

RESEARCH ARTICLE

Swin-PSO-SVM: A Novel Hybrid Model for Monkeypox Early Detection

HAGER SALEH^{1,2,3}, SHERIF MOSTAFA³, ALSHIMAA H. ISMAIL⁴, SUMIT SRIVASTAVA⁵,
AND SAEED HAMOOD ALSAMHI^{1,6,7}

¹Insight Research Ireland Centre for Data Analytics, University of Galway, Galway, H91 TK33 Ireland

²Atlantic Technological University, Letterkenny, F92 FC93 Ireland

³Faculty of Computers and Artificial Intelligence, Hurghada University, Hurghada 84511, Egypt

⁴Department of Information Technology, Faculty of Computer and Informatics., Tanta University, Tanta 31527, Egypt

⁵Department of Electronics and Communication Engineering, FET, MJP Rohilkhand University, Bareilly, Uttar Pradesh 243006, India

⁶Faculty of Engineering, IBB University, Ibb, Yemen

⁷Department of Computer Science and Engineering, College of Informatics, Korea University, Seongbuk-gu, Seoul 02841, Republic of Korea

Corresponding authors: Hager Saleh (hager.saleh@insight-centre.org) and Saeed Hamood Alsamhi (saeed.alsamhi@insight-centre.org)

This work was supported by the Science Foundation Ireland under Grant SFI/12/RC/2289_P2.

ABSTRACT A resurgent zoonotic disease called monkeypox presents serious public health issues, especially in locations with low resources where prompt and correct diagnosis is essential. Due to the health shortcomings in sensitivity, specificity, and adaptability to different clinical presentations, traditional diagnostic techniques frequently need to be revised. We introduce a novel hybrid Swin-PSO-SVM to improve the precision and dependability of monkeypox detection. The Swin-PSO-SVM model incorporates a Swin Transformer for complex feature extraction, Particle Swarm Optimization (PSO) for extracting the best features from complex feature extraction, and a Support Vector Machine (SVM) for accurate classification. The findings show how well the Swin-PSO-SVM model performs, attaining great diagnostic accuracy and resilience over various datasets and picture characteristics using two datasets: MSLD with two classes (monkeypox and others) and MSID with four classes (chickenpox, measles, monkeypox, and normal). Swin-PSO-SVM model has the highest accuracy of 95.556 and an F1 score of 95.569 on the MSLD dataset and a 96.429 accuracy and a 96.429 F1 score on the MSID dataset, outperforming several existing models. With an accuracy of 91.111 and an F1-score of 91.138, the experimental findings proved the model's exceptional performance and validated its dependability in real-world applications. The Swin-PSO-SVM provides a workable and understandable solution that can be easily applied in clinical settings and improves the generalizability of monkeypox detection. This paper contributes to worldwide efforts to limit monkeypox outbreaks through early and accurate identification.

INDEX TERMS Monkeypox, swin transformer, swin-PSO-SVM, particle swarm optimization, support vector machine, healthcare, early diagnosis.

I. INTRODUCTION

The monkeypox virus, a member of the same family as smallpox and belongs to the Orthopoxvirus genus, is the cause of this viral zoonotic illness. Monkeypox is less severe than smallpox, but it has nonetheless attracted much interest because of its increasing prevalence and potential for large-scale outbreaks, especially in areas with poor access to

The associate editor coordinating the review of this manuscript and approving it for publication was Muammar Muhammad Kabir^{id}.

healthcare. Controlling the spread of monkeypox requires accurate and timely diagnosis, particularly considering how similar its symptoms are to those of other illnesses that resemble the pox, such as chickenpox and measles. Even if they are effective, traditional diagnostic methods frequently call for specific laboratory equipment and skilled workers, which may not be easily accessible in environments with low resources. As a result, there is increasing interest in creating reliable and automated diagnostic models based on Deep Learning (DL) and Machine Learning (ML) approaches

that enable quick and accurate illness detection. Early classification and detection of monkeys can be crucial in controlling their spread. The classification of monkey tumors can aid in the diagnosis of disease severity and aid in appropriate treatment selection. Traditional classification is based on visual wound assessment by trained physicians. However, this method can be time-consuming, subjective, and error-prone [1].

DL techniques have shown promising results in various image classification tasks in recent years. Convolutional Neural Networks (CNNs) are a type of DL paradigm widely used in image classification tasks. CNNs can automatically recognize features from raw image data and use these features to classify images into different groups [2], [3], [4], [5], [6], [7], [8], [9], [10]. DL models for monkey analysis have been developed in many studies, using the Monkeypox Skin Lesion Dataset (MSLD) and Monkeypox Skin Image Dataset (MSID) for model validation and training [11], [12]. MSLD evaluated the accuracy of pre-trained DL models, where ResNet-18 achieved the most impressive accuracy; GoogleNet, SqueezeNet, ResNet-18, and similar Authors developed a model-based concept so that of seizures, seizures using the Squeeze and Excitation Network (SENet), performed better InceptionV3 and VGG16 for classifying monkeys in the three data sets. It was also tested in [13] that the ensemble system combines VGG-16, ResNet50, and InceptionV3 to classify monkey diarrhea and other diseases as ulcerative colitis ResNet50 had the best accuracy. Although such methods have shown power, models are needed for more successful interpretation and modification, primarily to address this when measuring variable image quality and limited data supply. Extraction of Naive Bayes (NB) and Support Vector Machine (SVM) features from VGG16Net, AlexNet, and other models.

Moreover, using the MSID—which consists of four classes: chickenpox, measles, monkeypox, and standard—has made it easier to create multi-class classification models. While [14] stated that MobileNetV2 surpassed VGG16 and VGG19, it showed how effective SqueezeNet is. Even with these developments, a more complete model that combines the benefits of DL, ML, and optimization methods is still required. To close this gap, we introduce a novel hybrid model called Swin-PSO-SVM, which combines PSO, SVM, and the Swin Transformer for early monkeypox detection. The Swin Transformer is an excellent tool for collecting local and global relationships in visual data because of its hierarchical processing and changing window self-attention processes [15]. We chose PSO for its simplicity, fast convergence, and effective global search capabilities, making it ideal for optimizing complex models like SVM. Compared to other algorithms such as GWO or AHA, PSO has fewer control parameters and is easier to implement, providing a practical balance between efficiency and performance. By integrating PSO, we can tune the SVM's hyperparameters, guaranteeing reliable and effective classification results on various datasets.

A. MOTIVATIONS AND CONTRIBUTIONS

There is an urgent need for more precise, understandable, and flexible diagnostic methods for monkeypox—especially in light of the growing dangers to world health. Although DL models have demonstrated potential in detecting monkeypox from skin lesion photos, DL models frequently have drawbacks such as a reliance on sizable, labeled datasets, poor interpretability, and poor performance in settings with fluctuating picture quality. Even if DL models are reliable, traditional methods often cannot be used in environments with limited resources since DL models need specialized tools and knowledge. Moreover, the healthcare sector has a notable vacuum due to the lack of a robust model that successfully combines the advantages of optimization, ML, and DL. Using the capabilities of hierarchical processing of the Swin, PSO, and SVM, the Swin-PSO-SVM model aims to address the above issues by improving diagnostic accuracy efficiency, and detection performance while ensuring the solution is interpretable and flexible enough to work with various datasets. The main contributions of this paper summaries as follows:

- We introduce a novel hybrid model, i.e., Swin-PSO-SVM, which effectively synergizes the Swin transformer for extracting local feature extraction, PSO for best features selection with reducing complexity, and SVM for accurate classification. The proposed Swin-PSO-SVM is designed for monkeypox detection and diagnosis, offering a robust solution for Monkeypox identification.
- The proposed Swin-PSO-SVM model provides disease detection with accuracy for monkeys and robustness in MSLD and MSID datasets. The swin-PSO-SVM model has high performance in terms of an F1-score of 95.24% for the monkeypox class on the MSLD dataset and 98.01% for the normal class on the MSID dataset, outperforming several existing models. Therefore, the proposed model can maintain high accuracy and balanced performance across multiple disease classes, ensuring reliability for real-world applications—especially in low-resource environments.
- The Swin-PSO-SVM model helps to promote globalization. Public health efforts that promote early detection and intervention in an epidemic of monkey interpretability and practicality in clinical environments.

B. PAPER STRUCTURE

The rest of the paper is organized as follows. In Section II, we provide a review of related studies. In Section III, we describe the proposed approach in detail (Methodology). Section IV presents the experimental results and analysis. Section V presents the discussion of the paper. Finally, in Section VI, we conclude the paper and highlight its contributions.

II. RELATED WORK

Related works that use the MSLD with two classes. For example, Nayak et al. [16] applied five pre-trained DL models:

GoogLeNet, Places365-GoogLeNet, SqueezeNet, AlexNet, and ResNet-18. The results showed that ResNet18 obtained the highest accuracy. GradCAM enables visual interpretation of the prediction, helping health professionals use the model. Surati et al. [17] proposed the SENet Attention model to classify three diseases Monkeypox, Chickenpox, and Measles. Using three datasets, the proposed model recorded the highest performance compared to InceptionV3, EfficientNet, and VGG16. Ali et al. [13] proposed an ensemble model based on VGG-16, ResNet50, and InceptionV3 and compared it using three models to classify monkeypox and other diseases. Data augmentation was applied to increase the number of images and enhance their quality. ResNet50 achieved the best overall accuracy. Kumar [18], extracted features from AlexNet, GoogleNet, and VGG16Net and then applied different ML models to extracted features, including SVM, NB, KNN, RF, and DT. The results showed that NB with VGG16Net features recorded the highest performance.

Haque et al. [19] proposed a hybrid model that combined DL models: VGG19, Xception, DenseNet121, EfficientNetB3, and MobileNetV2 with Convolutional Block Attention Module (CBAM). The results showed that the Xception-CBAM-Dense layers record the highest performance. Sahin et al. [20], applied DL models: ResNet18, GoogleNet, EfficientNetb0, NasnetMobile, ShuffleNet, and MobileNetv2. The results showed that MobileNetv2 recorded the highest accuracy. Related works that use MSID with four classes. For example, Nayak et al. [21] applied residual networks (ResNet) and SqueezeNet to classify four classes. The results showed that SqueezeNet recorded the highest performance. Irmak et al. [22] applied MobileNetV2, VGG16, and VGG19, and the highest performance scores were obtained with MobileNetV2.

A deep transfer learning method for identifying monkeypox disease using CNN was presented [16]. The study employed Generalization and Regularization-based Transfer Learning methods (GRA-TLA) for binary and multiclass classification. The proposed method was evaluated on 10 distinct CNN models in three investigations. According to the preliminary computational results, ResNet-101 performed best for multiclass classification. Sahin et al. [20] provided details of a mobile application that classifies skin lesions using deep learning, which helps identify and isolate human monkeypox. The program made use of a deep CNN that has been trained with pictures of skin lesions from both healthy individuals and those with monkeypox. The network used the TensorFlow Lite model to adapt to mobile devices and classify photos as positive or negative for monkeypox detection. Altun, Murat, et al. [23] used DL techniques to identify monkeypox disease from skin lesions in the event of a pandemic. After developing a customized CNN model, the authors employed it with hyper-parameter optimization and a customized hybrid function transfer learning model to obtain impressive results. Several transfer learning models were compared for performance, such as MobileNetV3-s,

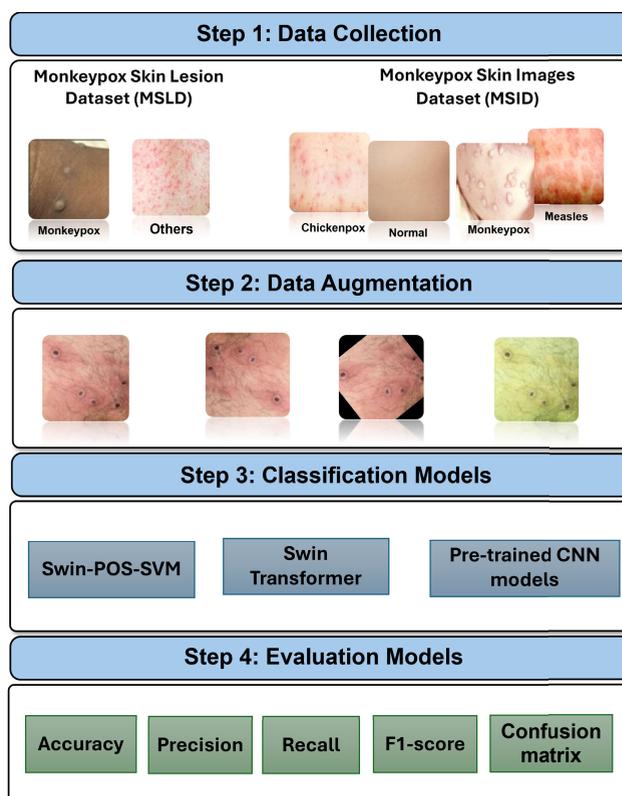


FIGURE 1. The steps of methods.

EfficientNetV2, ResNET50, Vgg19, DenseNet121, and Exception models.

The application of DL and attention mechanisms for image-based diagnosis of human monkeypox sickness is discussed [19]. The Xception-CBAM-Dense architecture outperformed the other five DL models with attention mechanisms, achieving the highest results. Uysal [24] proposed hybrid DL based on LSTM and pre-trained CNN models: CSPDarkNet, InceptionV4, MnasNet, MobileNetV3, RepVGG, SE-ResNet, and Xception. They applied different data augmentation techniques. Ahsan et al. [25] applied pre-trained CNN models: VGG16, InceptionResNetV2, ResNet50, ResNet101, MobileNetV2, VGG19, ViT. They modified versions of the VGG19 and MobileNetV2 models to enhance performance. The models were evaluated using an augmented testing set and a non-augmented testing set. Table 1 compares different models used in research. It outlines the methods, advantages, limitations, and datasets referenced in the studies.

III. METHODOLOGY

The procedure has six primary steps including a collection of image datasets, data augmentation, model training, and model evaluation as shown in Figure 1.

A. DATABASE DESCRIPTION

In this work, we use two datasets, i.e., MSLD [26] and MSID [27]. MSLD includes two classes: Monkeypox and

TABLE 1. Comparison of existing work.

Ref.	Methods	Advantages	Limitations	Dataset
[24]	Hybrid model CNN-LSTM	Hybrid model	Applying models on one only dataset, does not applied optimize methods	MSID
[25]	Different Pre-trained CNN models	Modifying VGG19 and MobileNetV2	—	
[16]	Different pretrained model CNN models	Applying XAI	It was evaluated models using augmented testing set	MSLD
[17]	Applying pre-trained CNN models with SENet	Applying pre-trained CNN models with SENet	Models were applied by three classes and one dataset.	MSLD
[13]	Applying VGG16, ResNet50, InceptionV3	proposing a new dataset	Only apply pre-trained CNN models	MSLD
[18]	Hybrid models: pre-trained CNN with ML	Combining pre-trained CNN with ML	Lower accuracy compared to other models; Requires further large-scale dataset exploration	MSLD
[20]	Developing android app that using pre-trained CNN models	Developing android app that using pre-trained CNN models	Limited to one dataset; needs broader dataset	MSLD
[21]	Applying pre-trained CNN models	—	Limited dataset size, only pre-trained CNN models	MSID
[19]	Applying CNN models with CBAM	Enhanced accuracy and feature focus using CBAM	Small number of original images in the dataset	MSLD
Our work	A novel hybrid model "Swin-PSO-SVM", which combines PSO, SVM, and the Swin Transformer for early monkeypox detection.	Early monkeypox detection .	—	MSLD And MSID

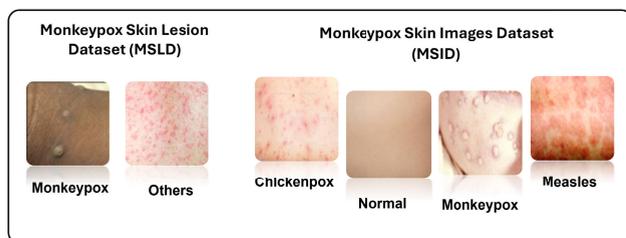


FIGURE 2. The sample image from each class.

Others. It contains a Fold1 folder with three folders: training with 2142 augmented images, validation with 420 augmented images, and testing with 45 images. Furthermore, MSID is used to test and develop DL algorithms to diagnose Monkeypox. Diponkor Bala and Md. Shamim Hossain of the School of Computer Science and Technology at the University of Science and Technology of China (USTC) assembled the dataset .it includes 770 images divided into four: 107 of Chickenpox, 91 of Measles, 279 of Monkeypox and 293 of Normal. Figure 2 shows one image from each class from the MSID and MSLD datasets.

B. IMAGE AUGMENTATION

Image augmentation is an ML and computer vision technique that uses manipulated images to increase the size of a training data artificially set [28]. This increases the resilience and generalization of the model, particularly when data availability is limited [29]. Many techniques are employed in Image Augmentation, such as rotation, which rotates

images to a particular degree. Flipping is inverting an image along a defined axis, typically horizontal or vertical, to introduce variance into the dataset and make the model adaptable to varied orientations [30]. Resizing via Adjusting an image's dimensions to a specific width and height helps guarantee that all images in the dataset are of comparable size, which is frequently required for model input [31]. Normalization aims to conform the values of the pixels to a standard scale, often by subtracting the mean while dividing by the standard deviation or scaling to a specific range, to accelerate convergence during training and ensure that different characteristics contribute equally to the learning process [32]. We perform augmentation to increase the training part of the MSID dataset. We apply the following methods to our dataset. RandomHorizontalFlip = 0.5, RandomVerticalFlip = 0.5, Resize = 256, and Normalization = Mean and standardization

C. PROPOSED MODEL

The proposed Swin-PSO-SVM uses a combination of approaches that use cutting-edge ML techniques to improve monkeypox identification efficiency and accuracy. Figure 3 shows the proposed Swin-PSO-SVM that combines Swin transformer, PSO, and replaces softmax with SVM.

1) SWIN TRANSFORMER

Our solution uses the Swin Transformer to extract hierarchical features at various scales. The Swin transformer extracts local and global characteristics from input pictures using a hierarchical structure with shifting windows, in contrast to

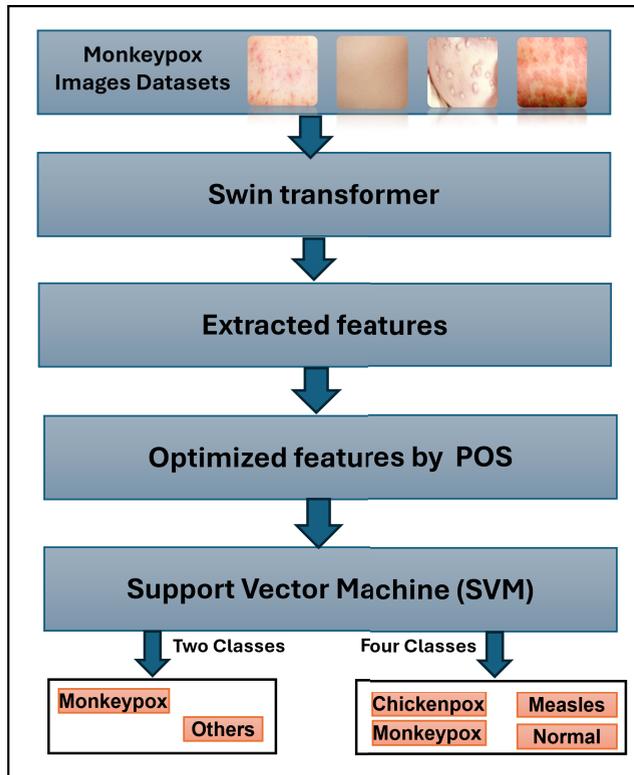


FIGURE 3. Proposed Swin-PSO-SVM model.

conventional CNNs. Because tiny details might be crucial for a successful diagnosis in medical image analysis. After processing the input images, the Swin Transformer extracts a rich set of highly descriptive features appropriate for classification. Swin Transformer is a deep learning system that uses an array of transformer encoder units [33]. Every encoder unit has a shifting window self-attention component, which stores regional dependencies across patches within a window, and a feed-forward neural network, enabling the model to recover local information from a patch [34]. Swin Transformer achieves process hierarchy by layering multiple Transformer encoder units. A single block's outcome serves as the information source for the next, enabling the model to record global interconnections throughout patches [35]. Right after hierarchical processing, an identification head is connected to the Swin Transformer to anticipate labels for classes during tasks related to image classification. The classification head typically comprises several fully connected layers that terminate with a softmax activation [35].

$$Z^l = W - MSA \left(LN \left(Z^{l-1} \right) \right) + Z^{l-1} \quad (1)$$

$$Z^l = MLP \left(LN \left(Z^l \right) \right) + Z^l \quad (2)$$

$$Z^{l+1} = SW - MSA \left(LN \left(Z^l \right) \right) + Z^l \quad (3)$$

$$Z^{l+1} = MLP \left(LN \left(Z^{l+1} \right) \right) + Z^{l+1} \quad (4)$$

LN and MLP refer to the linear layer and multi-layer perceptron, respectively, with Gaussian error linear unit

activation function. The typical multi-head self-attention layer found in most transformers is supplemented by a multi-head self-attention window (W-MSA) along with shifting window-based MSA layers. The architecture of the Swin model is shown in Figure 4

2) PSO

In the Swin-PSO-SVM model, our paper employs PSO [36] to optimize feature extraction from the Swin Transformer, thereby reducing dimensionality and focusing on the most relevant features. PSO is a metaheuristic that employs the population's optimization approach which can solve a variety of optimization issues, spanning continuous, discrete, and combinatorial ones [36]. The core idea underlying PSO is to create a swarm of particles, each reflecting a potential solution to the optimization issue. The particles traverse the search space, guided by the swarm's best-known positions. Steps in the basic PSO algorithm include Setting up a population of particles with random locations and velocities [37]. Then, evaluate each particle's fitness. Following that, Update each particle's individual best positioning, the swarm's global best position, along with its velocity and location. These phases repeat themselves until a requirement for termination is fulfilled, as a certain number of iterations or convergence happens. The basic PSO algorithm can be defined by the following equations [37]: The velocity update equation:

$$v_i^{t+1} = w \cdot v_i^t + c_1 \cdot r_1 \cdot (p_i^t - x_i^t) + c_2 \cdot r_2 \cdot (g^t - x_i^t) \quad (5)$$

where: v_i^{t+1} represents the updated velocity of particle i at time $t+1$, v_i^t shows the particle i current velocity at time t , w is the inertia weight, which governs the influence of the particle's past velocity, c_1 and c_2 are the cognitive and social acceleration coefficients, respectively, r_1 and r_2 are random numbers from 0 to 1, p_i^t is the particle i personal best position at time t , g^t is the global best position of the swarm at time t and x_i^t is the particle i current position at time t . The position update equation:

$$x_i^{t+1} = x_i^t + v_i^{t+1} \quad (6)$$

where: x_i^{t+1} is the updated position of particle i at time $t+1$, x_i^t is the particle i current position at time t , and v_i^{t+1} is the particle i updated velocity at time $t+1$. The global best update.

$$g^{t+1} = \arg \min_i \left(p_i^{t+1} \right) \quad (7)$$

where: g^{t+1} is the global best-updated position of the swarm in time $t+1$, p_i^{t+1} is the particle i updated personal best position at time $t+1$ and $f(x)$ is the objective function has to be optimized. These equation parameters can be tuned to regulate the appropriate level of exploration and exploitation during the search phase, which can have a substantial impact on the efficacy of the algorithm. PSO offers various advantages [37], including simplicity, ease of implementation, and the capacity to address non-differentiable, non-convex, and multi-modal optimization problems.

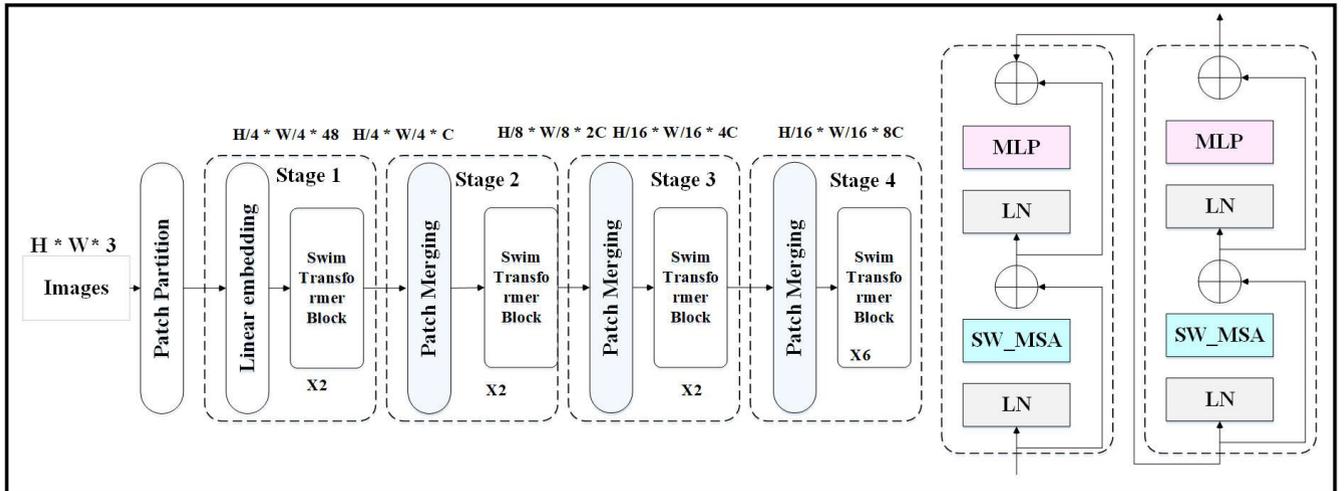


FIGURE 4. The architecture of Swin model.

Algorithm 1 Swin-PSO-SVM

1) **Input:** Set of images $X = \{x_1, x_2, \dots, x_n\}$ with labels $Y = \{y_1, y_2, \dots, y_n\}$.

2) **Output:** Predicted labels \hat{Y}_{test} .

3) **Feature extraction:**

- a) Divide each x_i into patches, embed to sequence Z_i .
- b) Apply $f_i = \text{SwinTransformer}(x_i)$

4) **PSO for feature selection:**

Input: Feature set $F = [f_1, f_2, \dots, f_p]$

- a) Define a binary particle vector $P = [p_1, p_2, \dots, p_p]$, where:

$$p_i \in \{0, 1\}, \quad i = 1, 2, \dots, p$$

Each p_i indicates whether feature f_i is selected ($p_i = 1$) or not ($p_i = 0$)

- b) Fitness function $\mathcal{F}(P)$:

$$\mathcal{F}(P) = \text{Accuracy}(\text{SVM}(F_P))$$

where F_P represents the subset of features selected P

- c) Optimize P using PSO to maximize $\mathcal{F}(P)$

5) **SVM Classification:**

- a) Train SVM with θ^* on F and Y :

$$\text{SVM Model} = \text{SVM}(F, Y; \gamma^*, C^*)$$

- b) Predict labels for X_{test} :

$$\hat{Y}_{test} = \text{SVM Model}(F_{test})$$

derived from the training set, while the testing set is used to train and evaluate the SVM, respectively. This combination ensures that the classifier is effectively generalized to handle unknown data while maintaining high accuracy.

IV. EXPERIMENT RESULTS

The MSLD and MSID results are shown in two subsections according to confusion matrices, recall, F1-score in each class, and precision. Moreover, each dataset’s average precision, recall, accuracy, and F1-score are shown in the discussion subsection V. Where TP represents the number of true positive predictions, TN is the whole true negative forecasts, FP is the total false positive estimates, and FN is the total false negative predictions.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{8}$$

$$\text{Recall} = \frac{TP}{TP + FN} \tag{9}$$

$$\text{Precision} = \frac{TP}{TP + FP} \tag{10}$$

$$\text{F1 - score} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}} \tag{11}$$

A. EXPERIMENTAL SETUP

The models were created using the Monai library, PyTorch, and Python. A laptop equipped with an Intel Core i7 10750H processor and 16GB of RAM was used for the experiment. An Nvidia GeForce GTX 1650 with 4 GB of VRAM was used as the single GPU for the networks’ training and validation. Table 2 includes the number of images in each class that were used in the dataset during the training, validation and testing of the model. The networks were trained and validated only on the augmented data set, with 80% of the combined images from both labels used for training and 20% used for validation. A summary of the POS parameters for selecting the best features can be found in Table 3.

3) SVM

We replace softmax with an SVM for classification in the final phase. The best-selected features, optimized by PSO, are

TABLE 2. The number of images in each class for MSID.

Datasets	Class label	Training	Testing	Validation	Total
MSLD	Monkeypox	980	20	168	1168
	Other	1162	25	252	1439
	Total	2142	45	420	2607
MSID	Chickenpox	74	28	5	107
	Measles	63	24	4	91
	Monkeypox	195	70	14	279
	Normal	205	74	14	293
	Total	537	196	37	770

TABLE 3. Parameters of PSO.

Parameter	Value
Population size	20
Max num of generation	30
Early stopping	True
Local best weight	1
Global best weight	1
Use local random seed	True

B. COMPARISON OF EXISTING MODELS AND Swin-POS-SVM PERFORMANCE IN MSLD

1) PERFORMANCE OF ADAM OPTIMIZER

Table 4 shows the precision, recall, and F1-score of DenseNet121, VGG16, ResNet18, AlexNet, Swin transformer, and the proposed model (Swin- POS-SVM) based on each class: monkeypox and others for MSLD dataset. We can see that Swin-POS-SVM records the highest precision, recall, and F1-score at 90.91, 100, and 95.24, respectively, for the monkeypox class compared to other models. DenseNet121 performs well in classifying others, with a recall of 88.00. AlexNet records the worst performance compared to other models, with 60 recalls for monkeypox. ResNet18 performs well for other pre-trained CNN models with 92.00 of recall. The Swin model achieves excellent results across two classes, with precision being 100 for monkeypox compared to pre-trained CNN.

Six confusion matrices are shown in Figure 5, each showing how various models performed when asked to categorize data into two classes: monkeypox and others. The number of samples classified into each class is indicated by the color intensity in each confusion matrix; darker colors indicate larger counts. Swin-POS-SVM has the highest TP, correctly identifying 20 cases as Monkeypox and 23 as Others.

Figure 6 includes several receiver operating characteristics (ROC) curves and models already applied and evaluated on the MSLD dataset. The existing models can be summarized as Swin-POS-SVM, Swin, ResNet18, Densenet121, AlexNet, and VGG. Each graph displays a subgraph of the ROC curve of one of the mentioned models with the AUC value on the opposite side. The Y-axis represents the true positive rates, while the X-axis represents the false positive rates. We can conclude the following observations and facts from the graphs: The Swin-POS-SVM model outperformed all other methods by achieving an area under the curve of 0.96. Swin,

ResNet18, Densenet121, and VGG also showed good results by achieving an area under the curve of 0.92, 0.86, 0.81, and 0.73, respectively. The model representing AlexNet’s performance showed the lowest results in the AUC curve, at about 0.68. In summary, the Swin-POS-SVM model achieved the best results among all models during comparison and evaluation, unlike AlexNet, which showed the lowest performance on the MSLD dataset.

2) PERFORMANCE OF MODELS WITH SGD OPTIMIZER

Table 5 shows the precision, recall, and F1-score of DenseNet121, VGG16, ResNet18, AlexNet, Swin transformer, and the proposed model (Swin-POS-SVM) based on each class: monkeypox and Others for MSLD dataset. We can see that Swin-POS-SVM records the highest results (95 recall and 90.48 of F1-score) for the monkeypox class compared to other models. DenseNet121 performs well in classifying others, with a recall of 88.00. AlexNet records the worst performance with 60.00 precision for monkeypox. ResNet18 performs well for others with 92.00 of recall. The Swin model achieves excellent results across two classes, with precision being 90 for monkeypox compared to pre-trained CNN.

Figure 7 shows six confusion matrices used to assess how well a binary classification task—possibly linked to differentiating between Monkeypox and Others is performed. ResNet18 has more mistakes in Monkeypox. DenseNet121 is comparable to ResNet18; however, it makes more errors in its monkeypox predictions. Overall, the models vary in effectiveness, with Swin-POS-SVM and Siwn showing better accuracy, particularly in distinguishing Monkeypox cases from others. The darker cells in the matrices indicate higher counts, with a scale bar showing the range of values.

Figure 8 shows the comparisons made on the MSLD dataset based on the SGD optimizer using ROC curves, which provide a visual representation that largely illustrates the balance between the true positive rates and the false positive rates, from which we conclude the following results: Swin-POS-SVM achieved the best performance with a result of 0.91. Although the algorithms Swin, RestNet18, and DenNet121 achieved acceptable results of 0.87, 0.81, and 0.77, respectively. While AlexNet and VGG achieved the lowest results with the same rate of 0.66.

C. COMPARISON OF EXISTING MODELS AND Swin-POS-SVM PERFORMANCE IN MSID

1) PERFORMANCE WITH ADAM OPTIMIZER

Table 6 shows the precision, recall, and F1-score of DenseNet121, VGG16, ResNet18, AlexNet, Swin, and the proposed model (Swin-POS-SVM) based on each class: Chickenpox, Measles, Monkeypox, and Normal for MSID dataset. We can see that Swin-POS-SVM recorded the highest precision, recall, and F1-score at 96.10, 100, and 98.01, respectively for the normal class. In addition, Swin-POS-SVM outperforms the pre-trained models across all categories, achieving the highest F1-scores, indicating a

TABLE 4. Performance models with Adam optimizer in each class for MSLD datasets.

Approaches	Models	Class Name	Precision	Recall	F1-score
Pre-trained models	DenseNet121	Monkeypox	83.33	75.00	78.95
		Others	81.48	88.00	84.62
	VGG16	Monkeypox	72.22	65.00	68.42
		Others	74.07	80.00	76.92
	ResNet18	Monkeypox	88.89	80.00	84.21
		Others	85.19	92.00	88.46
	AlexNet	Monkeypox	66.67	60.00	63.16
		Others	70.37	76.00	73.08
Transformer	Swin transformer	Monkeypox	83.33	100	90.91
		Others	100	84.00	91.30
The proposed models	Swin-POS-SVM	Monkeypox	90.91	100	95.24
		Others	100	92.00	95.83

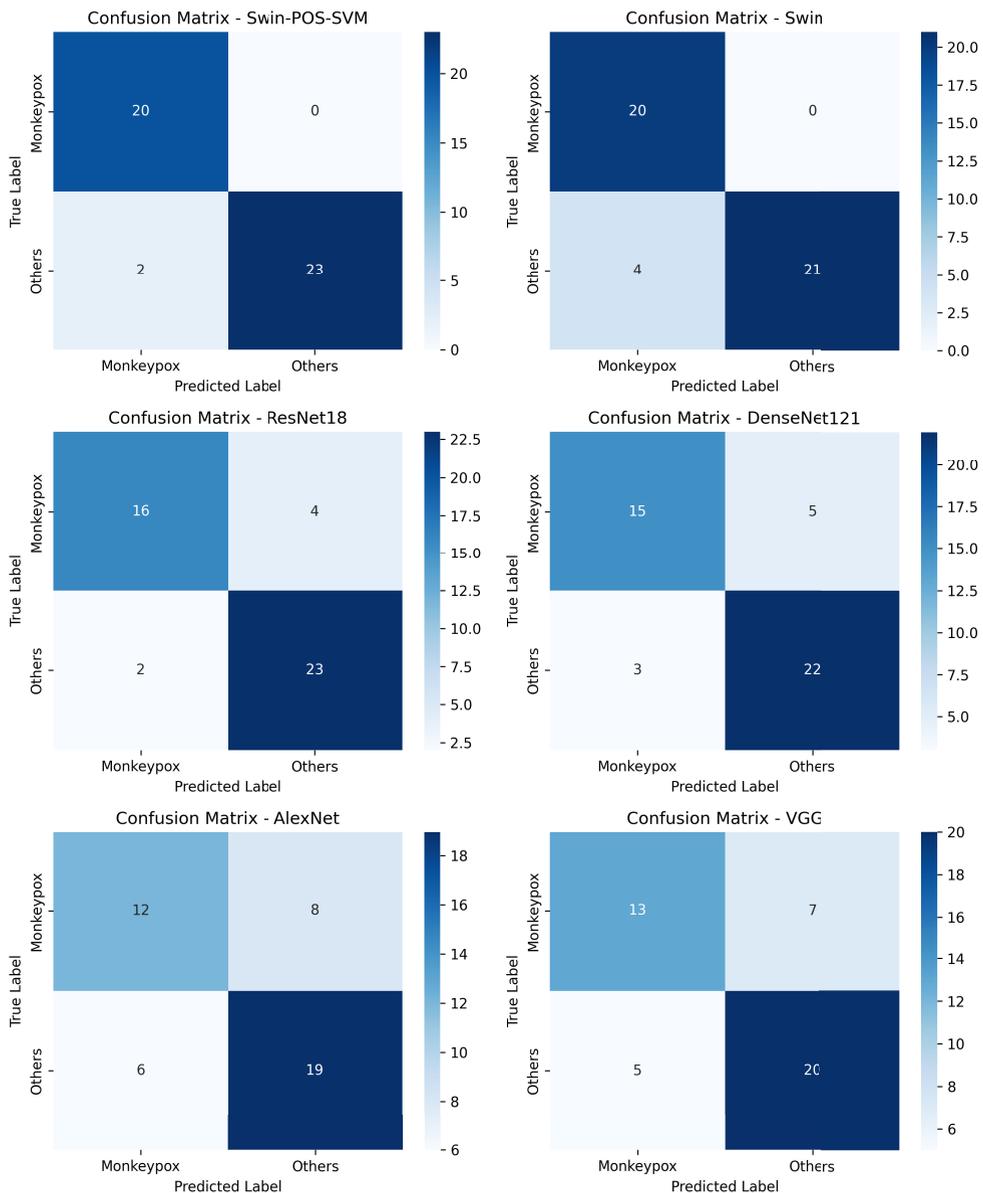


FIGURE 5. Confusion matrices of models with Adam optimizer for MSLD dataset.

more balanced performance regarding precision and recall. DenseNet121 performs well in classifying Monkeypox and

Normal, with F1-scores of 93.75, but it records the lowest recall for Measles with 62.50. ResNet18 performs extremely

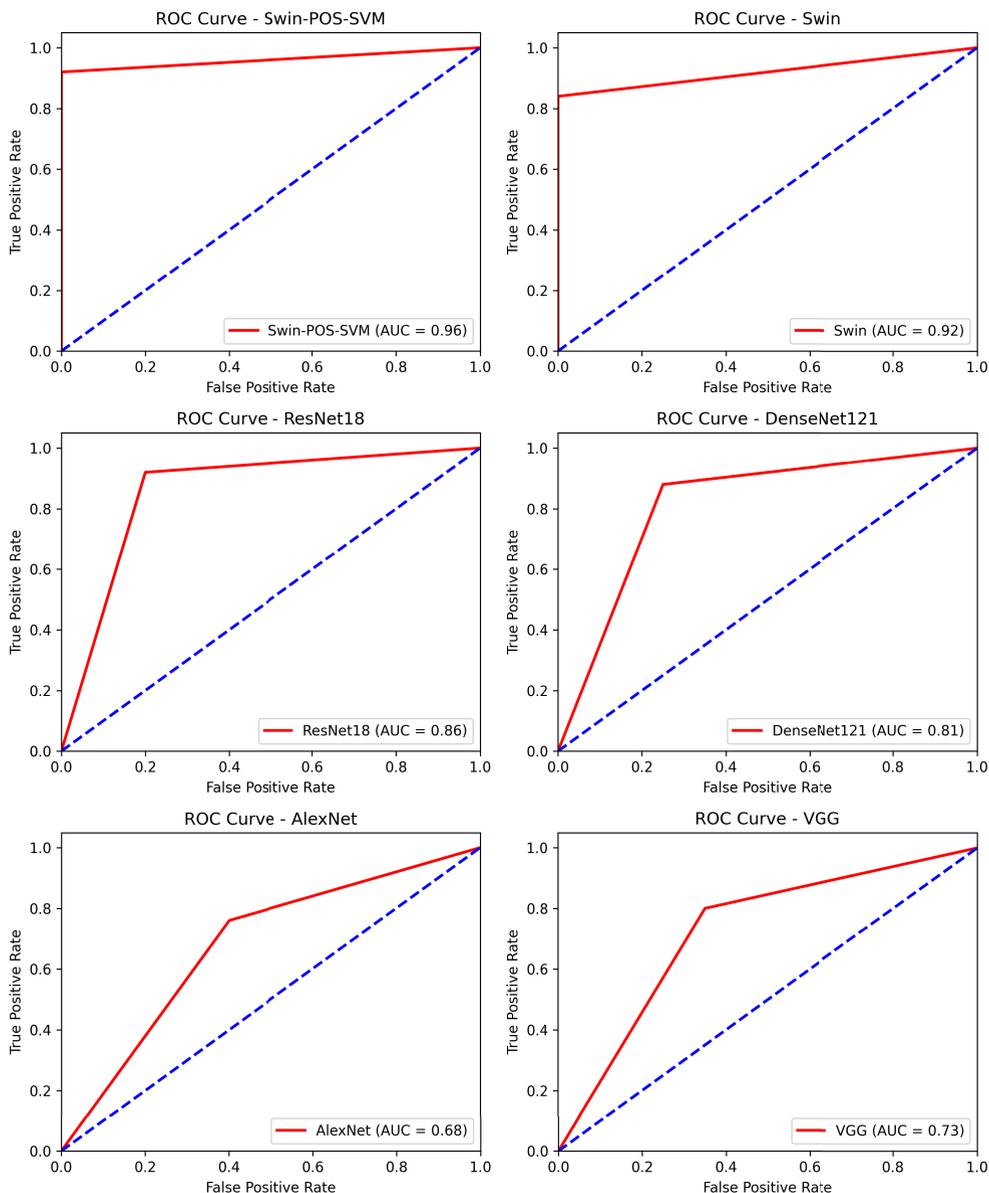


FIGURE 6. ROC curves of models with Adam optimizer for MSLD dataset.

TABLE 5. The results of MSLD datasets with SGD optimizer.

Approaches	Models	Class Name	Precision	Recall	F1-score
Pre-trained models	DenseNet121	Monkeypox	81.25	65.00	72.22
		Others	75.86	88.00	81.48
	VGG16	Monkeypox	80.00	40.00	53.33
		Others	65.71	92.00	76.67
	ResNet18	Monkeypox	87.50	70.00	77.78
		Others	79.31	92.00	85.19
	AlexNet	Monkeypox	63.16	60.00	61.54
		Others	69.23	72.00	70.59
Transformer	Swin transformer	Monkeypox	81.82	90.00	85.71
		Others	91.30	84.00	87.50
The proposed models	Swin-POS-SVM	Monkeypox	86.36	95.00	90.48
		Others	95.65	88.00	91.67

well for Monkeypox and Normal with F1-scores above 90. It also performs the best for Chickenpox and Measles

with more heightened recall results than DenseNet121 and VGG16. The Swin model achieves excellent results across

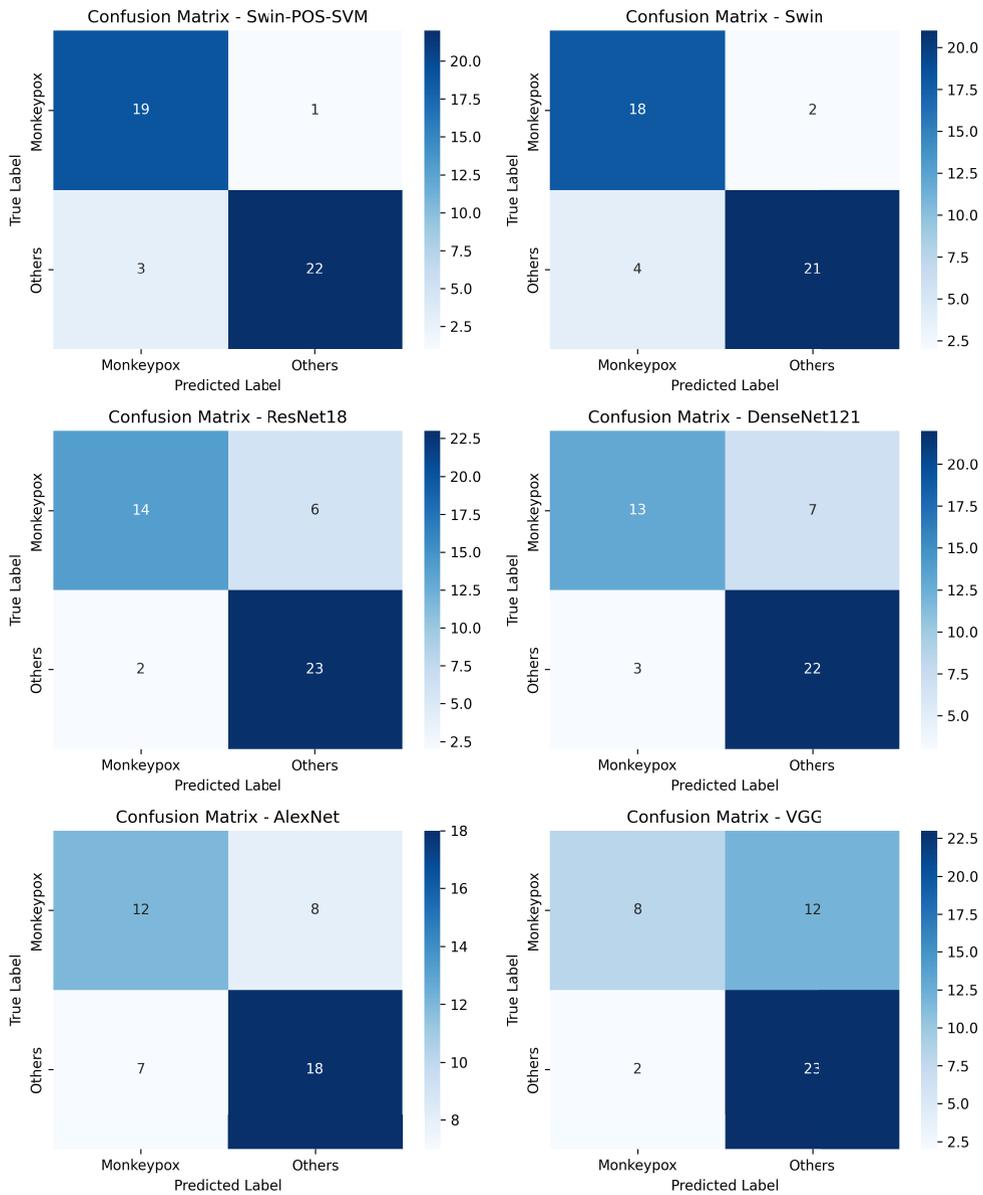


FIGURE 7. Confusion matrices with SGD optimizer for MSLD dataset.

all classes, with the highest F1-score being 97 and the highest recall being 100 compared to pre-trained CNN.

Figure 9 shows the confusion matrices for multiple models implementing the Adam optimizer on the MSID dataset present an in-depth overview of model performance across four classes: Chickenpox, Measles, Monkeypox, and Normal. The models vary in their efficacy, with Swin-POS-SVM and Swin Transformer exceeding the others in discriminating between classes, particularly Chickenpox and “Normal. The ResNet18 and DenseNet121 models function well, yet they have some flaws in Monkeypox while VGG16 and AlexNet confront the most severe challenges, particularly with Measles and Monkeypox, noticing areas for growth for these models.

Figure 10 shows ROC curves of models with the Adam MSLD dataset. The performance of the algorithms was compared in terms of their ability to identify the diseases chickenpox, measles, monkeypox, and normal. From the figure, we can deduce the following observations: For Chickenpox and Measles, Swin-POS-SVM achieved the best results with a rate of 0.96 for both, while AlexNet achieved the lowest with a rate not exceeding 0.74 and 0.68, respectively. As for Monkeypox, Swin-POS-SVM is still on top with a result of 0.97, and AlexNet is still the lowest, but this time it reached 0.86. As for the normal diagnosis, Swin-POS-SVM also had the best results, reaching 0.99, and the lowest results were for AlexNet, which achieved only 0.88.

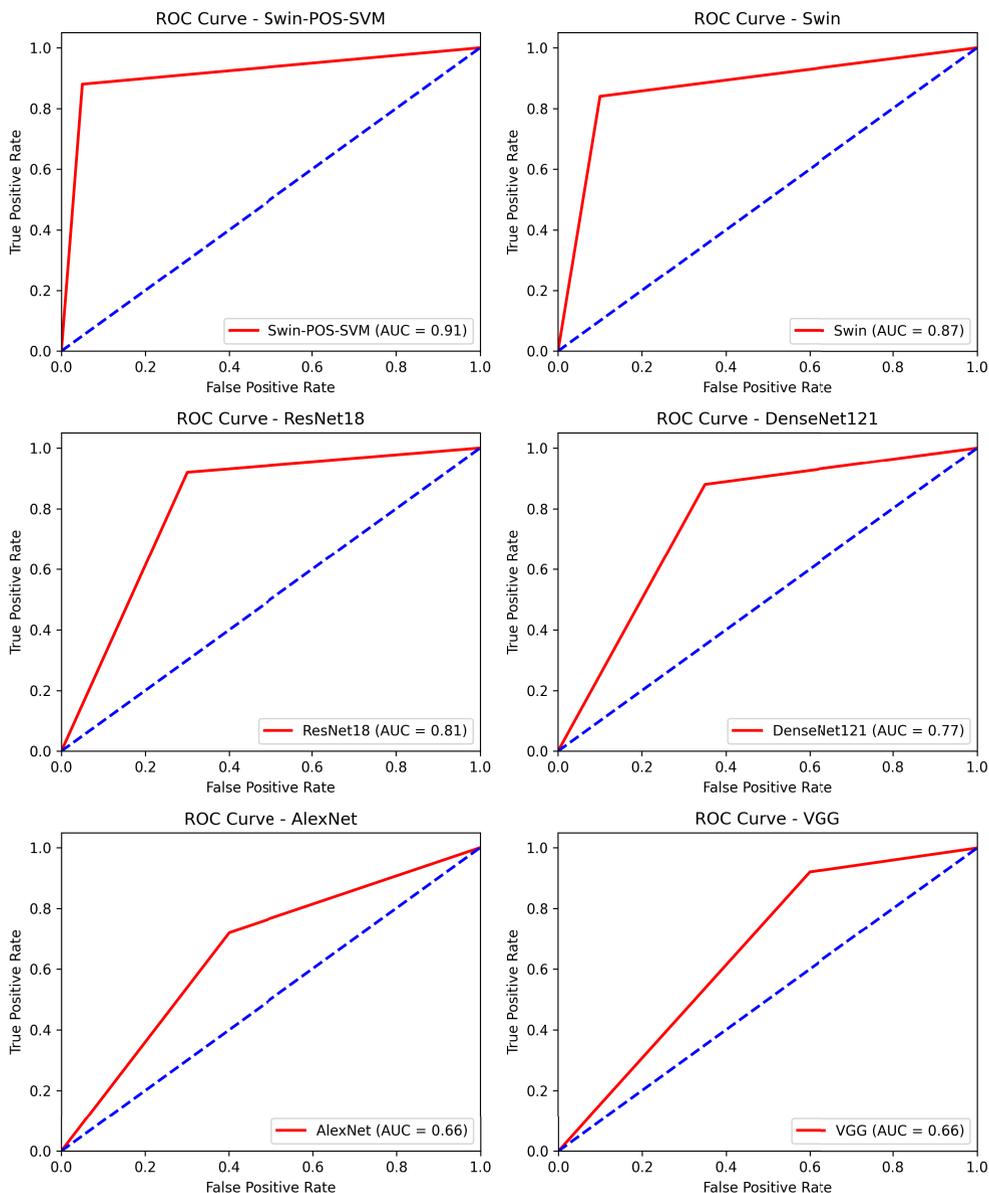


FIGURE 8. ROC curves with SGD optimizer for MSLD dataset.

2) PERFORMANCE WITH THE SGD OPTIMIZER

Table 7 shows the precision, recall, and F1-score of models based on each class: Chickenpox, Measles, Monkeypox, and Normal for MSID dataset. We can see that Swin-POS-SVM recorded the highest recall, and F1-score at 99, and 97, respectively for the normal class. DenseNet121 has the lowest recall for measles (58), but it does well in classifying monkeypox and normal (96 recall for normal). ResNet18 exhibits F1-scores above 90 and performs exceptionally well for Normal and Monkeypox. In addition, it outperforms DenseNet121 and VGG16 in terms of heightened recall results for measles and chickenpox. Comparing the Swin model to pre-trained CNN, it obtains superior performance

in all classes, with the greatest recall of 97 and the highest F1-score of 95.

Confusion matrices with models trained by the SDG optimizer on the MSID dataset, which involve the Swin-POS-SVM, Swin, ResNet18, VGG16, DenseNet121, and AlexNet models as shown in Figure 11, compared with the predicted labels for the four classes: chickenpox, measles, monkeypox, and normal. The values shown in the matrices revealed that Swin-POS-SVM and Swin perform well with high TP rates, especially for the Monkeypox and Normal classes, with TP = 25 and 22 for Monkeypox and TP = 73 and 72 for Normal classes. Again, both VGG16 and AlexNet got the lowest results for chickenpox, with TP = 19 and 15,

TABLE 6. The results of models with Adam optimizer in each class for MSID dataset.

Approaches	Model	Class Name	Precision	Recall	F1-score
Pre-trained models	DenseNet121	Chickenpox	70.97	78.57	74.58
		Measles	75.00	62.50	68.18
		Monkeypox	93.75	85.71	89.55
		Normal	87.65	95.95	91.61
	VGG16	Chickenpox	63.33	67.86	65.52
		Measles	72.22	54.17	61.90
		Monkeypox	90.77	84.29	87.41
		Normal	84.34	94.59	89.17
	ResNet18	Chickenpox	69.70	82.14	75.41
		Measles	85.00	70.83	77.27
		Monkeypox	95.38	88.57	91.85
		Normal	89.74	94.59	92.11
	AlexNet	Chickenpox	51.61	57.14	54.24
		Measles	69.23	37.50	48.65
		Monkeypox	88.52	77.14	82.44
		Normal	75.82	93.24	83.64
Transformer	Swin transformer	Chickenpox	85.19	82.14	83.64
		Measles	100.00	83.33	90.91
		Monkeypox	94.37	95.71	95.04
		Normal	94.87	100.00	97.37
The proposed models	Swin-POS-SVM	Chickenpox	89.66	92.86	91.23
		Measles	100.00	91.67	95.65
		Monkeypox	98.53	95.71	97.10
		Normal	96.10	100.00	98.01

TABLE 7. The results of models with SDG optimizer in each class for MSID dataset.

Approaches	Model	Class Name	Precision	Recall	F1-score
Pre-trained models	DenseNet121	Chickenpox	67.74	75.00	71.19
		Measles	70.00	58.33	63.64
		Monkeypox	90.62	82.86	86.57
		Normal	87.65	95.95	91.61
	VGG16	Chickenpox	59.38	67.86	63.33
		Measles	66.67	50.00	57.14
		Monkeypox	89.23	82.86	85.93
		Normal	83.95	91.89	87.74
	ResNet18	Chickenpox	65.71	82.14	73.02
		Measles	80.95	70.83	75.56
		Monkeypox	95.00	81.43	87.69
		Normal	88.75	95.95	92.21
	AlexNet	Chickenpox	50.00	53.57	51.72
		Measles	61.54	33.33	43.24
		Monkeypox	86.89	75.71	80.92
		Normal	75.00	93.24	83.13
Transformer	Swin transformer	Chickenpox	88.00	78.57	83.02
		Measles	95.24	83.33	88.89
		Monkeypox	91.67	94.29	92.96
		Normal	92.31	97.30	94.74
The proposed models	Swin-POS-SVM	Chickenpox	83.33	89.29	86.21
		Measles	100.00	87.50	93.33
		Monkeypox	97.06	94.29	95.65
		Normal	94.81	98.65	96.69

respectively, while for monkeypox, TP was 58 for VGG16 and 53 for AlexNet.

Figure 12 shows ROC curves of models with the SDG MSLD dataset. From the figure, we can deduce the following observations: For Chickenpox, Measles, and Monkeypox, Swin-POS-SVM achieved the best results with a rate of 0.93, 0.94, and 0.96, respectively, while AlexNet achieved the lowest with a rate not exceeding 0.72, 0.65, and 0.85, respectively. As for the diagnosis of Normal, Swin-POS-SVM also had the best results, reaching 0.98,

while AlexNet also had the lowest results, achieving only 0.87.

V. DISCUSSION

We conduct various experiments for two classifications and four classifications using two image datasets. Two classification tasks include compression between pre-trained CNN models, the Swin transformer, and the Swin-POS-SVM with Adam optimizer and SGD optimizer to classify Monkeypox and Others using the MSLD dataset.

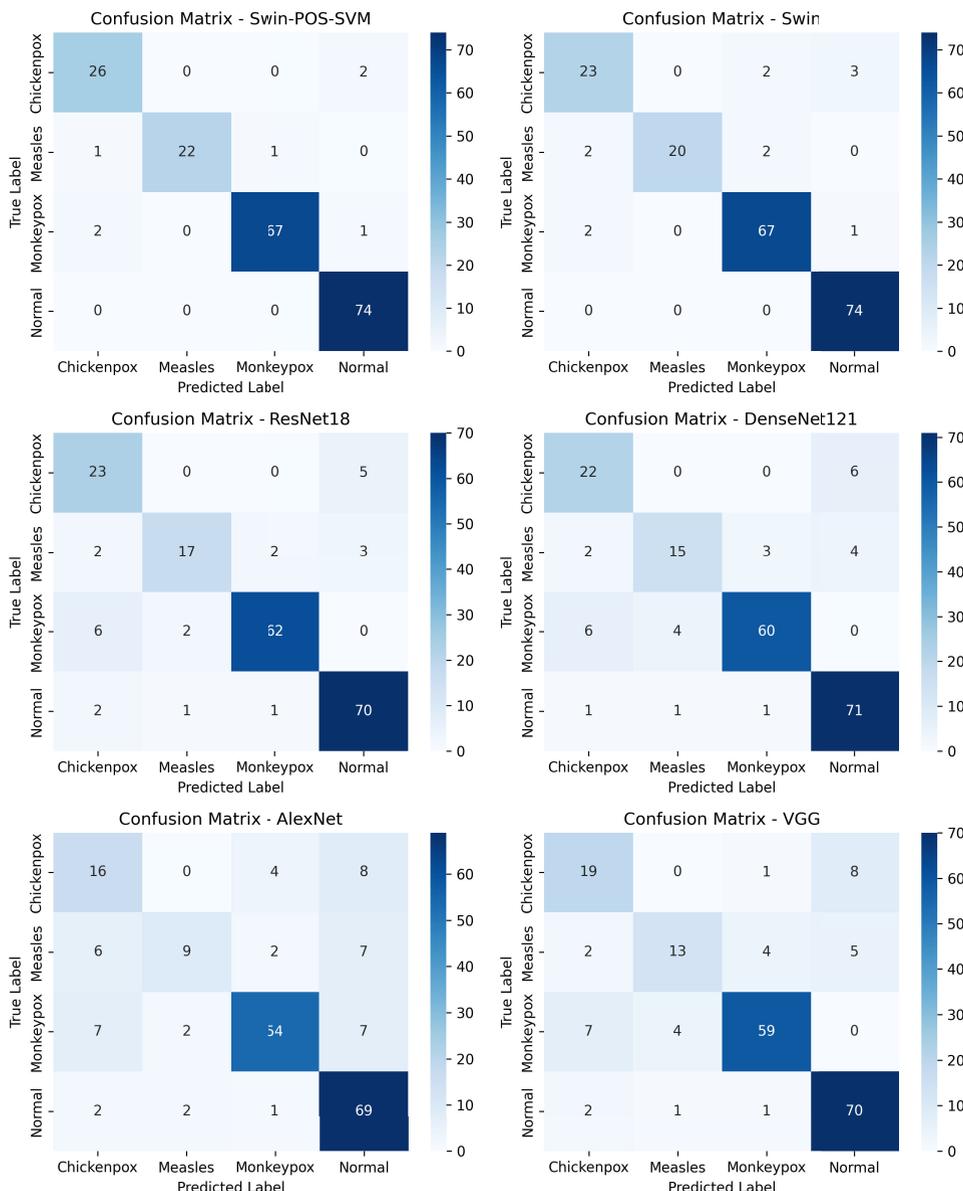


FIGURE 9. Confusion matrices of models with Adam optimizer for MSID dataset.

Swin-POS-SVM is built based on extracted features from Swin, selecting the best features from extracted features by POS and replacing softmax with SVM.

1) AVERAGE PERFORMANCE FOR MSLD

Figure 13 shows models’ average performance with Adam optimizer. Swin-POS-SVM records the highest performance (accuracy of 95.556 and F1-score of 95.569), indicating that the combination offers the best solution. Swin transformer has the second-highest performance with 91.111 accuracy and 93.783 F1-score. ResNet18 records the highest performance with 86.667 accuracy and 86.572 F1-score compared to pre-trained CNN models. AlexNet records the worst on all measures (accuracy of 68.889). Figure 14

shows models’ average performance with SGD optimizer. Swin-POS-SVM records the highest performance (accuracy of 91.111 and F1-score of 91.138), indicating that this combination offers the best solution. Swin transformer has the second-highest performance with 86.667 accuracy and 86.706 F1-score. ResNet18 records the highest performance with 82.222 accuracy and 81.893 F1-score compared to pre-trained CNN models. AlexNet records the worst on all measures (accuracy of 66.667).

2) AVERAGE PERFORMANCE FOR MSID

Figure 15 shows the average accuracy, precision, recall, and F1-score of models for MSID. The proposed model combines the Swin transformer with POS and SVM outperforms the

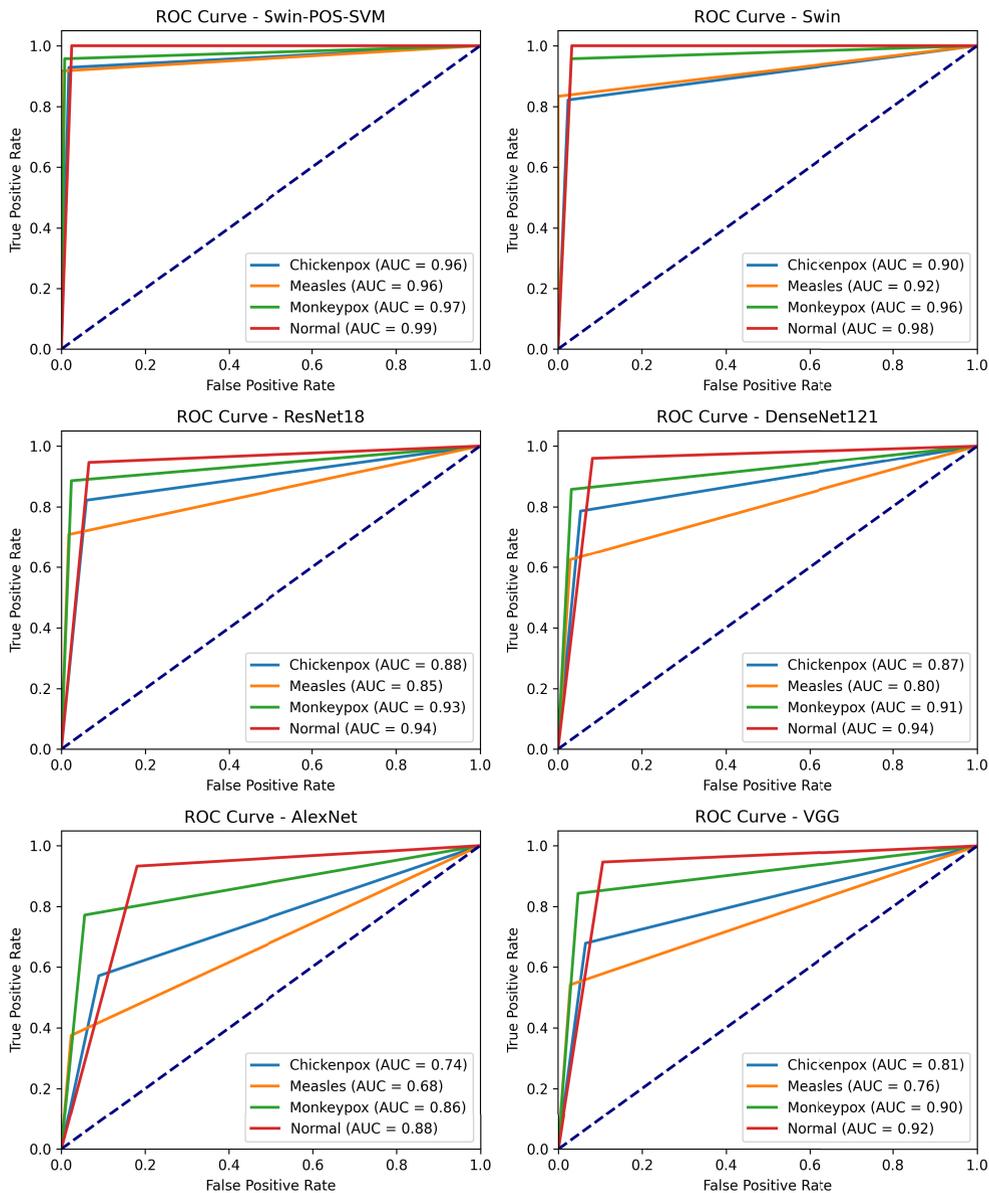


FIGURE 10. ROC curves of models with Adam optimizer for MSID dataset.

best on all measures (accuracy of 96.429 and F1-score of 96.429), indicating that the Swin-POS-SVM model offers the best solution. For pre-trained CNN models, ResNet18 recorded the highest performance with 87.755 accuracy and 87.813 F1-score. AlexNet recorded the worst on all measures (accuracy of 75.51). Swin transformer has the best performance compared to pre-trained CNN models with 93.878 accuracy and 93.783 of F1-score.

Figure 16 shows the average accuracy, precision, recall, and F1-score of models for MSID. The proposed model (Swin-POS-SVM) that combines the Swin transformer with POS and SVM outperforms all other models and records the best on all measures (accuracy of 94.388 and F1-score of 94.41), indicating that this combination offers

the best solution. For pre-trained CNN models, ResNet18 recorded the highest performance with 85.714 accuracy and 85.814 F1-score. AlexNet recorded the worst of all measures (accuracy of 73.98). Swin transformer has the best performance compared to pre-trained CNN models with 91.837 accuracy and 91.711 of F1-score. Figures 13, 14 and 15 show the average accuracy, precision, recall, and F1-score of models for MSLD and MSID, respectively. We can see that the proposed model (Swin-POS-SVM) that combines the Swin transformer with POS and SVM outperforms all other models and records the best on all measures. In addition, models with the Adam optimizer record the best performance compared to models with the SDG optimizer.

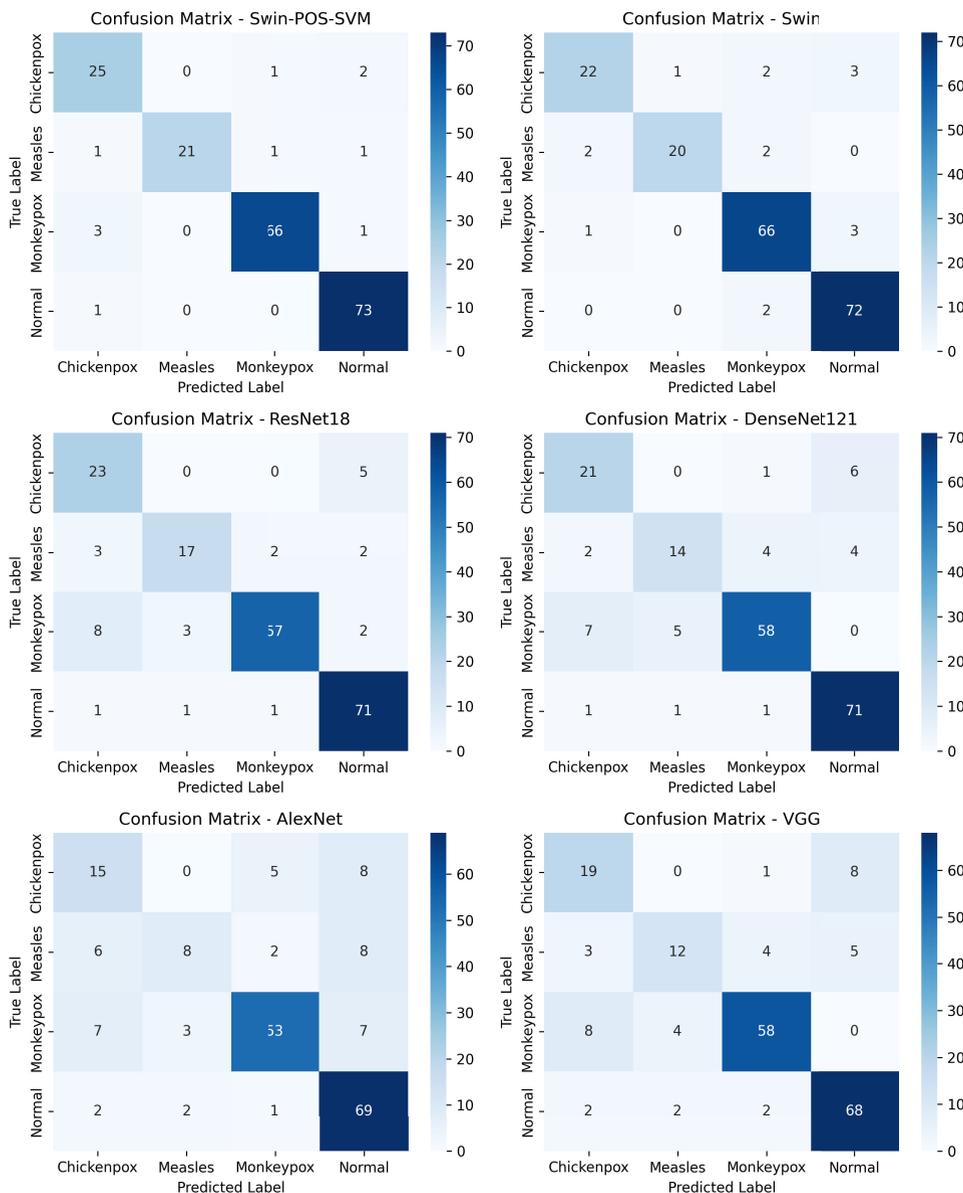


FIGURE 11. Confusion matrices of models with SGD optimizer for MSID dataset.

A. STATISTICAL ANALYSIS

The best results of two datasets were recorded by Adam optimizer; we approve the results by the Nemenyi test as statistical analysis. Specifically, the Nemenyi test allowed us to identify specific pairs of models that exhibited statistically significant differences in their performance, providing valuable insight into their relative strengths and weaknesses. The Nemenyi test represents the difference in terms of critical difference. To enhance the interpretability of the results, we created a critical difference diagram (CD diagram) as depicted in Figure 17 and 18. The CD diagram, a widely used visualization tool in multiple comparison analyses, highlighted the significant differences between the models and statistical methods often used to compare the performance of various classifiers or models.

By displaying the average ranks of the models, it provided a clear and concise representation of their performance disparities. Figure 17 shows the CD of models for MSLD. The x-axis shows ranks from 1 to 7. Lower ranks (closer to 1) indicate better performance, while higher ranks (closer to 7) indicate worse performance. Figure 18 shows the CD of models for MSID. We can see that Swin-POS-SVM has 1 indicate better performance, while AlexNet has 6 indicate worse performance.

B. COMPARISON WITH EXISTING WORK AND THE PROPOSED MODELS BASED ON MODELS AND PERFORMANCE

We compared the proposed models based on models with binary classification and non-augmented testing sets and

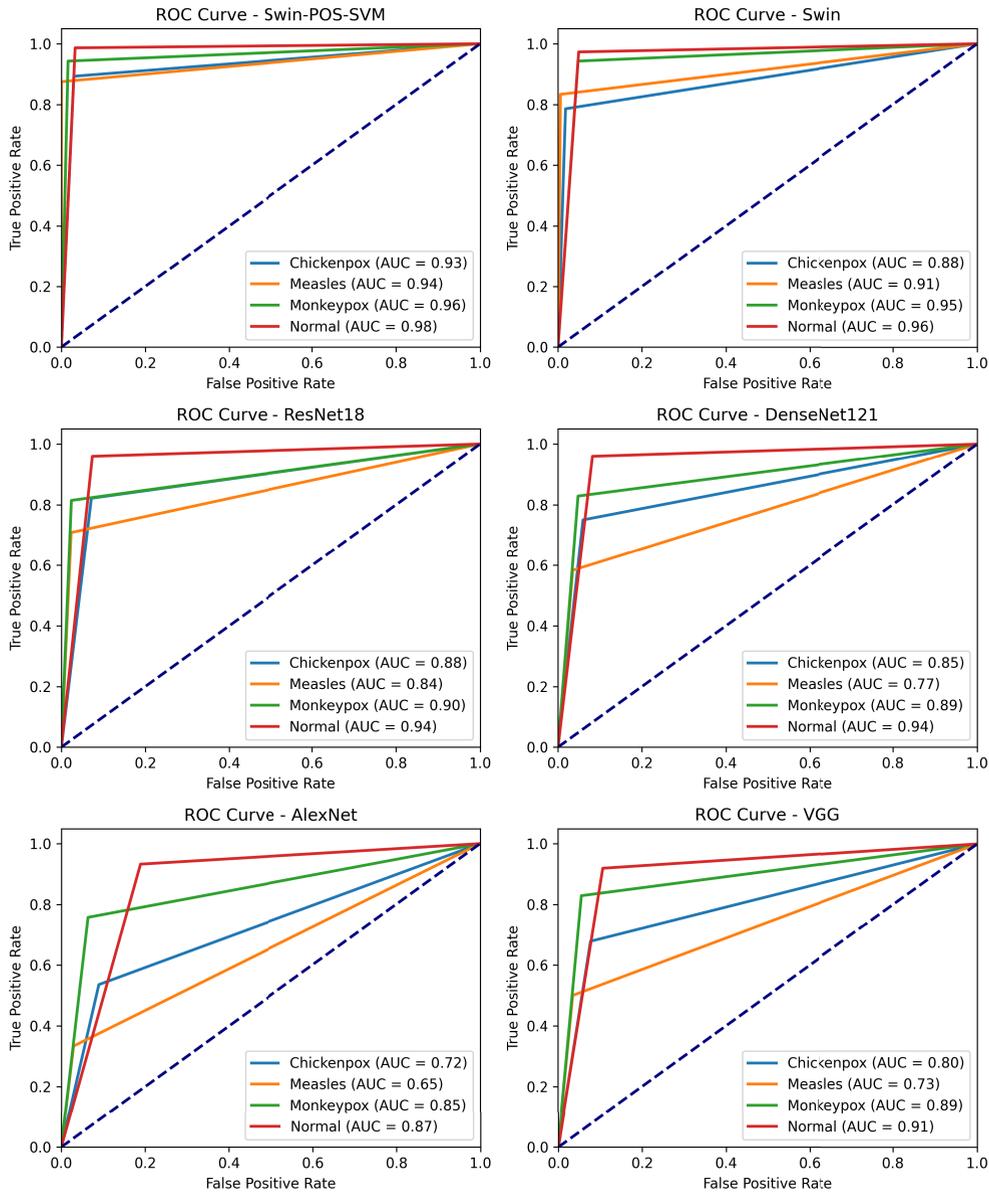


FIGURE 12. ROC curves of models with Adam optimizer for MSID dataset.

TABLE 8. Comparison with existing work and the proposed models based on models and performance.

Papers	Models	Datasets	Accuracy	Precision	Recall	F1-score
[13]	ResNet50	MSLD	82.96	87	84	83
[18]	NB with VGG16Net	MSLD	91.11	93.1	93.35	93.15
[19]	Xception-CBAM-Dense layers	MSLD	83.89	90.70	89.10	90.11
[20]	MobileNetv2	MSLD	91	90	90	90
[21]	SENet	MSID	91.19	90.43	90.43	92.55
[22]	MobileNetV2	MSID	91.38	90.5	86.75	88.25
Our work	Swin-POS-SVM	MSLD	95.556	95.96	95.556	95.569
Our work	Swin-POS-SVM	MSID	96.429	96.526	96.429	96.429

four classes with non-augmented testing sets. In addition, we do not compare our work with other studies that applied models with three classes and an augmented testing set. For example, The authors in [16] and [23] evaluated models on an

augmented testing set. In [17] were evaluated models using an augmented testing set. Table 8 compares our Swin-POS-SVM model with the existing work. Our experiments show that the proposed Swin-POS-SVM achieves an accuracy

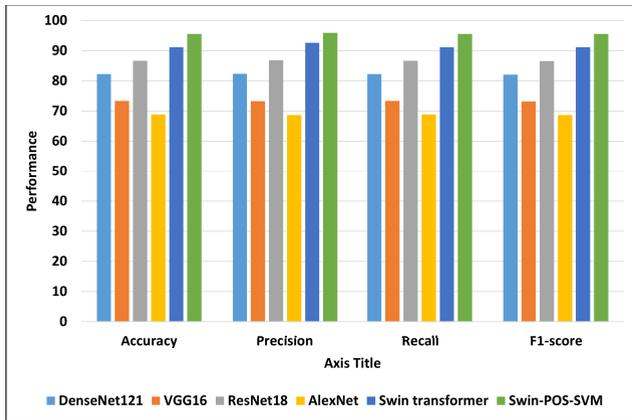


FIGURE 13. Models performance average with Adam optimizer for MSLD.

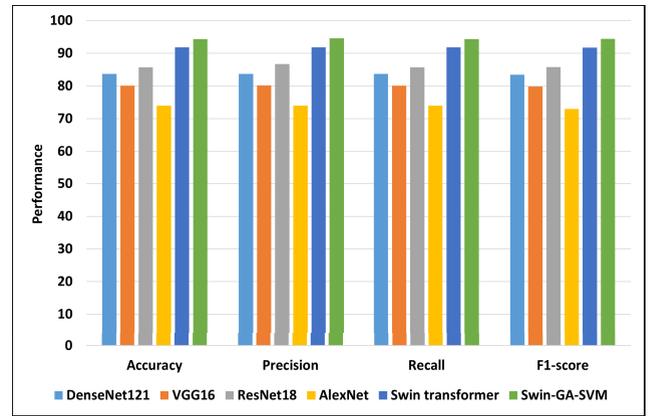


FIGURE 16. Models performance average with SDG optimizer for MSID.

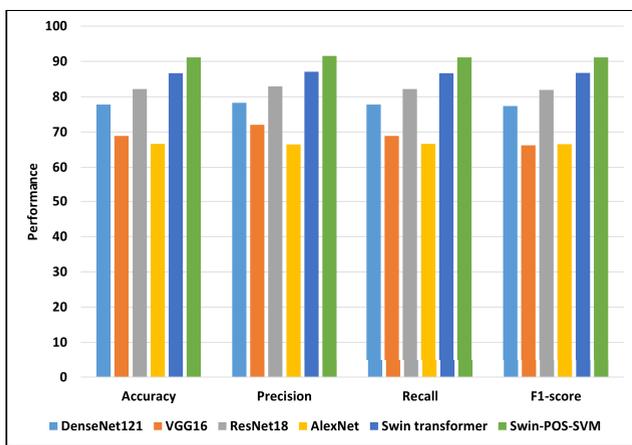


FIGURE 14. Models performance average with SGD optimizer for MSLD.

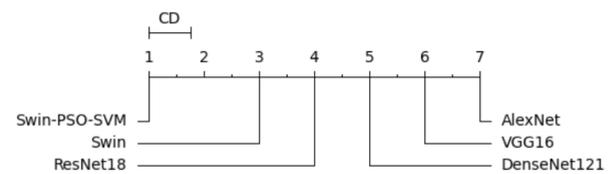


FIGURE 17. CD diagram of models for MSID.

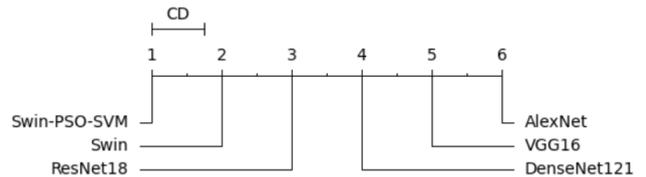


FIGURE 18. CD diagram of models for MSID.

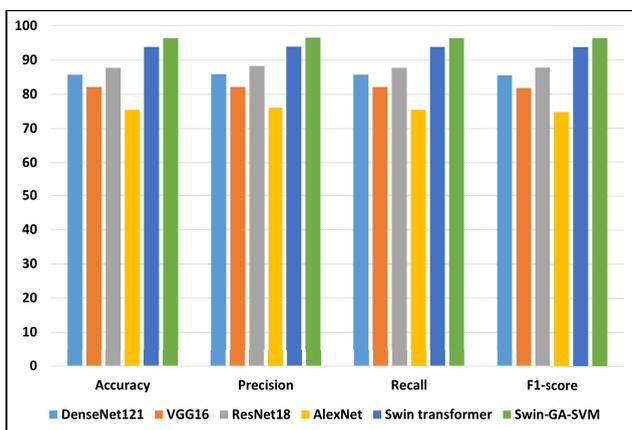


FIGURE 15. Models performance average with Adam optimizer for MSID.

of 95.556 in classifying two classes ,i.e., monkey and others in MSLD, and four classes, i.e., monkeypox, chickenpox, measles, and routine. For MSLD, ResNet50 was used by the authors in In [13], and its accuracy, precision, recall, and F1-score were 82.96, 87, 84, and 83, respectively. In [18], authors applied NB with VGG16Net and recorded

91.11 accuracy and 93.15 F1-score. Xception-CBAM-Dense layers in [19] and recorded with 83.89 precision. The authors used MobileNetv2 in [20], and its precision, precision, recall, and F1 score were 91, 90, 90, and 90, respectively. Swin-POS-SVM recorded the highest accuracy rate with 95.556. For MSID, In [21], SENet was obtained accuracy = 91.19, precision=90.43, recall = 90.43, F1-score = 92.55. In [22], MobileNetV2 recorded at 91.38, 90.5, 86.75, and 88.25, accuracy, precision, recall, and F1-score, respectively. Swin-POS-SVM recorded the highest accuracy rate with 95.556.

In summary, The proposed algorithm is the “hybrid Swin-PSO-SVM model” for diagnosing Monkeypox and improving classification performance. The Swin Transformer can extract both basic and complex features. POS reduces the dimension of Swin’s extracted features. When used for classification, SVM allows operating on the Swin Transformer outputs rich in basic features and thus. The proposed Swin-PSO-SVM is designed for monkeypox detection and diagnosis, offering a robust solution for Monkeypox. The proposed algorithm provides high diagnostic accuracy, flexibility, and improved performance. Nevertheless, the model requires a lot of computing power, might be challenging to use, runs the risk

of overfitting, and requires more extended training periods. Furthermore, the Swin Transformer increases interpretability, a critical component in medical applications.

VI. CONCLUSION

We introduced a novel hybrid Swin-PSO-SVM to improve the performance and dependability of monkeypox detection. The Swin-PSO-SVM model incorporates a Swin Transformer for complex feature extraction, a PSO for extracting the best features from complex feature extraction, and a SVM for accurate classification. Swin-PSO-SVM compared with different models: DenseNet121, VGG19, ResNet18, AlexNet, and Swin transformer using two datasets, which are MSLD with two classes (monkeypox and others) and MSID with four classes (chickenpox, measles, monkeypox, and normal). The Swin-PSO-SVM model has achieved the highest performance in terms of 95.556 accuracy, 95.96 precision, 95.556 recall, and 95.569 F1-score on the MSLD dataset and 96.429 accuracy, 96.526 precision, 96.429 recall, and 96.429 F1-score on the MSID dataset. The experimental findings proved the model's exceptional performance and dependability in real-world applications. The Swin-PSO-SVM enhanced the generalizability of monkeypox detection by offering a practical and understandable approach that is simple to implement in clinical settings. This study supports global initiatives to prevent monkeypox outbreaks through accurate and prompt identification. In the future, we plan to develop the proposed Swin-PSO-SVM for detecting multi-pox diseases

DATA AVAILABILITY

All datasets used to support the findings of this study are available from the direct link in the dataset citations

CONFLICT OF INTERESTS

All authors declare that they have no conflicts of interest.

ETHICAL APPROVAL

Not Applicable.

REFERENCES

- [1] A. Singh, C. Vashist, P. Gaurav, and A. Nigam, "A generic framework for deep incremental cancelable template generation," *Neurocomputing*, vol. 467, pp. 83–98, Jan. 2022.
- [2] S. Kumar, M. H. Alsamhi, S. Kumar, A. V. Shvetsov, and S. H. Alsamhi, "Early MTS forecasting for dynamic stock prediction: A double Q-learning ensemble approach," *IEEE Access*, vol. 12, pp. 69796–69811, 2024.
- [3] K. Sundararajan and D. L. Woodard, "Deep learning for biometrics: A survey," *ACM Comput. Surv.*, vol. 51, no. 3, pp. 1–34, May 2018.
- [4] S. Kumar, M. K. Chaube, S. H. Alsamhi, S. K. Gupta, M. Guizani, R. Gravina, and G. Fortino, "A novel multimodal fusion framework for early diagnosis and accurate classification of COVID-19 patients using X-ray images and speech signal processing techniques," *Comput. Methods Programs Biomed.*, vol. 226, Nov. 2022, Art. no. 107109.
- [5] A. AlMohimeed, R. M. A. Saad, S. Mostafa, N. M. El-Rashidy, S. Farrag, A. Gaballah, M. A. Elaziz, S. El-Sappagh, and H. Saleh, "Explainable artificial intelligence of multi-level stacking ensemble for detection of Alzheimer's disease based on particle swarm optimization and the sub-scores of cognitive biomarkers," *IEEE Access*, vol. 11, pp. 123173–123193, 2023.
- [6] S. Kumar, A. V. Shvetsov, and S. H. Alsamhi, "Fuzzyguard: A novel multimodal neuro-fuzzy framework for cpd early diagnosis," *IEEE Internet Things J.*, early access, Sep. 24, 2024, doi: [10.1109/JIOT.2024.3467176](https://doi.org/10.1109/JIOT.2024.3467176).
- [7] M. A. A. Al-Qaness, J. Zhu, D. Al-Alimi, A. Dahou, S. H. Alsamhi, M. A. Elaziz, and A. A. Ewees, "Chest X-ray images for lung disease detection using deep learning techniques: A comprehensive survey," *Arch. Comput. Methods Eng.*, vol. 31, no. 6, pp. 3267–3301, Aug. 2024.
- [8] A. AlMohimeed, H. Saleh, S. Mostafa, R. M. A. Saad, and A. S. Talaat, "Cervical cancer diagnosis using stacked ensemble model and optimized feature selection: An explainable artificial intelligence approach," *Computers*, vol. 12, no. 10, p. 200, Oct. 2023.
- [9] I. Arshad, S. H. Alsamhi, Y. Qiao, B. Lee, and Y. Ye, "A novel framework for smart cyber defence: A deep-dive into deep learning attacks and defences," *IEEE Access*, vol. 11, pp. 88527–88548, 2023.
- [10] J. Yu, A. V. Shvetsov, and S. H. Alsamhi, "Leveraging machine learning for cybersecurity resilience in industry 4.0: Challenges and future directions," *IEEE Access*, vol. 12, pp. 159579–159596, 2024.
- [11] D. Bala, M. S. Hossain, M. A. Hossain, M. I. Abdullah, M. M. Rahman, B. Manavalan, N. Gu, M. S. Islam, and Z. Huang, "MonkeyNet: A robust deep convolutional neural network for monkeypox disease detection and classification," *Neural Netw.*, vol. 161, pp. 757–775, Apr. 2023.
- [12] S. U. R. Khan, S. Asif, O. Bilal, and S. Ali, "Deep hybrid model for mpxo disease diagnosis from skin lesion images," *Int. J. Imag. Syst. Technol.*, vol. 34, no. 2, p. 23044, Mar. 2024.
- [13] S. N. Ali, M. T. Ahmed, J. Paul, T. Jahan, S. M. S. Sani, N. Noor, and T. Hasan, "Monkeypox skin lesion detection using deep learning models: A feasibility study," 2022, *arXiv:2207.03342*.
- [14] A. S. Jaradat, R. E. Al Mamlouk, N. Almkayee, N. Alharbe, A. S. Almuflhi, A. Nasayreh, H. Gharaibeh, M. Gharaibeh, A. Gharaibeh, and H. Bzizi, "Automated monkeypox skin lesion detection using deep learning and transfer learning techniques," *Int. J. Environ. Res. Public Health*, vol. 20, no. 5, p. 4422, Mar. 2023.
- [15] I. Pacal, "A novel swin transformer approach utilizing residual multi-layer perceptron for diagnosing brain tumors in MRI images," *Int. J. Mach. Learn. Cybern.*, vol. 15, no. 9, pp. 3579–3597, Sep. 2024.
- [16] T. Nayak, K. Chadaga, N. Sampathila, H. Mayrose, N. Gokulkrishnan, M. G. Bairy, S. Prabhu, K. S. Swathi, and S. Umakanth, "Deep learning based detection of monkeypox virus using skin lesion images," *Med. Novel Technol. Devices*, vol. 18, Jun. 2023, Art. no. 100243.
- [17] S. Surati, H. Trivedi, B. Shrimali, C. Bhatt, and C. M. Travieso-González, "An enhanced diagnosis of monkeypox disease using deep learning and a novel attention model senet on diversified dataset," *Multimodal Technol. Interact.*, vol. 7, no. 8, p. 75, Jul. 2023.
- [18] V. Kumar, "Analysis of CNN features with multiple machine learning classifiers in diagnosis of monkeypox from digital skin images," *MedRxiv*, pp. 1–4, Sep. 2022, doi: [10.1101/2022.09.11.22278797](https://doi.org/10.1101/2022.09.11.22278797).
- [19] M. E. Haque, M. R. Ahmed, R. S. Nila, and S. Islam, "Classification of human monkeypox disease using deep learning models and attention mechanisms," 2022, *arXiv:2211.15459*.
- [20] V. H. Sahin, I. Oztel, and G. Y. Oztel, "Human monkeypox classification from skin lesion images with deep pre-trained network using mobile application," *J. Med. Syst.*, vol. 46, no. 11, p. 79, Oct. 2022.
- [21] T. Nayak, K. Chadaga, N. Sampathila, H. Mayrose, G. M. Bairy, S. Prabhu, S. S. Katta, and S. Umakanth, "Detection of monkeypox from skin lesion images using deep learning networks and explainable artificial intelligence," *Appl. Math. Sci. Eng.*, vol. 31, no. 1, Dec. 2023, Art. no. 2225698.
- [22] M. C. Irmak, T. Aydin, and M. Yağanoğlu, "Monkeypox skin lesion detection with MobileNetV2 and VGGNet models," in *Proc. Med. Technol. Congr. (TIPEKNO)*, Oct. 2022, pp. 1–4.
- [23] M. Altun, H. Gürüler, O. Özkaraca, F. Khan, J. Khan, and Y. Lee, "Monkeypox detection using CNN with transfer learning," *Sensors*, vol. 23, no. 4, p. 1783, Feb. 2023.
- [24] F. Uysal, "Detection of monkeypox disease from human skin images with a hybrid deep learning model," *Diagnostics*, vol. 13, no. 10, p. 1772, May 2023.
- [25] M. M. Ahsan, M. S. Ali, M. M. Hassan, T. A. Abdullah, K. D. Gupta, U. Bagci, C. Kaushal, and N. F. Soliman, "Monkeypox diagnosis with interpretable deep learning," *IEEE Access*, vol. 11, pp. 81965–81980, 2023.
- [26] (2024). *Monkeypox Skin Lesion Dataset*. Accessed: 2024. [Online]. Available: <https://www.kaggle.com/datasets/nafin59/monkeypox-skin-lesion-dataset>

- [27] D. Bala and M. S. Hossain, "Monkeypox skin images dataset (MSID)," *Mendeley Data*, vol. 6, p. 2023, Jun. 2022.
- [28] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, pp. 1–48, Dec. 2019.
- [29] A. Mikolajczyk and M. Grochowski, "Data augmentation for improving deep learning in image classification problem," in *Proc. Int. Interdiscipl. PhD Workshop (IIPhDW)*, May 2018, pp. 117–122.
- [30] F. Garcea, A. Serra, F. Lamberti, and L. Morra, "Data augmentation for medical imaging: A systematic literature review," *Comput. Biol. Med.*, vol. 152, Jan. 2023, Art. no. 106391.
- [31] S. M. R. Hashemi and A. Broumandnia, "A new method for image resizing algorithm via object detection," *Int. J. Mechatronics, Electr. Comput. Technol.*, vol. 5, no. 16, pp. 213–234, 2015.
- [32] H. Morgan and M. Druckmüller, "Multi-scale Gaussian normalization for solar image processing," *Sol. Phys.*, vol. 289, no. 8, pp. 2945–2955, Aug. 2014.
- [33] K. Han, Y. Wang, H. Chen, X. Chen, J. Guo, Z. Liu, Y. Tang, A. Xiao, C. Xu, Y. Xu, Z. Yang, Y. Zhang, and D. Tao, "A survey on vision transformer," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 1, pp. 87–110, Jan. 2023.
- [34] Z. Liao, K. Xu, and N. Fan, "Swin transformer assisted prior attention network for medical image segmentation," in *Proc. 8th Int. Conf. Comput. Artif. Intell.*, Mar. 2022, pp. 491–497.
- [35] X. Zhou, L. Zhou, S. Gong, S. Zhong, W. Yan, and Y. Huang, "Swin transformer embedding dual-stream for semantic segmentation of remote sensing imagery," *IEEE J. Sel. Topics Appl. Earth Observat. Remote Sens.*, vol. 17, pp. 175–189, 2024.
- [36] A. G. Gad, "Particle swarm optimization algorithm and its applications: A systematic review," *Arch. Comput. Methods Eng.*, vol. 29, no. 5, pp. 2531–2561, Aug. 2022.
- [37] T. M. Shami, A. A. El-Saleh, M. Alswaiti, Q. Al-Tashi, M. A. Summakieh, and S. Mirjalili, "Particle swarm optimization: A comprehensive survey," *IEEE Access*, vol. 10, pp. 10031–10061, 2022.

HAGER SALEH was a Senior Machine Learning Engineer with Paxera-Health. She is currently a Postdoctoral Researcher with the Insight Research Centre for Data Analytics, University of Galway, Ireland. She is also an Assistant Professor with the Faculty of Computers and Artificial Intelligence, South Valley University, Egypt. She has extensive experience in developing and implementing machine learning and deep learning models across various applications, such as climate change and healthcare. Her expertise covers a broad range of areas, including artificial intelligence, explainable AI (XAI), transformer models, multimodal, generative AI, natural language processing, database management, image processing, streaming data analytics, and time series analysis. She has created an excellent research collaboration network nationally and internationally and published articles in journals with a high impact factor.



SHERIF MOSTAFA received the master's degree in securing wireless networks from the Computer Science Department, Faculty of Computers and Information, Helwan University, and the Ph.D. degree in computer science in medical image compression from the Computer Science Department, Faculty of Science, South Valley University, in 2020. He was with the Modern Sciences and Arts University MSA, for four years; the Department of Computer Engineering, Al Baha College, Saudi Arabia, for ten years; and a Software Engineer with the Central Agency for Organization and Administration. He is currently a Lecturer with the Faculty of Computers and Artificial Intelligence, Hurghada University. His research interests include image processing and the use of deep learning methods and computer vision.

ALSHIMAA H. ISMAIL received the B.S. degree in electronic engineering and the M.S. degree in computer science and engineering from Menofa University, in 2009 and 2015, respectively, and the Ph.D. degree in electrical engineering from Minia University, in 2019. Since 2019, she has been appointed as an Assistant Professor with the Delta Higher Institute for Engineering and Technology. She has been appointed as an Assistant Professor with the Faculty of Artificial Intelligence, Delta University, in 2021. She has been appointed as an Assistant Professor with the Faculty of Computer and Informatics, Tanta University, since June 2023. Her research interests include congestion control, cloud computing, mobile edge computing, deep learning, smart agriculture, climate change, and cognitive radio networks. She has served as a Reviewer for many high-quality journals, such as *Personal and Ubiquitous Computing* (Springer), *International Journal of Communication Systems* (Wiley), and *IEEE SYSTEMS JOURNAL*.



SUMIT SRIVASTAVA received the Ph.D. degree from the Indian Institute of Technology (BHU), Varanasi. He has more than 18 years of teaching experience. He is currently an Assistant Professor (Senior Scale) with the Department of Electronics and Communication Engineering, FET, MJP Rohilkhand University, Bareilly, India. He has published more than 30 research papers in different reputed journals and conferences. His research interests include the IoT, artificial intelligence, machine learning, smart sensors, microstrip patch antennas, and federated learning.



SAEED HAMOOD ALSAMHI received the M.Tech. and Ph.D. degrees in communication systems from the Department of Electronics Engineering, Indian Institute of Technology (Banaras Hindu University), IIT (BHU), Varanasi, India, in 2012 and 2015, respectively. In 2009, he was a Lecturer Assistant with the Engineering Faculty, IBB University. Afterward, he held a postdoctoral position with the School of Aerospace Engineering, Tsinghua University, Beijing, China. Since 2019, he has been an Assistant Professor with Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen. In 2020, he was an MSCA Smart 4.0 Fellow with Athlone Institute of Technology, Athlone, Ireland. He is currently a Senior Research Fellow with the Insight Centre for Data Analytics, University of Galway, Ireland; an Adjunct Professor with the Department of Computer Science and Engineering, College of Informatics, Korea University, Seongbuk-gu, Seoul, Republic of Korea; and an Assistant Professor with the Faculty of Engineering, IBB University, Ibb, Yemen. His research interests include green and semantic communication, the green Internet of Things, QoE, QoS, cybersecurity, multi-robot collaboration, blockchain technology, federated learning, and space technologies (high altitude platforms, drones, and tethered balloon technologies).