

Multiclass blood cancer classification using deep CNN with optimized features

Wahidur Rahman ^{a,c,*}, Mohammad Gazi Golam Faruque ^b, Kaniz Roksana ^c, A H M Saifullah Sadi ^c, Mohammad Motiur Rahman ^a, Mir Mohammad Azad ^d

^a Department of Computer Science and Engineering, Mawlana Bhashani Science and Technology University, Tangail, 1902, Bangladesh

^b Department of Computer Science and Engineering, Khwaja Yunus Ali University, Sirajganj, Bangladesh

^c Department of Computer Science and Engineering, Uttara University, Dhaka, Bangladesh

^d Department of Computer Science and Engineering, Hamdard University Bangladesh, Munshiganj, 1510, Bangladesh

ARTICLE INFO

Keywords:

Blood cancer

Convolutional neural network

Particle swarm optimization

Cat swarm optimization

Machine learning

ABSTRACT

Breast cancer, lung cancer, skin cancer, and blood malignancies such as leukemia and lymphoma are just a few instances of cancer, which is a collection of cells that proliferate uncontrollably within the body. Acute lymphoblastic leukemia is one of the significant forms of malignancy. The hematologists frequently make an oversight while determining a blood cancer diagnosis, which requires an excessive amount of time. Thus, this research reflects on a novel method for the grouping of the leukemia with the aid of the modern technologies like Machine Learning and Deep Learning. The proposed research pipeline is occupied into some interconnected parts like dataset building, feature extraction with pre-trained Convolutional Neural Network (CNN) architectures from each individual images of blood cells, and classification with the conventional classifiers. The dataset for this study is divided into two identical categories, Benign and Malignant, and then reshaped into four significant classes, each with three subtypes of malignant, namely, Benign, Early Pre-B, Pre-B, and Pro-B. The research first extracts the features from the individual images with CNN models and then transfers the extracted features to the features selections such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and SVC Feature Selectors along with two nature inspired algorithms like Particle Swarm Optimization (PSO) and Cat Swarm Optimization (CSO). After that, research has applied the seven Machine Learning classifiers to accomplish the multi-class malignant classification. To assess the efficacy of the proposed architecture a set of experimental data have been enumerated and interpreted accordingly. The study discovered a maximum accuracy of 98.43% when solely using pre-trained CNN and classifiers. Nevertheless, after incorporating PSO and CSO, the proposed model achieved the highest accuracy of 99.84% by integrating the ResNet50 CNN architecture, SVC feature selector, and LR classifiers. Although the model has a higher accuracy rate, it does have some drawbacks. However, the proposed model may also be helpful for real-world blood cancer classification.

Contributions of the paper

- To extract the features with pre-trained CNN models from the individual images of four significant stages of both the healthy and malignant tissues.
- To apply the feature selection algorithms to work with optimized deep features and track out the performance.

- To apply the nature inspired algorithms to find the best features from the extracted and apply the ML based classifiers and interprets the calculated experimental data accordingly.

1. Introduction

Cancer is a cluster of cells undergoing unchecked growth in the body [1], and it can quickly spread to any organ. Cancer comes in various forms; the most common are breast cancer [2], lung cancer, skin cancer,

* Corresponding author. Department of Computer Science and Engineering, Mawlana Bhashani Science and Technology University, Tangail, 1902, Bangladesh.

E-mail addresses: wahiduhin0@gmail.com (W. Rahman), golam.faruq@gmail.com (M.G.G. Faruque), kanizroksana96@gmail.com (K. Roksana), saifullah.cse@uttarauniversity.edu.bd (A.H.M.S. Sadi), motiurcse@mbstu.ac.bd (M.M. Rahman), csdrazad@hamdarduniversity.edu.bd (M.M. Azad).

and blood cancers like leukemia and lymphoma. There have been 9.2 million fatalities from lung cancer, 1.7 million from skin cancer, and 627,000 from breast cancer [3,4], according to reports from the World Health Organization (WHO) [5]. When it comes to cancers, leukemia has a remarkably high mortality rate. It's a malignant tumor that forms in the bone marrow when immature white blood cells are cloned in a destructive way. With lung, colon, breast, and prostate cancers, leukemia [6,7] is among the most frequently diagnosed cancers in the United States. According to projections made by the US government's cancer data collector, the Surveillance, Epidemiology, and End Results (SEER) Program, there were 60,650 newly diagnosed cases of leukemia, and 24,000 death occurred in the US in 2022. According to a review of the cancer database by the WHO, leukemia [8] incidence varies significantly by region and subtype. More than 20,000 cases of pediatric blood cancer are detected annually in India, with approximately 15,000 cases of leukemia only [9]. Around 61,780 instances of leukemia were diagnosed in the United States in 2019, with another 9900 cases being found in the United Kingdom. From 345,000 in 1990 to 518,000 in 2018, the number of newly diagnosed cases of leukemia increased, lowering the Annualized Survival Insusceptibility Rate (ASIR) by 0.43% per year [10,11].

The subtypes of leukemia are Acute Leukemia (AL) and Chronic Leukemia (CL). The progression of CL is usually gradual. On the other hand, without specialized care, the average life expectancy for those with AL is only 3 months [6]. There are two subtypes of AL recognized by the French-American-British (FAB) classification system: Acute Myeloid Leukemia (AML) and Acute Lymphoblastic Leukemia (ALL) [12]. Additionally, Chronic Myeloid Leukemia (CML) and Chronic Lymphocytic Leukemia (CLL) are the 2 different types of CL [12–15]. Acute Lymphocytic Leukemia (ALL) is a fatal malignancy that affects both adults and children, making up around 25% [16] of all malignancies diagnosed in kids.

Again, ALL is often referred to as acute lymphoblastic leukemia. The term "acute" denotes that, if neglected, leukemia can spread swiftly and be lethal within a matter of months. The term "lymphocytic" refers to the fact that it originates from lymphocyte precursors, a subset of white blood cells. Leukemia that begins in the bone marrow and spreads rapidly is called Acute Lymphoblastic Leukemia (ALL) [17,18] or acute lymphocytic leukemia (ALL) [19]. The rapid proliferation of leukemic cells in the blood and their subsequent spread to other organs and systems of the body [20], such as the spleen, lymph nodes, liver, brain, and neurological system, can produce a wide range of various symptoms [21]. Some of these symptoms include bruising, bleeding from the gums and nose, fever, swollen lymph nodes, sore joints, and infections [22]. Bone marrow and blood are the primary organs affected by Acute Lymphoblastic Leukemia [8,23,24]. Since it occurs more frequently in children than chronic or myeloid leukemia, the term "acute childhood leukemia" has been coined to describe this condition. It can be cured if caught early enough, but if not, it can kill in a matter of months if left untreated [25–27]. L1, L2, and L3 are ALL subtypes recognized by the FAB categorization system. The nuclei of the L1-type cells are tiny in size and have homogeneous chromatin, few nucleoli, and a modest amount of basophilic cytoplasm. Nevertheless, L2 of enormous size exhibits uneven nuclear structure and clefting. L3 are very large or medium-large, with prominent cytoplasmic vacuoles. The lifespan rate can be increased with appropriate therapy, but only if the cancer is detected early and correctly [21,28].

Hematologists perform blood smears or bone marrow examinations under the microscope to diagnose ALL and its subtypes. Nevertheless, the accuracy of these tests depends on the expertise of the examining pathologists and may be compromised by the microscope's extended use [21,29–31]. Besides, the hematologists frequently makes an oversight while determining a blood cancer diagnosis, which requires an excessive amount of time. Automatic diagnosis methods are urgently needed to decrease the reliance on manual examination, speed up the procedure, and improve the precision of leukemia identification. In recent years, human-centric clinical diagnosis powered by Machine Learning (ML),

Artificial Intelligence (AI), and Deep Learning (DL) has emerged as a significant tool for aiding clinicians in determining treatment decisions. Thus, many computer-aided diagnosis methods have been developed to detect ALL blood pictures without human intervention [8,32].

Thus, this research reflects on the ML based method to classify the type of Acute Lymphoblastic Leukemia (ALL) to classify the malignant tissues with optimized deep features. The contribution of this study is given as follows.

- The proposed research has performed image preprocessing techniques and extracted the features from the individual images with pre-trained CNN models of four significant stages of both the healthy and malignant tissues.
- This study then applied feature selection algorithms to work with optimized deep features and track out the performance for malign classification.
- To achieve the optimum results in multi-class leukemia classification, the proposed research has applied the nature inspired algorithms to find the best features from the extracted features and apply the ML based classifiers and interprets the calculated experimental data accordingly.

The manuscript is classified into five sections. The section two represents a comprehensive study of previous works. Section three provides proposed method and working principles. The section four illustrates the results with the relevant discussion. Finally, the section five represents the conclusion of this manuscript.

2. Literature review

Many machine learning and deep learning techniques were used to identify or classify the ALL (acute lymphoblastic leukemia) type. Some of the previous papers are described in this section.

Researchers offer a novel Bayesian-based optimized CNN method for identifying ALL in microscopic smeared images in study [8]. A hybrid dataset was formed to be used in this study by combining two subsets of ALL-IDB datasets (ALL-IDB1 and ALL-IDB2). The hybrid dataset consists of 368 blood smear photos. In the test set, the optimized CNN model for ALL identification was found using the Bayesian optimization technique, and it achieved maximum accuracy of 100%.

The article [11] suggested an approach of convolutional neural network called SK U-Net to perform the task of nucleus segmentation for ALL. All 198 input photos come from the publicly available database ALL-IDB2. The SK U-Net achieved a higher Dice score of 0.916 than the traditional U-Net, which only achieved a score of 0.320. A 98% accuracy is achieved with SVM, which is significantly higher than other methods. The proposed method has a higher accuracy of 0.97% than prior methods. Additionally, KNN and SVM achieved an accuracy of 0.85% and 0.98%, respectively.

The latest developments in ALL detection and categorization using deep and machine learning are presented in study [16] through a systematic review. This article thoroughly examines the advantages and disadvantages of many different AI-based ALL detection methods. Lastly, a range of tough topics and potential future scopes are presented, which may inspire readers to develop their own research questions in ALL areas.

For the categorization of ALL using microscopic pictures of white blood cells, a powerful and effective hybrid InceptionV3 XGBoost framework was developed in paper [19]. Both InceptionV3 and the XGBoost model are used in the proposed approach, while Inception is responsible for extracting image features, and XGBoost handles categorization. The proposed method used the transfer learning capability of pre-trained CNN architectures to train classifiers for the ISBI C-NMC 2019 dataset. The database contains white blood cell photos of 10,661. The proposed hybrid model obtains an F1 score of 0.986.

In order to acquire reliable classification results from the model's

training on the bone marrow pictures, a robust segmentation methodology along with deep learning techniques and a CNN are utilized in study [21]. Amreek Clinical Laboratory in Pakistan provided the dataset for this study. The experimental results indicate that the suggested approach has a high accuracy of 97.78%.

Article [22] proposed a new approach to categorizing and identifying ALL using SVM and CNN. In the proposed research, lymphocyte identification is followed by retrieving CNN features using the Alex-Net Model. Finally, SVM is used to classify the discovered cell as either normal or cancerous. In order to conduct this research, 4000 blood smears lymphocyte samples were collected from the Hayatabad Medical Complex in Peshawar, Pakistan. The accuracy of the proposed method is 98%.

The Acute lymphoblastic leukemia is diagnosed with the use of a ViT-CNN ensemble model in article [33]. Vision transformer and Convolutional Neural Network (CNN) models were combined together to generate the ensemble model. The study used a noisy, unbalanced ISBI 2019 dataset (consisting of 10,661 cell images). For the evaluation set, the proposed ViT-CNN model achieved an accuracy of 99.03% in its classification.

To divide ALL into two groups, a convolutional network with 10 layers and two-by-two max-pooling layers (with strides of 2) was proposed, and 6 widely used machine learning approaches were constructed in the research [34]. Both the ResNet50 and VGG16, which are both well-known deep learning networks, were utilized in this study. The dataset was gathered from a CodaLab competition. In addition, the validation accuracy of VGG16 is 84.62%, ResNet50 is 81.63%, and the proposed convolutional network is 82.10%. Moreover, the accuracy of machine learning classifiers is as follows: 81.72% for RF, 79.88% for LR, 79.28% for SVM, 77.89% for KNN, 68.91% for SGD, and 27.33% for MLP.

In order to classify photos of ALL and healthy cells, study [35] presented an attention-based CNN. The suggested method consists of a CNN-based model that employs a module named ECA (Efficient Channel Attention) in conjunction with the VGG16 to extract higher-quality feature information from the image dataset. The study utilized the C-NMC dataset which consists of 10,661 single cell pictures. The experimental outcomes demonstrate the effectiveness of the suggested CNN model in extracting deep features, with an accuracy of 91.1%.

A quick and accurate diagnosis of ALL can be aided by the instance segmentation proposed in article [36], accomplished by applying Mask R-CNN on microscope pictures of white blood cells. The present research used the transfer learning method to build Mask R-CNN in order to fit the instance segmentation problem on microscopy white blood cell images. To fix the issue of poor lighting in stained white blood cell microscope images, the proposed method applied a contrast enhancement method to the dataset. An actual dataset from Dr. Soetomo Hospital in Surabaya was used by the proposed method. The system achieved 83.72% accuracy.

Research [37] includes the following 3 proposed systems: the first consists of a feed forward neural network (FFNN), an ANN (artificial neural network), and an SVM. These three components are all based on hybrid features that were retrieved with the LBP (Local Binary Pattern), Gray Level Co-occurrence Matrix (GLCM), and FCH (Fuzzy Color Histogram) methods, respectively. CNN models AlexNet, GoogleNet, and ResNet-18, trained with the transfer learning approach, are the basis of the second suggested system. These architectures were used to successfully extract and classify deep feature maps. The third proposed method combines CNN and SVM algorithms to extract and categorize feature maps. The dataset of ALL IDB1 and ALL IDB2 were used in the study. A total of 108 pictures are in the ALL IDB1 dataset and the ALL-IDB2 dataset consists 260 images. The accuracy of the ANN and FFNN were 100%, where the SVM had an accuracy of 98.11. Additionally, for the ALL IDB1 and ALL IDB2 dataset, the AlexNet, GoogleNet and ResNet achieved an accuracy of 100%. ResNet-18 + SVM outperformed the competition on the ALL IDB1 and ALL IDB2 datasets (100%

accuracy). On the other hand, AlexNet + SVM, GoogLeNet + SVM, and ResNet50 + SVM hybrid models obtained 100%, 98.1% and 100% accuracy, respectively.

For leukemia diagnosis, research [38] offered an intelligent method for automating the process of detecting lymphoblasts (blast cells) in the single-celled image. A CNN is used in the method's implementation to distinguish malignant from healthy blood cells. The C NMC 2019 dataset was used to train and evaluate the proprietary ALLNET model. The dataset consists of 10,661 pictures. The highest accuracy achieved by the custom deep learning ALL-NET classifier was 95.54.

The goal of study [39] is to develop methods for rapid and accurate detection of ALL cells in order to stop cancer from spreading in children. CNN was utilized in this method to identify every cell that was present on the blood smears sample slide. The main focus of this research is to provide clinicians in hospital Hematology Laboratories with a better tool for detecting ALL utilizing CNN. The accuracy of CNN in detecting ALL cancer was determined to be 98.53%.

Study [40] proposes a method that uses CNNs (Convolutional Neural Networks) to identify WBC and then investigate ALL illnesses automatically. Using the ALL IDB dataset, the efficiency of three different pre-trained CNN models (VGG, Alexnet, and GoogleNet) was compared. The results show that GoogleNet and VGG were superior to AlexNet in terms of pre-trained models, with both obtaining 100% accuracy throughout training. According to the results of the tests, VGG has the highest performance, with an accuracy of 99.13%.

The purpose of study [41] is to develop a deep learning model for detecting acute leukemia from images of lymphocytes and monocytes, utilizing a customized framework. In order to classify pictures of acute leukemia, a CNN model was presented that combined the Tversky loss function and the Adam optimizer with the four dense layers, six convolution layers, and a Softmax activation function. The database was obtained from the Shahid Ghazi Tabatabai Cancer Center in Tabriz. The suggested approach correctly identified 99% of cases of acute leukemia, such as ALL and AML (Acute Myeloid Leukemia).

Using the EfficientNet-B3 CNN framework, study [42] categorized ALL as a model that automatically modifies its own learning rate. The suggested model was tested using the C-NMC Leukemia dataset which contains 27,558 pictures of RBCs. Recently developed classifiers were used to assess the proposed model. In general, the proposed model had an accuracy of 98.31%, and the Disc similarity coefficient (DSC) of 98.05%. The proposed methodology was also used to separate healthy and parasitized microscopic pictures with an average accuracy of 97.68%, proving its usefulness beyond detecting ALL.

The article [43] proposes an autonomous approach that gives users access to a widely-used classifier for improved facial expression recognition. The system is broken down into two primary machine-learning phases: feature selection and feature categorization. The Active Shape Model (ASM), which is made up of landmarks, is used to perform feature selection, and the classification of features has been evaluated using seven different well-known classifiers, namely KNN, NB, DT, Quadratic classifier, RF, MLP, and SVM. The experimental findings showed that the Quadratic classifier gives excellent performance and has the best accuracy of any classifier tested (92.42%).

Studies [44] have provided an overview of tools for detecting cancer and techniques for treating it. The study aimed to develop accurate colon cancer survival prediction models by analyzing data from the SEER program. The authors also evaluated different classification systems to estimate mortality rates within five years of diagnosis. Results showed that the deep autoencoder model yielded the best prediction performance (97%) and AUC-ROC (95%), respectively.

In the study [45], a unique hybrid AlexNet-gated recurrent unit (AlexNet-GRU) model was used to detect and classify breast cancer in lymph nodes (LNs). Three models, including the suggested AlexNet-GRU, the convolutional neural network GRU (CNN-GRU), and the CNN long short-term memory (CNN-LSTM), are evaluated and contrasted in the present study. The experimental results showed that

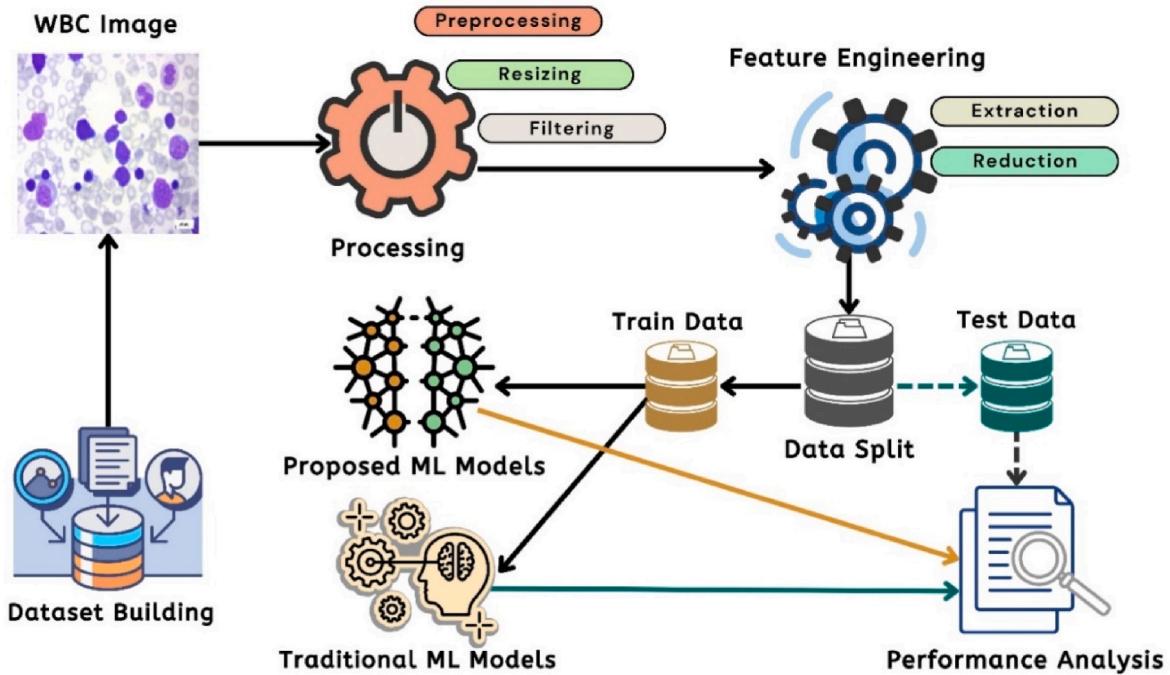


Fig. 1. Overall system illustration of the proposed system.

the proposed AlexNet-GRU model outperformed the CNN-GRU and CNN-LSTM models on all performance metrics, with an accuracy of 99.50%, respectively.

The research [46] introduces the Histogram of Directional Gradient (HDG) and the Histogram of Directional Gradient Generalized (HDGG), two revolutionary new descriptors for collecting discriminant facial expression features that outperform existing classifiers in terms of accuracy and efficiency in feature extraction. The proposed descriptors are grounded in linear classification using SVM and directional local gradients. Low-dimensional characteristics are employed to improve classification performance, allowing for more accurate face and expression recognition to be developed. Compared to other works already published, the experiment's findings demonstrate an accuracy of 92.12%.

Many previous work [36,38–41] only used DL techniques and on the other hand, some work [22,34,37] used both ML and DL techniques to detect ALL cases from the image dataset. Optimizer algorithm is used in the model to reduce the loss and improve the accuracy level of the model. Only paper [36–38,41] used optimizer algorithm. In this proposed framework both ML & DL techniques are used with 2 optimizer algorithm PSO & CSO which improvised the accuracy level of the proposed work. Additionally, some previous work [34,36,37,40,41] dataset size was very poor and it effects the model's performance. Besides, most of the previous studies worked on the binary classification of the infected blood cell to detect malignant. Multiclass classification was merely used in these studies with less significances. Thus, the proposed model used a large dataset that mainly focuses on the multiclass classification of the infected cells of blood tissues along with its noteworthy stages. Further, seven traditional ML algorithms with five DL methods are utilized to classify the Acute Lymphoblastic Leukemia (All). Moreover, PCA, LDA and SVC feature selector algorithms were used in the present study with PSO and CSO optimization algorithms.

3. Materials and methodology

In this section, methodology of the research is described. The section is classified into four interconnected subsections such as the research dataset, feature extraction with pre-trained Convolutional Neural Network (CNN) models, the extraction of the feature vectors and

classification with the conventional Machine Learning (ML) classifiers. Fig. 1 shows the overall proposed system illustration with existing components.

In this figure, first, the dataset is collected from the secondary sources. After that, image pre-processing has been applied to enrich the dataset. After that the images has been provided to the pre-trained CNN models to extract the significant features. Then, the research has applied the feature optimization techniques to find the best features from the collected image features. After that, the dataset is spitted into two identical parts training data and test data where train data is provided to the ML models with existing traditional models to find the experimental results. Also, test data has applied to evaluate performance analysis matrices. After finding the experimental results, a set of comparison have been performed (see Table 1).

3.1. Dataset

This research has been taken the dataset from the secondary sources more specifically from Kaggle [47]. The dataset is comprised with 3262 images of actual peripheral blood smear images. The images were included from the 89 patients where 25 patients were suspected as healthy individuals and rest of the 64 patients were suspected as Acute Lymphoblastic Leukemia (ALL). The dataset is classified into two identical classes such as Benign and malignant categories and further reshaped the dataset into four significant classes with three subtype of malignants namely, Benign, Early Pre-B, Pre-B and Pro-B. All of the images were captured with a Zeiss camera in a microscope at 100 \times magnification and stored the images as JPG format in the storage. The types and subtypes of these images were carefully and conclusively determined by a specialist using flow cytometry. Table 2 shows the corresponding sample images for this research.

3.2. Feature extraction

In this sub-section, a mechanism of feature extraction is described. Firstly, the pipeline of feature extraction will be presented with algorithmic annotations. Then, the mechanism of feature selection will be illustrated. After, this subsection will present the working principles of

Table 1

Comparison between the previous work and proposed framework.

Ref	Algorithms/Models	Techniques	Optimizer	Accuracy
[22]	SVM, CNN, Alex-Net Model	ML + DL	–	98%
[41]	CNN, Tversky loss function	DL	Binary cross-entropy optimizer, Adaptive movement estimation (Adam)	99%
[40]	CNN, VGG, Alexnet, and GoogleNet	DL	–	99.13%
[39]	CNN	DL	–	98.53%
[38]	CNN, cross-entropy loss function	DL	Adaptive movement estimation (Adam)	95.54%
[37]	Feed forward neural network (FFNN), ANN, SVM, AlexNet, GoogleNet, ResNet-18, CNN	ML + DL	Adaptive movement estimation (Adam)	100%
[36]	Mask R-CNN	DL	Stochastic gradient descent (SGD) with a momentum coefficient	83.72
[34]	CNN, ResNet-50, VGG-16, RF, LR, SVM, KNN, SGD, MLP	ML + DL	Adaptive movement estimation (Adam)	84.62%
Proposed Method	CNN, VGG19, ResNet50, InceptionV3, Xception, Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Naive Bayes (NB), Extreme Gradient Boosting (XGB), K-Nearest Neighbor (KNN), Logistic Regression (LR)	ML + DL	Particle Swarm Optimization (PSO), Cat Swarm Optimization (CSO)	99.84%

Particle Swarm Optimization (PSO) and Cat Swarm Optimization (CSO) along with the algorithmic annotations and interpretations. This subsection will present the feature vectors and the working mechanism of conventional classifiers.

3.2.1. Feature Extraction with Pre-trained CNN

In this research, four conventional pre-trained Convolutional Neural Network (CNN) models have been applied to extract the features from the single images. Fig. 2 shows the corresponding diagram of respective pipeline of feature extraction mechanism of this study. In this figure, initially, the system extracts the images from the dataset to feed the pre-trained models to extract the images. Four traditional pre-trained CNN architecture namely, VGG19, ResNet50, InceptionV3 and Xception have been applied sequentially to extract feature vectors from the images. After extracting the features, seven conventional classifiers have been implemented to classify the images. Due to work with best features, two nature inspired algorithms have been implemented and performed the feature selection method. After that the system measures the performance based on their significant classes. The whole pipeline follows the Algorithm 01 to extract the feature vectors from a particular image.

Algorithm 01

Working mechanism of proposed pipeline to extract feature vectors

Input: 2D Images
Output: Feature Vectors

Initialization :

1. $n = 2N-1$, Where $N = 1, 2, 3, 4 \dots \dots \dots$
2. $X \leftarrow \text{Input Image}$
3. $Y_n \leftarrow \text{Apply the median filter on the input image } X \text{ using the kernel size } n \times n$
4. $F_v \leftarrow \text{Respective Feature Vector}$

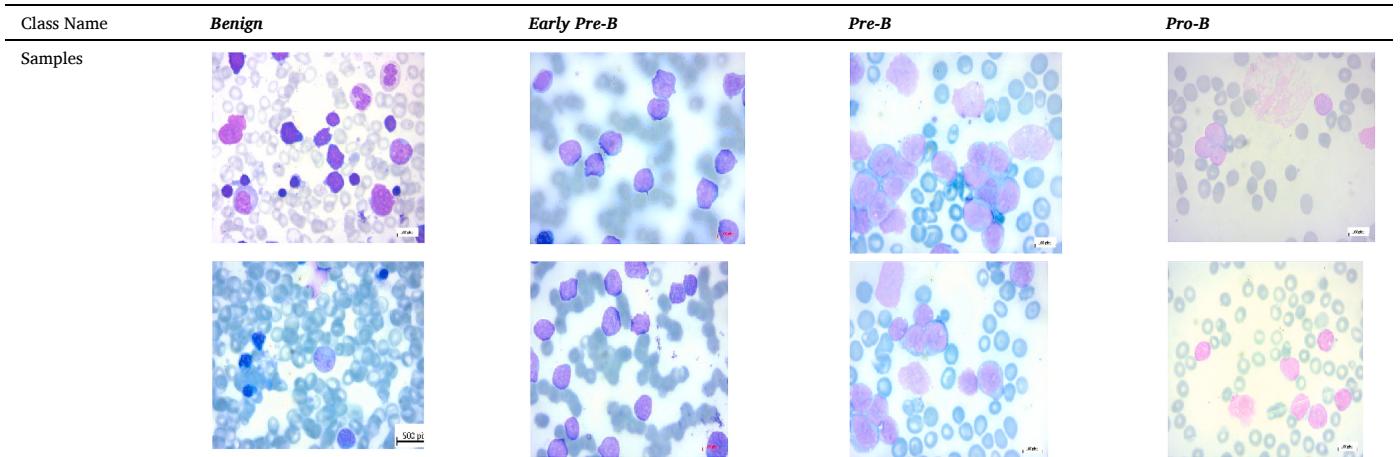
Start :

1. **for each N:**
2. Find Y_n
3. Use (X, Y_n) to get $F_n | F_n \{P_0, P_1, \dots, P_{14}\}$
4. $F_v \leftarrow F_n$
5. **End for**
6. Show F_v

End :

Table 2

Sample images of each classes of the dataset.



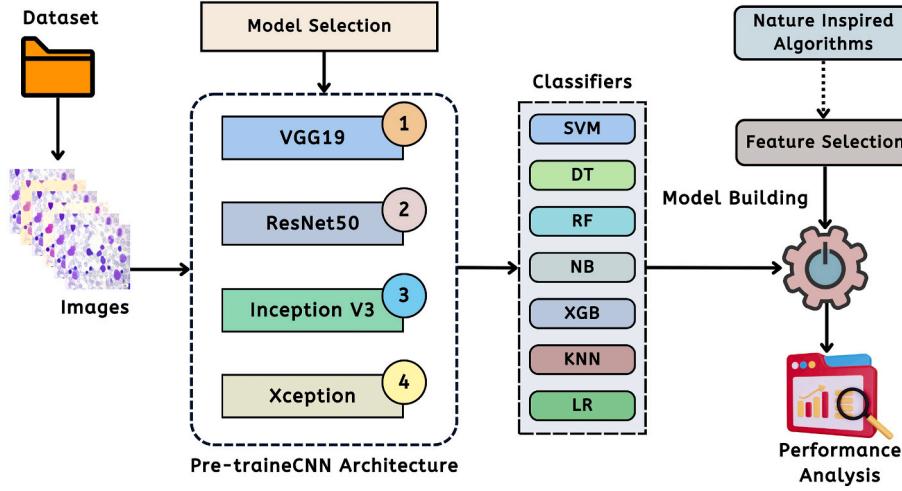


Fig. 2. Pipeline of the proposed Deep Learning method for feature extraction.

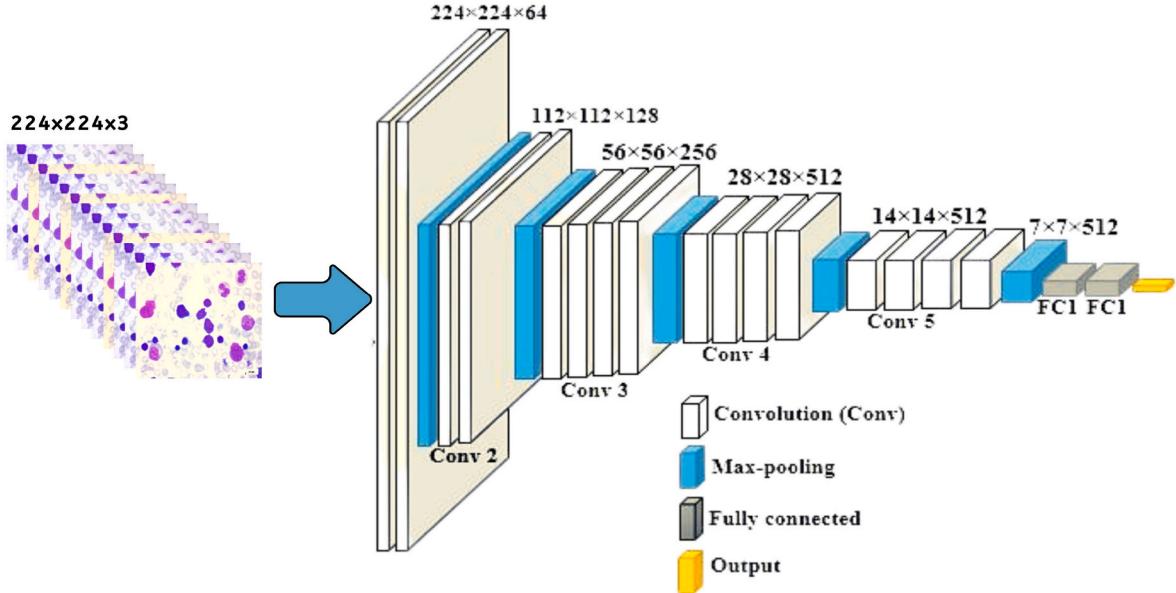


Fig. 3. The architecture of VGG19.

The Algorithm 01 shows efficient view of feature extraction from a particular image. In this procedure, the system first initialized the dataset based on the number of images. Where X represent the input image and Y is the output after applying the image filtering and resizing. Then, the pseudocode represents a loop structure to enumerate the feature vectors.

Fig. 3 shows the architecture of VGG19 pre-trained CNN architecture that has been included in the pipeline in Fig. 2. VGG19 is fine-tuned with some of the layers to ignore the overfitting issues for a small dataset. In this architecture, the pre-trained model is comprised of a series of Convolutional Layers (CL) and single or multiple Fully Connected (FC) layers. The model is identically classified into two interconnected parts. The first part denotes the feature extraction part from the input layer to the last max-pooling layer. The second part represents the residual network of the model, which is mainly responsible for the classification. The proposed solution mainly focuses the VGG19 model on feature extraction; thus, the classification part is declined in this study. The proposed model with VGG19 accepts the Blood cell images of $224 \times 224 \times 3$ and assembles 4096 features from the out of the last layer of feature extraction part for each image [48].

The research also utilizes the pre-trained ResNet50 CNN architecture [48] in the proposed pipeline to extract the features from a particular image. ResNet50 identically consists of 50 layers with having approximately 2M parameters. The ResNet50 architecture has several parts to constitute the model. The first part contains 64 kernels with a max-pooling layer, convolution layer and fully connected layer. The augmentation layer permits dilapidation problems and eliminates the disappearing problem. Besides, the skip connection act like a super-pathway. The research predominantly includes the ResNet50 for feature extraction and excludes the classifier part. The proposed model with ResNet50 CNN architecture accepts the Blood cell images of $224 \times 224 \times 3$ and assembles 2048 features from the out of the last layer of feature extraction part for each image. Fig. 4 shows the architecture of the corresponding ResNet50 model. After that, the research utilizes the Inception V3 architecture. InceptionV3 is known as GoogleNet, and the model itself is a pre-trained network model. The Inception model has 22 layers having 5 M parameters with a filter size of 1×1 , 3×3 and 5×5 to extract features at various scales through max pooling. After introducing InceptionV3, the 5×5 convolutional filters are replaced with two 3×3 filters to reduce computation discarding the performance of

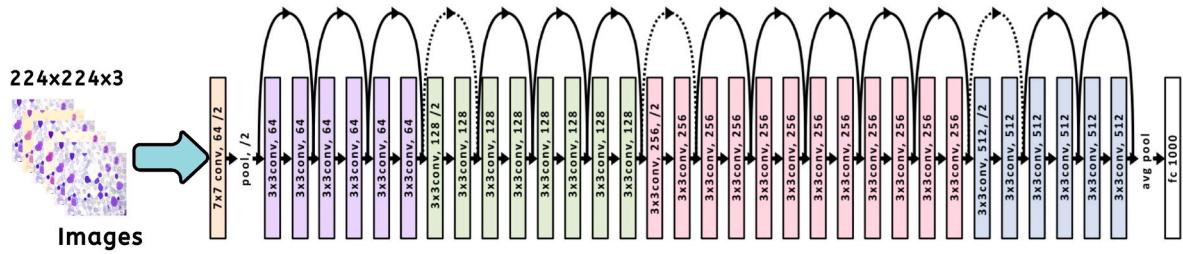


Fig. 4. The architecture of ResNet50.

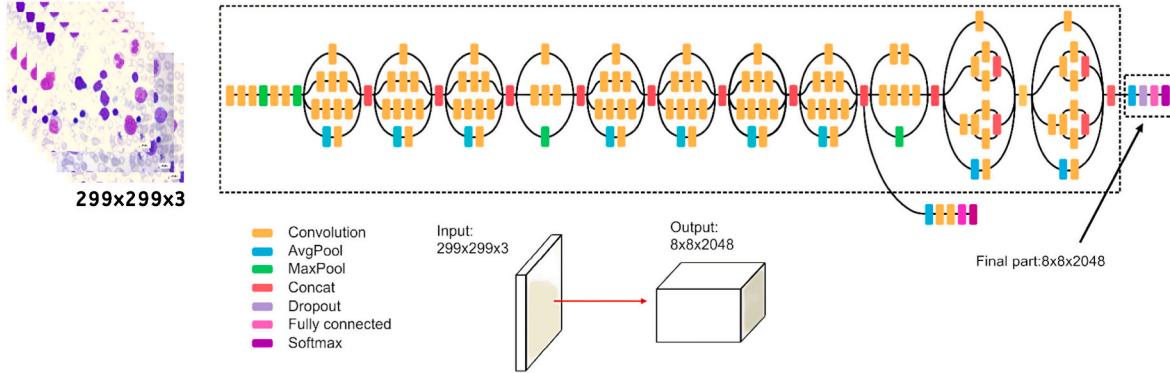


Fig. 5. The architecture of Inception V3.

networks. The InceptionV3 consists of 48 layers and fine-tuned structure to avoid overfitting. In the proposed pipeline, InceptionV3 model takes the Blood cell images of $299 \times 299 \times 3$ and extracts 2048 features from the out of the last layer of feature extraction part for each image [48]. Fig. 5 illustrates the respective diagram of proposed InceptionV3 architecture.

3.2.2. Feature Selection Models

This subsection provides the working principles of the feature selection models in our proposed system. In the research, the model mainly develop with three significant feature selection models namely, Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and Support Vector Classifier (SVC) feature selector. PCA and LDA [49,50] are favorable algorithms to diminish the feature vectors [50]. PCA is an unsupervised learning algorithm and mainly focuses on enhancing the variation in a particular dataset. On the contrary, LDA and SVC feature sector are the supervised learning method that focuses on a feature vector subspace that improves the separability between the groups. In the research, PCA, LDA, and utilize with some sort of following mathematical expression.

PCA can effectively work with the construction of covariance matrix. A Symmetric $d \times d$ -dimensional covariance is prerequisite to build PCA where d denotes the number of dimensions in a certain dataset and holds the pairwise covariance's between diverse noteworthy features [50]. For instance, assume X_j and X_k denotes the two features of the targeted population. Then, the covariance can be calculated by the following Equation (1).

$$\sigma_{jk} = \frac{1}{n} \sum_{i=1}^n (x_j^{(i)} - \mu_j)(x_k^{(i)} - \mu_k) \quad (1)$$

But, LDA works with the five interconnected steps. Initially, LDA calculates the respective d -dimensions of the mean vectors. Then LDA creates scatter matrices and computes eigenvectors. Then, LDA sorts the eigen vectors in decreasing order. Then, a matrix multiplication results in the corresponding features reduction or feature selection.

On the other hand, SVC works with the particular problem with

Linear SVC. The main objective of the SVC is to fit the data and return the "best feature" hyperplane that has the ability to categorize the data from a certain dataset [50]. In the proposed model, the SVC feature selector works as In Algorithm 02.

Algorithm 02

Working mechanism of the proposed SVC Feature Selector in feature selection

Input: Images Features
Output: Best Feature Vectors
Initialization :

1. $X = N-1$, Where $N = \text{Number of features of a particular CNN model}$.
2. $Y \leftarrow \text{No. of classes}$
3. $X_n \leftarrow \text{No. of training data}$
4. $Y_n \leftarrow \text{No. of testing data}$
5. $F_v \leftarrow \text{Respective Best Feature Vectors}$
6. $S_v \leftarrow \text{No. of Selected Best Feature Vectors}$

Start :

1. $\text{feature} = \text{SVMFeatureSelection}(X_n, Y_n)$
2. $\text{if } (\text{No. of feature} > 0.5)$
3. $S_v = \text{features}$
4. $F_v = \text{add}(\text{all the } S_v)$
6. **End if**
5. **Show** F_v

End :

Fig. 6 depicts the block diagram of the feature selection method used by these feature selectors to identify the "Best Feature." In this figure, the system obtains test data, converts the data into feature vectors, and then feeds these vectors to the algorithms in order to find the most appropriate features for dealing with conventional classifiers. Lastly, the model computes the level of precision with the assistance of ML classifiers.

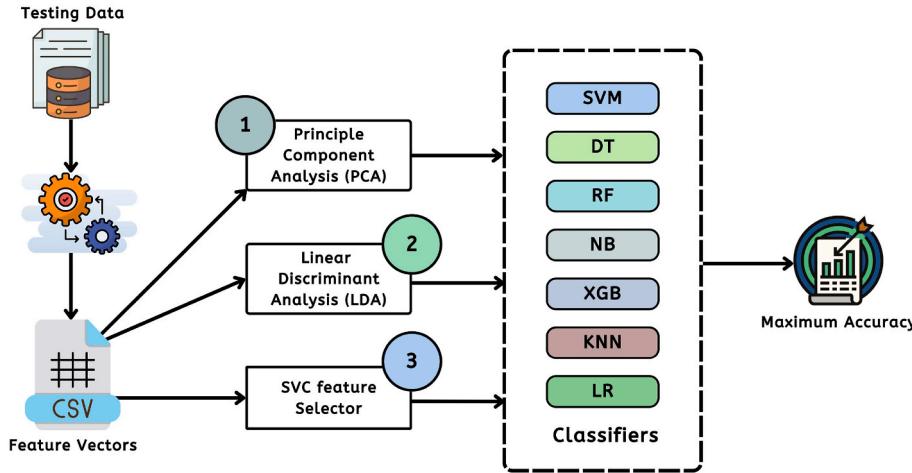


Fig. 6. The proposed architecture of the feature selector algorithms.

3.2.3. Working Principles of Proposed Particle Swarm Optimization (PSO) and Cat Swarm Optimization (CSO)

This subsection represents the two significant inspired algorithms for the selection of the best feature that can optimize the level of accuracy. In the research pipeline, the research includes PSO and CSO algorithm to find best features also work with the minimal number of best fittest features.

Particle Swarm Optimization (PSO) is the most prominent and influential meta-heuristic-based optimization model. This algorithm is mainly inspired by nature's significant behaviors, specifically in fish and bird schooling. The algorithm is called a heuristic solution because the model always tends toward the global optimal. In nature, any of the birds in a particular swarm has minimal observable proximity to the observer. But, more than one bird out of these birds allows all the swarm birds to be conscious of the more excellent surface of a fitness function [51].

Assume, P is the number of particles and i denotes the position of each iteration t as $X^i(t)$. Let's consider the $X_i(t)$ as a coordinate like $X^i(t) = (x^i(t), y^i(t))$. Also, consider the velocity of the each particle will be denoted as $V^i(t) = (v_x^i(t), v_y^i(t))$. Then the position of the particle will be like Equations (2)–(4).

$$X^i(t+1) = X^i(t) + V^i(t+1) \quad (2)$$

$$x^i(t+1) = x^i(t) + v_x^i(t+1) \quad (3)$$

$$y^i(t+1) = y^i(t) + v_y^i(t+1) \quad (4)$$

Then, the following equation can be written like Equation (5)

$$V^i(t+1) = wV^i(t) + c_1r_1(pbest^i - X^i(t)) + c_2r_2(gbest - X^i(t)) \quad (5)$$

In Equation (5), r_1, r_2 represents the number in a range $[0, 1]$. Also, w, c_1, c_2 are the one of the significant parameters of PSO algorithm and $pbest^i$ is the position that gives calculated best fit $F(x)$ value for the particle i . Finally, $gbest$ is the measured value explored by all the particles in the swarm. Algorithm 03 shows the corresponding steps of the algorithm and Fig. 7 show the summary of the algorithm with the flow chart.

Algorithm 03

PSO algorithm for best particle calculation

Input: Set of random particles
Output: Best fitted particles
Initialization :

1. Initialize random particle $P = \{P^1, P^2, P^3, \dots, P^i\}$.
2. $t \leftarrow$ Time
3. $c_1 \leftarrow$ cognitive factor
4. $c_2 \leftarrow$ Social factors
5. $r_1, r_2 \leftarrow$ the value in the interval $[0, 1]$
6. $w \leftarrow$ inertia weight
7. $pbest^i \leftarrow$ best position of P^i
8. $gbest \leftarrow$ global best position of the particle

Start :

1. **while** (best solution)
2. **for each** P^i
3. Update the velocity $V^i(t+1) = wV^i(t) + c_1r_1(pbest^i - X^i(t)) + c_2r_2(gbest - X^i(t))$
4. Update position $X^i(t+1) = X^i(t) + V^i(t+1)$
5. Use objective function f to evaluate the fitness value of P^i
6. Update $pbest^i(t) | pbest^i(t+1) = \begin{cases} pbest^i(t) & \text{if } f(pbest^i(t)) \leq f(p_i(t+1)) \\ p_i(t+1) & \text{if } f(pbest^i(t)) > f(p_i(t+1)) \end{cases}$
7. **End for**
- End while**

End :

Chu and Tsai first developed cat swarm optimization (CSO) in 2007, and it functions in both Seeking Mode (SM) and Tracking Mode (TM). However, in TM, cats move to their next location with some velocities, showing how cats pursue their target. In SM, cats do not move and remain in a specific position and feel for the next best move. Fig. 8 shows the working flow chart of the CSO algorithm [52].

To work with the CSO, some of the parameters need to be noticed like Number of Dimension of Variations (CDC), Mutative ration for the selected dimensions (SRD), Number of Duplicate Cats (SMP), and Passed position as one of the candidates (SPC). The position change of the duplicate cats can be enumerated with following Equation (6):

$$X_{mn} = 1 + SRD \times R \times X_m \quad (6)$$

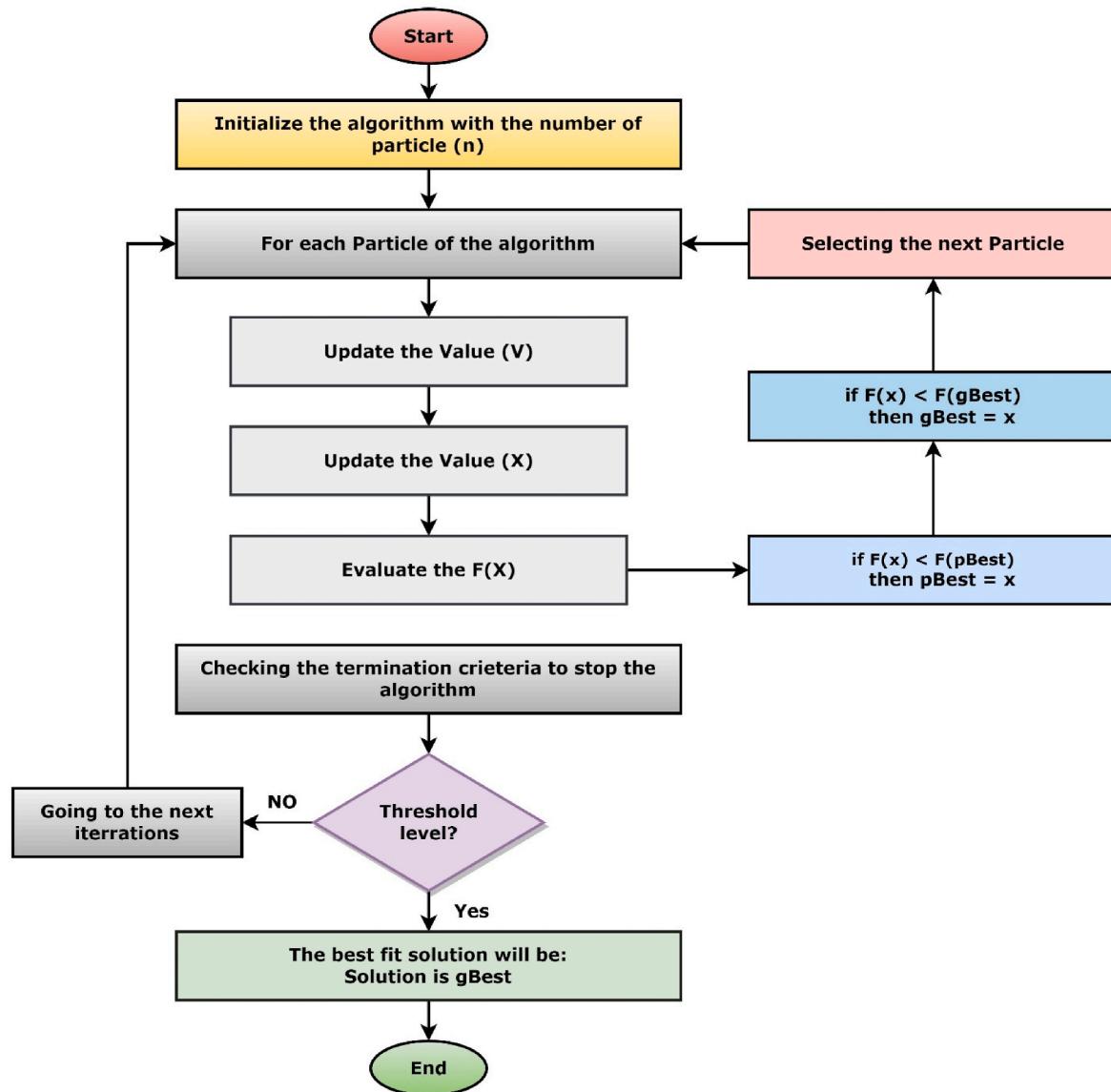


Fig. 7. The flowchart of PSO algorithm.

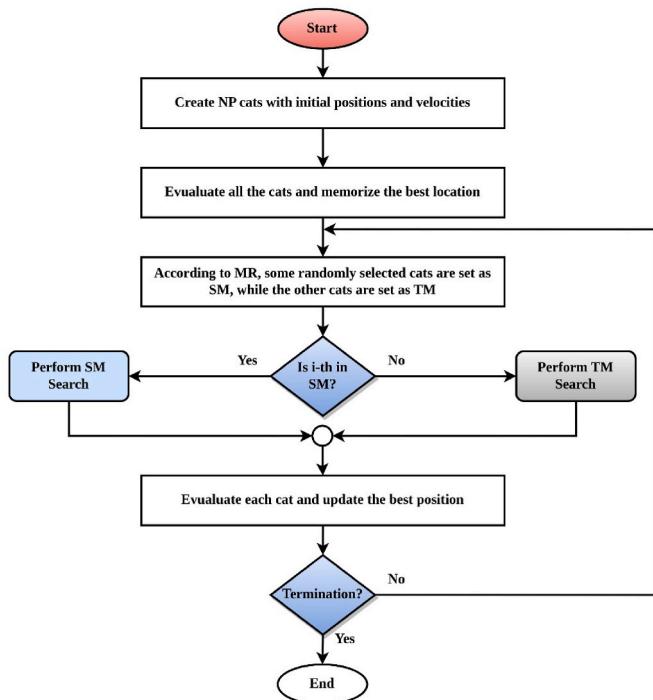


Fig. 8. The flowchart of CSO algorithmCat Swarm Optimization (CSO).

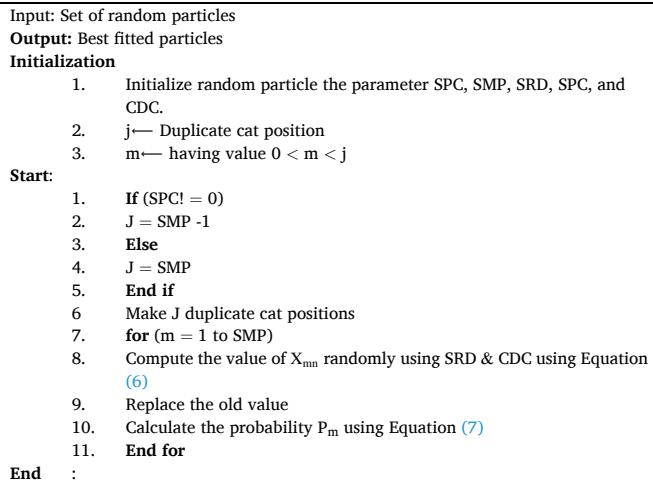
Where, X_{mn} = The current Position $R = [0, 1]$.

The algorithm will set the calculated probability of each candidate points if the Fitness Function (FS) is considered with Equation (7). Other the algorithm will set the value to 1.

$$P_m = \frac{|FS_m - FS_{min}|}{FS_{max} - FS_{min}} \quad (7)$$

Where, $0 < m < j$.Algorithm 04

CSO algorithm Seeking Mode (SM)



The algorithm mainly works with two type of search strategies such as SM and TM. The working procedures of SM is depicted in Algorithm 04 and The TM is illustrated in the **Algorithm 05**. In TM, the velocity of each individual cat is represented by the following Equation (8).

$$v_{l,i} = v_{l,i} + r \times q \times X_{best,i} - X_{l,i} \quad (8)$$

Where, $i = 1, 2, 3, 4 \dots M$, here M = cat number. r = random value in a

range of 0–1. $X_{best,i}$ = i-th best position of the cat. $X_{l,i}$ = The current Position $R = [0, 1]$.

Then the update position of each cat will represented by:

$$X_{l,i} = X_{l,i} + v_{l,i} \quad (9)$$

Algorithm 05. CSO algorithm in Tracking Mode (TM)

Input: Set of random particles
Output: Best fitted particles

Initialization	:
1.	Initialize random particle.
2.	$i \leftarrow$ Cat number $M = 1, 2, 3, 4 \dots$
3.	$r \leftarrow$ the value in a range 0–1
4.	$Mmax \leftarrow$ Max Speed
5.	$v =$ velocity of the cat
Start	:
1.	Update the current velocity by using Equation (8)
2.	for ($v \leq r$)
4.	Execute the subsequent part with v
5.	Else
6.	$v = Mmax$
7.	End if
End	:
8.	Update the position of the each cat by using Equation (9)

3.3. Classification with the conventional classifiers

In this study, ML methods are applied to the features data extracted from cases of acute lymphoblastic leukemia. To achieve this, the study makes use of established classifiers found in current ML practice. A total of seven different classifiers, including Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Naive Bayes (NB), Extreme Gradient Boosting (XGB), K-Nearest Neighbor (KNN), Logistic Regression (LR) use to construct the final model [53].

3.3.1. Support Vector Machine (SVM)

SVM is a classifier and regression algorithm that can classify any object based on a given set of data. It generates the optimal decision boundary, which divides n-dimensional spaces into distinct classes, thereby facilitating future data insertion. The classification was constructed using the kernel trick mechanism of SVM, and its corresponding equation is shown in Equation (2) for the kernel trick of SVM. For a two-dimensional non-linearly separable dataset, kernel trick transforms the two-dimensional data to a higher dimension, such as three, four, or ten dimensions. This method is known as the kernel trick.

$$\text{Kernel trick. } k(x_i, x_j) = x_i \cdot x_j \quad (10)$$

3.3.2. Random Forest (RF)

The machine learning technique known as Random Forest is used for both classification and regression. An ensemble classifier is another name for this algorithm, which incorporates several decision tree models. A classifier like this one takes the results of many decision trees applied to various subsets of a dataset. It averages them in order to improve the accuracy of the dataset's overall predictions. A more significant number of trees can be accurately assessed using this method.

3.3.3. Decision Tree (DT)

An example of a supervised learning algorithm is the decision tree (DT). A tree-based data structure represents the algorithm. The Decision Tree is a graphical representation of a collection of rules for making decisions based on a dataset's features. The tree's internal nodes reflect the features of the dataset, while the tree's branches represent the rules. The DT is a helpful tool for visually representing any Boolean value (true or false). The equation for DT can be seen in Equation (10).

$$\text{Information Gain}, S = -P\left(\frac{1}{\text{true}}\right)\log_2 P\left(\frac{1}{\text{true}}\right) - P\left(\frac{0}{\text{false}}\right)\log_2 P\left(\frac{0}{\text{false}}\right) \quad (11)$$

3.3.4. Naïve Bayes (NB)

The probability of a hypothesis can be calculated using Bayes' theorem, often known as Bayes' Rule or Bayes' law. Bayes's theorem can be expressed as the following Equation (11). The NB in data classification works with this equation to find the results.

$$P(A | B) = \frac{P(B | A) P(A)}{P(B)} \quad (12)$$

3.3.5. XGBoost Algorithm (XGB)

The powerful XGBoost machine-learning algorithm can aid in data analysis and decision-making. To be precise, XGBoost is a program that uses decision trees that are boosted by a gradient. Several researchers and data scientists around the world have utilized it to fine-tune their machine learning models.

3.3.6. K-Nearest Neighbor (KNN)

The Supervised Learning method ML techniques use K-Nearest Neighbor. This algorithm assumes that the new data and existing data are similar. When placing new data into the class most comparable to the available categories. The Manhattan equation is utilized for distance calculation in this approach for KNN. Equation (12) gives the Manhattan distance formula for the KNN classifier.

Manhattan distance

$$\sum_{i=1}^k |x_i - y_i| \quad (13)$$

3.3.7. Logistic Regression (LR)

A supervised classification approach is logistic regression. This is used to compute the target variable's possibility. Because the target variable is binary, there are only two potential classes (1/success/yes) or (0/failure/no). It is classified into three types: binary or binomial, multivariable, and ordinal. Hence equation for LR can be represented as Equation (13).

$$y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_3 + \dots + b_n x_n \quad (14)$$

4. Results and discussion

This section provides the experimental data analysis with Machine Learning (ML) based models. Firstly, this section presents the performance analysis with only pre-trained Convolutional Neural Network (CNN) architecture along with ML based classifiers. Secondly, this section illustrates experimental data analysis with the nature inspired algorithms as well as the feature selection algorithms like Linear Discriminant Analysis (LDA), Principal Component Analysis (PCA) and Support Vector Feature Selector. Finally, this section provides the performance analysis with different classifiers.

To evaluate the efficacy of the proposed model a set of performance evaluation matrices have been enumerated such as Accuracy (A), Precision (P), Recall (R) and F1-Score. Table 3 shows the descriptions of the performance evaluation matrices with some mathematical annotations. On the other hand, Fig. 9 shows the orientation of the confusion matrix for Blood cancer cell classification.

4.1. Performance analysis without feature selection and PSO

This subsection provides the results analysis with ML based classifiers only with the pre-trained CNN architecture. This study ran all of the programs on the Google Colab, which has 53 GB of RAM and a dedicated Graphics Processor Unit (GPU). This setup's subscription was 'pro subscription'. The research initially retrieved the important elements from a

Table 3
Description of performance evaluation matrices.

Metrics	Description
Accuracy (A)	Displays the overall right prediction percentage. $A = \frac{TP + TN}{TP + TN + FP + FN} \times 100$
Precision (P)	Describes as a way to assess a model's quality. $P = \frac{TP}{TP + FP} \times 100$
Recall (R)	Defines as a measurement of model quantity. $R = \frac{TP}{TP + FN} \times 100$
F1-score (F1)	Demonstrates how reliable and accurate a model is. $F1 = 2 \times \frac{R \times P}{R + P} \times 100$
NCM	Gives a tabular representation of the rates of correct and incorrect detection for various classifications.

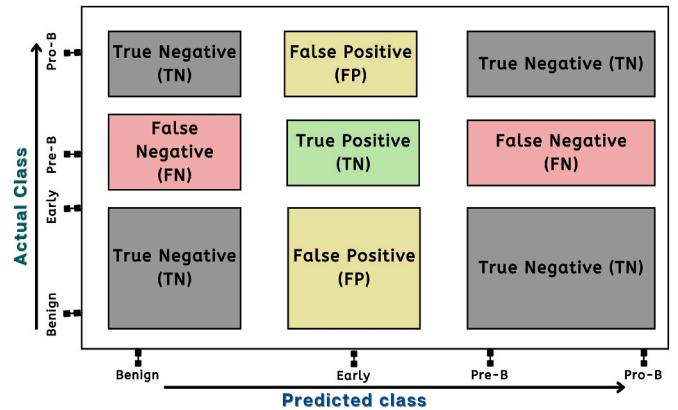


Fig. 9. The structure of the confusion matrix of the results.

certain image. The gathered data was then fed into ML-based models to assess performance. We divided the dataset into 80% training data and 20% testing data for our study. Table 4 shows a summary of the models after extracting the features from each individual images before performing the classification.

Initially, we have chosen the CNN models with Support Vector Machine (SVM) classifier. To find the better result we have also combined the features to perform the feature fusion of InceptionV3 and Xception models. Table 5 shows the summary of the results achieved from the pertained CNN model and SVM classifiers. We have achieved the highest accuracy of 98.43% with ResNet50 architecture where the closest

Table 4
The overall summary of memory consumption of each model and time.

Model Name	Memory Usage For Extracted Features (MB)	No. of Features	Time Required to complete the extraction (Apo. Hour)
VGG19	63.2260	4096	3.50
ResNet50	47.1710	2048	2.10
InceptionV3	61.6590	4048	2.30
InceptionV3 + Xception	91.2230	4096	4.00

Table 5
The overall performance of different Pre-trained CNN models.

CNN Model	A	P	R	F1
VGG19	97.18	96.64	96.87	96.75
ResNet50	98.43	98.28	98.32	98.30
Inception V3	79.15	77.82	75.68	75.76
Inception + Xception	87.62	87.71	84.27	85.06

Table 6

The fold-wise performance measurement of the different Pre-trained CNN models.

CNN Model	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
VGG19	95.61	96.08	96.07	96.08	96.23
ResNet50	97.95	97.80	97.64	97.95	97.80
Inception V3	77.89	80.41	80.72	81.03	80.87
Inception + Xception	87.77	88.71	90.12	90.12	90.28

accuracy was 97.18% with VGG19 model. We also ran 5-fold cross validation on each model to see whether it could detect overfitting concerns. [Table 6](#) shows the summary of cross-validation scores.

On the contrary, [Fig. 10](#) shows a comparison on accuracy of different pre-trained CNN architectures. This figure clearly indicates that ResNet50 provides the nearly the optimum results with the SVM classifiers where InceptionV3 gives very poor performances than the feature fusion and the other models.

Besides, we have performed the seven conventional ML based classifier such as Support Vector Machine (SVM), Random Forest (RF), Decision Tree (DT), Naive Bayes (NB), Extreme Gradient Boosting (XGB), K-Nearest Neighbor (KNN), and Logistic Regression (LR). We have interpreted the results the sequentially based on the performance matrices. In these experiments, we have found the satisfactory result from VGG19 with LR classifier but the accuracy was not very convincing. Because, we have tracked highest accuracy of 99.53% accuracy with ResNet50 architecture with LR classifier. The corresponding results are shown in [Tables 7–10](#). [Table 11](#) shows the summary of the result analysis with InceptionV3 and Xception models. The all the experiments are performed without explicitly using the nature inspired algorithms.

4.2. Performance analysis with feature selection and PSO and CSO

This subsection presents the experiment data analysis with nature inspired algorithm like Particle Swarm Optimization (PSO) and Cat Swarm Optimization (CSO). Initially, we have applied the PSO on the extracted features with the pre-trained CNN architectures. After we have applied the CSO.

1) Data Analysis with PSO

The summary of data analysis with PSO are illustrated in [Table 10](#). This table, clearly shows the research achieved the highest accuracy of 99.68% in Acute Lymphoblastic Leukemia (All) while working with ResNet50 architecture and SVM classifier along with the PSO algorithm. We have also enumerated the results with the cross validation scores along with the learning curves and confusion matrix and Area Under

Table 7

The overall performance of different classifier with VGG19.

Classifiers	A	P	R	F1
SVC	94.98	95.17	93.73	94.35
RF	93.57	93.74	91.89	92.64
DT	84.17	82.78	83.52	83.09
NB	78.21	78.51	78.70	78.08
XGB	93.42	93.03	92.57	92.78
KNC	93.89	94.14	92.40	93.12
LR	95.61	95.44	94.98	95.20

Table 8

The fold-wise performance of different classifiers with VGG19.

Classifiers	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
SVC	89.91	90.59	91.22	91.22	91.53
RF	89.65	89.65	89.81	89.96	89.96
DT	77.27	78.84	79.62	80.25	78.37
NB	75.86	76.95	76.17	76.01	70.00
XGB	89.18	89.18	89.65	90.59	89.65
KNC	89.49	91.22	90.59	90.27	90.74
LR	92.16	93.26	93.10	93.57	92.94

Table 9

The overall performance of different classifier with ResNet50.

Classifiers	A	P	R	F1
SVC	98.43	98.28	98.32	98.30
RF	96.86	97.04	96.57	96.78
DT	91.68	90.95	91.40	91.15
NB	90.25	90.58	90.32	90.23
XGB	98.27	98.15	98.31	98.23
KNC	98.27	98.46	97.49	97.93
LR	99.53	99.39	99.57	99.48

Table 10

The fold-wise performance of different classifiers with ResNet50.

Classifiers	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
SVC	95.44	96.39	95.91	96.23	96.38
RF	94.82	95.60	94.81	95.29	95.60
DT	86.50	86.34	89.01	86.65	86.82
NB	70.79	75.04	78.65	78.90	78.96
XGB	95.60	96.70	97.01	97.48	96.85
KNC	95.60	95.29	95.76	95.61	95.92
LR	97.95	97.80	97.64	97.95	97.80

ROC Curve. [Table 12](#) shows the summary of the corresponding results with the PSO and pre-trained CNN architectures. [Figs. 11–14](#) shows the learning curves, the calculated confusion matrix and AUC-ROC curve.

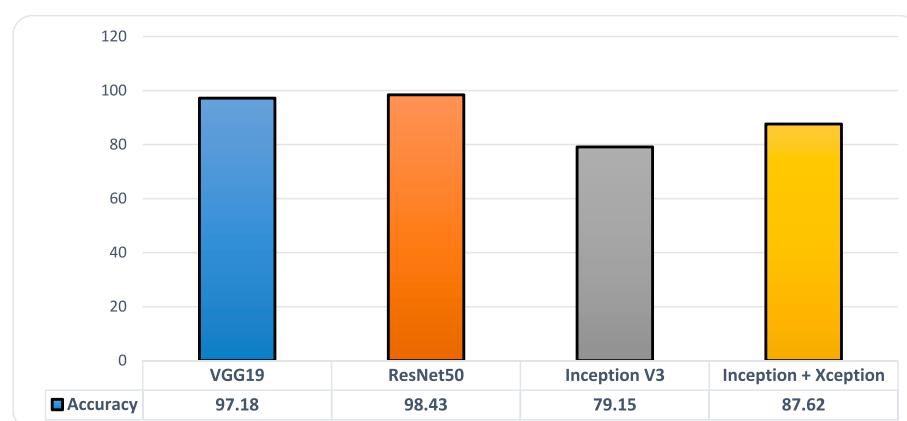
**Fig. 10.** Comparison on Accuracy of different CNN models.

Table 11

The overall performance of different classifier with Inception V3 + Xception.

Classifiers	A	P	R	F1
SVC	87.62	87.71	84.27	85.06
RF	87.93	88.11	84.65	85.46
DT	79.62	78.38	76.80	77.37
NB	72.88	73.17	69.25	69.51
XGB	92.95	92.43	91.29	91.74
KNC	88.40	88.69	85.33	85.83
LR	92.48	92.05	91.66	91.84

Also, we have illustrated the comparison between the achieved results from with PSO and without PSO. **Table 13** shows the respective findings from the different techniques. This table clearly represented that after embedding the PSO the model has achieved the better performance than previous techniques. Also, we have accomplished the better performance in measured performance evaluation matrices.

2) Data Analysis with CSO

Table 12

The overall performance of different Pre-trained CNN models with PSO.

CNN Model	Subset Accuracy	A	P	R	F1	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
VGG19	97.64	97.81	97.69	97.43	97.55	92.63	92.94	92.95	93.57	93.26
ResNet50	99.68	99.84	99.75	99.87	99.81	98.11	98.43	98.11	97.96	98.27
Inception V3	80.87	79.93	77.31	75.98	76.08	71.47	72.57	72.73	73.67	73.66
Inception + Xception	88.55	87.93	87.80	83.65	84.48	76.80	78.83	79.63	80.24	80.24

Table 14 shows an overview of data analysis with CSO. This table plainly demonstrates that the study got the preeminent accuracy of 99.68% in Acute Lymphoblastic Leukemia (All) while using the ResNet50 architecture, SVM classifier, and CSO algorithm. We also listed the findings with the cross validation ratings, learning curves, confusion matrix, and Area Under ROC Curve. **Table 14** summarize the equivalent findings with the CSO and pre-trained CNN architectures. **Figs. 15–18** depict the learning curves, the computed confusion matrix, and the AUC-ROC curve.

We have also shown how the outcomes obtained with and without CSO can be compared. **Table 15** displays the relevant results from the various methods. This chart made it abundantly obvious that the model performed better than earlier methods after the CSO was embedded. Additionally, we have improved our performance in matrices used to evaluate measured performance.

4.3. Performance analysis of different feature selection techniques

This subsection presents the experimental data analysis with several feature selector or feature reduction techniques such as PCA, LDA and

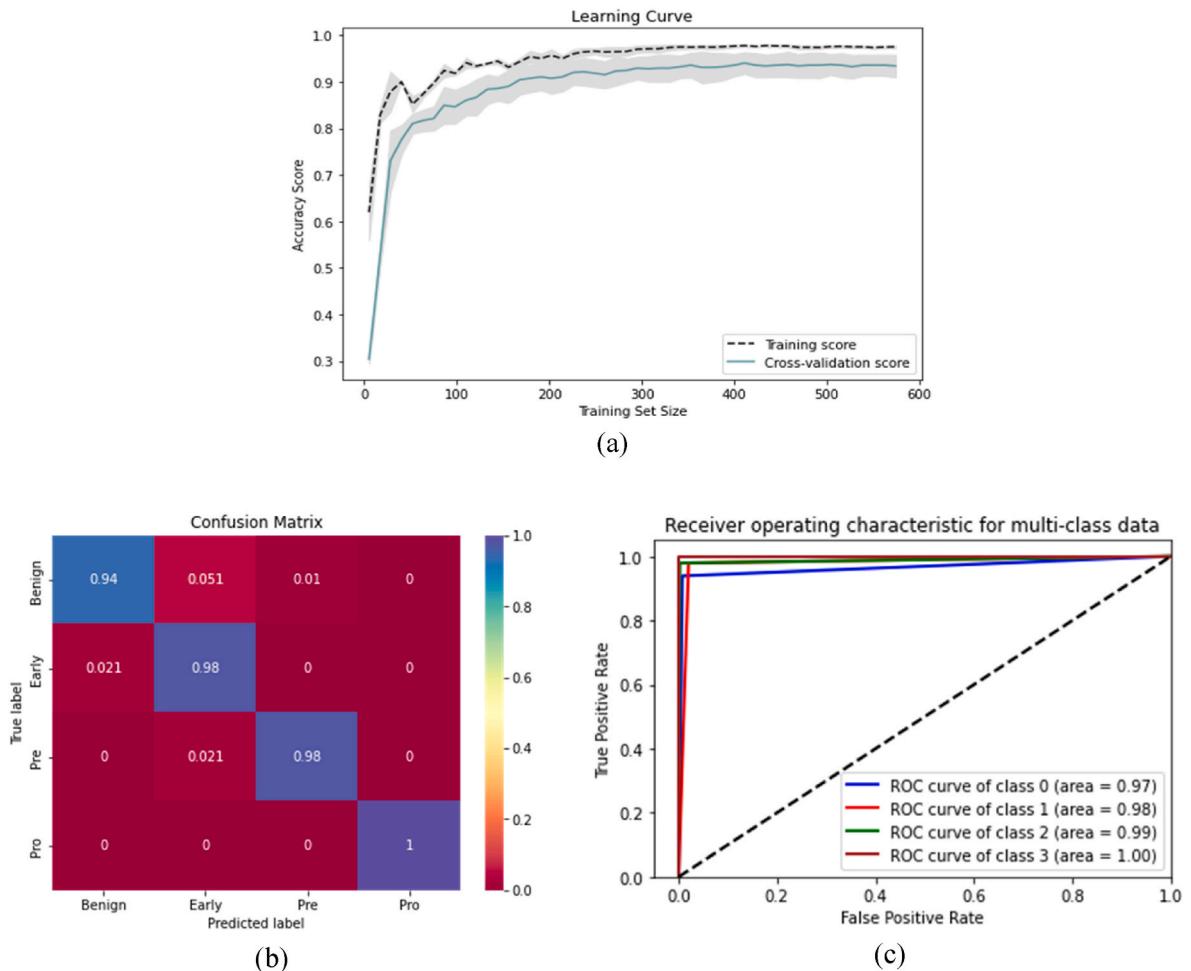


Fig. 11. The performance measurement of VGG19 and PSO with (a) Learning curve (b) Confusion Matrix (c) AUC-ROC curve with PSO.

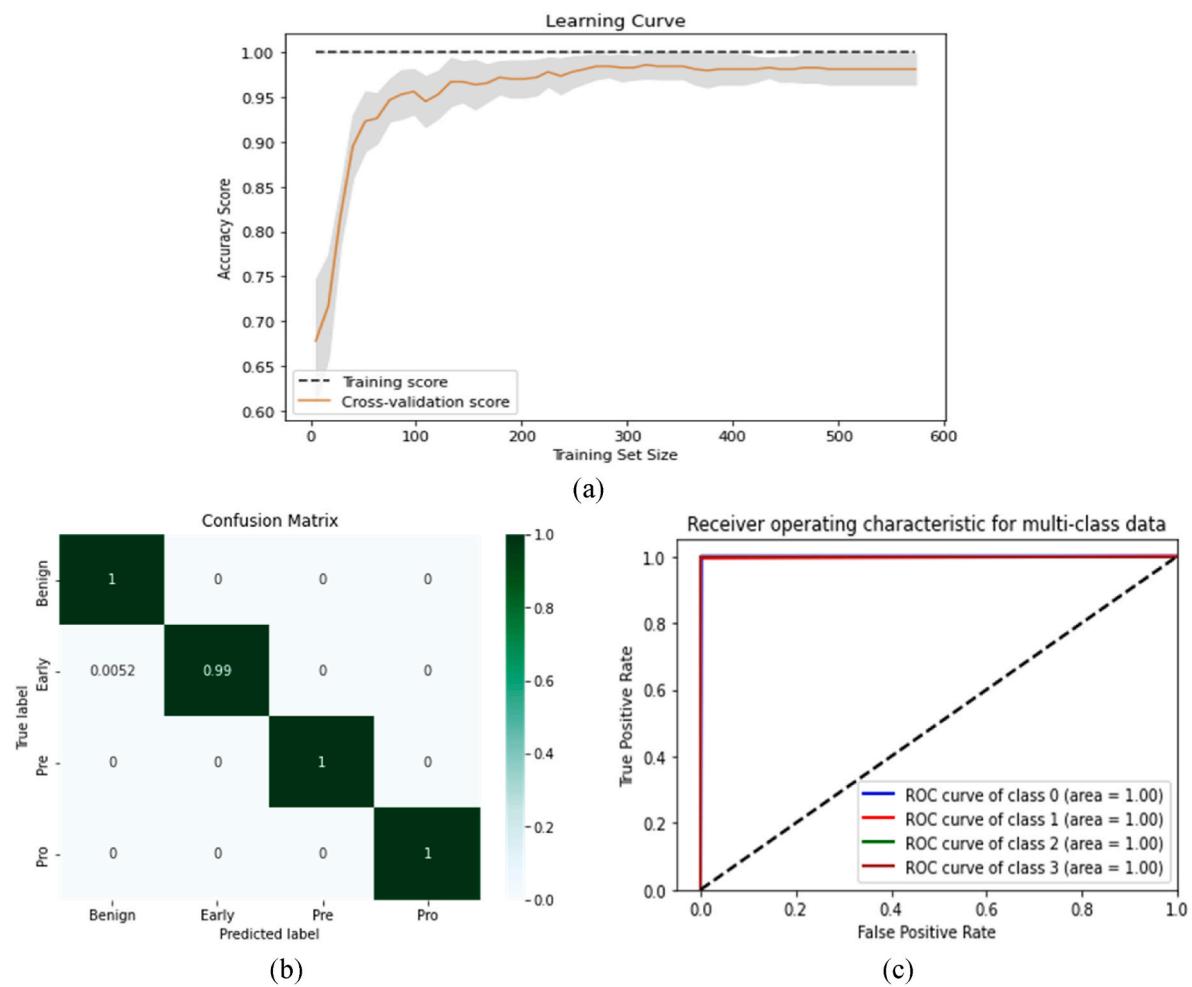


Fig. 12. The performance measurement of ResNet50 and PSO with (a) Learning curve (b) Confusion Matrix (c) AUC-ROC curve.

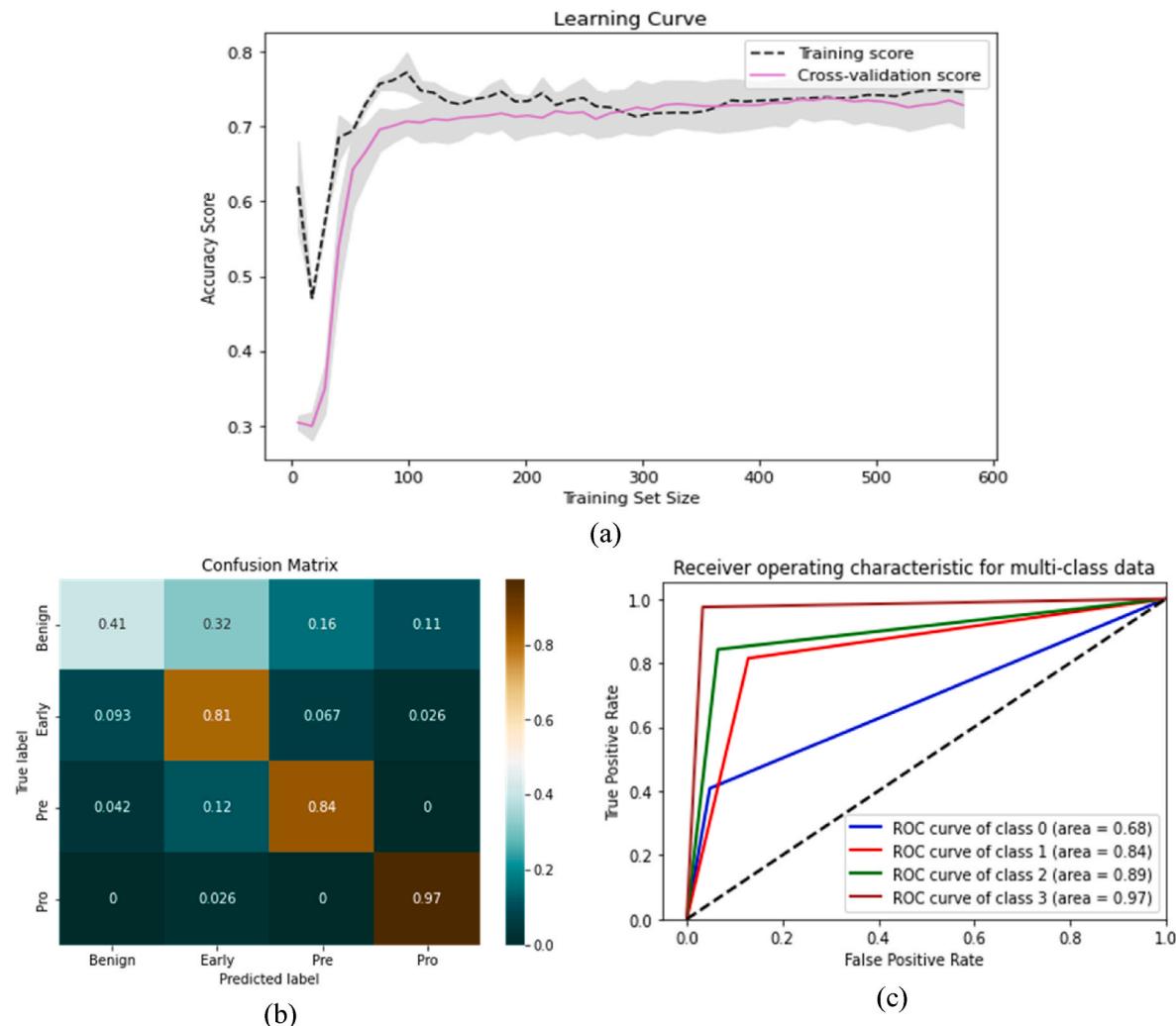


Fig. 13. The performance measurement of Inception V3 and PSO with (a) Learning curve (b) Confusion Matrix (c) AUC-ROC curve.

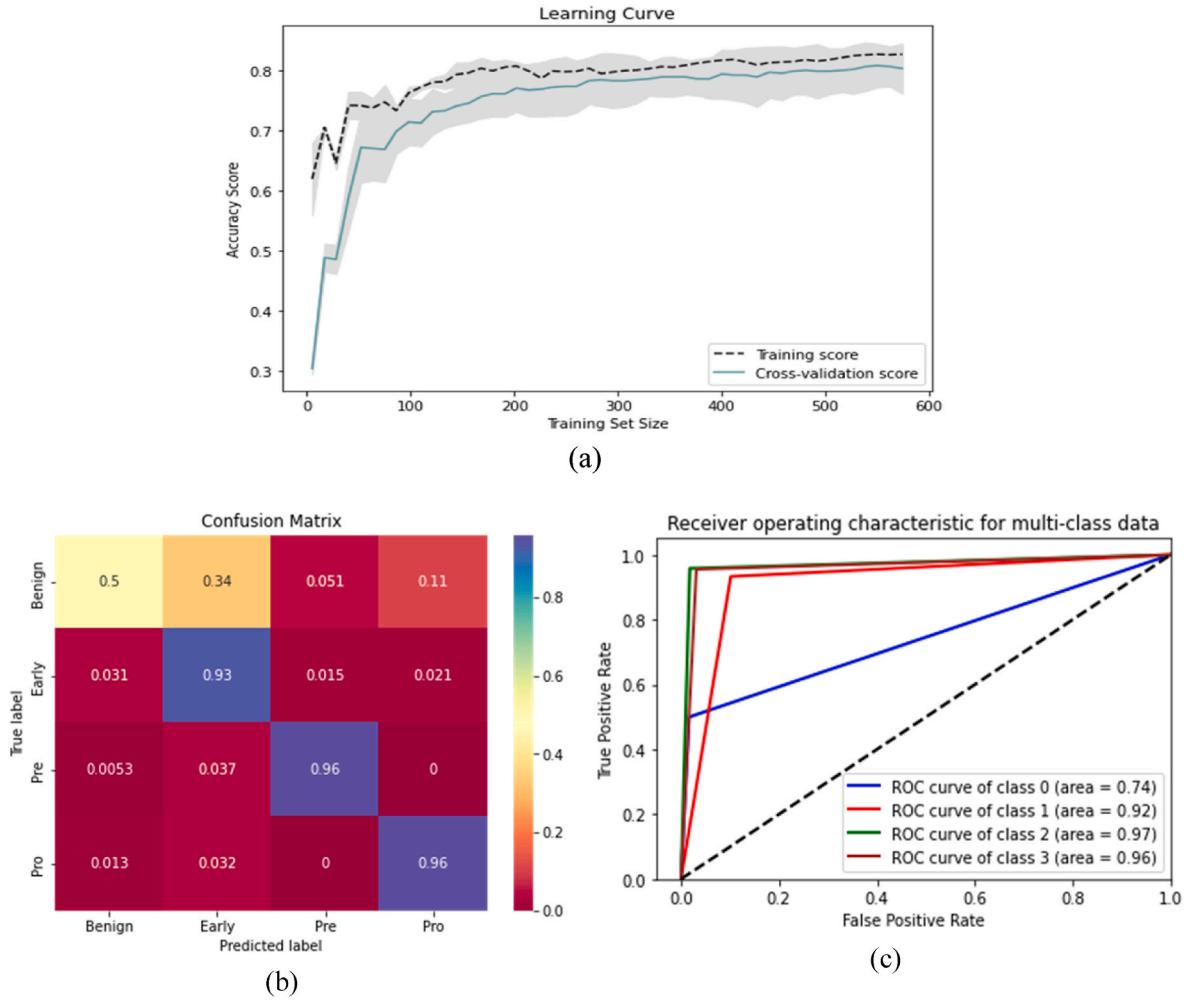


Fig. 14. The performance measurement of Inception V3 + Xception and PSO with (a) Learning curve (b) Confusion Matrix (c) AUC-ROC curve.

Table 13
Comparison of the performance of ResNet50 with and without PSO.

	Techniques	A	P	R	F1
Without PSO	Extracted Feature Vectors + ResNet50 + LR	99.53	99.39	99.57	99.48
With PSO	Extracted Feature Vectors + ResNet50 + PSO + LR	99.84	99.75	99.87	99.81

SVC Feature Selector. We have accomplished the data analysis with PCA with seven conventional ML classifiers on different pre-trained architectures. Initially, we have performed the operations with VGG19 features and with these feature selection techniques. After applying the SVC, we have found massive improvement in accuracy. Table 16 shows the corresponding results achieved from the different feature selector techniques. This table shows the highest accuracy of 98.59% in Acute Lymphoblastic Leukemia (All) with SVC Feature Selectors and LR classifiers.

After that, we have executed the operations with ResNet50 features and with these feature selection techniques. With the SVC, we have found immense improvement in accuracy. Table 17 shows the corresponding results achieved from the different feature selector techniques. This table shows the highest accuracy of 98.84% in Acute Lymphoblastic Leukemia (All) while dealing with ResNet50 architecture along with SVC Feature Selectors and LR classifiers.

Further, we have applied the feature selection techniques on the fusion features from InceptionV3 and Xception models. This orientation provides better results than previous where the highest accuracy in blood cancer cell classification was 94.50%. Table 18 shows the corresponding data for Inception V3+Xception and different feature selectors.

To assess the efficacy of the proposed model we have given a comparison on their performances and interpreted the results accordingly. Table 19 provides a short overview of the performances found from ResNet50 and LR classifier. In this table, we have taken the ResNet50 architecture with LR because this model provides much better results

Table 14
The overall performance of different Pre-trained CNN models with CSO.

CNN Model	Subset Accuracy	A	P	R	F1	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5
VGG19	97.80	97.81	97.69	97.43	97.55	92.63	92.94	92.95	93.57	92.26
ResNet50	99.68	99.84	99.75	99.87	99.81	98.11	98.43	98.11	97.96	98.27
Inception V3	88.55	88.56	87.85	87.51	87.67	81.91	81.50	81.50	82.29	81.44
Inception + Xception	88.87	87.93	87.93	83.65	84.48	76.80	78.83	79.63	80.24	84.24

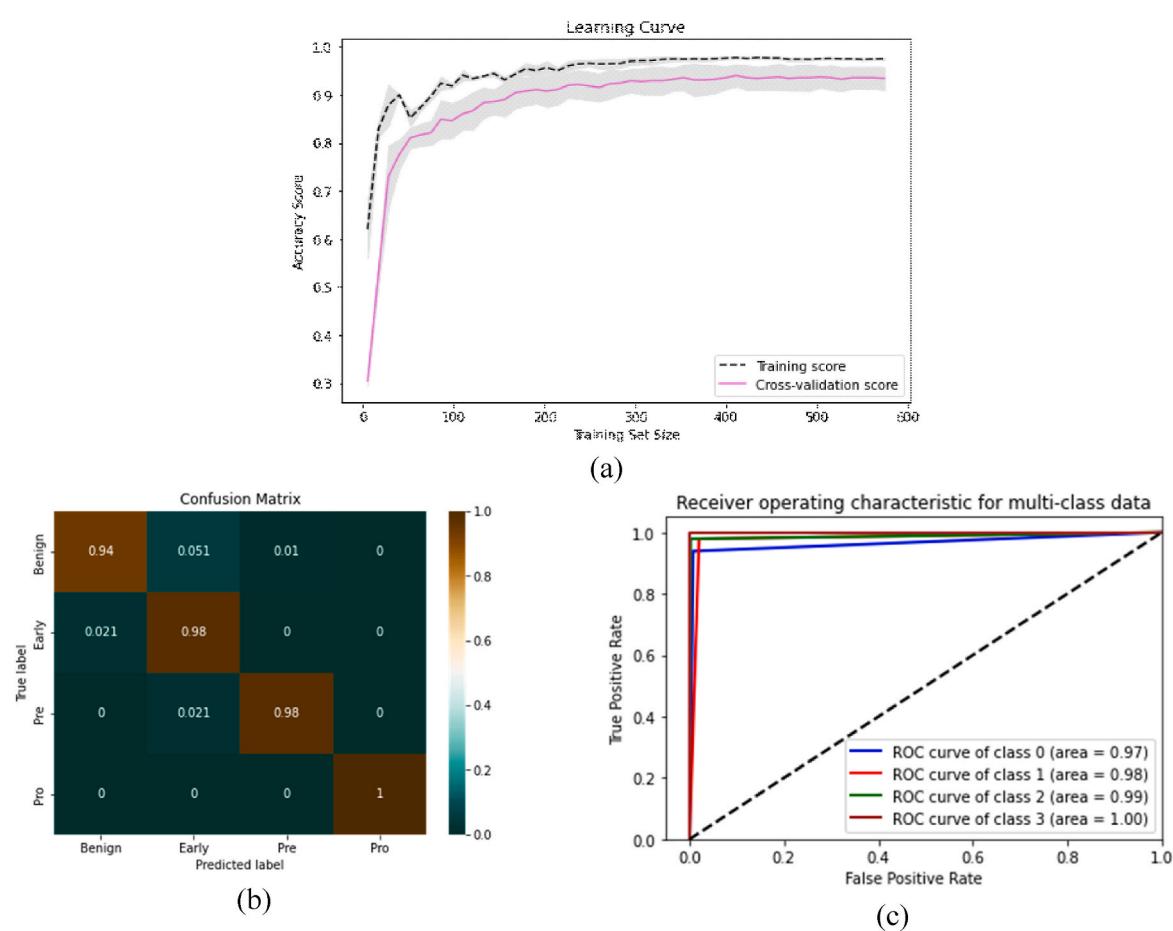
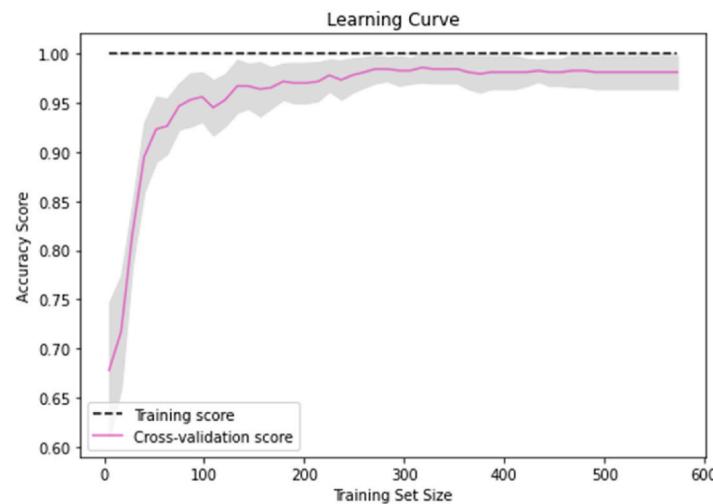
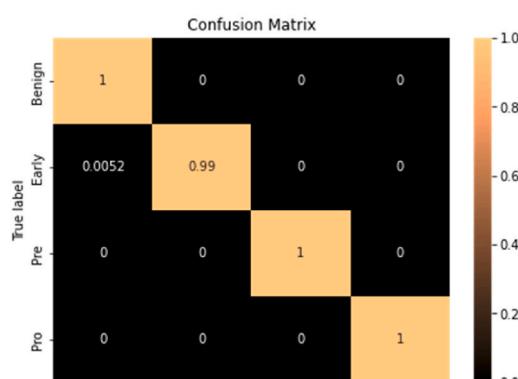


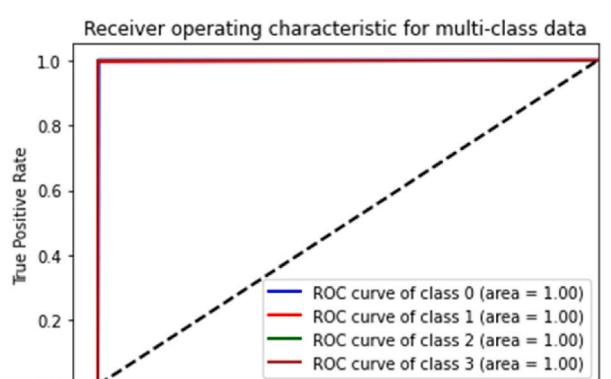
Fig. 15. The performance measurement of VGG19 and CSO with (a) Learning curve (b) Confusion Matrix (c) AUC-ROC curve.



(a)



(b)



(c)

Fig. 16. The performance measurement of ResNet50 and CSO with (a) Learning curve (b) Confusion Matrix (c) AUC-ROC curve.

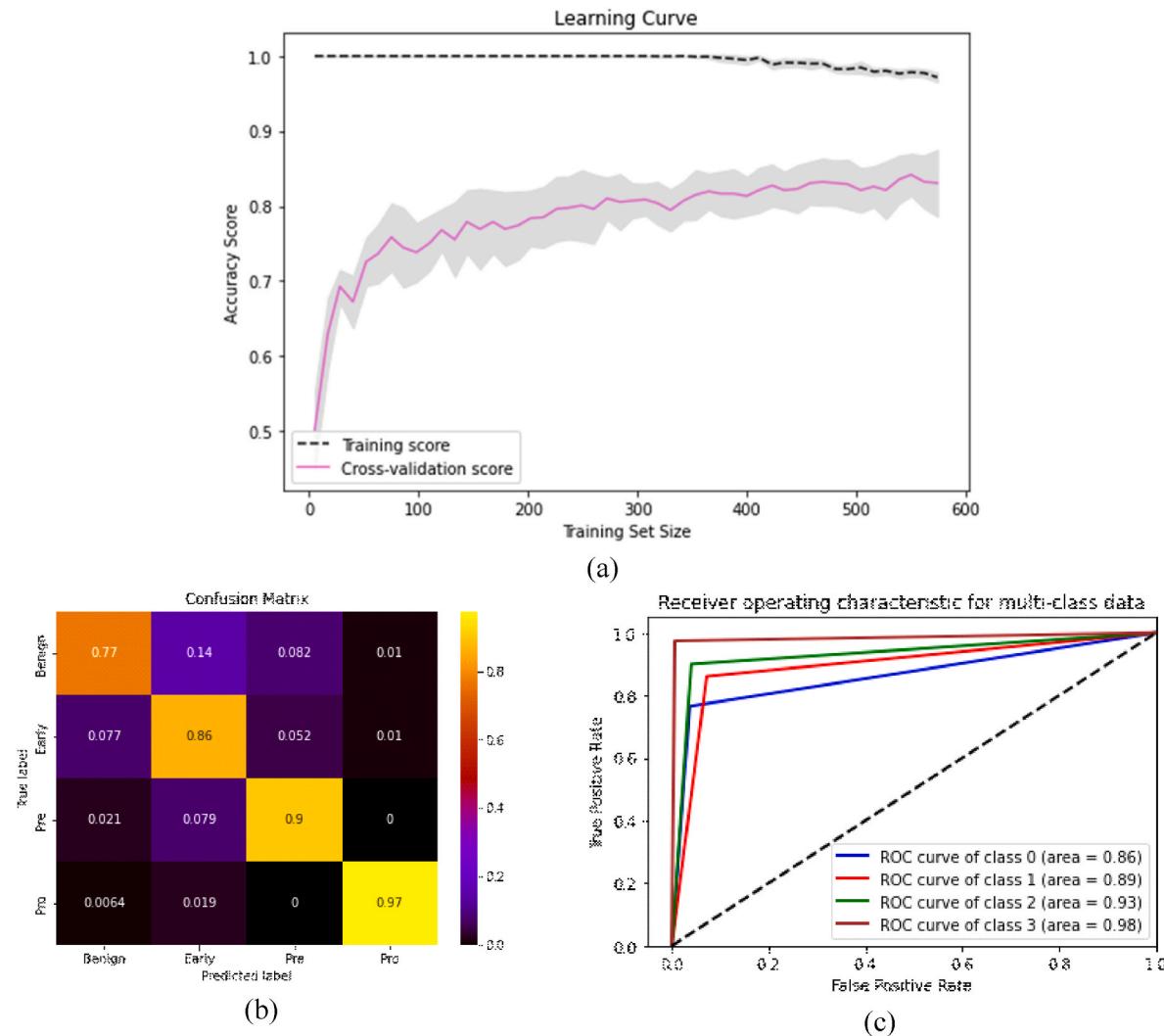


Fig. 17. The performance measurement of InceptionV3 and CSO with (a) Learning curve (b) Confusion Matrix (c) AUC-ROC curve.

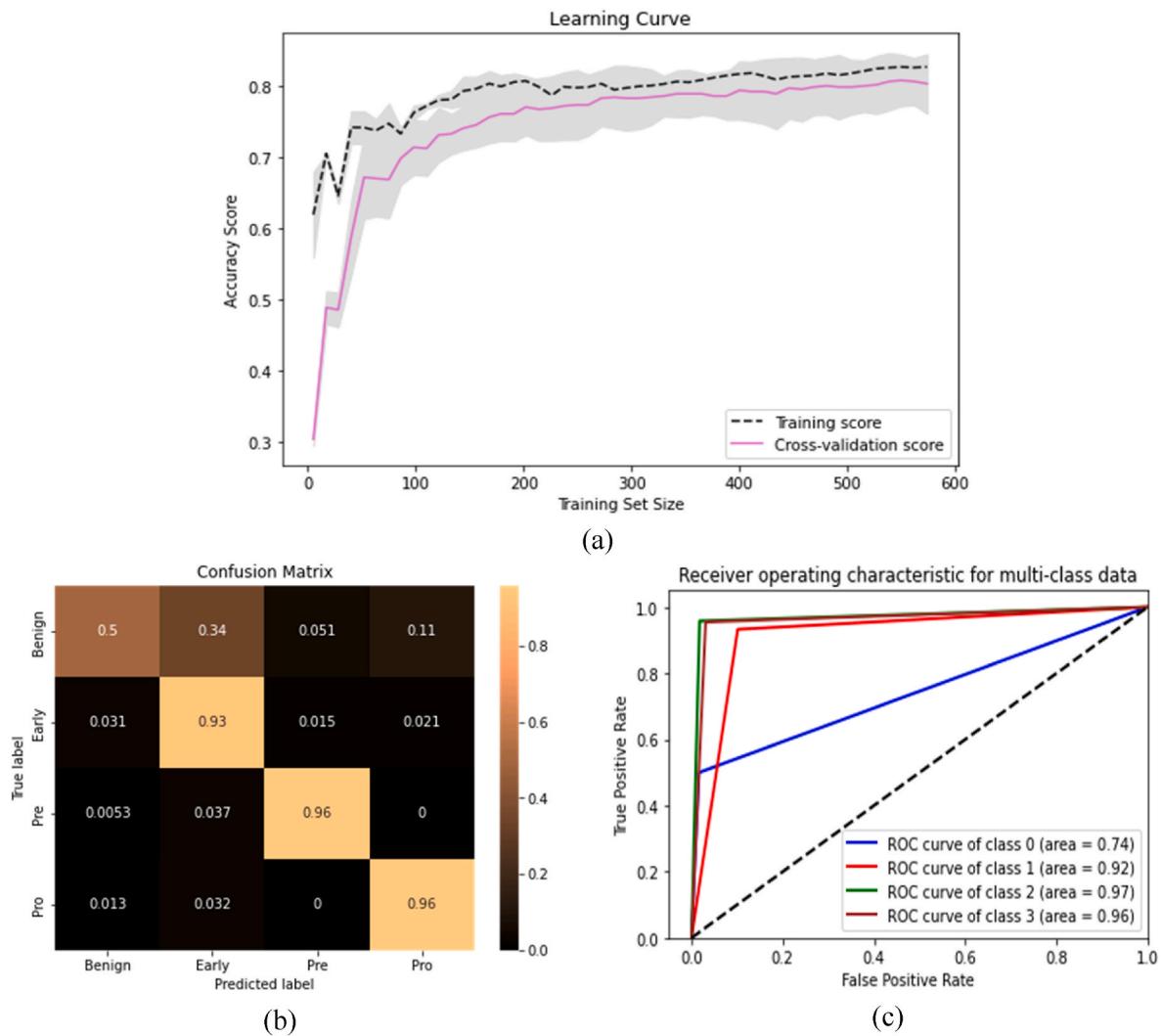


Fig. 18. The performance measurement of Inception V3 + Xception and CSO with (a) Learning curve (b) Confusion Matrix (c) AUC-ROC curve.

Table 15

Comparison of the performance of ResNet50 with and without CSO.

	Techniques	A	P	R	F1
Without PSO	Extracted Feature Vectors + ResNet50 + LR	99.53	99.39	99.57	99.48
With CSO	Extracted Feature Vectors + ResNet50 + CSO + LR	99.84	99.75	99.87	99.81

than any other CNN model in this research. When we applied the PCA algorithm the accuracy was 98.90%. When we worked with the LDA the

accuracy was slightly boosted and was 99.375. But when we included the nature inspired algorithm with SVC feature selector, the accuracy was optimized to 99.84%. This the highest accuracy we have achieved from both PSO and CSO. Fig. 19 shows a bar chart of the comparison among different techniques.

This research also reflects a comparison on the performances between two nature inspired algorithms. Though these two algorithm provides same results, there is a difference in number of feature selection. To work with these algorithm, we have occupied same number of iterations, population size as well as number of seeds. The PSO algorithm found the number of 1019 best selected features. Whereas the CSO

Table 16

Classifiers with the Feature Selection techniques for VGG19.

Classifiers	Feature Selector								SVC Feature Selector			
	PCA				LDA				SVC Feature Selector			
	A	P	R	F1	A	P	R	F1	A	P	R	F1
SVC	95.30	94.50	94.67	94.57	93.73	92.81	93.85	93.10	97.81	97.69	97.43	97.55
RF	85.74	87.26	80.69	81.77	94.98	95.87	93.77	94.57	95.77	96.42	94.36	95.17
DT	73.82	72.06	72.64	72.29	85.11	83.95	83.87	83.88	86.52	85.56	85.20	85.33
NB	39.18	49.52	32.22	25.61	76.02	75.86	76.52	75.18	87.30	85.69	86.23	85.57
XGB	93.73	93.69	92.71	93.14	96.08	96.29	95.57	95.87	96.55	96.96	95.48	96.10
KNC	93.26	94.71	90.99	92.12	92.01	93.69	89.37	90.53	93.73	94.63	91.77	92.77
LR	97.02	96.93	96.53	96.72	97.49	97.30	97.18	97.23	98.59	98.43	98.45	98.44

Table 17

Classifiers with the Feature Selection techniques for ResNet50.

Classifiers	Feature Selector											
	PCA				LDA				SVC Feature Selector			
	A	P	R	F1	A	P	R	F1	A	P	R	F1
SVC	97.33	96.03	97.55	96.67	97.33	96.55	97.64	96.97	98.90	98.85	98.71	98.78
RF	86.34	86.21	78.71	79.26	97.96	98.33	97.63	97.94	97.33	97.80	96.41	97.00
DT	85.56	84.18	80.08	81.09	88.54	87.33	87.63	87.47	88.54	87.63	86.84	87.17
NB	44.27	30.15	39.62	31.81	89.01	88.32	88.95	88.37	91.05	89.86	90.96	90.11
XGB	97.96	97.47	97.48	97.47	98.12	97.91	98.14	98.01	98.90	98.80	98.56	98.68
KNC	96.08	96.62	94.14	95.10	94.66	95.42	93.03	93.91	97.80	98.10	96.93	97.44
LR	99.84	99.87	99.71	99.79	99.69	99.63	99.74	99.68	99.84	99.75	99.87	99.81

Table 18

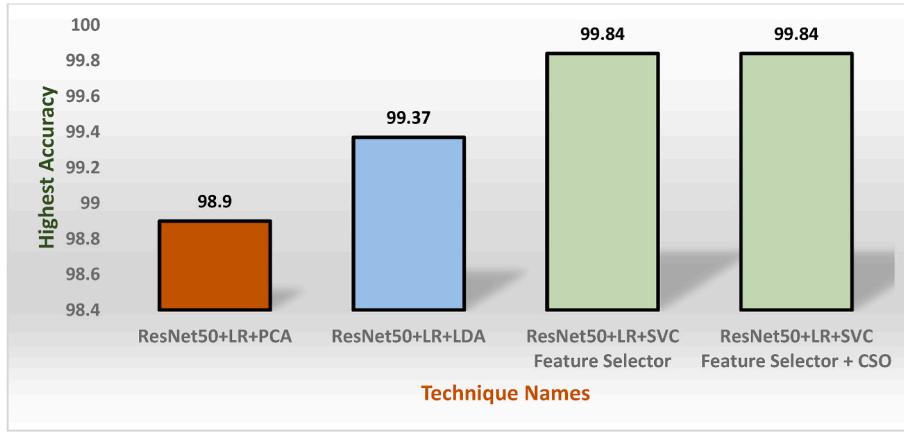
Classifiers with the Feature Selection techniques for InceptionV3+Xception.

Classifiers	Feature Selector											
	PCA				LDA				SVC Feature Selector			
	A	P	R	F1	A	P	R	F1	A	P	R	F1
SVC	90.28	90.12	87.91	88.89	87.46	87.56	84.81	85.63	87.93	87.80	83.65	84.48
RF	70.69	70.97	65.49	62.49	87.62	87.59	84.28	85.16	89.50	89.25	89.26	87.10
DT	60.42	63.68	62.89	62.92	78.53	76.88	76.48	76.66	76.33	73.04	72.91	72.91
NB	42.16	34.69	38.05	32.01	65.20	65.46	61.99	62.63	74.92	74.27	70.77	71.26
XGB	86.52	86.73	83.50	84.33	92.48	92.03	90.74	91.24	93.83	93.68	92.00	92.67
KNC	80.72	80.45	77.38	77.45	81.03	79.99	76.79	76.90	89.97	90.11	86.77	87.64
LR	91.85	91.12	90.92	90.99	93.10	92.18	91.78	91.95	94.04	93.63	92.95	93.26

Table 19

Comparison of different Feature Selection Models with ResNet50's features and LR.

Techniques	A	P	R	F1
ResNet50 + LR + PCA	98.90	98.64	98.91	98.77
ResNet50 + LR + LDA	99.37	99.46	99.11	99.28
ResNet50 + LR + SVC Feature Selector + PSO	99.84	99.75	99.87	99.81
ResNet50 + LR + SVC Feature Selector + CSO	99.84	99.75	99.87	99.81

**Fig. 19.** Bar chart of comparison among the existing best techniques.**Table 20**

Among the performance of PSO and CSO.

Techniques	No. of Iterations	Population Size	No. of Seeds	A	P	R	F1	No. Best Selected Features
ResNet50 + LR + SVC Feature Selector + PSO	05	10	1234	99.84	99.75	99.87	99.81	1019
ResNet50 + LR + SVC Feature Selector + CSO				99.84	99.75	99.87	99.81	909

algorithm tracked out the number of features of 909. Table 20 shows the corresponding comparison between these two algorithms.

5. Conclusion

The many different varieties of cancer, which are the collection of cells that are developing uncontrollably within the body, include breast cancer, lung cancer, skin cancer, and blood cancers like leukemia and lymphoma. One of the most important types of cancer is Acute Lymphoblastic Leukemia (ALL). This study examines the application of a novel technique for categorizing Acute Lymphoblastic Leukemia using cutting-edge technologies like Machine Learning (ML) and Deep Learning (DL). The major components of the proposed research pipeline include dataset construction, feature extraction using Convolutional Neural Network (CNN) architectures that have been pre-trained from each individual image of a blood cell, and classification using traditional ML-based classifiers. The dataset is split into two similar categories—benign and malignant—and then reconfigured into four significant classes, each of which has three subtypes of malignant, namely benign, early pre-B, pre-B, and pro-B. The research first extracts the features from CNN models, and then feeds the extracted features to feature selectors such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), and SVC Feature Selectors, along with two nature-inspired algorithms such as Particle Swarm Optimization (PSO) and Cat Swarm Optimization (CSO). The seven ML classifiers have thereafter been used in research. A collection of experimental data has been compiled and analyzed in order to evaluate the effectiveness of the suggested architecture. The research first worked with pre-trained CNN models with conventional ML classifiers and found the highest accuracy of 98.43% accuracy without explicitly using the feature selection algorithms and nature inspired algorithms. The research has executed the proposed model with ResNet50 architecture with feature selection algorithms and PSO & CSO. Then, we have tracked out the highest accuracy up to 99.84%. This is very remarkable improvement in multi-class classification in malignant with the feature fusion and nature inspired algorithms.

The model has certain shortcomings even though the outcomes are optimum. Firstly, the proposed is not applied in real-time malignant classification. Secondly, we have only applied the PSO and CSO on the extracted features. If these algorithms are applied in the deep layers of customized CNN, then the model can be embedded in the small IoT based devices like smart watch or smartphone. In future, this research will overcome these issues to create the proposed system more reliable in real-life applications. Firstly, this study will develop and android application. Then, the proposed model will be applied to android application in order to compile with phone camera for real-time classification. Secondly, this study will be implemented in small IoT device to aid hematologist in the leukemia classification. However, the proposed model may also be helpful for real-world Acute Lymphoblastic Leukemia (ALL) classification.

Contribution of the authors

Wahidur Rahman: Conceptualization, Methodology, Software Formal analysis, **Mohammad Gazi Golam Faruque:** Data Collection, Data Optimization, **Kaniz Roksana:** Visualization, Investigation, **A H M Saifullah Sadi:** Investigation, Data Analysis and Figure Drawing, **Mohammad Motiur Rahman:** Supervision, and Writing- Reviewing and Editing, **Mir Mohammad Azad:** Supervision and Reviewing

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

References

- [1] Ou F-S, et al. Biomarker discovery and validation: statistical considerations 2021; 16(4):537–45.
- [2] Chithrakkannan R, et al. Breast cancer detection using machine learning 2019;8 (11):3123–6.
- [3] Chaurasia V, et al. Prediction of benign and malignant breast cancer using data mining techniques 2018;12(2):119–26.
- [4] Solanki YS, et al. A hybrid supervised machine learning classifier system for breast cancer prognosis using feature selection and data imbalance handling approaches 2021;10(6):699.
- [5] Ahmad S, et al. A novel hybrid deep learning model for metastatic cancer detection. 2022. p. 2022.
- [6] Fujita TC, et al. Acute lymphoid leukemia etiopathogenesis 2021;48:817–22.
- [7] Leukemia—Cancer Stat Facts. [cited 2023 14 March]; Available from: <https://seer.cancer.gov/statfacts/html/leuk.html>.
- [8] Atteia G, Alhussan AA, Samee NAJS, Bo-Allcnn. Bayesian-Based Optimized CNN for Acute Lymphoblastic Leukemia Detection in Microscopic Blood Smear Images 2022;22(15):5520.
- [9] Cancer Research Uk [cited 2023 17 March]; Available from: <http://www.cancerr esearchuk.org/>.
- [10] Dong Y, et al. Leukemia incidence trends at the global, regional, and national level between 1990;9:1–11. 2017. 2020.
- [11] Cranaf, R., G. Kavitha, and S. Alagu, A decision Support system for the identification of acute lymphoblastic leukemia in microscopic blood smear images.
- [12] Mishra S, et al. Texture feature based classification on microscopic blood smear for acute lymphoblastic leukemia detection 2019;47:303–11.
- [13] Das PK, Jadoun P, Meher S. Detection and classification of acute lymphocytic leukemia. IEEE-HYDCON; 2020. 2020. IEEE.
- [14] Patel N, Mishra AJPCS. Automated leukaemia detection using microscopic images 2015;58:635–42.
- [15] Bennett JM, et al. Proposals for the classification of the acute leukaemias French-American-British (FAB) co-operative group 1976;33(4):451–8.
- [16] Das PK, et al. A systematic review on recent advancements in deep and machine learning based detection and classification of acute lymphoblastic leukemia. 2022.
- [17] Abbas N, et al. Nuclei segmentation of leukocytes in blood smear digital images 2015;28(5).
- [18] Abbas N, et al. Machine aided malaria parasitemia detection in Giemsa-stained thin blood smears 2018;29:803–18.
- [19] Ramaneswaran S, et al. Hybrid inception v3 XGBoost model for acute lymphoblastic leukemia classification, vol. 2021; 2021. p. 1–10.
- [20] Brownlee J. Deep learning with Python: develop deep learning models on Theano and TensorFlow using Keras. Machine Learning Mastery; 2016.
- [21] Rehman A, et al. Classification of acute lymphoblastic leukemia using deep learning 2018;81(11):1310–7.
- [22] Arbab Q, Khan MQ, Ali HJTS. Automatic Detection and Classification of Acute Lymphoblastic Leukemia Using Convolution Neural Network 2022;3. 4.
- [23] Mughal B, et al. Removal of pectoral muscle based on topographic map and shape-shifting silhouette 2018;18(1):1–14.
- [24] Norouzi A, et al. Medical image segmentation methods, algorithms, and applications 2014;31(3):199–213.
- [25] Mughal B, et al. Extraction of breast border and removal of pectoral muscle in wavelet domain. 2017. p. 5041–3. 28(11).
- [26] Mughal B, Sharif M, N.J.T.E.P.J.P. Muhammad. Bi-model processing for early detection of breast tumor in CAD system 2017;132:1–14.
- [27] Mughal B, et al. A novel classification scheme to decline the mortality rate among women due to breast tumor. 2018. p. 171–80. 81(2).
- [28] Mohapatra S, et al. An ensemble classifier system for early diagnosis of acute lymphoblastic leukemia in blood microscopic images 2014;24:1887–904.
- [29] Rehman A, et al. Rouleaux red blood cells splitting in microscopic thin blood smear images via local maxima, circles drawing, and mapping with original RBCs. 2018. p. 737–44. 81(7).
- [30] Saba TJBR. Halal food identification with neural assisted enhanced RFID antenna. 2017. p. 7760–2. 28(18).
- [31] Waheed SR, et al. Multifocus watermarking approach based on discrete cosine transform. 2016. p. 431–7. 79(5).
- [32] Ahmed N, et al. Identification of leukemia subtypes from microscopic images using convolutional neural network. 2019. p. 104. 9(3).
- [33] Jiang Z, et al. Method for diagnosis of acute lymphoblastic leukemia based on ViT-CNN ensemble model. 2021. p. 2021.
- [34] Rezayi S, et al. Timely diagnosis of acute lymphoblastic leukemia using artificial intelligence-oriented deep learning methods. 2021. p. 2021.
- [35] Zakir Ullah M, et al. An attention-based convolutional neural network for acute lymphoblastic leukemia classification. 2021, 10662. 11(22).
- [36] Revanda AR, et al. Classification of acute lymphoblastic leukemia on white blood cell microscopy images based on instance segmentation using Mask R-CNN. 2022. 15(5).
- [37] Abunadi I, E.M.J.S. Senan. Multi-method diagnosis of blood microscopic sample for early detection of acute lymphoblastic leukemia based on deep learning and hybrid techniques. 2022. p. 1629. 22(4).

- [38] Sampathila N, et al. Customized deep learning classifier for detection of acute lymphoblastic leukemia using blood smear images. *Healthcare*; 2022 [MDPI].
- [39] Pallegama R, et al. Acute lymphoblastic leukemia detection using convolutional neural network. 2020, 26529, 10(6).
- [40] Safuan SNM, et al. Investigation of white blood cell biomarker model for acute lymphoblastic leukemia detection based on convolutional neural network. 2020. p. 611–8. 9(2).
- [41] Ansari S, et al. A customized efficient deep learning model for the diagnosis of acute leukemia cells based on lymphocyte and monocyte images. 2023. p. 322. 12 (2).
- [42] Abd El-Ghany S, Elmogy M, El-Aziz AJD. Computer-Aided Diagnosis System for Blood Diseases Using EfficientNet-B3 Based on a Dynamic Learning Algorithm 2023;13(3):404.
- [43] Ayache F, Alti A. Performance evaluation of machine learning for recognizing human facial emotions. *Rev. d'Intelligence Artif.* 2020;34(3):267–75.
- [44] Gupta S, et al. Prediction performance of deep learning for colon cancer survival prediction on SEER data. *BioMed Res Int* 2022;2022.
- [45] Ahmad S, et al. A novel hybrid deep learning model for metastatic cancer detection. *Comput Intell Neurosci* 2022;2022.
- [46] Ayache F, Alti A. HDG and HDGG: an extensible feature extraction descriptor for effective face and facial expressions recognition. *Pattern Anal Appl* 2021;24: 1095–110.
- [47] Aria M. Acute lymphoblastic leukemia (ALL) image dataset. 2021 [cited 2023 March, 10]; Available from: <https://www.kaggle.com/datasets/mehradaria/leukemia>.
- [48] Khan HA, et al. Brain tumor classification in MRI image using convolutional neural network. *Math Biosci Eng* 2020;17(5):6203–16.
- [49] Dobilas SLDA. Linear discriminant analysis — how to improve your models with supervised dimensionality reduction. 2021 [cited 2023 March, 10]; Available from: <https://towardsdatascience.com/lda-linear-discriminant-analysis-how-to-improve-your-models-with-supervised-dimensionality-52464e73930f>.
- [50] Tao Z, et al. GA-SVM based feature selection and parameter optimization in hospitalization expense modeling. *Appl Soft Comput* 2019;75:323–32.
- [51] Sharif M, et al. An integrated design of particle swarm optimization (PSO) with fusion of features for detection of brain tumor. *Pattern Recogn Lett* 2020;129: 150–7.
- [52] Sikkandar H, Thiagarajan R. Deep learning based facial expression recognition using improved Cat Swarm Optimization. *J Ambient Intell Hum Comput* 2021;12: 3037–53.
- [53] Kurban H, Kurban M. Building Machine Learning systems for multi-atoms structures: CH₃NH₃PbI₃ perovskite nanoparticles. *Comput Mater Sci* 2021;195: 110490.