World Journal of Advanced Engineering Technology and Sciences
eISSN: 2582-8266
Cross Ref DOI: 10.30574/wjaets
Journal homepage: https://wjaets.com/

(RESEARCH ARTICLE)

Check for updates

Automated white blood cell diagnostics using federated learning and distributed deep learning

Md Fakrul islam Polash [1], Shakil Khan [2], Mohammed Imam Hossain Tarek [3], Mehedi Hasan [3], Mostafizur Rahman Shakil [1] and Istiak Kabir [4,*]

[1] *Department of Engineering Project Management, Westcliff University, Irvine, CA 92614, USA.*
[2] *Department of Business Analytics, International American University, Los Angeles, CA 90010, USA.*
[3] *Department of Business Administration, International American University, Los Angeles, CA 90010, USA.*
[4] *Department of Computer Science, Westcliff University, Irvine, CA 92614, USA.*

World Journal of Advanced Engineering Technology and Sciences, 2025, 17(01), 218-232

Publication history: Received on 31 August 2025; revised on 06 October 2025; accepted on 08 October 2025

Article DOI: https://doi.org/10.30574/wjaets.2025.17.1.1391

Abstract

The rise of big data has transformed the way complex problems in fields such as medicine and biology are solved. In the medical field, analyzing White Blood Cells (WBCs) is crucial for diagnosing diseases and evaluating the immune system. While automated tools like cell counters can quickly generate results, manual blood smear analysis remains critical for accuracy and patient monitoring. Unfortunately, this manual process is slow, labour-intensive, and prone to errors, making it challenging to manage large-scale data efficiently. This study combines the strengths of Federated Learning and Big Data to tackle these problems. The authors propose a new approach for classifying WBCs by leveraging Federated Learning (FL) for privacy-preserving, distributed training on large datasets, while utilizing Apache Spark's tools for big data management and processing. Additionally, advanced deep learning models, such as ResNet50, VGG19, and U-Net, enhance WBC classification accuracy by creating five RDDs and training each of the three models on each of the five RDDs. The ResNet50 model achieved the highest accuracy of 94.06% in RDD2 and RDD5, followed by VGG19 with 94.27% in RDD1, and U- Net with 85.99% in RDD4. RDD has its validation accuracy. This study addresses the dual challenges of scalability and privacy. Additionally, distributed data on five Nodes of Resilient Distributed Datasets (RDD) demonstrated that both VGG19 and ResNet50 achieved higher accuracy compared to U-Net, while training each deep-learning model to enhance diagnostic accuracy by integrating Federated Learning with Big Data frameworks and recent deep-learning techniques. This innovative technique highlights the potential of combining these technologies to advance healthcare and biomedical research.

Keywords: Federated Learning; White Blood Cell Classification; Big Data Processing; Deep Learning Models; Resilient Distributed Datasets (RDDS)

1. Introduction

The 10 Vs. of Big Data (BD)—volume, velocity, variety, veracity, value, validity, volatility, variability, visualization, and vulnerability. These qualities are essential for Big Data projects to be implemented successfully. Inconsistencies in outcomes resulting from various data dimensions and sources are called variability [1]. Graphically presenting data for simple interpretation is known as visualization. Security methods for data processing in compliance with laws and client preferences are vulnerable [2]. Additionally, Figure 1 draws attention to every industry's varying vulnerability.

Ultimately, whereas big data poses obstacles, it simultaneously provides significant opportunities for those capable of leveraging it proficiently. Big companies now store vast amounts of data in various formats, from structured relational

* Corresponding author: Istiak Kabir

databases to unstructured flat files, due to the ease of digitalization, the diversity of data descriptions, and individual preferences [3]. Researchers state that the volume of data worldwide will Reach 163 zettabytes by 2025, with businesses producing half of that data [4]. Decision- makers must consider data dispersed across several locations and in various forms to obtain insightful information that helps them in their day-to-day tasks, as it gains value when enriched and combined with other data. However, the explosion of data in volume, variety, and velocity—the "3Vs" of Big Data increases complexity and renders traditional methods of data integration like data warehousing. It is more expensive in terms of time and money and less able to ensure the freshness of data.

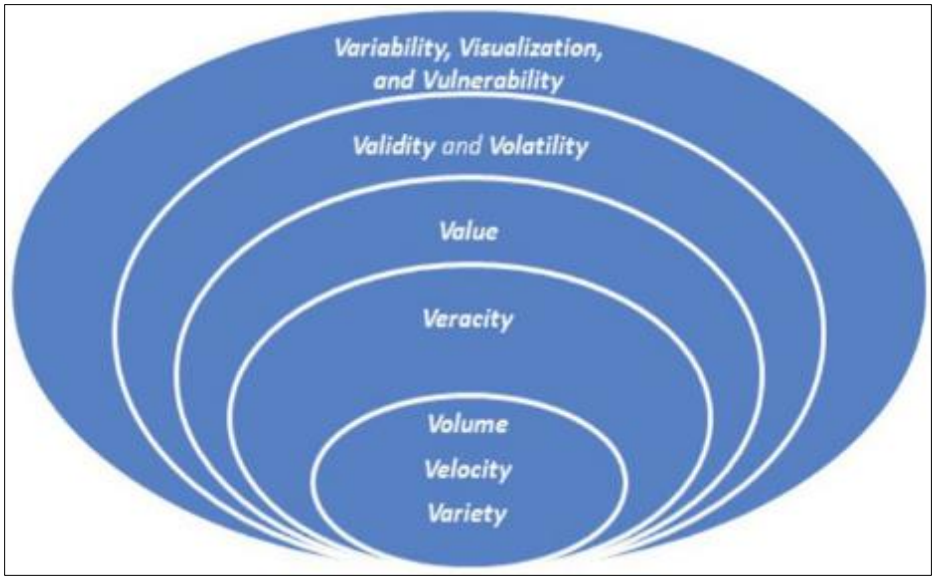

Figure 1 General Concept of 10 Characteristic Growth Big Data

Millions of sensors are needed for the Internet of Things to create and collect data, which is complicated and needs to be processed by conventional applications; big data is processed, stored, and analyzed using tools like Hadoop MapReduce and RDDs. However, MapReduce's time-consuming calculations make it ineffective for iterative algorithms. Distributed data collection, fault tolerance, parallel processes, and the utilization of numerous data sources are the four primary characteristics that RDDs provide [5]. They do not need HDFS storage, which speeds up operations and improves memory management. Compared to other contexts, RDDs are more suited to host iterative algorithms [6].

WBCs include neutrophils, monocytes, basophils, eosinophils, and lymphocytes (T, B, and Killer natural cells). These WBCs work together to detect and fight infections or diseases in the bloodstream, which is vital in maintaining overall health [7]. The number and diversity of WBCs in the body can provide important insights into a person's health status. To accurately classify WBCs, the authors [8] utilized CNN-based models, such as U- Net, VGG19, and ResNet50, which excel at extracting high-level features from image data. These models were evaluated using both transfer learning and full- parameter learning. Additionally, it leveraged Spark to handle the large dataset efficiently and created five Resilient Distributed Datasets to distribute the data, enabling scalable processing and improving the classification workflow. So, one problem with WBC classification is that low resolution, noise, and poor image quality in WBC images make classification difficult [9]. Staining methods can alter color, cytoplasm, and nucleus, while light strength affects appearance. WBCs' development and size vary, and various imaging systems can cause noise and contrast. Data imbalance is a common issue, necessitating improved visual quality for accurate classification. WBC classification is crucial for diagnosing various blood diseases and conditions like lymphoma and leukemia [10]. Additionally, it aids in the detection of blood illnesses caused by bacteria, viruses, or parasites.

However, the challenges and limitations of the classification of White blood cells may result in a less accurate estimation, and manual classification is time-consuming and requires extensive expertise. Additionally, the varying size of WBC and a massive amount of data lead to difficult control [11]. So, to address this limitation, the study used federated learning by using RDD to increase memory and speed. WBC classification is also essential for tracking a patient's health after chemotherapy, organ transplants, etc. Inaccurate WBC counts and classifications can cause serious health problems, including death. WBC function is essential for the correct diagnosis and course of treatment. WBC images with low quality, however, may be misdiagnosed, which could worsen health consequences or even cause death. This paper

addresses how large datasets can be efficiently managed by storing them across multiple locations instead of a single centralized location. Therefore, this study's primary contributions are:

- RDDs allow multiple nodes to efficiently process large data sets through distributed preprocessing, achieving scalability and fault tolerance.
- RDD nodes perform parallel execution of deep learning models, including VGG19, ResNet, and U-Net, which enables specialized feature extraction and flexible data modality adaptation.
- Several distributed datasets collaborate under a federated learning scheme by aggregating their models into a unified global model. The decentralized raw data collection requirement decreases through this approach, strengthening privacy protection and regulatory compliance.
- Data ingestion and processing must operate quickly through the pipeline using node-based workload distribution. The system performs time-efficient processes of large datasets during critical operational deadlines.

2. Related works

Recent biomedical research has significantly improved treatments, such as precision wound healing [12], which uses regenerative therapies for tailored recovery. Innovations like hybrid nanoconjugates of temozolomide enhance drug stability and effectiveness in glioblastoma [13], providing new options beyond traditional therapies. Molecular erasers also enable targeted protein degradation, transforming cancer immunity and creating new avenues in immuno-oncology [14]. In parallel, Artificial Intelligence (AI) is transforming various sectors. In agriculture, AI-powered navigation systems in farming equipment boost productivity and resource use [15]. In transportation, computer vision with attention mechanisms improves road segmentation, enhancing the safety of autonomous vehicles [16]. In healthcare, hybrid models like CNN-SVM applied to MRI data show promise for accurately classifying Alzheimer's disease [17]. Additionally, multimodal deep learning frameworks like MultiSenseNet improve predictive maintenance in industrial settings by assessing machine failure risks [18]. In plant science, deep stacking models combining CNNs with gradient boosting enhance the detection of plant leaf diseases, promoting early intervention [19].

Big data management involves collecting and organizing vast amounts of data from various sources. Its main goals are to ensure data quality, ownership, documentation, and accessibility. High-quality data is essential for gaining valuable insights. Veracity, or data accuracy, is crucial, as erroneous data can lead to meaningless insights. Managing the large volume of data requires significant resources. By 2020, the amount of digital data was projected to reach 44 trillion gigabytes, with 1.7 MB created every second. Traditional RDBMS struggles to handle these volumes, necessitating modern solutions like NoSQL, Apache Drill, and MapReduce. The variety of data formats complicates storage and processing, often requiring costly processing power and frequent updates. Once collected, data must be cleaned and refined for accuracy and relevance, which can be automated or done manually. Advanced tools leveraging machine learning help reduce costs and time while maintaining data integrity, preventing failures in data initiatives.

Recent advancements in deep learning and multimodal fusion have significantly contributed to medical image analysis and sentiment classification across diverse domains. Hybrid and tensor fusion strategies have been employed to enhance multimodal object recognition by integrating vision and audio signals [20]. In the medical domain, several studies have introduced transformer-based and ensemble learning approaches for cancer detection and diagnosis, including a hybrid vision transformer with attention mechanisms for explainable lung cancer diagnosis [21], a hierarchical Swin Transformer ensemble for decentralized breast cancer diagnosis [22], a deep stacking ensemble for transparent brain tumor detection [23], and a novel stacking ensemble for accurate cervical cancer diagnosis [24]. Similarly, innovative data-driven solutions have been proposed for healthcare-related sentiment analysis, such as the identification of patient sentiments from online drug reviews [25]. Other works have advanced diagnostic capabilities for leukemia by integrating image processing with transfer learning [26], applied transformer-based ensembles for oral cancer segmentation [27], developed ensemble learning strategies for recognizing monkeypox [28], and introduced hybrid vision transformers for prostate cancer classification from MRI images [29]. Collectively, these studies demonstrate the growing importance of ensemble methods, transformers, and explainable AI in improving diagnostic accuracy, interpretability, and reliability across multiple healthcare applications.

Federated learning enables autonomous vehicles to train models collaboratively without sharing raw data; however, diverse, non-iid data across vehicles poses challenges to accuracy and convergence. The FedRAV framework addresses this by dividing vehicles into sub-regions and creating personalized models for each, selectively adopting valuable updates [30]. Experiments on real-world datasets show that FedRAV outperforms existing methods, offering a more effective solution for federated learning in autonomous driving.

Authors tackled strict and flexible fraud prevention and security rules for the Internet of Medical Things (IoMT). They introduced a privacy-focused strategy called FL-BEPP (Federated Learning with Blockchain-Enabled Privacy Preservation) to enhance data security and privacy in healthcare systems [31]. Authors used IDS-based anomaly detection to avoid cyberattacks on IoT networks. To be more precise, suggest applying Federated Deep Learning (FDL). Authors [32] describe a fog-based intrusion detection system (IDS) architecture that uses the Bot-IoT dataset and the Lost Short-Term Memory (LSTM) model. Because the solution uses a local learning strategy, devices. Furthermore, authors [33] introduced a differentially private FL approach for intent classification using text data. They implemented privacy protection at the sample level, updating local parameters with gradients and introducing Gaussian noise to preserve privacy. The sampling ratio affected the model's performance. Authors [34] proposed a blockchain-based Federated Learning Intrusion Detection System (FL IDS) for IoT healthcare, featuring an Artificial Neural Network (ANN) model for detecting attacks and monitoring sensors. While this approach goes beyond existing methods, it still faces challenges like privacy concerns and managing decentralized patient data. Authors [35] examined the application of clustered federated learning (CFL) and edge computing in healthcare settings to address security, privacy, and latency challenges in COVID-19 diagnosis, highlighting their potential and addressing these challenges.

Several studies have explored robust architectures and ensemble strategies for improving accuracy and interpretability in crop disease detection. For instance, a MaxViT-based framework has been proposed for accelerated and precise identification of soybean leaf and seed diseases [36], while the ViX-MangoEFormer model integrates Vision Transformer, EfficientFormer, and stacking ensembles to achieve reliable mango leaf disease recognition with explainable AI [37]. Similarly, an ensemble-driven explainable approach has been developed for the recognition and conservation of rare medicinal plants [38], and deep learning has been applied for automated weed species classification in rice cultivation [39]. Other contributions include a web-based transfer learning application for effective cucumber disease recognition [40] and a robust ensemble of transfer learning architectures for accurate diagnosis of eggplant diseases [41]. Collectively, these works highlight the growing role of advanced neural architectures and ensemble methods in enhancing precision, efficiency, and interpretability in agricultural disease detection.

To address model poisoning, the authors [42] developed FL-WBC, a client-based protection method to address model poisoning attacks in Federated Learning, ensuring model convergence and preventing attacks, with high model accuracy demonstrated in tests. Recent studies [43], [44] have shown that FL can enhance privacy and security in machine learning models, mitigating model poisoning attacks and improving smartphone prediction performance while maintaining user data privacy. Several Convolutional Neural Network (CNN) models have been investigated recently by several researchers [44], [45] for the classification of WBCs in complete blood cell counting (CBC). Authors proposed a bilinear CNN that reduces the number of learnable parameters while capturing interactions between local pair features using two separate CNNs. Authors [45] used sophisticated machine-learning methods for WBC segmentation and classification along with a deep contour-aware CNN. Authors [46] proposed a hybrid method that uses SVM classification after feature extraction from WBC pictures using GoogleNet and AlexNet. It deals with blood cell categorization, essential for evaluating health, using a CNN and RNN merging model that uses Canonical Correlation Analysis (CCA) to improve. In contrast to traditional techniques, by effectively managing overlapping cells, CCA accelerates network convergence, reduces classification time, and compresses input dimensions. Authors [47] focused on identifying high-performing machine learning (ML), deep learning (DL), and CNN-based methods for white blood cell (WBC) analysis, highlighting their potential to improve early diagnosis and clinical decision-making. The study [48] proposed an ensemble model for deep learning to achieve high accuracy. The ensemble combined three models, including EfficientNet and DenseNet-201 ConvNeXt, for the classification of WBCs.

3. Methodology

3.1. Dataset

A Blood Cell Classification dataset containing five categories—eosinophil, lymphocyte, monocyte, basophil, and neutrophil images—is used in the suggested WBC classification method. The authors collected 53,385 images from public and private sources. The author utilized 48000 images to make balanced classes. The public datasets used are Kaggle for WBC, which has (28,000) images with a pixel size (64x64x3); Rabin, which has (3,787) images with pixel size (112x112x3), Mendeley, which has (10,229) images with pixel size (360x360x3), IEEE, which has (1,408) images with pixel size (722x722x3), and the private Hiwa hospital dataset, which has (10,299) images with pixel size (64x64x3). Two months were needed to collect the dataset. Applying oil to the slide beneath the microscope is required for Hiwa Hospital—there are two sections to this dataset. The data was split into the validation and training sections, respectively. There is an 80:20 split between the training and testing components.

3.2. Methods

The study proposed an architecture for classifying and analyzing the white blood cell dataset into five RDDs (Resilient Distributed Datasets). For each RDD, apply three model architectures of deep learning, and for each RDD, make it by taking 20% of each of the five classes of white blood cells and then training each model of deep learning. Since U-Net generates a segmentation mask, this study converts the output to a single label for classification tasks. Replace the final segmentation layer with a Global Average Pooling (GAP) layer. Add fully connected (Dense) layers for classification. The author utilizes methods that aren't yet widely used in big data, such as U-Net, ResNet50, and VGG 19, as illustrated in Figure 2.

FL is pivotal in this context, as it offers several benefits compared to conventional centralized training, including addressing privacy concerns, reducing computational costs, and enhancing model robustness. This differs from centralized learning, which involves the centralization of essential medical data, thereby increasing privacy risks and potentially violating laws. FL also includes the collaborative training of the model without exchanging the training data, therefore minimizing costs in terms of bandwidth and storage. Besides, centralized models are sensitive to domain shifts with variations of imaging devices, staining techniques or even patient population, and thus, poor in generalization. FL addresses this systematically by enabling the construction of models on different diverse real datasets across multiple institutions, making the models more robust and adaptive. From a computational perspective, centralized training involves a large number of computations at a single point. In contrast, FL decentralizes the training process across multiple parties, thereby increasing the scalability and efficiency of the training phase. Based on these factors, FL is not just an option, but a requirement for achieving large-scale, non-intrusive, and efficient WBC classification. Additionally, 80% of the preprocessed dataset is used for training, and the remaining 20% is used for testing. The batch size is set to 32 for the three models, and the learning rate is set to 0.0001. Using a preprocessed and normalized dataset, first choose three pre-trained models (ResNet50, U-Net, and VGG19). Then, resize the images to be different to be consistent. To improve model robustness, ImageDataGenerator will then be used for data augmentation. Random rotations, translations, shearing, zooming, and flipping were all included in the augmentation. Following training, each model predicted the class and improved classification accuracy. Once the model is trained, its performance will be assessed using accuracy, precision, recall, and F1-score. After that, generate a confusion matrix to see how well the model classifies each category.

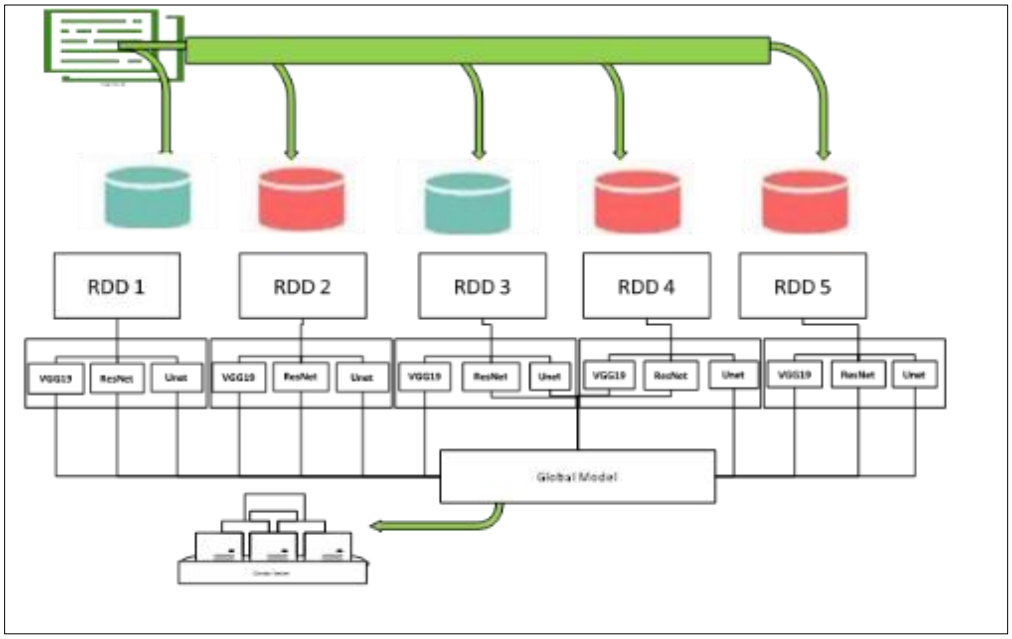

Figure 2 Proposed framework for WBC classification

3.3. Dataset Preprocessing

The suggested method utilizes a customized dataset consisting of images of blood cells. The collection, which consists of 48,000 images, is divided into five categories: eosinophil (E.P.), lymphocyte (L.C.), monocyte (M.C.), neutrophil (N.P.), and basophil (B.C.). They are all downsized to 64 x 64 x 3. There are two sections to this dataset. There is an 80:20 split between the training and testing components. During the preparation stage, several steps are taken to clean the dataset, including eliminating inaccurate or unnecessary images, detecting duplicates, and resizing images to 71 by 71 pixels.

Apart from normalizing images by scaling them to the (0,1) range, data augmentation methods like rotation, flipping, and dataset balancing were used to ensure that all five classes had 10,000 samples, except for the basophile class, which has 8,000 samples because only 1% of human cells are basophils.

3.4. U-Net

The U-Net model is a type of neural network designed initially for image segmentation but adapted here for classifying WBCs into five categories: Basophil (B), Neutrophil (N), Eosinophil (E, Monocyte (M), and Lymphocyte (L). It has an encoder that captures essential features through convolution and pooling layers and a decoder that restores spatial details using upsampling and skip connections. To classify images, the model includes a final layer that uses global average pooling and a softmax activation to predict the class. The dataset, processed with PySpark, contains 9,600 images per RDD (1,920 from each class) to manage the large-scale data efficiently. Images are resized, normalized, and split into training (80%) and testing (20%) sets, with the learning rate set to 0.0001 and the batch size set to 32. The model is trained for 70 epochs using the categorical cross-entropy loss and Adam optimizer, with performance tracked throughout. This approach combines U-Net's powerful feature extraction with scalable data processing, making it practical for WBC classification. As illustrated in the figure below, the designing U-Net in each RDD with an illustration of the Pseudo code for the U-Net structure. Table 1 summarizes the architecture.

Table 1 Summary of the U-Net Architecture

Layer	**Details**	**Output Shape (128x128x3 input)**
Input Layer	Input image (RGB)	128x 128x3
Encoder		
Conv2D + ReLU	64 filter, 3x3 kernel, same padding	128x128x64
MaxPooling2D	2x2 pool size	64x64x64
Conv2D +ReLU	128 filter,3x3 Kernel, same padding	64x64x128
MaxPooling2D	2x2 pool size	32x32x256
Conv2D +ReLU	256 filter,3x3 kernel, same padding	32x32x256
Decoder		
Conv2DTranspose	128 filter, 2x2 Kernal,2x2 strides	64x64x128
Concatenate	Skip connection	64x64x256
Conv2DTranspose	128 filter, 3x3 kernel, same padding	64x64x128
Concatenate	64 filter,2x2 kernal, 2x2 strides	128x128x64
Conv2D +ReLU	Skip connection	128x128x 128
Classification Head	64 filter, 3x3 kernel, same padding	128x128x64
Classification Head		
GlobalAveragePooling2D	Reduce spatial dimensions	64
Dense	5 units, SoftMax activation	5

3.5. ResNet50

The ResNet50 model is a pre-trained feature extractor with its top layer removed. A GlobalAveragePooling2D layer and custom fully connected layers are then added for classification. The data is split into training and testing sets, and the model is trained using the Adam optimizer. After training, its performance is evaluated on the test set, and validation accuracy is displayed. Finally, a plot shows how training and validation accuracy change over time. Spark speeds up data loading and preprocessing by distributing the work across multiple RDDs, making it easier to handle large datasets efficiently. This code utilizes Spark to process large image datasets in parallel for training a ResNet-50 model on five classes of white blood cells. First, it initializes a Spark Context and loads 1,920 images from each class (9,600 images per RDD) into separate RDDs. Each RDD is processed by sampling images and then preprocessing them for the ResNet50 model. Images are resized and normalized before being labeled according to the class, as shown in Figure 3 below.

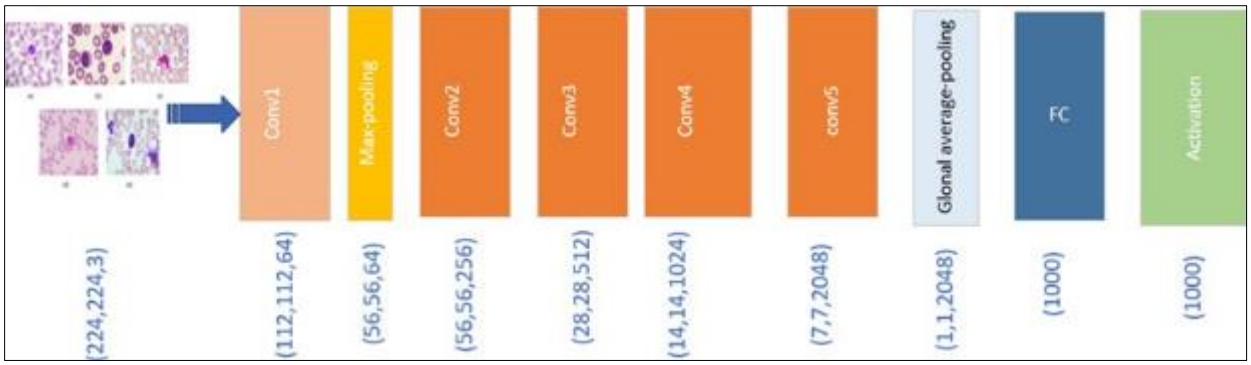

Figure 3 Model architecture of ResNet50

3.6. VGG19

VGG19 is a deep convolutional neural network (CNN) architecture used to classify images into five categories: Basophil, Eosinophil, Neutrophil, Monocyte, and Lymphocyte. It consists of 19 layers and is known for its simplicity and effectiveness. The workflow for classifying WBCs involves loading images, preprocessing them, building the model, training it using the Adam optimizer, and evaluating its performance. The approach uses PySpark to handle the large dataset, which contains 9,600 images per RDD, with a batch size of 32 and a learning rate of 0.0001, ensuring efficient data distribution across multiple nodes in a cluster. The model's performance is then evaluated and plotted to visualize the accuracy over epochs. Illustrate the architecture of VGG19, as shown in the table below.

Table 2 Proposed VGG19 Architecture for the last 20 layers

Layer Type	Output Shape	Parameters	Description
Input Layer	(224, 244,3)	0	Input RGB image resized to 224,224 pixels
VGG19(pre-trained)	Variable	20,024,384	Convolution layers of VGG19 with include_top=false, preserving feature extraction layers.
Flatten	(7,7,512-25088)	0	Flattens the output of VGG19 into a single vector
Dense	(256)	6,422,784	Fully connected layer with 256 units and ReLU activation.
Dense	(5)	1,285	Output layer with 5 units (one for each class) and softmax activation.

Pooling layers simplify the data by reducing its size, allowing the model to focus on essential features while minimizing computation. Max pooling, often with a 2×2 or 3×3 filter, selects the highest value in each section, making it more effective than average pooling at keeping essential details. Dense (fully connected) layers come after convolutional and pooling layers to perform the final classification. They typically have 512–1024 neurons to handle complex patterns. ReLU is commonly used in hidden layers for its efficiency, while the output layer uses Sigmoid for binary classification or Softmax for multi-class problems.

4. Result and Discussion

The human body consists of WBC, RBC, and PBC, which are very important for the immune system of the human body; the thousands of people around the world have died because of leukemia and lack of medical devices, so through some advanced algorithms of deep learning can reduce the percentage of human death by classifying and detecting human blood cell. The proposed model is to classify White blood cells into five sub-classes, which are Neutrophil, Basophil, Lymphocyte, Monocyte, and Eosinophil. The proposed work collected data about five types of (WBC) White blood cells through public site and private sites; public sites include(mendeley, Rabin, IEEE, and Kaggle) and on our private site, Hiwa Hospital, the author collected about 10.000 images through a microscope and put oil on the slide and take images of cell that contains blood cell type and then crop each cell alone for each class to put in one folder, as well as some of the data it was not completed or scratch and cleaned the data through deleted images or some it was not clear. Hence, all of them are the same size (64x64) in RGB and augmented images and normalized data between (0,1). They used a learning rate to prevent overfitting from becoming a vast dataset and used learning rate to prevent overfitting, so the

customized dataset is about (48,000) images; so after collecting the data, preprocessed the data utilized five RDD (Resilient distributed database), and made balanced classes, that each RDD by taking 20% of each of the five classes (B, E, N, M, L), created 5 RDDs, and then trained deep learning models like (ResNet50, U-Net, and vgg19) on each RDD. This experiment has been done on Kaggle by using (GPUx100) and compared five values of accuracy for five RDDs by training a deep learning model. The comparison of the confusion matrix for precision, recall, and F1 score is illustrated in Table 3. Utilizing RDD differs from using a deep learning model because some algorithms, such as VGG19 and ResNet50, achieve high accuracy with 30 epochs, while U-Net achieves high accuracy at epoch 70.

Table 3 The 5 RDDs' Val-accuracy for five classes

	U-Net	**VGG19**	**ResNet50**
RDD 1(Accuracy)	83.23%	94.27%	93.59%
RDD 2(Accuracy)	85.42%	92.34%	94.06%
RDD 3(Accuracy)	85.05%	92.29%	93.12%
RDD 4(Accuracy)	85.99%	89.84%	93.80%
RDD 5(Accuracy)	84.32%	93.23%	94.06%

The performance evaluation of three deep learning algorithms (ResNet50 and VGG19, alongside U-Net) on various RDD (Red Blood Cell Disorders) diagnosis types is presented in Table 3, which displays their recall scores and precision metrics, resulting in F1-score calculations. Multiple RDD datasets evaluate models based on their capability to correctly detect B, E, N, M, and L class types. The evaluation analysis reveals how each model performs well or poorly in terms of its classification output. VGG19 demonstrates superior performance over ResNet50 and U-Net among the three deep learning models during most testing scenarios, including precision-based examinations and F1-score analyses. Numerical evidence supports the claim that VGG19 provides a superior capacity to recognize positive test examples with fewer errors correctly. VGG19 demonstrates superior performance over ResNet50 and U-Net among the three deep learning models during most testing scenarios, including precision-based examinations and F1-score analyses. Numerical evidence supports the claim that VGG19 provides a superior capacity to recognize positive test examples with fewer errors correctly.

The recall outcomes of ResNet50 are high, which indicates that this model detects many important instances. Some model cases reveal that the precision rate has lower numbers while the trade-off between sensitivity and specificity exists. U-Net delivers the worst results in every metric evaluation, showing specific underperformance when identifying class M samples across multiple RDD datasets. The segmentation-focused structure of U-Net leads to performance problems in specific medical class categories.

- Class (B) and Class (E): The performance of ResNet50 and VGG19 excels in the detection task since their recall and precision values remain above 0.90 throughout most evaluation tests. U-Net demonstrates acceptable scoring though it falls short behind other models, especially in precision evaluation metrics.
- Class (N): The F1-score remains high for VGG19 because it demonstrates near- perfect precision and recall in multiple examined cases. ResNet50 demonstrates consistent performance in different RDDs, although it performs at a slightly different level than its peers. U-Net maintains acceptable results, although its performance remains below VGG19 and ResNet50.
- Class (M): All models struggle to overcome this class with the highest difficulty level. U-Net demonstrates significant difficulty achieving its tasks as recall reaches an all-time low of 0.17 within RDD 1, followed by overall low performance across datasets. ResNet50 and VGG19 have imprecise results since precision varies from 0.68 to 0.98.
- Class (L): The VGG19 model establishes itself as high-performing by regularly surpassing precision and recall values of 0.95. ResNet50 demonstrates efficient performance while displaying different prediction variations on tasks. U-Net demonstrates inferior precision, with evidence pointing to problems in identifying positive samples correctly.

U-Net fails to correctly classify type M samples, thus demonstrating its inability to process the dataset effectively. Select ResNet50 when recall achievement is the most critical factor because it performs better in detecting all potential cases. The VGG19 model offers optimal performance when precision needs to be balanced with recall for medical diagnosis applications. The performance metrics of U-Net indicate it struggles with classifying data because its original design

targets segmentation tasks. The U-Net model should be improved through hybrid techniques and beneficial fine- tuning methods to enhance its classification abilities. It should examine DenseNet alongside transformer-based models to establish their superiority level against VGG19—analysis of why the M class proves challenging to classify and test various preprocessing and feature extraction methods. The data presented in Tables 4 and 5 demonstrates the capabilities of different deep- learning models when classifying RDD entities. The robust performance of VGG19 makes it the preferable choice due to its superior recall metrics F1-score, and precision levels. The ResNet50 model proves itself as a solid replacement for specific classification needs where recall demands become primary. The classification competency of U-Net appears insufficient when tackling specific tasks. The studies confirmed that selecting the appropriate model becomes critical when dealing with various classification requirements.

Table 4 Comparison between Recall, Precision, and F1 scores for each (RDD) while training three deep learning models

Classes of (WBC)	**ResNet50 (Recall / Precision / F1 Score)**	**VGG19 (Recall / Precision / F1 Score)**	**U-Net (Recall / Precision / F1 Score)**
RDD 1 - Class(B)	0.96 / 0.93 / 0.95	0.92 / 0.96 / 0.94	0.79 / 0.90/ 0.84
RDD 1 - Class (E)	0.95 / 0.97 / 0.96	0.94 / 0.96/ 0.95	0.81 / 0.87 / 0.84
RDD 1 - Class(N)	0.97 / 0.91 / 0.94	0.95 / 0.93/ 0.94	0.87 / 0.96/ 0.91
RDD 1 - Class(M)	0.88/ 0.91 / 0.89	0.92 / 0.92 / 0.92	0.82 / 0.66 / 0.73
RDD 1 - Class (L)	0.93 / 0.96 / 0.94	0.97 / 0.94 / 0.96	0.87 / 0.84 / 0.85
RDD 2 - Class(B)	0.94 / 0.96 / 0.95	0.90 / 0.97 / 0.93	0.81/ 0.90 / 0.85
RDD 2 - Class (E)	0.96 / 0.95 / 0.95	0.95 / 0.93 / 0.94	0.86 / 0.92 / 0.89
RDD 2 - Class(N)	0.97 / 0.92 / 0.95	0.94 / 0.96 / 0.95	0.92 / 0.89 / 0.91
RDD 2 - Class(M)	0.86 / 0.96 / 0.91	0.89 / 0.89 / 0.88	0.76 / 0.78 / 0.77
RDD 2 - Class (L)	0.97 / 0.92 / 0.94	0.93 / 0.92 / 0.93	0.90 / 0.81 / 0.85
RDD 3 - Class(B)	0.94 / 0.96 / 0.95	0.91 / 0.93/ 0.92	0.83 / 0.91/ 0.87
RDD 3 - Class (E)	0.93 / 0.96 / 0.94	0.95 / 0.92 / 0.94	0.83 / 0.95 / 0.88
RDD 3 - Class(N)	0.92/ 0.98/ 0.95	0.95/ 0.95 / 0.95	0.92 / 0.85 / 0.89
RDD 3 - Class(M)	0.93/ 0.83 / 0.88	0.85 / 0.87 / 0.86	0.80 / 0.71 / 0.75
RDD 3 - Class (L)	0.94 / 0.95 / 0.94	0.95 / 0.94 / 0.94	0.88 / 0.85 / 0.87
RDD 4 - Class(B)	0.95 / 0.93 / 0.94	0.89 / 0.93 / 0.91	0.88 / 0.82 / 0.85
RDD 4 - Class (E)	0.92 / 0.98 / 0.95	0.89 / 0.94 / 0.92	0.90 / 0.93 / 0.91

RDD 4 - Class(N)	0.96 / 0.92 / 0.94	0.91 / 0.97 / 0.94	0.89 / 0.96 / 0.92
RDD 4 - Class(M)	0.96 / 0.92/ 0.94	0.92 / 0.75 / 0.83	0.68 / 0.85/ 0.75
RDD 4 - Class (L)	0.94 / 0.96 / 0.95	0.88 / 0.94 / 0.91	0.95 / 0.77 / 0.85
RDD 5 - Class(B)	0.93 / 0.97 / 0.95	0.92/ 0.95 / 0.94	0.84 / 0.91 / 0.88
RDD 5 - Class (E)	0.98 / 0.91 / 0.94	0.97 / 0.92 / 0.94	0.78 / 0.92 / 0.84
RDD 5 - Class(N)	0.97 / 0.95 / 0.96	0.91 / 0.96 / 0.94	0.91 / 0.92 / 0.91
RDD 5 -Class (M)	0.88 / 0.91 / 0.89	0.92 / 0.88 / 0.90	0.75 / 0.75 / 0.75
RDD 5 - Class (L)	0.94 / 0.97 / 0.95	0.93 / 0.97 / 0.95	0.93 / 0.75 / 0.83

The training and testing accuracy curves with corresponding confusion matrices of each RDD (RDD 1 to RDD 5) appear in Figure 4 to show the U-Net deep learning model's performance during training and evaluation**.** The statistical test results are shown in Table 5.

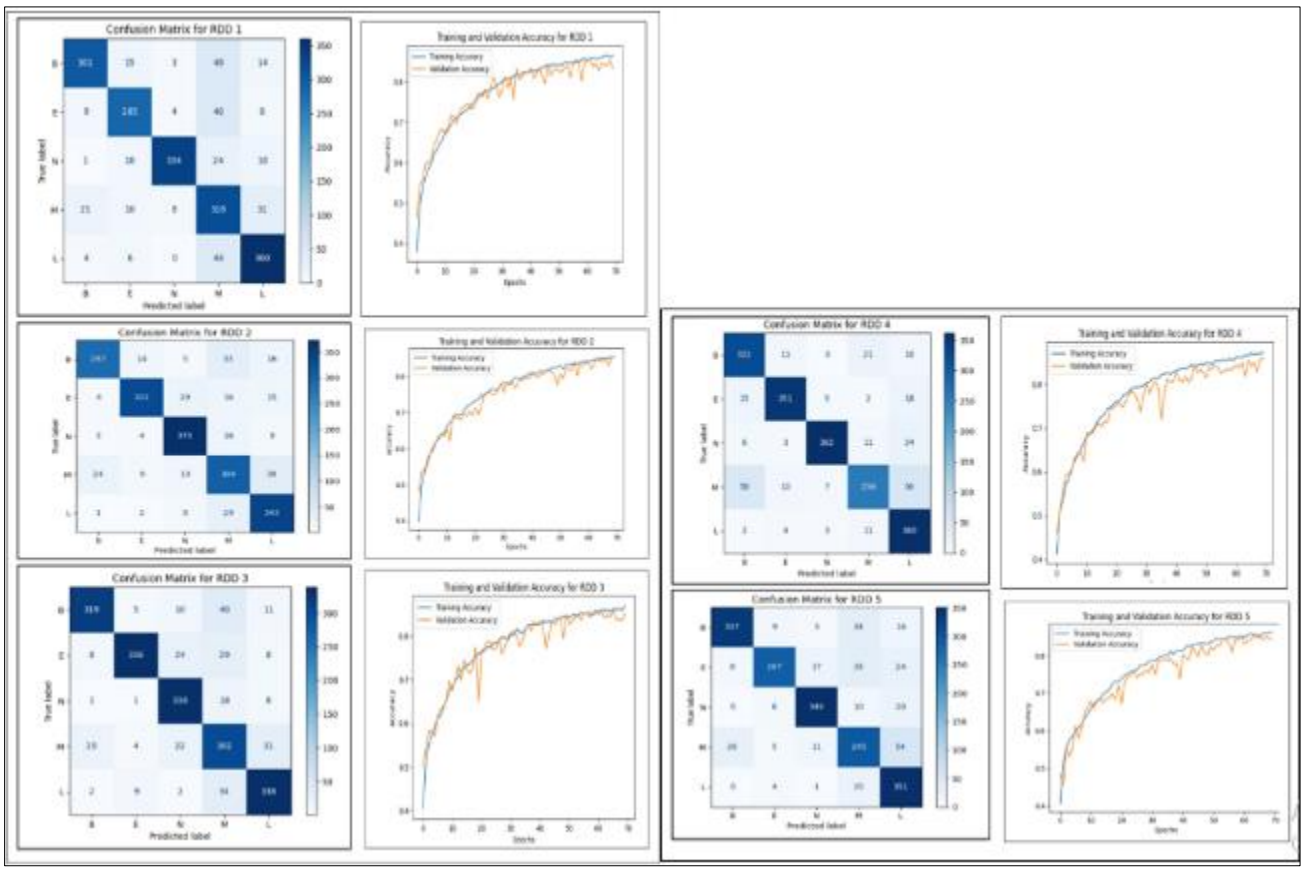

Figure 4 Plot for accuracy and confusion matrix for the U-Net model

Table 5 Statistical tests for each class in the RDDs

	WBC classes	ResNet50/support	VGG19/support	U-Net/support
	B	383	377	381
RDD1	E	386	397	352
	N	368	369	385
	M	385	390	388
	L	398	387	414
	B	351	381	355
RDD2	E	402	390	388
	N	375	385	405
	M	391	396	389
	L	401	368	383
	B	401	370	385
	E	356	398	405
RDD3	N	394	371	366
	M	361	390	379
	L	408	391	385
	B	393	371	364
RDD4	E	384	387	391
	N	381	392	406
	M	373	375	379
	L	389	395	380
	B	391	370	399
RDD5	E	406	391	369
	N	361	369	385
	M	359	383	391
	L	403	407	376

4.1. Training and Testing Accuracy

All RDDs show learning improvement through increasing accuracy in training and validation results during the epochs. At the beginning of training, the models show quick upward trends in accuracy, while validation accuracy follows training accuracy at a slight distance. The deep learning process displays sufficient convergence, but testing accuracy exhibits some tiny variations, which point to possible overfitting, especially with RDD 1 and RDD 2. Training accuracy for RDD 1 and RDD 2 demonstrates a persistent development pattern because it steadily rises throughout the process. The peak in testing accuracy is unstable as the results start to vary slightly afterwards, potentially due to data variability and possible overfitting issues. The learning process of RDD 3 and RDD 4 produces orderly convergence as training and testing accuracy levels rise together steadily, thus indicating balanced model training. The accuracy evolution in RDD 5 remains steady because both curves attain high accuracy levels, demonstrating reliable generalization capabilities.

The model classification results for various RDDs can be found in detail through the confusion matrix analysis. Key observations include: High Classification Accuracy for Most Classes: Strong classification achievements emerge from the diagonal elements, which control most spaces in each confusion matrix. Recall and precision performance of the B, E, and N classes demonstrate high stability when evaluating all RDDs. The model performs poorly with Class M, as

evidenced by noticeable incorrect classifications that primarily affect RDD 1 and RDD 2, since these rounds exhibited many Class M samples being misclassified as Class N. The misclassification in these classes stems from their overlapping features, which causes trouble in their distinction. The Class L prediction achieves success rates, except for minor cases where some samples in RDD 4 and RDD 5 are misclassified into the wrong categories. Class L requires better modifications to its feature extraction techniques. The model demonstrates uniform success throughout RDDs, regardless of a few incorrect predictions. The deep learning method demonstrates robustness through its steady classification outcomes. The deep learning models demonstrate robust learning capabilities through their high accuracy scores in classifying most instances, as shown in Figure 4. Further improvements are required to overcome classification difficulties with Class M objects and to decrease the number of wrong assignments for Class L types. The performance and error rate of the classification system are expected to improve with additional parameter refinements, specialized feature extraction methods, and class-specific modifications. To avoid unstable validation accuracy and overfitting in some of the RDDs, the author employs the following approach: First, the technique of augmenting the current dataset through random operations, such as rotation, translation, and scaling, is used to improve model generalization. Second, the concept of early stopping is employed about validation loss to prevent overfitting as training progresses. Third, dropout layers and L2 regularization are introduced in deep learning models to avoid dependency on specific features. Additionally, it accelerates learning by normalizing activations and reducing covariance at each layer during the training process. Therefore, the use of the concept of learning rate scheduling involves gradually adjusting the learning rate to improve convergence rates. Using multiple subsets of RDD validation also enables estimation of model reliability. Finally, federated averaging helps ensure that all models trained using the various sections in the RDDs converge to a single global model, thereby eliminating biases present in a specific RDD. The techniques mentioned above together make the model more robust and prevent the model from overfitting in the distributed data.

The evaluation focused on five distinct RDDs, utilizing deep learning models such as VGG19, ResNet50, and U-Net. The results, analyzed through the confusion matrix, showed that the performance of VGG19 on the 5 RDDs achieved higher accuracy compared to ResNet50 and U-Net. VGG19 demonstrated superior metrics such as precision, recall, and F1 scores, highlighting its ability to handle the complexities of WBC classification more effectively than the other models. In contrast, ResNet50 and U-Net performed well but did not match the accuracy levels achieved by VGG19 in this setup. These findings demonstrate the potential of leveraging advanced deep learning models with the distributed architecture of RDDs to manage large-scale data efficiently. This approach provides a reliable solution for medical applications requiring precise classification. While the data is divided among 5 RDDs and the data size is smaller than the entire dataset, utilizing RDD in big data has 9600 images and takes less memory. FL is an important aspect of this study because centralized training faces challenges of scalability and memory, especially when used in the classification of WBC in this study. Training VGG19 is not possible for a classical system given that 50,000 training images can cause out-of-memory errors and require much time. RDDs of Apache Spark support distributed computation, and although training the VGG19 architecture inside RDDs, it is slower due to processes such as shuffling, serialization, and transformation, which took approximately 2,275.088 sec per RDD across 70 epochs. This is because although designing and implementing these operations for RDDs takes a longer time compared to standardized formats, RDDs help avoid cases of memory overload on large-scale models, thus making FL with RDDs feasible where centralized training (from personal computers) would not be feasible.

5. Conclusion

The problems of scalability, accuracy, and privacy in medical diagnostics can be improved by combining FL, Big Data, and sophisticated deep learning models. To categorize WBCs, this study employed Apache Spark for effective data processing and FL for safe, distributed training across big datasets. The analysis utilized deep learning models, including VGG19, ResNet50, and U-Net, across five different RDDs. VGG19 regularly outperformed ResNet50 and U-Net, achieving the highest accuracy among the models, according to the data examined using the confusion matrix. VGG19 proved its capacity to manage the intricacies of WBC categorization more successfully by exhibiting superior metrics, including precision, recall, and F1 scores. U-Net underperformed compared to ResNet50, which also produced competitive results, demonstrating the different capabilities of these models for the given job. This method shows how the latest methods can be combined to increase diagnosis accuracy by using sophisticated deep-learning models trained on distributed RDDs. The results underscore the importance of utilizing both FL and Big Data frameworks to efficiently and securely manage massive datasets. This novel approach has great potential to improve biomedical research and healthcare by opening the door to more precise and scalable medical diagnostics.

Compliance with ethical standards

Disclosure of conflict of interest

There is no conflict of interest.

References

[1] Z. Zhang, "Early warning model of adolescent mental health based on big data and machine learning," Soft comput, vol. 28, no. 1, pp. 811–828, Jan. 2024, doi: 10.1007/S00500-023-09422-Z.

[2] I. Lee, "Big data: Dimensions, evolution, impacts, and challenges," Bus Horiz, vol. 60, no. 3, pp. 293–303, May 2017, doi: 10.1016/J.BUSHOR.2017.01.004.

[3] A. Mumuni and F. Mumuni, "Automated data processing and feature engineering for deep learning and big data applications: A survey," Journal of Information and Intelligence, vol. 3, no. 2, pp. 113–153, Mar. 2025, doi: 10.1016/J.JIIXD.2024.01.002.

[4] L. Theodorakopoulos, A. Theodoropoulou, and Y. Stamatiou, "A State-of-the-Art Review in Big Data Management Engineering: Real-Life Case Studies, Challenges, and Future Research Directions," Eng 2024, Vol. 5, Pages 1266-1297, vol. 5, no. 3, pp. 1266–1297, Jul. 2024, doi: 10.3390/ENG5030068.

[5] L. Jin et al., "Big data, machine learning, and digital twin assisted additive manufacturing: A review," Mater Des, vol. 244, p. 113086, Aug. 2024, doi: 10.1016/J.MATDES.2024.113086.

[6] J. Luo, W. Zhuo, S. Liu, and B. Xu, "The Optimization of Carbon Emission Prediction in Low Carbon Energy Economy under Big Data," IEEE Access, vol. 12, pp. 14690–14702, 2024, doi: 10.1109/ACCESS.2024.3351468.

[7] T. A. M. Devi and P. Thangaselvi, "Ancient blood cell classification on explication using convolutional neural networks," Multimed Tools Appl, vol. 84, no. 19, pp. 21105–21115, Jun. 2025, doi: 10.1007/S11042-024-19865-7.

[8] Y. Guo, A. I. Shahin, and H. Garg, "An indeterminacy fusion of encoder-decoder network based on neutrosophic set for white blood cells segmentation," Expert Syst Appl, vol. 246, p. 123156, Jul. 2024, doi: 10.1016/J.ESWA.2024.123156.

[9] N. S. Jagtap et al., "Deep learning-based blood cell classification from microscopic images for haematological disorder identification," Multimed Tools Appl, vol. 84, no. 20, pp. 21917–21944, Jun. 2025, doi: 10.1007/S11042-024-19900-7.

[10] R. Asghar, S. Kumar, and P. Hynds, "Automatic classification of 10 blood cell subtypes using transfer learning via pre-trained convolutional neural networks," Inform Med Unlocked, vol. 49, p. 101542, Jan. 2024, doi: 10.1016/J.IMU.2024.101542.

[11] O. Islam, M. Assaduzzaman, and M. Z. Hasan, "An explainable AI-based blood cell classification using optimized convolutional neural network," J Pathol Inform, vol. 15, p. 100389, Dec. 2024, doi: 10.1016/J.JPI.2024.100389.

[12] A. H. Malik and S. Rahman, "Toward precision wound healing: Integrating regenerative therapies and smart technologies," International Journal of Science and Research Archive, vol. 16, no. 3, pp. 244–257, Sep. 2025, doi: 10.30574/IJSRA.2025.16.3.2492.

[13] A. H. Malik and S. Rahman, "Hybrid Temozolomide Nanoconjugates: A polymer–drug strategy for enhanced stability and glioblastoma therapy," International Journal of Science and Research Archive, vol. 16, no. 3, pp. 258–268, Sep. 2025, doi: 10.30574/IJSRA.2025.16.3.2493.

[14] A. H. Malik and S. Rahman, "Molecular erasers: Reprogramming cancer immunity through protein degradation," World Journal of Advanced Engineering Technology and Sciences, vol. 16, no. 3, pp. 277–291, Sep. 2025, doi: 10.30574/WJAETS.2025.16.3.1335.

[15] M. Sabbir Hossain et al., "Automatic Navigation and Self-Driving Technology in Agricultural Machinery: A State-of-the-Art Systematic Review," IEEE Access, vol. 13, pp. 94370–94401, 2025, doi: 10.1109/ACCESS.2025.3573324.

[16] M. S. Hossain, M. Rahman, M. Ahmed, M. M. Kabir, M. F. Mridha, and J. Shin, "DeepLabv3Att: Integrating Attention Mechanisms in DeepLabv3 for Enhanced Road Segmentation," 2024 International Conference on Innovation and

Intelligence for Informatics, Computing, and Technologies, 3ICT 2024, pp. 711–718, 2024, doi: 10.1109/3ICT64318.2024.10824285.

[17] M. Rahman, M. S. Hossain, A. A. Eva, M. M. Kabir, M. F. Mridha, and J. Shin, "Alzheimers Disease Classification with a Hybrid CNN-SVM Approach on Enhanced MRI Data," 2024 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, 3ICT 2024, pp. 50–57, 2024, doi: 10.1109/3ICT64318.2024.10824633.

[18] M. Rahman, M. S. Hossain, U. Rozario, S. Roy, M. F. Mridha, and N. Dey, "MultiSenseNet: Multi-Modal Deep Learning for Machine Failure Risk Prediction," IEEE Access, vol. 13, pp. 120404–120416, 2025, doi: 10.1109/ACCESS.2025.3586978.

[19] M. S. Hossain, M. Rahman, M. G. R. Abir, J. Maua, and A. Rahman, "Plant Leaf Disease Detection Using Deep Stacking: Integrating CNNs and Gradient Boosting for Enhanced Classification Accuracy," Studies in Computational Intelligence, vol. 1202, pp. 15–26, 2025, doi: 10.1007/978-981-96-4520-6_2.

[20] M. R. Ahmed, R. Haque, S. M. A. Rahman, A. W. Reza, N. Siddique, and H. Wang, "Vision-audio multimodal object recognition using hybrid and tensor fusion techniques," Information Fusion, vol. 126, p. 103667, Feb. 2026, doi: 10.1016/J.INFFUS.2025.103667.

[21] J. Debnath et al., "LMVT: A hybrid vision transformer with attention mechanisms for efficient and explainable lung cancer diagnosis," Inform Med Unlocked, vol. 57, p. 101669, Jan. 2025, doi: 10.1016/J.IMU.2025.101669.

[22] Md. R. Ahmed et al., "Hierarchical Swin Transformer Ensemble with Explainable AI for Robust and Decentralized Breast Cancer Diagnosis," Bioengineering 2025, Vol. 12, Page 651, vol. 12, no. 6, p. 651, Jun. 2025, doi: 10.3390/BIOENGINEERING12060651.

[23] R. Haque et al., "Explainable deep stacking ensemble model for accurate and transparent brain tumor diagnosis," Comput Biol Med, vol. 191, p. 110166, Jun. 2025, doi: 10.1016/J.COMPBIOMED.2025.110166.

[24] M. I. H. Siddiqui et al., "Accelerated and accurate cervical cancer diagnosis using a novel stacking ensemble method with explainable AI," Inform Med Unlocked, vol. 56, p. 101657, Jan. 2025, doi: 10.1016/J.IMU.2025.101657.

[25] R. Haque, S. H. Laskar, K. G. Khushbu, M. J. Hasan, and J. Uddin, "Data-Driven Solution to Identify Sentiments from Online Drug Reviews," Computers 2023, Vol. 12, Page 87, vol. 12, no. 4, p. 87, Apr. 2023, doi: 10.3390/COMPUTERS12040087.

[26] R. Haque et al., "Advancing Early Leukemia Diagnostics: A Comprehensive Study Incorporating Image Processing and Transfer Learning," BioMedInformatics 2024, Vol. 4, Pages 966-991, vol. 4, no. 2, pp. 966–991, Apr. 2024, doi: 10.3390/BIOMEDINFORMATICS4020054.

[27] A. Hossain et al., "Transformer-Based Ensemble Model for Binary and Multiclass Oral Cancer Segmentation," 2025 International Conference on Electrical, Computer and Communication Engineering, ECCE 2025, 2025, doi: 10.1109/ECCE64574.2025.11012921.

[28] A. Al Noman et al., "Monkeypox Lesion Classification: A Transfer Learning Approach for Early Diagnosis and Intervention," 2024 7th International Conference on Contemporary Computing and Informatics (IC3I), pp. 247–254, Sep. 2024, doi: 10.1109/IC3I61595.2024.10828678.

[29] J. Debnath et al., "Hybrid Vision Transformer Model for Accurate Prostate Cancer Classification in MRI Images," 2025 International Conference on Electrical, Computer and Communication Engineering, ECCE 2025, 2025, doi: 10.1109/ECCE64574.2025.11013952.

[30] Y. Zhai et al., "FedRAV: Hierarchically Federated Region-Learning for Traffic Object Classification of Autonomous Vehicles," Nov. 2024, Accessed: Oct. 03, 2025. [Online]. Available: https://arxiv.org/pdf/2411.13979

[31] K. Begum, M. A. I. Mozumder, M. Il Joo, and H. C. Kim, "BFLIDS: Blockchain-Driven Federated Learning for Intrusion Detection in IoMT Networks," Sensors 2024, Vol. 24, Page 4591, vol. 24, no. 14, p. 4591, Jul. 2024, doi: 10.3390/S24144591.

[32] N. F. Syed, M. Ge, and Z. Baig, "Fog-cloud based intrusion detection system using Recurrent Neural Networks and feature selection for IoT networks," Computer Networks, vol. 225, p. 109662, Apr. 2023, doi: 10.1016/J.COMNET.2023.109662.

[33] P. Basu, T. S. Roy, R. Naidu, and Z. Muftuoglu, "Privacy enabled Financial Text Classification using Differential Privacy and Federated Learning," Proceedings of the 3rd Workshop on Economics and Natural Language Processing, ECONLP 2021, pp. 50–55, Oct. 2021, doi: 10.18653/v1/2021.econlp-1.7.

[34] E. Ashraf, N. F. F. Areed, H. Salem, E. H. Abdelhay, and A. Farouk, "FIDChain: Federated Intrusion Detection System for Blockchain-Enabled IoT Healthcare Applications," Healthcare 2022, Vol. 10, Page 1110, vol. 10, no. 6, p. 1110, Jun. 2022, doi: 10.3390/HEALTHCARE10061110.

[35] A. Qayyum, K. Ahmad, M. A. Ahsan, A. Al-Fuqaha, and J. Qadir, "Collaborative Federated Learning for Healthcare: Multi-Modal COVID-19 Diagnosis at the Edge," IEEE Open Journal of the Computer Society, vol. 3, pp. 172–184, 2022, doi: 10.1109/OJCS.2022.3206407.

[36] A. S. U. K. Pranta et al., "A Novel MaxViT Model for Accelerated and Precise Soybean Leaf and Seed Disease Identification," Computers 2025, Vol. 14, Page 197, vol. 14, no. 5, p. 197, May 2025, doi: 10.3390/COMPUTERS14050197.

[37] A. Al Noman et al., "ViX-MangoEFormer: An Enhanced Vision Transformer–EfficientFormer and Stacking Ensemble Approach for Mango Leaf Disease Recognition with Explainable Artificial Intelligence," Computers 2025, Vol. 14, Page 171, vol. 14, no. 5, p. 171, May 2025, doi: 10.3390/COMPUTERS14050171.

[38] S. Khan et al., "Ensemble-Based Explainable Approach for Rare Medicinal Plant Recognition and Conservation," 2025 10th International Conference on Information and Network Technologies, ICINT 2025, pp. 88–93, 2025, doi: 10.1109/ICINT65528.2025.11030872.

[39] H. Rahman et al., "Automated Weed Species Classification in Rice Cultivation Using Deep Learning," 2025 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1–6, Feb. 2025, doi: 10.1109/ECCE64574.2025.11014047.

[40] M. S. Rahman et al., "Effective Disease Recognition in Cucumbers: A Web-Based Application Using Transfer Learning Models," 2024 IEEE 3rd International Conference on Robotics, Automation, Artificial-Intelligence and Internet-of-Things, RAAICON 2024 - Proceedings, pp. 59–64, 2024, doi: 10.1109/RAAICON64172.2024.10928353.

[41] M. I. Hossain Siddiqui et al., "Eggplant Disease Diagnosis Using a Robust Ensemble of Transfer Learning Architectures," 2025 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1–6, Feb. 2025, doi: 10.1109/ECCE64574.2025.11013918.

[42] J. Ferdousi, S. I. Lincoln, M. K. Alom, and M. Foysal, "A deep learning approach for white blood cells image generation and classification using SRGAN and VGG19," Telematics and Informatics Reports, vol. 16, p. 100163, Dec. 2024, doi: 10.1016/J.TELER.2024.100163.

[43] S. A. Tarimo, M. A. Jang, E. E. Ngasa, H. B. Shin, H. J. Shin, and J. Woo, "WBC YOLO-ViT: 2 Way - 2 stage white blood cell detection and classification with a combination of YOLOv5 and vision transformer," Comput Biol Med, vol. 169, p. 107875, Feb. 2024, doi: 10.1016/J.COMPBIOMED.2023.107875.

[44] S. Ratheesh and A. A. Breethi, "Deep learning based Non-Local k-best renyi entropy for classification of white blood cell subtypes," Biomed Signal Process Control, vol. 90, p. 105812, Apr. 2024, doi: 10.1016/J.BSPC.2023.105812.

[45] A. Kanavos, O. Papadimitriou, K. Al-Hussaeni, M. Maragoudakis, and I. Karamitsos, "Advanced Convolutional Neural Networks for Precise White Blood Cell Subtype Classification in Medical Diagnostics," Electronics 2024, Vol. 13, Page 2818, vol. 13, no. 14, p. 2818, Jul. 2024, doi: 10.3390/ELECTRONICS13142818.

[46] Ş. N. Özcan, T. Uyar, and G. Karayeğen, "Comprehensive data analysis of white blood cells with classification and segmentation by using deep learning approaches," Cytometry Part A, vol. 105, no. 7, pp. 501–520, Jul. 2024, doi: 10.1002/CYTO.A.24839.

[47] L. W. Wu, C. R. Huang, and Y. C. Ouyang, "Segmentation domain adaptation for enhanced white blood cell classification using spatial co-occurrence feature embedding," Biomed Signal Process Control, vol. 112, p. 108529, Feb. 2026, doi: 10.1016/J.BSPC.2025.108529.

[48] H. Song and Z. Wang, "Automatic Classification of White Blood Cells Using a Semi-Supervised Convolutional Neural Network," IEEE Access, vol. 12, pp. 44972–44983, 2024, doi: 10.1109/ACCESS.2024.3380896.