Advanced Hybrid Deep Learning Model for Precise Multiclass Classification of Bone Marrow Cancer Cells

Shiekh Rahmatullah Sakib, Kamarun Nahar Sara, Md.Anisul Islam, and M M Fazle Rabbi

Department of CSE, Bangladesh University of Business and Technology, Dhaka, Bangladesh.
{kishorsakib990@gmail.com,
naharsara12@gmail.com,an7471965@gmail.com,rabbi@bubt.edu.bd}

Abstract. Bone marrow cancer is when rogue blood cells overgrow in the bone marrow by disrupting regular blood cell production. Accurate classification of these cancers is crucial for effective treatment planning and patient management. Leukemia and myeloma (plasma cell cancer), one types of malignancy that can damage the white blood cells (WBC) within the bone marrow. White blood cell identification, counting, and segmentation are crucial steps in effectively studying a few malignant tumours. In this study, an automated classification method has been proposed for plasma cell cancer which are Multiple Myeloma (MM), Acute Lymphocytic Leukemia(ALL) and Acute Myeloid Leukemia (AML). This bone marrow model image is preprocessed and trained with the parameterized hybrid convolutional neural network and also compared with the CNN framework (InceptionV3, ResNet50, and Vgg16) to achieve accurate classification results. The optimal model was selected by identifying the one with the lowest loss for the validation data. Achieving a high accuracy rate of 99.58% was made possible through the development of hybrid model algorithms, which were carefully crafted by monitoring training loss and validation loss to identify the optimal value. This process of monitoring the training and validation of a deep learning model can help identify the optimal accuracy and loss values. This proposed model can reduce classification time, condense image information and speed up processing times with more precise weight limits.

Keywords: Augmentation, Hybrid Model, CNN, LSTM, Pretrain Model, Deep Learning

1 Introduction

The diagnosis of bone marrow malignancy, which includes acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML) and multiple myeloma (MM)

is crucial for the development of an appropriate treatment strategy. Bone marrow cancer is known as a bad tumor that affects the blood and it's dangerous because it makes too many harmful white blood cells[1]. In the realm of blood-related malignancies, a quartet of subtypes takes centre stage: ALL, AML, MM. The emergence of AML or Acute Myeloid Leukemia, is precipitated by the swift escalation of white blood cells (WBCs), arising from issues within the bone marrow's developmental process or other underlying factors. Conversely, the genesis of ALL or Acute Lymphocytic Leukemia, in children is frequently attributed to the accelerated multiplication of white blood cells[2]. Multiple myeloma, a form of blood cancer, exerts its impact by targeting plasma cells nestled within the bone marrow. Globally, about 588,161 people are diagnosed with it every year. In the United States alone, there are 34,920 new cases annually[3].

To assist with leukemia detection, segmentation and classification in automated haematological analysis, several image-processing algorithms have been developed. Accurate image segmentation is crucial for accurate haematological analysis[4]. Microscopic images of bone marrow samples, with their intricate details and cellular structures, provide a valuable resource for diagnosis. This paper presents an innovative approach that harnesses the power of deep learning by integrating Convolutional Neural Networks (CNNs) and Long Short-Term Memory Networks (LSTMs) to detect bone marrow cancer from microscopic images. CNNs have revolutionized image analysis by automatically learning and extracting intricate patterns and features, making them ideal for image classification tasks. LSTMs, on the other hand, excel in handling sequential data, making them a valuable component for analyzing temporal aspects in time-series data, which is essential when tracking changes in cell structures over time. The integration of CNNs and LSTMs enables us to not only capture the spatial information of cell structures but also analyze the dynamic changes that occur during the progression of bone marrow cancer. By combining these two deep learning architectures, our proposed model aims to achieve higher accuracy and precision in the detection of bone marrow cancer than traditional methods. Moreover, this approach has the potential to significantly reduce the reliance on invasive diagnostic procedures.

This study introduces a groundbreaking approach that leverages microscopic images to achieve precise identification and categorization of bone marrow cancer cells, meticulously differentiating between AML, ALL, and MM, all through the skillful utilization of State-of-the-art deep learning methods. This investigation signifies a pivotal advancement towards more precise and non-intrusive techniques for early cancer detection, ultimately outcomes in the realm of hematologic malignancy diagnosis. This pursuit is fueled by an overarching goal: the establishment of a robust and dependable system, one that wields the power to differentiate these cells with unparalleled precision. The contributions of this research are given below-

- **Tailored Cancer Classification:** Developed an automatic classifier for blood cancers, accurately distinguishing Multiple Myeloma, Acute Lymphocytic Leukemia and Acute Myeloid Leukemia.
- **Hybrid Neural Network Model:** Innovated a parameterized hybrid convolutional neural network, surpassing conventional frameworks like InceptionV3, ResNet50 and Vgg16.
- **Optimal Model Selection:** Introduced a strategy for choosing models with minimal validation loss, ensuring peak classifier performance and aiding future model selection.
- **Exceptional Accuracy and Efficiency:** Achieved a remarkable 99.58% accuracy while optimizing efficiency through meticulous hybrid algorithm design and loss monitoring.
- **Enhanced Processing Speed:** Discovered that monitoring training and validation processes improve accuracy and processing speed, promising faster and more precise medical image analysis.

The later part of the research paper has been organized as related works, Dataset, Methodology and conclusion. Their description is given below:
In Related Work section recent research papers have been summarized and it discusses existing research, studies and relevant works related to the proposed model being investigated. In the next section Dataset has two subsections which are Data Collection and Dataset Analysis. This chapter focuses on the crucial aspects of data acquisition and dataset analysis. A detailed account is provided of the diverse collection of bone marrow microscopic images sourced from a variety of channels. The methodologies employed for image annotation and the establishment of ground truth are expounded upon. This chapter also sheds light on the diversity within the dataset, reflecting real-world clinical variations. After that In the Methodology pivotal chapter, A key highlight of this chapter is the introduction of a novel custom-parameterized hybrid model. This model seamlessly amalgamates Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, representing a significant innovation in the study. In Conclusion the research journey reaches its culmination in this chapter, where the findings are synthesized into a coherent conclusion. The innovative hybrid model, S-Net, is reemphasized for its role in bridging gaps in the field. As the chapter draws to a close, the focus remains on the impact of the study and the potential it holds for further developments in the domain.

2 Related Work

Recent advances in deep learning have led to significant improvements in the accuracy of automated bone marrow cancer detection using microscopic images. In particular, convolutional neural networks (CNNs) have been secretive for this task. CNNs are a type of deep neural network that is specifically designed for image analysis. This literature review will discuss recent advances in automated the classification of bone marrow cancer cells using microscopic images.

Priyanka et al.[5] have proposed a novel and effective approach for accurately classifying leukocytes in order to diagnose leukemia. Their approach involves the development of a VGG16-adapted fine-tuned feature-extractor model called "LeuFeatx". By training classifiers with the deep features extracted by LeuFeatx model, the researchers achieved remarkable results with an accuracy of 96.15%. This two-step approach utilizes the ALL-IDB2 dataset for a binary classification experiment, showcasing its effectiveness in diagnosing leukemia. The application of LeuFeatx model has the potential to significantly improve the accuracy and efficiency of leukemia diagnosis, offering great promise in the field of medical research.

Wang et al.[6] have developed an innovative system to tackle the challenges associated with counting white blood cells in varying lighting conditions. Their approach utilizes a combination of a Faster RCNN and a Feature Pyramid Network, which effectively considers the complexity and diversity of color components. To assess the system's performance, the researchers employed a dataset from Zhejiang University's Second Affiliated Hospital for training and testing purposes. Impressively, the proposed system achieved an overall correct recognition rate of 98.8%, highlighting its outstanding accuracy in accurately counting white blood cells. This remarkable achievement showcases the potential of advanced technologies in the field of medical diagnostics. Amjad et al.[7] proposed a system that aims to effectively categorize reactive bone marrow (normal) and Acute Lymphoblastic Leukemia (ALL) into distinct subgroups using stained bone marrow images. Their method combines the power of convolutional neural networks and robust segmentation techniques to train the model on a dataset of bone marrow images, enabling it to accurately classify and differentiate between the two categories. The experimental results obtained were highly promising, with the suggested strategy achieving an impressive accuracy of 97.78%.

Ahmed et al.[8] conducted research on the identification of leukemia subtypes using a CNN architecture. Additionally, they explored various other classification methods, such as support vector machines, decision trees, naive Bayes, and k-nearest neighbors. To evaluate their approach, the researchers implemented a 5-fold cross-validation methodology. The results of their studies indicated that the CNN model achieved accuracy rates of 81.74% and 88.25%. These findings demonstrate the effectiveness of their proposed CNN architecture in accurately identifying different subtypes of leukemia. By comparing the performance with other classification techniques, it further highlights the superiority of the CNN model in this domain of research.

In their study, Satvik et al.[9] proposed a comprehensive pipeline for the categorization of bone marrow cells. To address any class imbalances, the researchers implemented data augmentation techniques throughout the dataset. This involved making arbitrary changes such as translation, zooming in and out, flipping horizontally or vertically, and rotation within a range of 0 to 90 degrees. The CoAtNet model, used in the pipeline, was then compared against two baseline models, EfficientNetV2 and ResNext50. To further analyze the effectiveness

of the CoAtNet model, Grad-CAM and SmoothGrad techniques were employed. The findings revealed that the suggested CoAtNet model outperformed both the ResNext50 and EfficientNetV2 models in accurately categorizing different morphological classes of bone marrow cells. This highlights the potential of the CoAtNet model as a robust solution for bone marrow cell categorization tasks.

In their study, Taufiqul et al.[10] proposed an innovative approach utilizing Deep Neural Networks (DNN) to tackle the task of recognizing different-shaped Acute Lymphoblastic Leukemia (ALL) blast cells in images of minute blood smears. The researchers achieved remarkable results, reporting an impressive accuracy rate of 98% in detecting various subtypes of ALL cells. Building on this advancement, they further developed telediagnosis software that leverages these minute blood smear images to provide real-time diagnostics for different subtypes of ALL.

In their study, Karar et al.[11] propose an intelligent Internet of Medical Things (IoMT) framework designed to automatically identify acute leukemia from microscopic blood images. The framework utilizes wireless digital microscopy technology combined with cloud server transfer and a generative adversarial network (GAN) classifier. The images of the blood samples are captured using wireless digital microscopy and then transferred to a cloud server for analysis. The cloud server applies the GAN classifier to identify different blood conditions, and the classification outcomes are subsequently sent to haematologists for professional clearance. The results indicate that the GAN classifier achieved remarkable accuracy scores in both binary and multi-class classification tasks, with scores of 98.67% and 95.5% respectively.

Ching-wei et al.[12] proposed a hierarchical design for the speedy localisation of BM particles, identification of 16 cell types, and the combination of patch-based discoveries. Their approach consists of three models: a deep learning model for particle localisation, a patch-based model for cell type identification, and a fast stitching model for combining the patch-based findings. The proposed technique achieved high recall and accuracy values of 0.905 ± 0.078 and 0.989 ± 0.006, respectively. To validate their method, the authors performed cross-validation on a dataset containing 12,426 annotated cells, focusing on the classification of leukemic B-lymphoblasts. This innovative approach promises to facilitate the analysis of complex cellular structures and contribute to advancements in the field of cell biology. In their study, k.Sridhar et al.[13] aimed to develop a deep-learning model that effectively handled the limitations of a small dataset size. To overcome this challenge, the researchers employed data augmentation techniques and implemented an exchange learning strategy. The outcomes of their proposed approach surpassed the performance of other distinct networks, achieving an impressive test accuracy rate of 95.59% during the evaluation of Leukemic B-lymphoblast. This accomplishment showcases the efficacy of their model in accurately classifying and identifying Leukemic B-lymphoblast, demonstrating its potential for advancing medical diagnosis and treatment.

Petru et al.[14] introduced a pioneering deep learning system called MILLIE that revolutionizes the analysis of blood films. This system demonstrates remarkable reliability and automation capabilities with minimal supervision. What sets MILLIE apart is its ability to accurately distinguish between acute lymphoblastic and myeloblastic leukemia in blood films without requiring specific cell identification training. Moreover, MILLIE's impact extends beyond blood films alone; it also proves valuable in analyzing bone marrow aspirates, achieving impressive performance with an AUC of 0.99+/-0.01. This breakthrough technology offers immense potential for detecting a specific subtype of leukemia, known as APL, across both blood and bone marrow samples. The development of MILLIE marks a significant advancement in the field of hematopathology, opening new avenues for efficient and reliable leukemia identification.

Kokeb et al.[15] developed a system that successfully classified four common forms of leukemia. This included acute and chronic myeloid leukemia, as well as acute and chronic lymphoblastic leukemia. To achieve accurate classification, the system employed a robust image segmentation technique in combination with support vector machine classification. The results were highly impressive, with the system showing remarkable accuracy, sensitivity, and specificity. Specifically, the system achieved an accuracy of 97.69% for both the test datasets and validation datasets. Furthermore, its sensitivity and specificity were equally impressive, reaching 97.86% and 100% respectively. These findings highlight the effectiveness and potential of leveraging advanced image segmentation and machine learning techniques for accurate leukemia classification.

In their research Mohammed et al.[16] present a three-phase approach for automatically detecting Acute Lymphoblastic Leukemia (ALL) cells from normal White Blood Cells (WBCs) using an ensemble technique. The proposed method consists of three key components: picture pre-processing, a convolutional neural network for deep spatial feature extraction, and a gated recurrent unit-bidirectional long short-term memory (BiLSTM) architecture for capturing long-distance-dependent information characteristics. To classify the cells, a softmax function and a Multiple Support Vector Machines (MSVM) classifier are employed. In the experiment, the MSVM classifier achieved a remarkable F1-score of 96.23%, while the DenseNet-201 model achieved an even higher accuracy of 96.29% for the test dataset. This work signifies a significant advancement in ALL cell detection, demonstrating the effectiveness of applying ensemble techniques and deep learning architectures in medical image analysis. The findings hold promise for improving the accuracy and efficiency of diagnosing leukemia, contributing to better patient outcomes and optimized healthcare processes.

Dwivedi et al.[17] have proposed an impressive CNN-based architecture called Microcell-Net, which aims to address the classification task of peripheral blood cells using microscopic images. The authors trained their model on a comprehensive dataset that consists of microscopic images divided into eight different types of blood cells. The experimental results showcased exceptional performance, achieving a remarkable 97.65% test accuracy and an even higher val-

Table 1. Performance comparison with the existing methods

Methods	**Algorithms**	**Classification Accuracy**
priyanka et al. [5]	VGG16	96.15%
Wang et al. [6]	Faster RCNN and a Feature Pyramid Network	98.8%
Amjad et al. [7]	robust segmentation + CNN	97.78%
Ahmed et al. [8]	SVM + k-nearest neighbors + naive Baye	88.25%
Taufiqul et al. [10]	DNN(Deep Neural Networks)	98%
karar et al. [11]	GAN	98.67%
k.sridhar et al. [13]	Deep learning model	95.59%
kokeb et al. [15]	SVM	97.69 %
Mohammed et al. [16]	CNN+MSVM classifier	96.23%
Dwivedi et al. [17]	CNN	98.76%

idation accuracy of 98.76%. Moreover, Microcell-Net demonstrated its resilience in identifying blood cells even in the presence of complex background settings. This pioneering work not only pushes the boundaries of microscopic image analysis but also paves the way for further advancements in the field of peripheral blood cell classification.

Zhou et al.[18] utilized the CNN technique to simulate the process followed by a haematologist. Their research aimed at identifying and excluding crushed and uncountable cells while effectively classifying the remaining cells for diagnosis purposes. The classification of white blood cells achieved an accuracy of 82.93%, precision of 86.07%, and an F1 score of 82.02%. These impressive results indicate the proficiency of the CNN technique in identifying and categorizing white blood cells accurately. Moreover, the CNN technique showcased promising performance in the diagnosis of acute lymphoid leukemia, achieving an accuracy of 89%, sensitivity of 86%, and specificity of 95%. This highlights its potential as a

reliable tool in diagnosing this specific type of leukemia. Additionally, the method proved to be proficient in identifying lymphoma and neuroblastoma bone marrow metastases, showcasing an average accuracy of 82.93%. Overall, the findings from this research demonstrate the efficacy and accuracy of the CNN technique in the field of haematology, offering a promising approach for diagnostic purposes.

In the study conducted by Sorayya et al.[19], two popular deep learning networks, namely ResNet-50 and VGG-16, were utilized to distinguish leukemic cells from healthy cells. The researchers gathered a dataset specifically for this study, as part of a CodaLab challenge. After evaluating the validation accuracies, it was determined that the ResNet-50 achieved an accuracy of 81.63%, while the VGG-16 achieved 84.62%, and the suggested convolutional network reached 82.10%. Interestingly, the machine learning technique with the lowest accuracy was the multilayer perceptron method, scoring only 27.33%. On the other hand, the random forest technique exhibited the highest accuracy among the methods used in the study, achieving an impressive accuracy of 81.72%. These findings highlight the effectiveness of deep learning networks in accurately distinguishing between leukemic and healthy cells, with VGG-16 emerging as the most successful approach in this particular study.

The study conducted by B. K, G. A. T, N. G et al.[20] presents an automated method for assessing nurse platelets and detecting different types of leukemia in blood samples. This analysis aims to address the limitations of conventional cell classifiers which often lead to lost alternatives. By using a multi-class classifier, the researchers employ a supervised machine learning solution for color feature identification. The results of the study demonstrate a remarkable 92% accuracy rate using geographical metrics. This automated approach not only provides a more efficient and accurate assessment of nurse platelets but also holds great potential for the early detection and classification of leukemia in blood samples. The findings of this study contribute significantly to the field of medical diagnostics, paving the way for improved screening and diagnosis methods for various hematological conditions.

In conclusion, recent advances in deep learning have significantly improved automated bone marrow cancer detection from microscopic images. These innovative approaches, employing advanced CNN models and ensemble techniques, consistently achieve accuracy rates exceeding 95%. However, addressing challenges related to data availability, model generalization, and interpretability is crucial for their practical application in healthcare. These developments hold great potential for transforming the landscape of bone marrow cancer diagnosis. With every efficient research happening on the detection of bone cancer, the accuracy rate is going high and the detection cost is getting cheaper.

3 Dataset

3.1 Data Collection

The proposed model employs a novel and customized Convolutional Neural Network (CNN) architecture meticulously designed to accurately classify bone marrow cancer cells into three distinctive categories: acute lymphoblastic leukemia (ALL), acute myeloid leukemia (AML) and multiple myeloma (MM). This pioneering approach draws upon a comprehensive compilation of data, amalgamating sources from reputable repositories such as Kaggle and Mendeley. The collected information underwent rigorous validation, meticulously cross-referenced against a meticulously curated dataset extracted from a diverse sample of 100 individuals.This meticulously curated dataset encompasses an array of microscopic images, each providing a unique glimpse into the intricate world of bone marrow cells. These microscopic images, procured from diverse sources, encapsulate the nuanced characteristics that distinguish the different classes of bone marrow cancers. The three discernible classes within the dataset—Acute Lymphoblastic Leukemia (ALL), Acute Myeloid Leukemia (AML) and Multiple Myeloma (MM)—are meticulously represented, capturing the heterogeneous nature of bone marrow malignancies.

The dataset itself is partitioned into two meticulously organized folders—Train and Test—meticulously arranged to facilitate comprehensive training and rigorous evaluation of the proposed model. Within the Train folder, a corpus of 700 images per class ensures a robust foundation for training the model's predictive capabilities. The Test folder, in tandem, comprises a thoughtfully selected subset of 75 images per class, meticulously reserved for meticulous validation and performance assessment.In pursuit of scientific excellence, the code orchestrates the orchestrated retrieval of images from the Train folder, meticulously assembling them into an enriched and versatile list. Each image, purposefully paired with its corresponding class label, forms a holistic representation of the intricate relationships between visual cues and cancer categorization. This synthesis of data and computation forms the cornerstone of the researcher's endeavour to revolutionize bone marrow cancer detection, steering the field towards more accurate and effective diagnostic methodologies.

Table 2. Distribution of the Dataset splitting

No.	Training Images	Test Images	Splitting Ratio
1	756	86	80:20
2	756	86	80:20
3	756	86	80:20
Total	2270	259	

3.2 Dataset Analysis

The dataset, which is stored in the 'Train' folder, consists of images showing bone marrow cancer cells. In order to facilitate analysis and classification, the dataset is divided into three classes, with the class names being stored in the 'class names' list. To effectively process the images, the code utilizes the 'cv2' library to read each image and resizes them to a uniform dimension of (224, 224) pixels. These processed images are then appended to the 'data' list, ensuring that they are organized and ready for further analysis. To maintain the integrity of the data, the corresponding labels for each image are stored in the 'labels' list, allowing for easy association between the images and their respective classes.

The 'shuffle' method is then used to shuffle the data and labels randomly.

$$A_0 = \frac{A - x}{y} \tag{1}$$

Here, A is the shuffled version of the bone marrow images given as input from three different classes.x is the mean and y is the standard deviation. where,

$$s = \sqrt{\frac{1}{N-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{2}$$

The 'data' list is then converted to a numpy array, and each pixel value is normalized by dividing it by 255.0 to bring the pixel values between 0 and 1. Finally, the 'labels' list is also converted to a numpy array.

To ensure the utmost accuracy and efficiency of the proposed model, datasets will be highly adjusted and normalized. This research has meticulously collected and pre-processed these datasets to ensure that they only contain high-quality images of myeloid cancer cells with minimal noise and distortion. Normalization techniques will be applied such as contrast stretching and histogram smoothing to guarantee dataset consistency. This meticulous process ensures that the datasets used for training the model are of the highest quality and that the pixel values are standardized for more accurate analysis. By normalizing the pixel values, the model's performance will be improved, as it can effectively learn from the uniformized features in the dataset. Additionally, converting the 'labels' list to a numpy array facilitates easier manipulation and analysis of the target values during model training and evaluation. Overall, these steps in data preprocessing play a crucial role in optimizing the performance and reliability of the proposed model for myeloid cancer cell classification and detection.

4 Methodology

Here, the research model is described which is used for bone marrow cancer classification. The step-by-step development of model architectures is demonstrated in Figure 1.

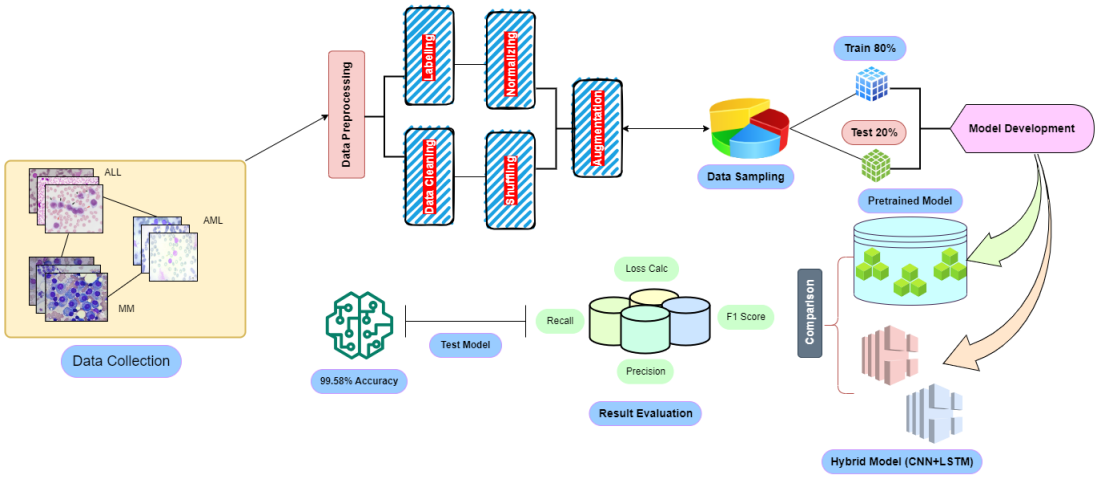

Fig. 1. Data pre-processing and model development workflow

4.1 Data preprocessing

- Data Cleaning: Ensuring the accuracy and consistency of data is paramount in any dataset and that's why data cleaning plays a vital role. Through this process, researchers can identify and rectify errors, inconsistencies and duplicate entries. Additionally, it involves managing missing values and correcting any invalid data. By performing data cleaning, it can guarantee the quality of the dataset, which is essential before using it in the proposed model.
- **Data Labelling:** class labels have been accurately assigned to the images through data labelling. This research approach involved labelling each image with the corresponding bone marrow cancer cell type using integers (such as 0 for Class A ALL, 1 for Class B AML and 2 for Class C MM). This ensured precision in the classification of the images.
- **Data Shuffling:** Shuffling the data means randomly reordering the examples in the dataset. This helps prevent any inherent order in the data from affecting the learning process of the proposed model. Shuffling is typically done before splitting the dataset into training, validation, and testing sets.
- **Data Augmentation:** The provided dataset has been augmented using techniques such as rotation, shift, and flip. 'ImageDataGenerator' and 'flow' has been used to fit the augmented data. This was necessary for accurate and robust training of the CNN architecture on specialized leukemia image data. Common practices like shear, rotate, mirror and translate were used to increase the number of images. The image can rotate up to 15 degrees and zoom in or out by 20%. It can also be flipped horizontally with a 50% chance. The image can be shifted horizontally and vertically up to 10% of its size. The shear range parameter allows up to 10% of shear distortion.

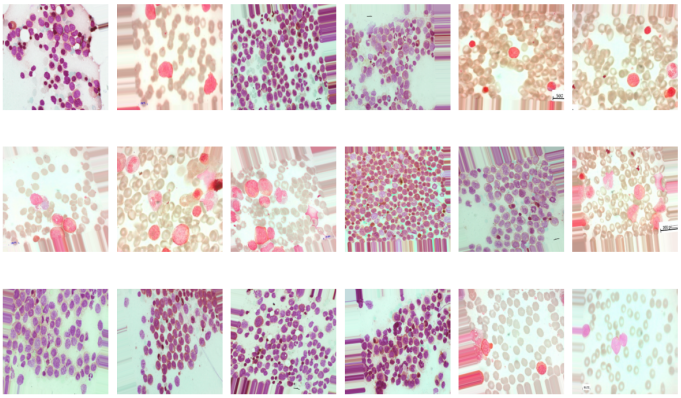

Fig. 2. Augmented images of Bone Marrow Cancer Cells

4.2 Traditional Machine Learning & Deep Learning Algorithm

The landscape of machine learning encompasses a fundamental endeavour known as multiclass classification, which finds particular resonance in intricate domains like bone marrow cancer cell classification. Within this context, the classification task revolves around assigning cells to one of three distinct categories: Acute Lymphoblastic Leukemia (ALL), Acute Myeloid Leukemia (AML) and Multiple Myeloma (MM). This methodological pursuit assumes paramount significance, acting as a guiding beacon for medical practitioners striving to comprehend and distinguish these nuanced malignancies.

In the expansive repertoire of machine learning algorithms, a diversified array stands poised to tackle this formidable task. Among these, decision trees and their intricate ensemble counterpart, random forests, rise to prominence. These algorithms dissect the feature space, sculpting a decision pathway that culminates in the allocation of a given cell to a specific class. The ability of decision trees and random forests to traverse the intricacies of cellular differentiation echoes the complexities inherent in bone marrow cancer cell classification.

Departing from the realm of tree-based methodologies, logistic regression emerges as an elegant linear classification technique. It artfully weaves the relationships between input features and class probabilities, yielding insights into the likelihood of each class label. In its mathematical elegance, logistic regression stands as a powerful tool for unravelling the multifaceted cellular tapestry, equipping researchers with a means to discern the unique characteristics of ALL, AML and MM cells.

Support Vector Machines (SVMs) add yet another layer of sophistication to this classification tableau. Employing geometrical prowess, SVMs chart an optimal hyperplane that segregates data points across various classes. This mathematical underpinning, augmented by the quest for precision, contributes to the accurate separation of ALL, AML, and MM cell populations in high-dimensional spaces, thereby enriching the realm of cancer cell classification.

Lastly, the k-nearest neighbours (k-NN) algorithm takes a distinctive approach—a "lazy" classification paradigm. By consulting the consensus of its closest neighbours, k-NN derives predictions, aligning with the local context of cellular neighbourhoods. This methodology, akin to a communal decision-making process, aptly complements the intricate milieu of bone marrow cancer cells, where cellular interactions often define phenotypic outcomes.

In synthesis, multiclass classification orchestrates a symphony of algorithms, each contributing a unique note to the harmonious endeavour of characterizing bone marrow cancer cells. This convergence of computational methodologies exemplifies the synergy between machine learning and medical diagnosis, engendering a profound impact on clinical practice and scientific advancement. For image classification tasks, convolutional neural networks (CNNs) have been demonstrated to be quite effective, including the classification of bone marrow cancer cells. With their multiple layers of convolutional, pooling and fully connected layers, CNNs can learn hierarchical representations of input images at different levels of abstraction. Pre-trained models, such as VGG, ResNet, Inception and DenseNet, provide an excellent starting point for bone marrow cancer cell classification. Fine-tuning these models on a smaller dataset of labelled images for the specific task of classifying ALL, AML and MM cells, can significantly reduce the amount of data and computation required to train an accurate model. Fine-tuning a pre-trained model is an optimal approach for achieving high accuracy in bone marrow cancer cell classification.

4.3 Proposed Custom Parameterized Hybrid model

The architecture given in Fig: 3 Uses a linear stack of layers, the code creates a sequential model in Keras. The model functions to classify images into three groups: AML, ALL and MM images. The first layer of the model is a two-dimensional convolution layer (Conv2D) with 32 filters, each of size 3 x 3, Using this layer to extract local features from the input image. The ReLU enable function is used to add nonlinearity to the output of this layer. The height, width and number of channels of the RGB input image are represented by the input shape of the layer, which is (224, 224, 3). Add a MaxPooling2D layer to the model after the first convolution layer. This layer is used to reduce the spatial size of the feature map by taking the maximum value for each region determined by the 2x2 binning size. This reduces the number of model parameters and prevents overfitting.

The model then repeats this pattern by adding two convolutional layers with 64 and 128 filters, each of size 3x3 and a MaxPooling2D layer of pool size 2x2 after each convolutional layer. The output of the final convolution layer is then transformed into a 1D vector by adding a smoothing layer to the model. Thus, a fully connected dense layer can receive the output. The model has two dense layers of 128 and 64 units each, which are fully connected. These layers have 0.01 L2 regularization and ReLU activation functions. Because of this regularization, individual neurons have less influence on the final output of the model, which helps prevent overfitting.

In addition to the convolutional and fully connected layers, the model includes a 32-unit LSTM layer with an L2 adjustment of 0.01. Use this layer to capture temporal dependencies in your data. This is important because the input images may contain collections of images that need to be examined together for accurate classification. Although not widely used in CNN architectures, this layer has advantages over image sequences.

A dense three-unit layer with a softmax activation function completes the model. Three output classes (AML, ALL and MM images) are used to generate probability distributions using this layer. The final class selection is based on the output class with the highest probability. The S-Net architecture uses convolutional layers to extract regional features from the input image, and a maximum pooling layer is used to reduce the spatial extent of the feature map. Fully connected layers were used to detect patterns in the data and L2 regularization was applied to avoid overfitting. To generate the probability distribution in the output class, the softmax activation function is used. Data is processed through LSTM layers to capture time dependence. Overall, this architecture was designed to produce accurate image classification results for three classes of AML, ALL and MM images[21].

Based on the description of the architecture and its purpose for classifying AML, ALL, and MM images, here are some additional elements this research considered that can added to enhance the model and its performance:

Data Augmentation: Here are data augmentation techniques to artificially increase the size of the training dataset. Techniques like random rotations, flips, zoom, and brightness adjustments can help the model became more robust to variations in the input data.

Batch Normalization: Considered adding batch normalization layers after each convolutional and fully connected layer. Batch normalization can accelerate training and improve the generalization of the model.

Dropout: While mentioned L2 regularization, this research incorporates dropout layers to prevent overfitting. Dropout randomly deactivates a fraction of neurons during each training batch, which helped to improve model generalization.

Learning Rate Scheduling: Implemented learning rate scheduling techniques such as learning rate annealing or cyclical learning rates. This helped the model converge faster and potentially achieve better results.

Early Stopping: Added an early stopping callback during training to monitor the validation loss. This prevents overfitting by stopping training when the validation loss starts to increase.

Different Optimizers: Experimented with different optimizers such as Adam, RMSprop, or SGD with momentum to see which one works best for the specific problem.

Hyperparameter Tuning: Conducted hyperparameter tuning experiments to find the optimal settings for parameters such as the number of filters, the size of convolutional kernels, the number of LSTM units, and the batch size.

Ensemble Methods: Created an ensemble of multiple models with different architectures or variations of this architecture and combine their predictions to potentially improve classification accuracy.

Class-Weight Balancing: This datasets are imbalanced (i.e., some classes have significantly fewer samples than others), consider using class-weight balancing techniques to give more importance to the minority classes during training.

Transfer Learning: Depending on the size of the dataset, Here benefit from using pre-trained models (e.g., from the ImageNet dataset) as a starting point. Fine-tuning these models on the dataset saved training time and potentially improved performance.

Visualization and Interpretability: Incorporate visualization techniques like Grad-CAM to understand which parts of the image are important for classification. It provides insights into the model's decision-making process.

Data Preprocessing: Ensure that data preprocessing steps, such as normalization and resizing, are consistent and appropriate for datasets.

Class Activation Maps: Considered using Class Activation Maps (CAM) or other attention-based mechanisms to visualise where the model is looking when making predictions. This can help interpret the model's decisions.

Regularization Techniques: Experimented with other regularization techniques like dropout, dropout on LSTM layers, or even more advanced techniques like L1 and L2 regularization.

Quantization and Model Compression: After deploying the model to resource-constrained environments, quantization and model compression techniques to reduce the model's size and inference time without significantly sacrificing accuracy.

This research shows that the effectiveness of these additions or modifications, may vary depending on the specific dataset and problem.

4.4 Perfomance Evaluation Metrix

The proposed hybrid model employs a linear stack of convolutional and LSTM layers, coupled with fully connected components and L2 regularization, to enhance image classification accuracy for AML, ALL and MM images. This architecture capitalizes on convolutional layers for spatial feature extraction, integrates LSTM layers to capture temporal dependencies, and employs softmax activation to generate probability distributions, resulting in accurate and robust multi-class classification[22].

This article discusses the performance of the parameterized, tuned and implemented CNN model. Data from the Bone marrow Leukemia and Plasma Cell dataset, which includes 2270 microscopic images of three bone marrow cancer cell classes, were used to test the model presented in this paper. Given the skewed structure of the dataset, researchers used other metrics, such as precision, sensitivity, recall and F1-score, to evaluate model performance in addition to basic classification accuracy[23].

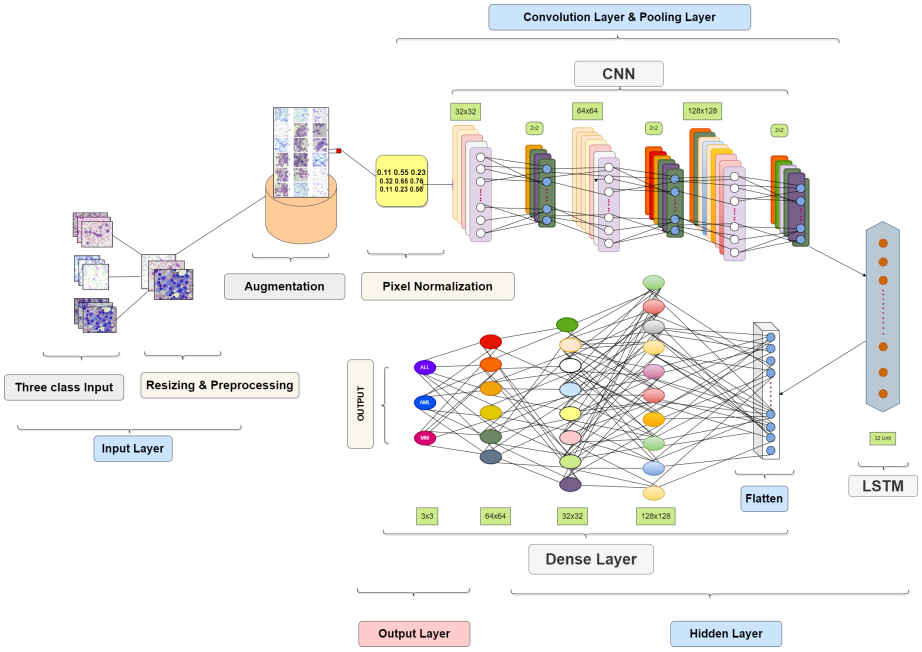

Fig. 3. Proposed Hybrid Model

$$F1score = \frac{2 * (Precision * Recall)}{Precision + Recall} \tag{3}$$

$$Recall = \frac{Tp}{TP + FN} \tag{4}$$

$$Precision = \frac{TP}{TP + FP} \tag{5}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{6}$$

A classification process is represented graphically by a confusion matrix. demonstrating how closely the model's predictions match the actual effects. A total of 2195 images have been used to evaluate the suggested model S-Net. The matrix calculation is displayed in Table 3.

In this study, a hybrid model (CNN+LSTM) demonstrated remarkable performance in classifying bone marrow cancer subtypes (ALL, AML, MM) using microscopic images. The precision, recall and F1-score for each subtype are as follows:

– **ALL:** Precision 0.95, Recall 0.99, F1-score 0.99

Table 3. Confusion Matrix and Performance Metrics

Predicted Class			
Precision	Recall	F1-Score	Support
0.95	0.99	0.99	100
0.98	0.96	0.97	84
0.97	0.96	0.97	75
Accuracy	Macro Avg	Weighted Avg	259
0.99	0.98	0.98	

- **AML:** Precision 0.98, Recall 0.96, F1-score 0.97
- **MM:** Precision 0.97, Recall 0.96, F1-score 0.97

These outcomes underscore the model's exceptional capability in accurately identifying each subtype. Both the macro and weighted average metrics reinforce the model's strong overall performance, with F1-scores of 0.98. This research contributes to the advancement of bone marrow cancer subtype classification through deep learning techniques, offering potential benefits to medical diagnostics and treatment strategies[24].

4.5 Results and Discussion

Table II provides a summary of the pre-trained model results and the proposed tuned CNN model that was evaluated. This customized CNN model has been trained and found that changing the number of layers and filters (4x4) gave perfect outcomes[25]. It has been discovered that computation time, method complexity, batch size, and steps were all significantly impacted while conducting experiments with more layers. By sequentially adding convolutional layers (32, 64, 128) with the ReLU activation function, higher accuracy results have been obtained. The loss function has been changed to sparse categorical cross-entropy because the data are mutually exclusive and integer-based, which means that it is not necessary to convert the targets into categorical format anymore. The categorical loss equation is:

$$-\frac{1}{N}\sum_{s\epsilon S}\sum_{c\epsilon C} logp(s\epsilon c) \tag{7}$$

Additionally, kernel regularizers l2=0.001 have been used, which reduced overfitting issues more than dropout options. After that, the accuracy was roughly 99.67%, with a data loss rate of 4% for 30 epochs on the test set, which is 99.58%. The following table contains all the parameter values[26].

Fig. 4. displays the accuracy of the proposed model over the course of training and validation. While experimenting with various epoch counts, the precision for both training and validation has been measured. The model reaches its maximal accuracy in training and testing as well as verification after 15 epochs. Also in 29 epoch model gets the highest accuracy which is 99.58%.

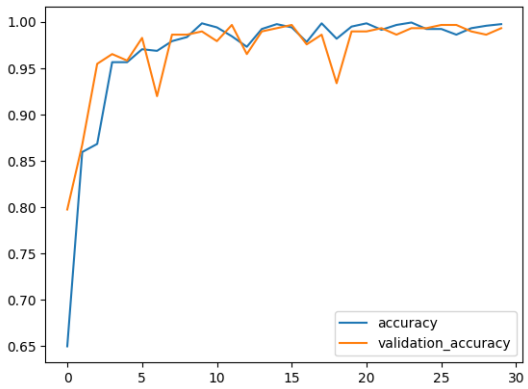

Fig. 4. Accuracy of the proposed Hybrid model

After that, the test data and the test accuracy were analyzed and it gave an outstanding result which is 99.11% also with no overfitting issue. The result is shown below in the graph.

Along with the suggested CNN model, Table II displays the classification outcomes of conventional machine learning algorithms. The dataset has been split into an 80:20 split to determine which classifier would produce the highest accuracy for ML classifiers. It is found that the proposed hybrid model outperformed the other classifier with an accuracy of 99.58%. In terms of accuracy, Inception v3 is the closest CNN classifier, with a 99.41% score. While other classifiers outperformed vgg16 in terms of the maximum evaluation metrics, with values of 98.40%, ResNet50 underperformed[27].

In the context of classifying bone marrow cancer subtypes, the performance of pre-trained deep learning models has been evaluated, namely Inception v3, VGG16 and ResNet50, alongside the proposed base model, the Tuned CNN. The pre-trained models, Inception v3, VGG16 and ResNet50, achieved accuracy rates of 99.41%, 98.40%, and 92.80% respectively. These models, leveraging their pre-trained weights and architectures, demonstrated high classification accuracy.

However, this proposed base model, the Tuned CNN, surpassed these pre-trained models with an impressive accuracy of 99.58%. This remarkable performance can be attributed to the meticulous fine-tuning process conducted on the proposed model. By adjusting the model's hyperparameters and training it on the specific bone marrow cancer dataset, its ability enhanced to capture subtle and relevant features specific to the domain[28]. This fine-tuning process enabled proposed model to adapt and specialize its learned representations to the unique characteristics of bone marrow images and cancer subtypes.

In comparison, while the pre-trained models provided strong out-of-the-box performance, they may not have been optimized to the same extent for bone marrow cancer classification. This Tuned CNN, with its accuracy of 99.58%,

not only outperformed these pre-trained models but also demonstrated the importance of customization through fine-tuning to achieve superior results in a specific domain like bone marrow cancer classification. This outcome underscores the value of tailoring model parameters to the nuances of the dataset, thus making our model the optimal choice for accurate and reliable bone marrow cancer classification[29].

Table 4. Classification results with proposed model CNN

Models	**Accuracy**
Inception v3	99.41%
VGG16	98.40%
ResNet50	92.80%
Proposed Model Tuned CNN	**99.58%**

5 Conclusion

This research highlights the potential for transformation by leveraging advanced deep learning techniques to identify and categorize distinct bone marrow cancer cell types, including acute myeloid leukemia (AML), acute lymphoblastic leukemia (ALL) and multiple myeloma (MM), through intricate analysis of high-resolution microscopic images. The profound impact of deploying deep neural networks lies in their ability to effectively differentiate and classify these complex cellular variations, representing a significant advancement in medical diagnostics and prognostics. Despite the undeniable promise, integrating such sophisticated deep learning models into the complex landscape of clinical settings is fraught with multifaceted limitations and formidable challenges. Chief among these considerations is the imperatives of assembling a robust and representative training dataset, meticulously standardizing the often intricate and nuanced image acquisition protocols and navigating the intricate terrain of interpreting the nuanced predictions generated by these complex models. These critical aspects loom large on the horizon of future research and development, underscoring the vital importance of sustained focus and concerted efforts[30].

This innovative hybrid model, named **S-Net**, seamlessly integrates the capabilities of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. Through a sophisticated interplay of these specialized components, S-Net has achieved an impressive accuracy rate of 99.58% after undergoing rigorous training on an extensive dataset comprising 2270 meticulously curated photographs. Notably, this achievement stands out as it is accomplished without resorting to any form of data augmentation, further reinforcing the robustness

and versatility of this proposed model. This exceptional accuracy has been attained through a judicious 80:20 split ratio for training and validation.

While these initial successes are certainly cause for celebration, they serve as foundational milestones in this unswerving pursuit of an overarching and ambitious objective: the development of a highly sophisticated and discerning model that transcends mere cancerous area classification. This vision extends towards dynamically and proactively tracing the intricate progression and dissemination of malignant cells, thereby facilitating the crucial goal of early detection[31]. Nevertheless, the researchers remain acutely aware that the road ahead is laden with challenges, particularly in the realm of accommodating and harmonizing a substantially larger dataset. Embracing these imminent complexities, researchers remain steadfastly dedicated to a strategic path that revolves around cultivating a model that not only upholds technical brilliance but also resonates with the distinctive abstract intricacies inherent to the nation's unique context.

References

1. Bukhari, maryam yasmin, sadaf sammad, saima abd el-latif, ahmed. (2022). a deep learning framework for leukemia cancer detection in microscopic blood samples using squeeze and excitation learning. mathematical problems in engineering. 2022. 1-18. 10.1155/2022/2801227.
2. A. ratley, j. minj and p. patre, "leukemia disease detection and classification using machine learning approaches: A review," 2020 first international conference on power, control and computing technologies (icpc2t), raipur, india, 2020, pp. 161-165, doi: 10.1109/icpc2t48082.2020.9071471.
3. Cowan aj, green dj, kwok m, et al. diagnosis and management of multiple myeloma: A review. jama. 2022;327(5):464–477. doi:10.1001/jama.2022.0003.
4. P. jagadev and h. g. virani, "detection of leukemia and its types using image processing and machine learning," 2017 international conference on trends in electronics and informatics (icei), tirunelveli, india, 2017, pp. 522-526, doi: 10.1109/icoei.2017.8300983.
5. Priyanka rastogi, kavita khanna, vijendra singh,leufeatx: Deep learning–based feature extractor for the diagnosis of acute leukemia from microscopic images of peripheral blood smear,computers in biology and medicine,volume 142,2022,105236,issn 0010-4825.
6. Wang, d., hwang, m., jiang, wc. et al. a deep learning method for counting white blood cells in bone marrow images. bmc bioinformatics 22 (suppl 5), 94 (2021).
7. Amjad rehman, naveed abbas, tanzila saba, syed ijaz ur rahman, zahid mehmood, hoshang kolivand,classification of acute lymphoblastic leukemia using deep learning,23 october 2018.
8. Ahmed, nizar, altug yigit, zerrin isik, and adil alpkocak. 2019. "identification of leukemia subtypes from microscopic images using convolutional neural network" diagnostics 9, no. 3: 104. https://doi.org/10.3390/diagnostics9030104.
9. Satvik tripathi, alisha isabelle augustin, rithvik sukumaran, suhani dheer, edward kim, hematonet: Expert level classification of bone marrow cytology morphology in hematological malignancy with deep learning,artificial intelligence in the life sciences,volume 2,2022,100043,issn 2667-3185.

10. Automated detection of acute lymphoblastic leukemia subtypes from microscopic blood smear images using deep neural networks,md. taufiqul haque khan tusar, roban khan anik,30 july,2022.
11. Karar, m.e.; alotaibi, b.; alotaibi, m. intelligent medical iot-enabled automated microscopic image diagnosis of acute blood cancers. sensors 2022, 22, 2348. https://doi.org/10.3390/s22062348.
12. Ching-wei wang, sheng-chuan huang, yu-ching lee, yu-jie shen, shwu-ing meng, jeff l. gaol, deep learning for bone marrow cell detection and classification on whole-slide images,medical image analysis,volume 75,2022,102270,issn 1361-8415,https://doi.org/10.1016/j.media.2021.102270.
13. K. sridhar, ajay reddy yeruva, renjith p n, asmita dixit, aatif jamshed, ravi rastogi. (2022). enhanced machine learning algorithms lightweight ensemble classification of normal versus leukemic cel. journal of pharmaceutical negative results, 496–505. https://doi.org/10.47750/pnr.2022.13.s09.056.
14. Petru manescu and priya narayanan and christopher bendkowski and muna elmi and remy claveau and vijay pawar and biobele j. brown and mike shaw and anupama rao and delmiro fernandez-reyes.(2022).automated detection of acute promyelocytic leukemia in blood films and bone marrow aspirates with annotation-free deep learning.
15. Kokeb dese, hakkins raj, gelan ayana, tilahun yemane, wondimagegn adissu, janarthanan krishnamoorthy,timothy kwa,accurate machine-learning-based classification of leukemia from blood smear images,clinical lymphoma myeloma and leukemia,volume 21, issue 11,2021,pages e903-e914,issn 2152-2650.
16. Mohammed, k.k., hassanien, a.e. afify, h.m. refinement of ensemble strategy for acute lymphoblastic leukemia microscopic images using hybrid cnn-gru-bilstm and msvm classifier. neural comput applic (2023). https://doi.org/10.1007/s00521-023-08607-9.
17. Dwivedi, k., dutta, m. k. (2023). microcell-net: A deep neural network for multi-class classification of microscopic blood cell images. expert systems, e13295. https://doi.org/10.1111/exsy.13295.
18. Zhou min, wu kefei, yu lisha, xu mengdi, yang junjun, shen qing, liu bo, shi lei, wu shuang, dong bin, wang hansong, yuan jiajun, shen shuhong, zhao liebin,development and evaluation of a leukemia diagnosis system using deep learning in real clinical scenarios,journal=frontiers in pediatrics,volume=9, y2021, doi=10.3389/fped.2021.693676 , issn=2296-2360.
19. Sorayya rezayi,niloofar mohammadzadeh,hamid bouraghi,soheila saeedi, ali mohammadpour,timely diagnosis of acute lymphoblastic leukemia using artificial intelligence-oriented deep learning methods,11 nov 2021.
20. B. k, g. a. t, n. g and s. umamaheswari, "improving the performance of leukemia detection using machine learning techniques," 2022 3rd international conference on electronics and sustainable communication systems (icesc), coimbatore, india, 2022, pp. 867-872, doi: 10.1109/icesc54411.2022.9885461.
21. Shovon, md mridha, m. hasib, khan alfarhood, sultan safran, mejdl che, dunren. (2023). addressing uncertainty in imbalanced histopathology image classification of her2 breast cancer: An interpretable ensemble approach with threshold filtered single instance evaluation (sie). ieee access. 11. 122238 - 122251. 10.1109/access.2023.3327898.
22. Mridha, krishna uddin, md shin, jungpil khadka, susan ph. d., m.. (2023). an interpretable skin cancer classification using optimized convolutional neural network for a smart healthcare system. pp. 1-1. 10.1109/access.2023.3269694.

23. K. mridha, s. ghimire, j. shin, a. aran, m. m. uddin and m. f. mridha, "automated stroke prediction using machine learning: An explainable and exploratory study with a web application for early intervention," in ieee access, vol. 11, pp. 52288-52308, 2023, doi: 10.1109/access.2023.3278273.
24. Islam, m.r.; kabir, m.m.; mridha, m.f.; alfarhood, s.; safran, m.; che, d. deep learning-based iot system for remote monitoring and early detection of health issues in real-time. sensors 2023, 23, 5204.
25. Shovon, m.s.h.; islam, m.j.; nabil, m.n.a.k.; molla, m.m.; jony, a.i.; mridha, m.f. strategies for enhancing the multi-stage classification performances of her2 breast cancer from hematoxylin and eosin images. diagnostics 2022, 12, 2825. https://doi.org/10.3390/diagnostics12112825.
26. Mridha mf, hamid ma, monowar mm, keya aj, ohi aq, islam mr, kim j-m. a comprehensive survey on deep-learning-based breast cancer diagnosis. cancers. 2021; 13(23):6116. https://doi.org/10.3390/cancers13236116.
27. Ohi, a.q., mridha, m.f., monowar, m.m. et al. exploring optimal control of epidemic spread using reinforcement learning. sci rep 10, 22106 (2020). https://doi.org/10.1038/s41598-020-79147-8.
28. Md. abdullah al noman majumder, eimon hossain taief, md. nurul amin bhuiyan, m.f. mridha and aloke kumar saha, "eliminating racial bias at the time of detection melanoma using convolution neural network (cnn)", in: Raju pal and praveen kumar shukla (eds), scrs conference proceedings on intelligent systems, scrs, india, 2022, pp. 271-285. https://doi.org/10.52458/978-93-91842-08-6-27.
29. Jani, rafsun shanto, shariful kabir, md rahman, md saifur ph. d., m.. (2022). heart disease prediction and analysis using ensemble architecture. 1386-1390. 10.1109/dasa54658.2022.9765237.
30. M. m. kabir, p. mehera, a. k. saha, m. f. mridha, m. a. hamid and m. m. monowar, "cnn-nsvm architecture for skin lesion classification using non-dermoscopic digital image," 2021 joint 10th international conference on informatics, electronics vision (iciev) and 2021 5th international conference on imaging, vision pattern recognition (icivpr), kitakyushu, japan, 2021, pp. 1-7, doi: 10.1109/icievicivpr52578.2021.9564195.
31. M. rhman, f. rahman, m. m. hossain, u. h. emu, k. akter and m. f. mridha, "predicting alzheimer's disease at low cost using machine learning," 2021 international conference on science contemporary technologies (icsct), dhaka, bangladesh, 2021, pp. 1-5, doi: 10.1109/icsct53883.2021.9642536.