can greatly impact the performance of the model. Finding the right balance between model complexity and generalization capability is crucial. Addressing these challenges requires careful consideration and experimentation. Researchers must ensure access to diverse and well-labeled datasets, explore techniques for model optimization and regularization, and make efficient use of computational resources to train and evaluate CNNs effectively. By tackling these challenges, researchers can

harness the power of CNNs to achieve accurate and robust image identification and classification results.

A convolutional layer with a few filters, each of which is applied to a small portion of the input image, makes up the first layer of a conventional CNN architecture. A non-linear function, such as Rectified Linear Unit (ReLU), is then used to activate the output of the convolutional layer, introducing non-linearity into the network, and enhancing performance. As shown in Figure 1, the spatial dimensions of the output are often decreased by employing a max pooling layer after the convolutional layer. This layer decreases the spatial resolution of the output by a factor of two and picks the maximum value in a restricted area of the output. Theoutput is flattened into a one-dimensional vector after several convolutional and pooling layers before being input into one or more fully connected layers with ReLU activation [14]. To create a final representation of theinput image, these layers learn to merge the information extracted from the convolutional layers. The output layer is then given a softmax activation function, which produces probability values for the outputs of the various classes that the model is taught to recognize. Usually, the number of classes in the dataset equals the number of neurons in the output layer. CNNs are effective tools for classifying and identifying images in research [15]. They can achieve cutting-edge performance on a range of image identification tasks because theyare made up of many layers that learn to extract progressively more complicated information from the input image.

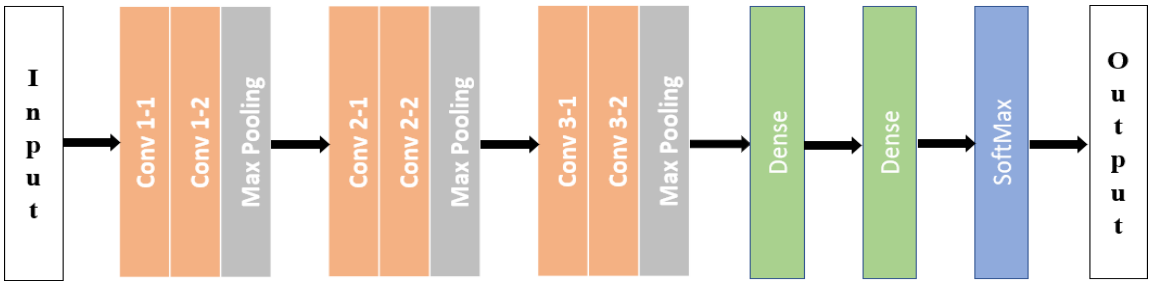

Figure 1. Basic Convolutional Neural Network (CNN)

2.2. Visual Geometry Group 16 Layers Network:

VGG16, a deep convolutional neural network (CNN), wasdeveloped by the Visual Geometry Group (VGG) at the University of Oxford. During the ImageNet Large Scale Visual Recognition Challenge in 2014, the 16-layer network showcased exceptional performance, reaching its peak capabilities. VGG16 is often used as a pre-trained model [16] for numerous computer vision applications, such as object detection, image classification, and segmentation. The VGG16 architecture is composed mostly of convolutional layers, fully connected layers, and softmax layers. Convolutional layers areemployed to extract characteristics from the input image, while fully connected layers are utilized to classify the image. The softmax layer is used to calculate the probabilities of the various classes.

As shown in Figure 2, five blocks, each made up of two or three convolutional layers followed by a max-pooling layer, make up the convolutional layers of VGG16. One portion of the three blocks that follow consists of three convolutional layers, in contrast to the first two blocks which have two convolutional layers each. Each layer contains 64, 128, 256, and 512 filters, respectively. A stride of 1 and a filter size of 3x3 are shared by all convolutional layers. Two hidden layers with 4,096 neurons each make up the VGG16's fully connected layers, which are followed by a softmax layer with 1,000 neurons to represent the ImageNet dataset encompasses a specific count of classes. The retrieved features are mapped to the various classes using the fully linked layers.

In the VGG16 algorithm, every convolutional layer and the initial two fully connected layers are succeeded by the application of a Rectified Linear Unit (ReLU) activation function. The class probabilities aredetermined via a softmax activation function in the last fully linked layer. Dropout regularization is another technique used by VGG16 to avoid overfitting when training [17]. VGG16 is a potent deep learning architecture that, all things considered, has produced outstanding results on several computer vision applications [18]. It is simple to create and modify for many applications thanks to its modular structure and use of common convolutional layers.

As a pre-trained model for transfer learning in the fields of deep learning and computer vision, VGG-16 is quite significant [19]. They are very adaptable to a variety of applications thanks to its deep convolutional layers, which excel in extracting hierarchical and useful features from images. Practitioners can save a significant amount of time and computing resources by utilizing pre-trained VGG-16 [20] models

because the model has already learned the necessary characteristics from a sizable dataset like ImageNet. This functionality is notably useful for applications where it would be impractical to gather labeled data and conduct training from the start. Additionally, VGG-16 is a great place for academics and developers to start because of its broad generalization across several object categories, its cutting-edge performance at the time of release, and its accessibility because of strong community support. Additionally, its interpretability and function as a benchmark model for comparing more recent architectures further reinforce its significance in the field of computer vision, guaranteeing its long-lasting influence on the development of deep learning models.

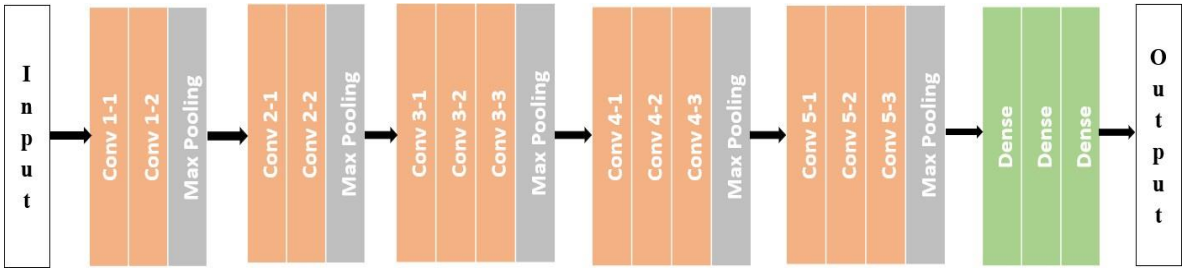

Figure 2. Visual Geometry Group (VGG) 16 Layers Network

One of the limitations of VGG16 is that it is computationally expensive and time-consuming to train because of its enormous depth and number of parameters, which makes it less helpful in real-time and resource-constrained contexts. The 224x224 pixel input requirement for VGG16 necessitates preprocessing steps that may cause information loss or distortion when working with photos of various sizes. Another significant shortcoming is the absence of skip connections or residual connections, which have become crucial for enabling the training of very deep networks and solving the vanishing gradient problem. Due to this omission, VGG16 is less able to produce cutting-edge results on difficult computer vision tasks like object detection and image segmentation. As a result, more modern and successful designs have emerged that outperform VGG16 in a variety of ways, including computational efficiency and task-specific performance.

2.3. Inception-Net:

Inception-Net is a deep learning architecture that Google researchers unveiled in 2014. Itwas created to use a special combination of convolutional layers with various filter sizes, pooling layers, and dimensionality reduction to enhance the accuracy and effectiveness of CNNs. The fundamental component of Inception Net is the inception module as shown in Figure 3, which allows the network to read representationsof pictures at various scales and resolutions, enhancing its capacity to categorize objects in images with variousdegrees of complexity [21]. Each of the stacking inception modules that make up the Inception Net architectureextracts and aggregates data from the input image at various scales and resolutions. Four parallel routes makeup the inception module, which performs 1x1, 3x3, and 5x5 convolutions, and pooling operations, respectively. It is very important to employ 1x1, 3x3, and 5x5 filters together with concurrent pooling procedures. It enables the network to capture a broad range of visual information, from minute local details to more extensive global context. The extraction of fine local characteristics is made easier by the 1x1 convolutions, which increase nonlinearity and channel mixing. The medium- to large-scale spatial information can be effectively captured by the 3x3 and 5x5 convolutions, in contrast. Pooling operations down sample while maintaining critical global features simultaneously. Concatenating these paths enables the network to smoothly integrate local and global features, which is the novelty of the Inception module. As a result, Inception-Net [22] is a strong option for image classification jobs since it is excellent at identifying objects in images of various complexity. In essence, the Network's parallel use of filtering and pooling procedures promotes an in-depth knowledge of images. This multi-scale method improves the network's ability to classify and analyze images by enabling it to simultaneously recognize minute local patterns and broad global structures. The output of the module is created by concatenating these paths, and it is then supplied into the subsequent module in the network.

The number of input channels is decreased, and the network's nonlinearity is increased, using the 1x1 convolution pathway in the inception module. The pooling pathway aids in capturing the most important characteristics of the image at a lower resolution while the 3x3 and 5x5 convolution pathways are intended tocollect spatial information in the image. Along with the inception module, Inception Net also has additional architectural elements that assist the network to perform better during training and generalization [23]. Examples of these include batch normalization and auxiliary classifiers. The activations of the network at each layer are normalized using batch normalization, which expedites training and enhances the network's capacity

to generalize to new data. To encourage the network to learn additional discriminative features and avoid overfitting, auxiliary classifiers are incorporated into the network at intermediate levels. On large-scale picture datasets like ImageNet, Inception-Net has demonstrated state-of-the-art performance as a proficient deep learning structure for the application of the classification of images.

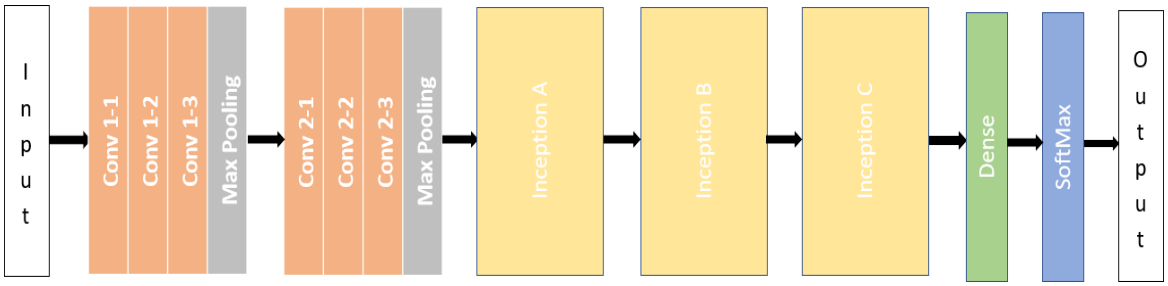

Figure 3. Inception Network

It can learn extremely expressive representations of pictures due to its distinctive blend of convolutional layers with various filter sizes, pooling layers, and dimensionality reduction, and the effective design that makes it possible to train on huge datasets.

3. RESULTS AND DISCUSSION

From Kaggle, we obtained a dataset of 3,242 photos. These pictures were taken in Tehran, Iran's Taleqani Hospital's bone marrow lab. The dataset is divided into benign and malignant categories. Hematogenous pictures are included in the benign class, whereas the malignant class is made up of the Early Pre-B, Pre-B, and Pro-B ALL subtypes of malignant lymphoblasts. 512 photos in the benign class, 979 in the early Pre-B class, 796 in the Pro-B class, and 955 in the Pre-B class are included in this dataset. All the photos were taken with a Zeiss camera mounted on a microscope at 100x magnification, and they were all saved as JPG files. The subtypes and types of the cells were precisely identified and categorized by a professional using flow cytometry technology. For our research, we took 2592 images for training, 322 images for validation, and 328 images for testing the models.

Based on a variety of performance metrics, we have assessed the model's output. Performance metrics are crucial measurements for assessing how well a machine learning algorithm is working [24]. Performance metrics are used to compare how well several models perform on a given data set and to identify the model that yields the best results. Performance indicators can be used to pinpoint areas for model improvement and assess the model's overall accuracy.

The most common performance measures used for deep learning and transfer learning techniques are:

1. Accuracy Score
2. Sensitivity Score
3. Specificity Score
4. Precision Score

Accuracy Score: The accuracy score gauges how accurate a categorization model is all in all. From Eq. 1,we can say that the accuracy score is the ratio of accurate forecasts to all predictions is calculated.

$$AS = \frac{TP+TN}{TP+FN+TN+FP} \quad (1)$$

Sensitivity Score: The sensitivity score, sometimes referred to as recall or true positive rate (TPR), gauges how well a model can identify positive cases. From Eq. 2, we can say that the sensitivity score is the ratio of accurate positive forecasts to all actual positive cases.

$$SS = \frac{TP}{TP+FN} \quad (2)$$

Specificity Score: The specificity score gauges how well a model can recognize negative cases. From Eq.

3,we can say that the specificity score is the ratio of accurate negative predictions to all actual negative cases.

$$SS = \frac{TN}{TN+FP} \quad (3)$$

Precision Score: The fraction of accurate positive predictions and overall positive predictions is measured bythe precision score as shown in Eq. 4. Instead, then concentrating on the model's overall accuracy, it emphasizes the accuracy of positive predictions.

$$PS = \frac{TP}{TP+FP} \quad (4)$$

We have compared all three algorithms on the basis of their performance measures accuracy score, sensitivity score, specificity score, and precision score. These scores will determine how the model will perform. They provide us with the suitability and effectiveness of algorithms for specific tasks and problems. This will help in accessing the main factors like:

1. Evaluation & Comparison
2. Objective Assessment
3. Selection & Optimization Trade-Off Analysis
4. Performance Monitoring

Evaluation & Comparison: Through comparing the performance measures of the algorithms, you gain the ability to assess their strengths and weaknesses relative to each other. This empowers you to make well-informed decisions regarding which algorithm performs better overall or for specific metrics.

Objective Assessment: The performance measures offer an objective means of evaluating the algorithms' performance. By relying on quantitative measures rather than subjective judgments, these scores provide a solid foundation for comparison and analysis.

Selection & Optimization: By considering the performance measures, you can choose the most appropriate algorithm for a given task or problem. The evaluation process helps identify the algorithm that achieves the desired level of performance or meets specific criteria. If none of the algorithms meet the desired performance, optimization techniques can be employed to enhance their effectiveness.

Trade-Off Analysis: Performance measures such as Sensitivity Score, Specificity Score, and Precision Score enable you to analyze trade-offs between different metrics. For instance, you may need to determine whether it's more important to minimize false positives (thus increasing specificity) or false negatives (thus increasing sensitivity). This analysis supports informed decision-making based on the specific requirements and priorities of the problem at hand.

Performance Monitoring: Continuous evaluation of the algorithm performance over time allows you to monitor their effectiveness and track any changes or improvements. This ongoing performance monitoring assists in identifying potential issues, detecting shifts in performance, and making necessary adjustments or optimizations.

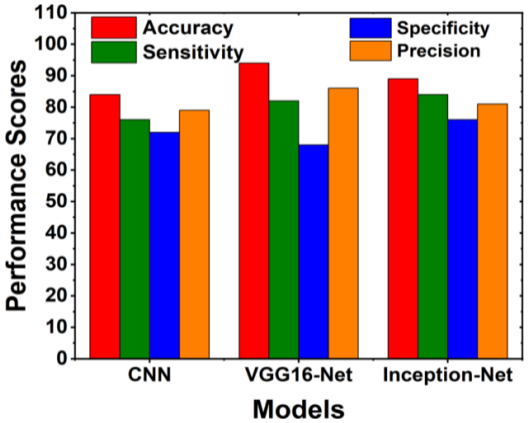

Figure 4. Comparison of accuracy, sensitivity, specificity, and precision scores on CNN, VGG16-Net, and Inception-Net models.

Evaluating and comparing algorithms using performance measures provides valuable insights into their performance, facilitating objective assessment, selection, optimization, trade-off analysis, and performance monitoring.

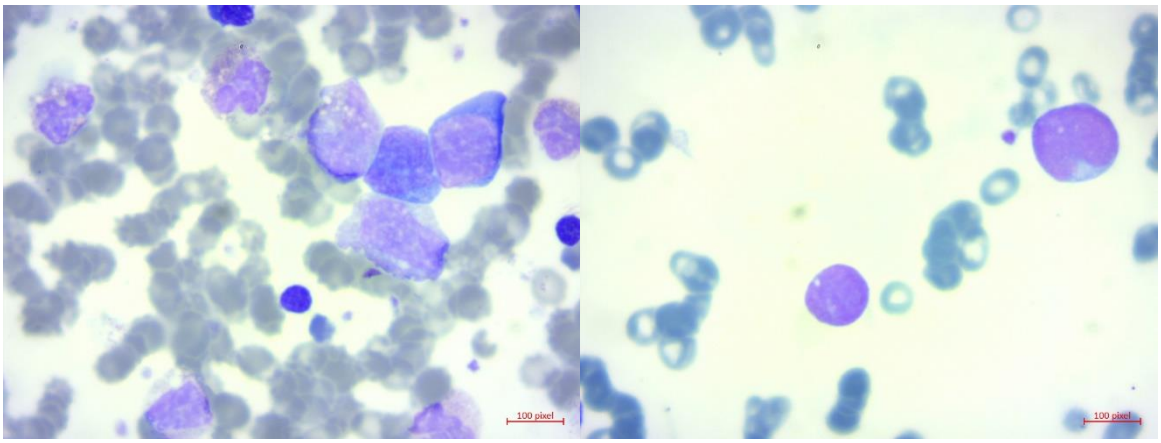

Figure 5. Detected Images by the Models

Our model has been rigorously trained using a dataset organized into four distinct folders, each corresponding to different stages of Acute Lymphoblastic Leukemia (ALL). This approach equips our model with the capability to detect ALL regardless of the specific disease stage, enabling it to analyze and make predictions on images from any stage of the condition. In addition to this, we've conducted a comprehensive comparative analysis involving three prominent algorithms: Convolutional Neural Networks (CNN), VGG-16 Net [23], and Inception Net [24], to assess their performance in the context of ALL detection. This thorough evaluation provides valuable insights into the effectiveness of these pretrained neural networks in enhancing our model's predictive accuracy and overall robustness for this critical medical application.

What sets our approach apart from others in the field is its holistic exploration of multiple algorithms, including CNN, VGG-16 Net, and Inception Net, for the detection of Acute Lymphoblastic Leukemia (ALL). While many prior studies have predominantly focused on a single algorithm, our research embraces a more comprehensive investigation, meticulously evaluating a range of algorithms to identify which achieves the highest accuracy. Furthermore, our methodology benefits from a unique dataset organization, with separate folders representing distinct phases of Acute Lymphoblastic Leukemia. This organizational structure naturally introduces data variability, potentially influencing the accuracy of our models. This distinctive dataset characteristic encourages us to account for nuanced differences in performance across various ALL phases, thereby providing a more refined and realistic evaluation of our algorithms' capabilities within the realm of this critical medical application.

To compare the performance of deep learning and transfer learning techniques for predictingacute lymphoblastic leukemia (ALL) blood cells, a comprehensive study was conducted, and the results are presented in Figure 4 using a bar chart. This chart illustrates the performance metrics that are used to evaluate the models, highlighting the accuracy, sensitivity, specificity, and precision scores. The accuracy scores, represented by red, demonstrate the overall performance of each model. Among the models compared, VGG16-Net achieved the highest accuracy score, indicating its ability to accurately classify ALL blood cells.On the other hand, the CNN model obtained the lowest accuracy score, suggesting that it may have struggled to classify the cells effectively. The sensitivity scores, represented by green, assess the ability of the models to correctly identify positive cases of ALL. In this regard, the Inception-Net model achieved the highestsensitivity score, indicating its proficiency in detecting the presence of ALL blood cells. Conversely, the CNNmodel obtained the lowest sensitivity score, suggesting that it may have exhibited limitations in correctly identifying positive cases. The specificity scores, represented by blue, evaluate the models' capacity to accurately identify negative cases of ALL. Here, the Inception-Net model attained the highest specificity score, indicating its capability to correctly classify non-ALL blood cells. On the contrary, the VGG16-Net model achieved the lowest specificity score, implying that it may have encountered challenges in accurately identifying negative cases. Lastly, the precision scores, represented by yellow, measure the models' precision in correctly classifying positive cases of ALL. VGG16-Net demonstrated the highest precision score,indicating its ability to precisely identify ALL blood cells. Conversely, the CNN model obtained the lowest precision score, suggesting that it may have exhibited some misclassification errors in positive cases.

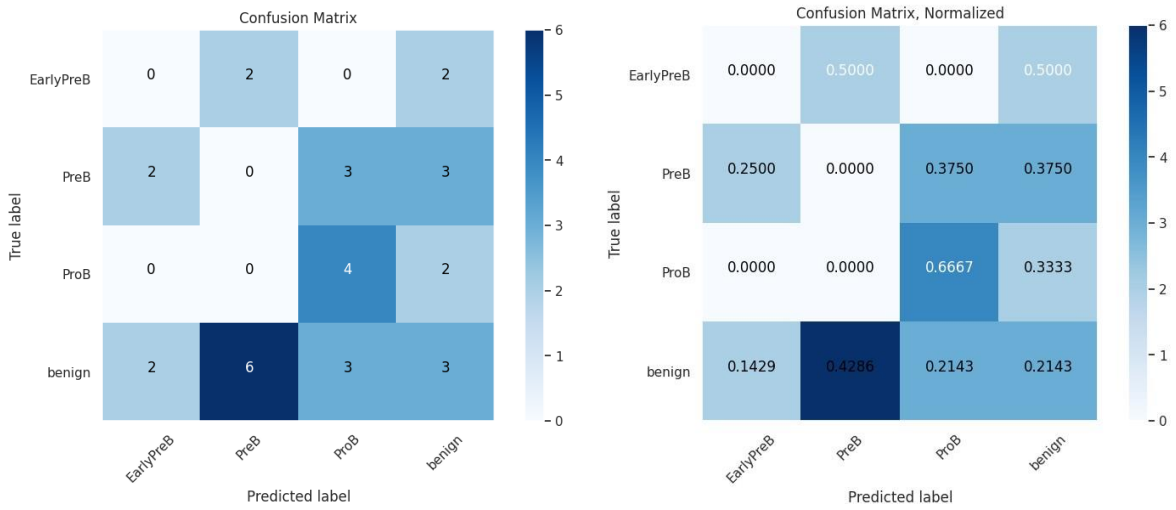

Figure 6. Confusion Matrix and Normalized Confusion Matrix for CNN

Notably, VGG16-Net achieves a 94% accuracy score and an 86% precision score, surpassing the scores of the deep learning model (CNN). VGG16 achieves higher accuracy and precision scores compared toa generic CNN due to its deeper architecture with 16 convolutional layers. The additional layers allow for morecomplex feature extraction, capturing intricate patterns. Furthermore, VGG16 is often pretrained on large-scaleimage datasets like ImageNet, enabling it to leverage the knowledge gained from extensive training. This pretrained weight initialization helps in generalizing well to other image classification tasks. Additionally, VGG16's specific design, optimized for image classification, contributes to its superior performance. The combination of these factors, including the depth, pretrained weights, and architecture design, collectively result in VGG16's improved accuracy and precision compared to a CNN.

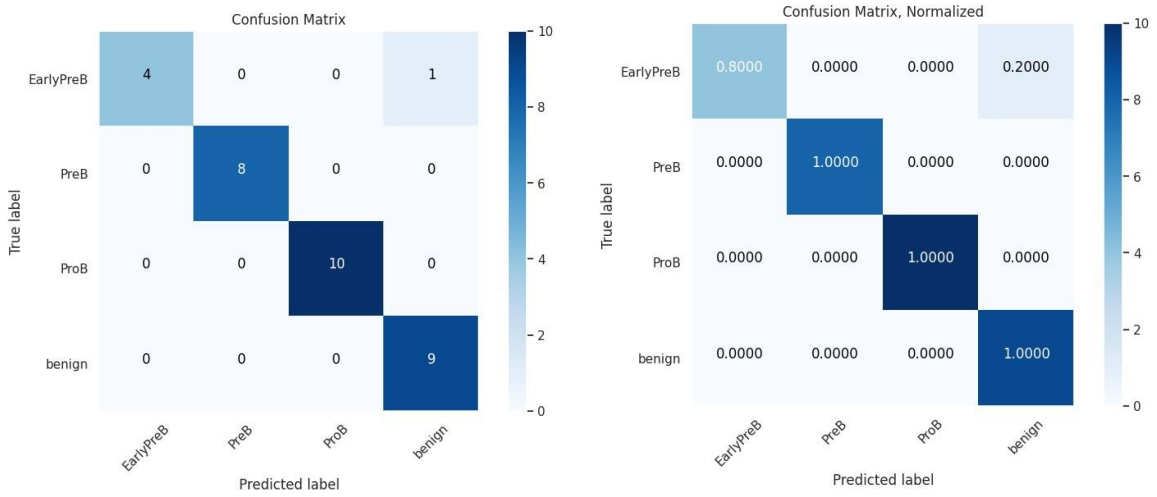

Figure 7. Confusion Matrix and Normalized Confusion Matrix for VGG-16 Net

Notably, Inception-Net achieves an 84% sensitivity score and an 76% specificity score, surpassing the scores of the deep learning model (CNN). Inception-Net achieves higher specificity and sensitivity scores compared to a generic CNN due to its unique architecture. Inception-Net incorporates the concept of "Inception modules" that utilize multiple filter sizes within each layer, allowing it to capture both local and global features effectively. The proposed design of themodel aims to enhance its ability to distinguish between different classes, thereby improving its specificity andsensitivity. In the case of Inception-Net, it benefits from pretrained weights obtained from large-scale datasets, which provide it with prior knowledge for more effective feature extraction. This utilization of pretrained weights, combined with the design enhancements, contributes to Inception-Net's superior performance in termsof specificity and sensitivity compared to a basic CNN.

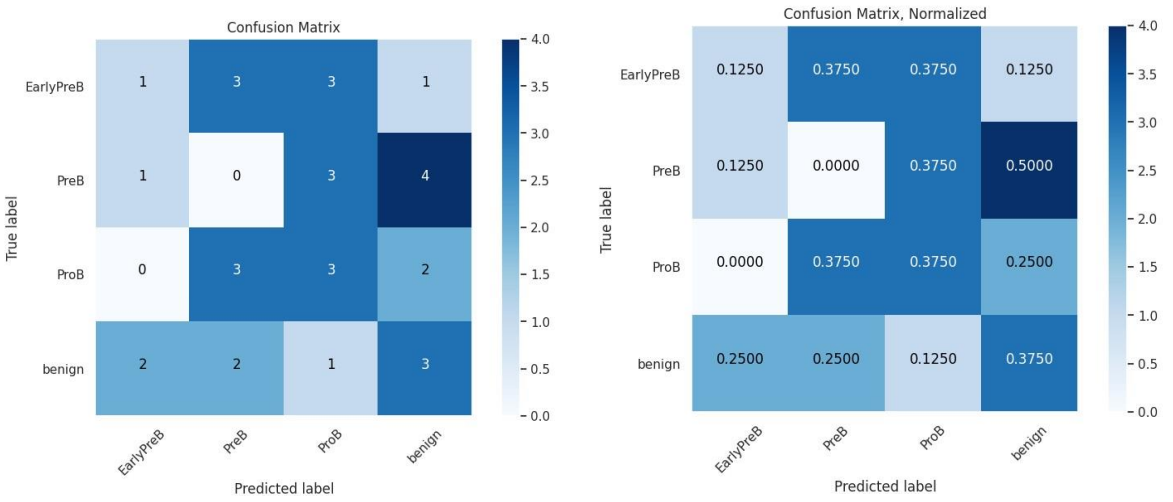

Figure 8. Confusion Matrix and Normalized Confusion Matrix for Inception Net

Furthermore, the transfer learning model VGG16-Net demonstrated better accuracy and precision scores compared to the other models examined. The VGG16-Net model leverages the knowledge learned from a large dataset, enabling it to achieve higher accuracy in classifying the blood cells accurately and with precision. On the other hand, the transfer learning model Inception V3 exhibited good scores in terms of sensitivity and specificity, indicating its effectiveness in correctly identifying positive cases and negative cases of acute lymphoblastic leukemia (ALL). Overall, the comparative analysis highlights the advantages of utilizing transfer learning models in this study. The transfer learning models, such as VGG16-Net and Inception V3, outperformed the basic CNN in terms of accuracy, precision, sensitivity, and specificity. These models leverage pretrained weights and the knowledge acquired from large datasets, allowing them to extract relevant features and improve the accuracy of classification. Consequently, the transfer learning models exhibit superior performance and efficiency compared to the deep learning models employed in this study.

4. CONCLUSION

This study proposes a technique for the timely detection of cancer using small images of white blood cells, employing Basic CNN, VGG16, and Inception V3. Developing a successful deep neural network necessitates a substantial number of labeled images. Among the models compared, VGG16 outperforms Basic CNN and Inception V3, achieving rates of 94% accuracy, 82% sensitivity, 68% specificity, and 86% precision. Firstly, transfer learning models are pretrained on large-scale datasets, enabling them to learn general features and patterns that are applicable across a wide range of tasks. This pretrained knowledge provides a head start, leading to faster convergence and improved performance. Secondly, transfer learning models can effectively leverage the learned representations from the pretrained layers, allowing for efficient feature extraction and better generalization to new data. Lastly, transfer learning models are particularly useful when data is limited, as they can adapt and fine-tune the pretrained weights on smaller datasets, reducing the need for extensive labeled data. To address the challenges of poor image resolution and identifying specific image regions, future explorations can include deep residual networks and U-Net. Another approach worth considering is the Convolutional Network in Density, which exhibits a low error rate and high accuracy in classifying blood cells. Additionally, federated learning can be employed for classification tasks, enhancing security, reducing training time, and yielding favorable accuracy levels.

REFERENCES

[1] R. Zemouri, N. Zerhouni, and D. Racoceanu, "Deep learning in the biomedical applications: recent andfuture status," *Applied Sciences*, vol. 9, no. 8, p. 1526, 2019.

[2] P. H. Kasani, S.-W. Park, and J.-W. Jang, "An aggregated-based deep learning method for leukemic B-lymphoblast classification," *Diagnostics*, vol. 10, no. 12, p. 1064, 2020.

[3] S. N. M. Safuan, M. R. M. Tomari, W. N. W. Zakaria, N. Othman, and N. S. Suriani, "Computer aidedsystem (CAS) of lymphoblast classification for acute lymphoblastic leukemia (ALL) detection using various pre-trained models," in *Proceedings of the IEEE Student Conference on Research and Development (SCOReD)*, pp. 411–415, Batu Pahat, Malaysia, September 2020.

[4] Boldú, L., Merino, A., Acevedo, A., Molina, A., & Rodellar, J. (2021). A deep learning model (ALNet)for the diagnosis of acute leukaemia lineage using peripheral blood cell images. *Computer Methods andPrograms in Biomedicine*, *202*, 105999.

[5] M. Loey, M. Naman, and H. Zayed, "Deep transfer learning InDiagnosing leukemia in blood cells," *Computers*, vol. 9, no. 2, 2020.

[6] M. Macawile, V. Quiñones, A. Ballado, J. Cruz, and M. Caya, "White blood cell classification and counting using convolutional neural network," in *Proceedings of the 2018 3rd International Conferenceon Control and Robotics Engineering (ICCRE)*, pp. 259–263, Nagoya, Japan, April 2018.

[7] Gerecke C, Fuhrmann S, Strifler S, Schmidt-Hieber M, Einsele H, Knop S. The Diagnosis and Treatment of Multiple Myeloma. Dtsch Arztebl Int. 2016 Jul 11;113(27-28):470-6. doi: 10.3238/arztebl.2016.0470. PMID: 27476706; PMCID: PMC4973001.

[8] Yanada, M., Jinnai, I., Takeuchi, J., Ueda, T., Miyawaki, S., Tsuzuki, M., ... & Naoe, T. (2007). Clinicalfeatures and outcome of T-lineage acute lymphoblastic leukemia in adults: a low initial white blood cellcount, as well as a high count predict decreased survival rates. *Leukemia research*, *31*(7), 907-914.

[9] M. Hasan, M. Fatemi, M. Khan, M. Kaur, and A. Zaguia, "Comparative analysis of skin cancer (benignvs. malignant) detection using convolutional neural networks," *Journal of Healthcare Engineering*, vol.2021, 17 pages, 2021.

[10] Renuka, T.V., Surekha, B. (2021). Acute-Lymphoblastic Leukemia Detection Through Deep Transfer Learning Approach of Neural Network. In: Patil, V.H., Dey, N., N. Mahalle, P., Shafi Pathan, M., Kimbahune, V.V. (eds) Proceeding of First Doctoral Symposium on Natural Computing Research. Lecture Notes in Networks and Systems, vol 169. Springer, Singapore. https://doi.org/10.1007/978-981-33-4073-2_17

[11] A. Ramcharan, K. Baranowski, P. McCloskey, B. Ahmed, Legg, and D. P. Hughes, "Deep learning forimage-based cassava disease detection," *Frontiers of Plant Science*, vol. 8, p. 1852, 2017.

[12] Weng, R., Yu, H., Huang, S., Cheng, S., & Luo, W. (2020). Acquiring Knowledge from Pre-TrainedModel to Neural Machine Translation. *Proceedings of the AAAI Conference on Artificial Intelligence*, *34*(05), 9266-9273.

[13] A. Lavric and P. Valentin, "KeratoDetect: keratoconus detection algorithm using convolutional neuralnetworks," *Computational Intelligence and Neuroscience*, vol. 9, 2019.

[14] T. Pansombut, S. Wikaisuksakul, K. Khongkraphan, and A. Phon-On, "Convolutional neural networks for recognition of lymphoblast cell images," *Computational Intelligence and Neuroscience*, vol. 2019, Article ID 7519603, 2019.

[15] U. Pawar, D. O'Shea, S. Rea, and R. O'Reilly, "Explainable AI in Healthcare," in *Proceedings of the2020 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (CyberSA)*, Dublin, Ireland, June 2020.

[16] L. H. S. Vogado, R. D. M. S. Veras, A. R. Andrade, F. H. D. de Araujo, R. R. V. Silva, and K. R. T. Aires, "Diagnosing leukemia in blood smear images using an ensemble of classifiers and pre-trained convolutional neural networks," in *Proceedings of the 30th SIBGRAPI Conference On Graphics, Patterns And Images (SIBGRAPI)*, pp. 367–373, Niteroi, Brazil, October 2017.

[17] Rohaziat, N., Tomari, M. R. M., Zakaria, W. N. W., & Othman, N. (2020). White Blood Cells Detection using YOLOv3 with CNN Feature Extraction Models. *International Journal of Advanced Computer Science and Applications*, *11*(10).

[18] K. Suzuki, "Overview of deep learning in medical imaging," *Radiological Physics and Technology*, vol. 10, no. 3, pp. 257–273, 2017.

[19] A. Genovese, M. S. Hosseini, V. Piuri, K. N. Plataniotis and F. Scotti, "Histopathological Transfer Learning for Acute Lymphoblastic Leukemia Detection," 2021 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA), Hong Kong, China, 2021, pp. 1-6, doi: 10.1109/CIVEMSA52099.2021.9493677.

[20] M. Y. Al-khuzaie, S. A. Zearah and N. J. Mohammed, "Developing an efficient VGG19-based model and transfer learning for detecting acute lymphoblastic leukemia (ALL)," 2023 5th International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), Istanbul, Turkiye, 2023, pp. 1-5, doi: 10.1109/HORA58378.2023.10156679.

[21] Alam, M. M., & Islam, M. T. (2019). Machine learning approach of automatic identification and countingof blood cells. *Healthcare technology letters*, *6*(4), 103-108.

[22] R. R. Maaliw et al., "A Multistage Transfer Learning Approach for Acute Lymphoblastic Leukemia Classification," 2022 IEEE 13th Annual Ubiquitous Computing, Electronics & Mobile Communication Conference (UEMCON), New York, NY, NY, USA, 2022, pp. 0488-0495, doi: 10.1109/UEMCON54665.2022.9965679.

[23] Dwivedi, K., & Dutta, M. K. (2023). Microcell-Net: A deep neural network for multi-class classification of microscopic blood cell images. *Expert Systems*, e13295.

[24] O. S. Ghongade, S. K. S. Reddy, S. Tokala, K. Hajarathaiah, M. K. Enduri and S. Anamalamudi, "A Comparison of Neural Networks and Machine Learning Methods for Prediction of Heart Disease," 20233rd International Conference on Intelligent Communication and Computational Techniques (ICCT), Jaipur, India, 2023, pp. 1-7, doi: 10.1109/ICCT56969.2023.10076174.

BIOGRAPHY OF AUTHORS

Omkar Subhash Ghongade is pursuing his Bachelors in Computer Science at SRM University, AP. He has published few research papers in conferences related to machine learning. Currentlyhe is also pursuing his internship at a multinational Fin-Tech Company. He also did his researchrelated to Deep Learning, Transfer Learning, Computer Vision & Secure Authentication Systems.

S Kiran Sai Reddy is pursuing his B.Tech in Computer Science at SRM University-AP, India. He has published few research articles in reputed International Journals and Conferences. He worked as a SDE Intern at EION Inc, Canada. He secured 1st Runner Up position in the Hackathonbased on Deep Learning and Machine Learning held at IIIT, Sri City. Also, he secured Gold Medal for his Research work in the 5th Ed. Of Research Day held at SRM-AP. His research interests include Machine Learning, Deep Learning, Transfer Learning, Image Processing,Artificial Intelligence and natural language processing.

Gavini Yaswanth Chowdary is currently pursuing the Bachelor of Technology (B.TECH) degree with the Department of Computer Science and Engineering, SRM University, Amaravati , Andhra Pradesh, India. I had also worked on project related to Machine Learning i.eSentiment Analysis.
His research interests include Machine Learning, Deep Learning, and Transfer Learning.

Srilatha Tokala received B.Tech Degree in Computer Science and Engineering from Acharya Nagarjuna University, Guntur, Andhra Pradesh, India in 2013, and M.Tech degree in Computer Science and Engineering from Acharya Nagarjuna University, Guntur, Andhra Pradesh, India in 2015. She is currently pursuing Ph.D. degree in Computer Science and Engineering from SRM University-AP, Amaravati, India. Her research interests include Recommender Systems, Algorithms, Machine Learning, and Complex Networks.

Dr. Murali Krishna Enduri is an Assistant Professor at the Department of Computer Science and Engineering at SRM University-AP, Amaravati, India. He holds a M.Sc. in Mathematics fromthe Acharya Nagarjuna University, Andhra Pradesh, India in 2008. He obtained his M. Tech in Systems Analysis and Computer Applications from the National Institute of Technology Karnataka, India in 2011. He received his Ph. D in Computer Science and Engineering from the Indian Institute of Technology Gandhinagar, Gujarat, India in 2018. He has research experience/post Doc at Indian Institute of Technology Madras, Tamil Nādu, India in 2018. Murali Krishna Enduri's research interest includes Recommender Systems, Algorithms, Complexity Theory, Machine Learning, and Complex Networks.