Computers & Security 133 (2023) 103385

Contents lists available at ScienceDirect

Computers & Security

journal homepage: www.elsevier.com/locate/cose

A new deep boosted CNN and ensemble learning based IoT malware detection

Saddam Hussain Khan [a,*], Tahani Jaser Alahmadi [b,*], Wasi Ullah [a], Javed Iqbal [a], Azizur Rahim [a], Hend Khalid Alkahtani [b], Wajdi Alghamdi [c], Alaa Omran Almagrabi [c]

[a] Department of Computer Systems Engineering, University of Engineering and Applied Science (UEAS), Swat, Pakistan
[b] Department of Information Systems, Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia
[c] Department of Information Systems, Faculty of Computing and Information Technology, King Abdulaziz University, Jeddah 21589, Saudi Arabia

ARTICLE INFO

Keywords:
Malware
IoT
Ensemble learning
Deep learning
CNN
Detection

ABSTRACT

Security issues are threatened in various types of networks, especially in the Internet of Things (IoT) environment that requires early detection. IoT is the network of real-time devices like home automation systems and can be controlled by open-source android devices, which can be an open ground for attackers. Attackers can access the network credentials, initiate a different kind of security breach, and compromises network control. Therefore, timely detecting the increasing number of sophisticated malware attacks is the challenge to ensure the credibility of network protection. In this regard, we have developed a new malware detection framework, Deep Squeezed-Boosted and Ensemble Learning (DSBEL), comprised of novel Squeezed-Boosted Boundary-Region Split-Transform-Merge (SB-BR-STM) CNN and ensemble learning. The proposed STM block employs multi-path dilated convolutional, Boundary, and regional operations to capture the homogenous and heterogeneous global malicious patterns. Moreover, diverse feature maps are achieved using transfer learning and multi-path-based squeezing and boosting at initial and final levels to learn minute pattern variations. Finally, the boosted discriminative features are extracted from the developed deep SB-BR-STM CNN and provided to the ensemble classifiers (SVM, MLP, and AdabooSTM1) to improve the hybrid learning generalization. The performance analysis of the proposed DSBEL framework and SB-BR-STM CNN against the existing techniques have been evaluated by the IOT_Malware dataset on standard performance measures. Evaluation results show progressive performance as 98.50% accuracy, 97.12% F1-Score, 91.91% MCC, 95.97 % Recall, and 98.42 % Precision. The proposed malware analysis framework is robust and helpful for the timely detection of malicious activity and suggests future strategies.

1. Introduction

Malware is an undesired software that can harm digital devices like computers, android, and especially the Internet of Things (IoT) devices. IoT has gained popularity expeditiously in the digital market due to its robust features and applications. The IoT devices improved human life quality and will increase to 43 billion in 2023. The concept of IoT is to transform real objects into virtual objects having unique addresses and can be driven by the popular open-source android devices. In this emerging technology, intelligent devices share their information and resources accordingly (Madakam et al., 2015). Interconnected devices have become an integral part of our daily lives, playing important roles in areas such as healthcare, home automation, education, and industry. Numerous applications are available in different industries, including the monitoring of soil conditions in agriculture (Vuran et al., 2018), e-health (Zafar et al., 2022; SM et al., 2015; Zahoor et al., 2022; Khan, 2022; Khan et al., 2023; Rauf et al., 2023), and military application (Iyer and Patil, 2018; Qamar et al., 2022; Arshad et al., 2022; Zahoor and Khan, 2022). The concept of industry 4.0 has been utilized to establish a connection between manufacturers and consumers, resulting in a more efficient supply chain (Mikhalevich and Trapeznikov, 2019). Industrial IoT (IIoT) has undoubtedly contributed well towards products and innovations in improving industrial infrastructure.

IoT devices are heterogeneous in both structures and network protocols, where each heterogeneous device has a unique microprocessor characteristic (Vignau et al., 2021). So, this is the major cause that the

* Corresponding authors.
E-mail addresses: saddamhkhan@ueas.edu.pk (S.H. Khan), tjalahmadi@pnu.edu.sa (T.J. Alahmadi).

https://doi.org/10.1016/j.cose.2023.103385
Received 14 January 2023; Received in revised form 24 May 2023; Accepted 6 July 2023
Available online 7 July 2023

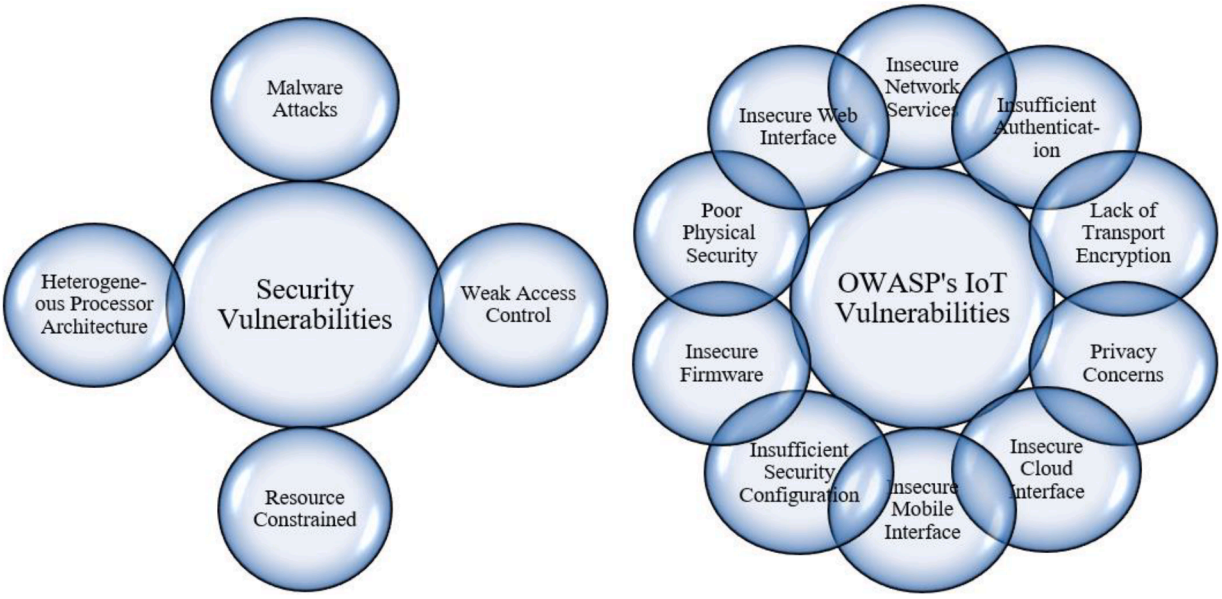

Fig. 1. Security breaches and OWASP IoT s vulnerabilities.

IoT industry is lagging in security protocols and becoming enlarged attack surface, leading to security breaches. This provides tunnels for cyber criminals to exploit the vulnerabilities and utilize the attacks for their illegal actions. IoT devices are vulnerable to security attacks, easily exploited, and compromise network control. Recently, more than 178 million IoT devices, like webcams, medical devices, routers, etc., have been exposed to attackers because new technology is the key entry point (Chaganti et al., 2022). Therefore, it is highly desired to secure IoT devices, and security countermeasures are required to protect them from cyberattacks.

Major cyber security concerns include malware attacks, DDoS, botnets, rootkits, intrusions, ransomware, and compromise nodes (Zahoora et al., 2022). Malware is software that includes viruses, adware, Trojan horses, spyware, etc., and can harm computers and web devices. In a malware attack, the attacker can gain access to the network and take complete control without any awareness. It is becoming a massive barrier to malware analysts and making the ground interest for security researchers. Compared to other digital devices, there is no regular patching in IoT devices because of their embedded nature (Vignau et al., 2021), and it is impossible to implement the security protocols on all IoT devices uniformly. These security breaches and OWASP IoT IoT Security Vulnerabilities security are interpreted in detail in Fig. 1. IoT vulnerabilities are extremely important because IoT devices are becoming increasingly prevalent in our daily lives, and they can have a significant impact on our privacy and security. Malware attacks are a potential threat to the security of IoT devices due to their widespread use and insufficient security measures. These vulnerabilities list serves as a critical resource for developers and security experts to pinpoint and resolve the most significant security threats facing IoT devices. Android malware detection reached 26.61 million in 2018 and noticed a 520,000 monthly increase (Ngo et al., 2020). Therefore, there may be a mechanism for detecting malware attacks under these issues in IoT devices to take immediate action and secure the system or device before compromising.

IoT Malware analysis comes under the umbrella of static and dynamic analysis. The static malware detection method is the way of detection by signature-based, permission-based, and bytecode-based methods. However, static malware analysis is simple and can be easily fooled by obfuscation, and runtime vulnerabilities lead unnoticed. On the other side, the dynamic method is the way of detection in which the applications are executed in isolated platforms such as (simulators, sandbox, and virtual machines). The environment is secured, trusted, and undetectable, tracking their behaviors during the execution of the suspicious file, whether normal or malicious. The traditional detection techniques are relied on built-in signature libraries and mainly on human involvement, and it is hard to detect malware grown very extensively (Asam et al., 2021; Asam et al., 2022). In addition, a technique has been used to convert malware binary files into images by assigning bytes to pixels, and Machine learning (ML) methods have been employed to detect ELF-based malware (R. and DeepMalNet, 2018). However, ML methods required additional effort for feature extraction from images to get domain expert knowledge for malware detection. Lately, deep learning (DL) and deep CNN models have been considered for IoT malware detection (Vinayakumar et al., 2019; Shalaginov and Øverlier, 2021).

To the best of our knowledge, The IOT_Malware detection using deep hybrid and ensemble learning incorporated in this study has not been used in any previous studies. In this study, IoT malware is utilized by its visual image representation and benign files as by the observations, deep CNNs have shown extraordinary performance for the visual representation of challenges (Bendiab et al., 2020). A novel split, transform and merge (STM) block and squeezed-boosted channel (S.B.) is introduced in the novel model SB-BR-STM for analyzing the feature space and further for malware detection accurately and efficiently in the field of IoT ensemble learning classifiers are used. Additionally, the proposed classification architecture STM block of deep CNN) exploits the idea of region-heterogeneity and homogeneity. The main contributions from our side in the current studies are depicted below:

1 A new DSBEL framework is proposed for detecting malware-infected packets in an IoT environment. The framework comprises the stacking newly developed SB-BR-STM CNN and ensemble classifiers.
2 The novel STM block and channel-SB ideas are incorporated in the new SB-BR-STM dilated-CNN. Moreover, Max-Pooling, average-pooling, and dilated-convolutional operations are incorporated at various levels in the new STM block for extracting the diverse feature-set, especially consisting of the intensity-homogeneity and heterogeneity, global malicious patterns.
3 TL-based auxiliary-channel extraction and multi-path-based new S.B. idea for achieving various feature map to help improve the proposed SB-BR-STM CNN performance. These operations are employed at

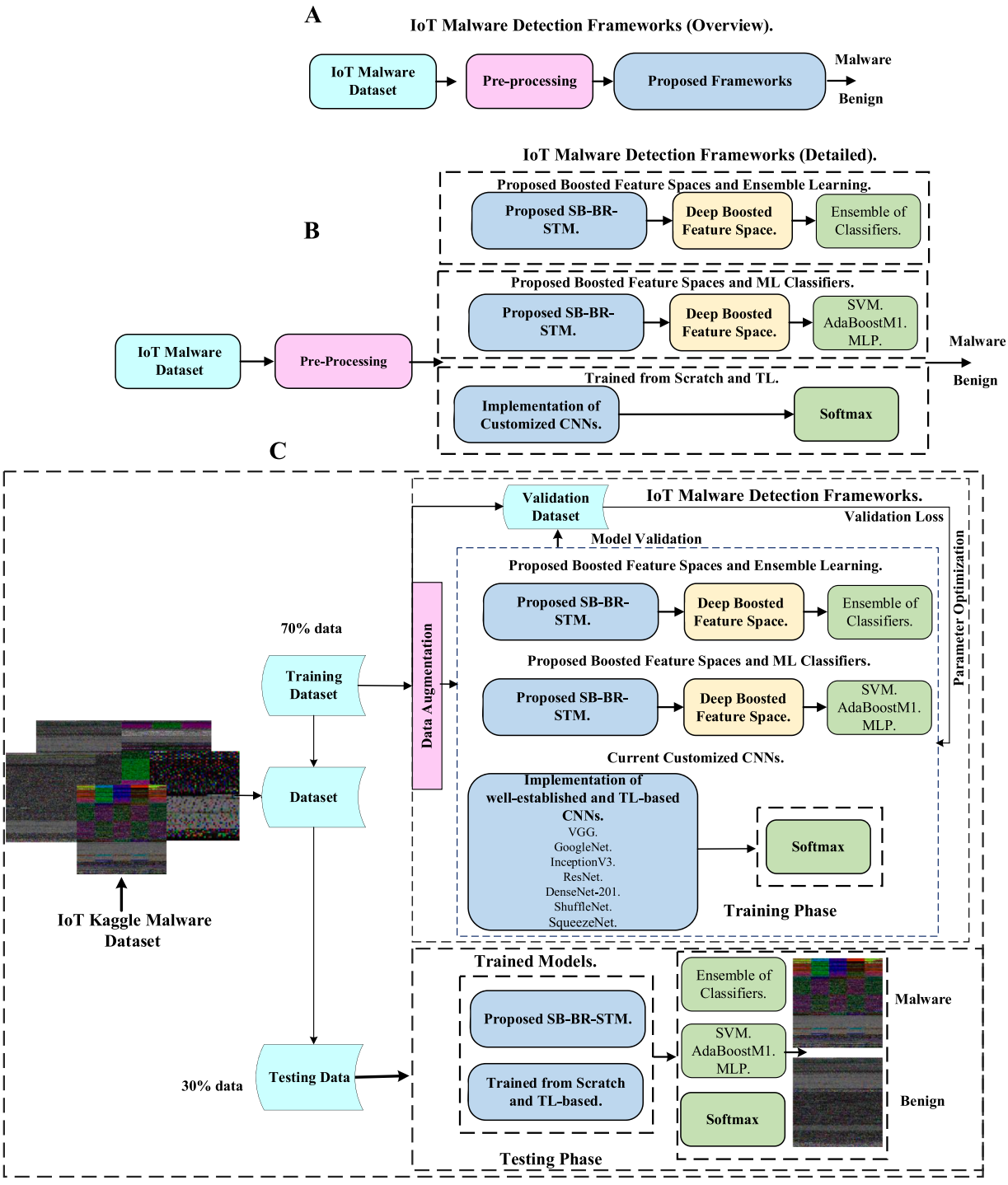

Fig. 2. The proposed malware detection framework.

initial, mid, and conclusion levels for capturing minute texture variation.

4 The proposed DSBEL framework grants the boosted discriminative features from SB-BR-STM CNN and provides the ensemble classifiers for improving the hybrid learning discrimination.

5 The proposed DSBEL framework's and SB-BR-STM CNN's performance is compared with the current techniques and evaluated on the IOT_Malware dataset using standard performance measures.

Onwards the paper is presented in the subsections as Section 2 represents the previous study, our proposed framework is presented in Section 3, Section 4 covers the experimental setup, result discussions are given in Section 5, and Section 6 will conclude the paper.

2. Background and related work

Malware attacks on IoT devices are growing, and detection using traditional methods is difficult, as these techniques adopted the traditional signatory libraries and interactions expertise of malware analysts. On the other hand, ML and DL techniques can apply to detect malware, which is automatic and adaptable in any discipline (Muzaffar et al., 2022; Deng et al., 2023). ML-based malware detection method involves four steps: construction of the dataset, feature engineering, training of the model, and evaluating the model. Feature engineering calculates the

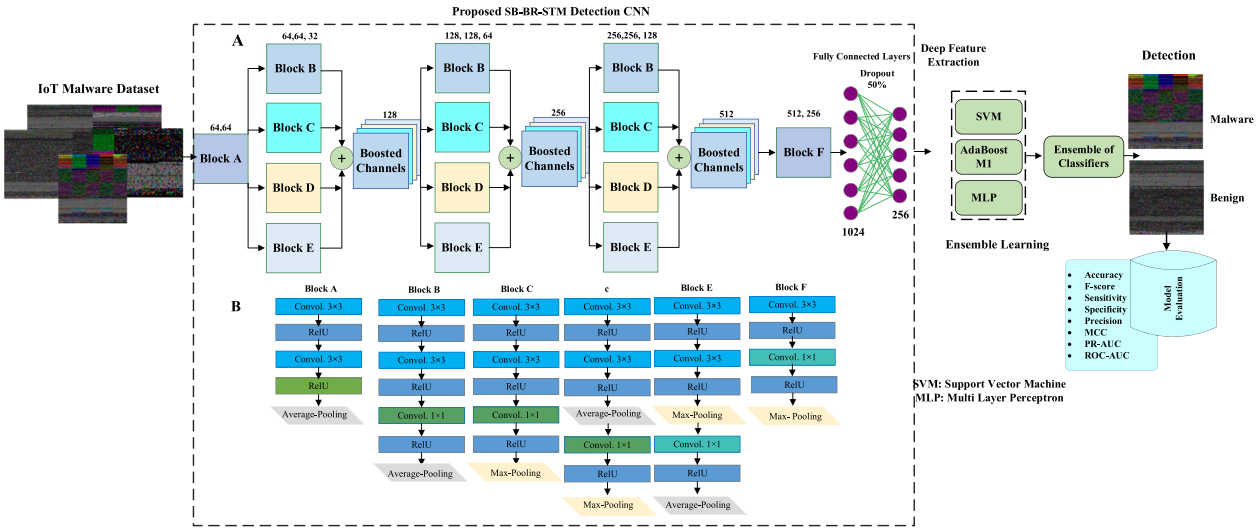

Fig. 3. The proposed DSBEL framework comprised of deep SB-BR-STM CNN.

Table 1
Benchmark IOT_Malware dataset details.

Properties	Description
Total	3,959 images
Benign ware	2,486 images
Malware	1,473 images
Train and validation (70%)	(1741, 1032)
Test (30 %)	(745, 441)

model's validity and characterizes the A.P.K.s by extracting robust and informative features. AndroidManifest.xml file and classes.dex file is the main feature used to characterize the A.P.K.s. Basic information about an A.P.K. is recorded in AndroidManifest.xml, such as requested permissions, hardware information, A.P.K. component, and filtered intents. vClasses. dex is transformed into a small format that consists of Dalvik commands (includes operands and opcode). Extracting advanced features such as flow diagram control can be achieved by disassembling classes.dex files (Su et al., 2018) and API dependency graph (Ren et al., 2020), can also apply to train the malware detection models. Dynamic behavioral features like network operations, operations, calls, and encryption operations can obtain by running the applications on isolated platforms, which has been discussed in the reported literature (Hussain et al., 2019). Combining dynamic features with static features can result in a more precise model and improved detection performance.

Traditional ML models (such as Random Forest (Shafiq et al., 2021), SVMs (Xu et al., 2018)) and DL (such as CNN (Zhang et al., 2018), Long Short-Term Memory (LSTM) (Xu et al., 2018; Alzaylaee et al., 2020)) have been extensively used for malware detection. Several ML and DL algorithms provided promising and robust performance for IoT malware detection (Ye et al., 2018). These tools employ vulnerability mining in the firmware and applications of IoT, which can infect the whole network or the edge devices of the network (Asam et al., 2022). During recent research advancements, an inclination toward ML tools and computational power has increased due to their anti-malware applications. Cozzi et al. (2018) used malware detection based on ML under the Linux-based platform malware of IoT by using of data set provided. They also used clustering techniques for malware detection. To detect Mirai botnet attacks in IoTs, Ganesh et al. (Palla and Tayeb, 2021) use ML techniques; they applied the approach of A.A.N. by using of N-BaIot dataset.

Bendiab et al. (2020) used the pre-trained ResNet50 for malware traffic analysis in IoT using a 1000 network (pcap) file. A lightweight CNN malware detection approach compared with existing VGG-16 for IoT was reported by Su et al. (2018). In their studies, the central work theme like DDoS malware and IoTPOT used the malware images. Considerable better performance of 95% accuracy were achieved for malware of type DDoS and good ware in their experimental setup (Pa et al., 2016). Ren et al. (2020) developed an end-to-end malware detection mechanism for Android IoT devices by gathering 8000 malicious and 8000 benign A.P.K. files from VirusShare and the Google Play Store. The authors employed various deep learning methods on the Mobile dataset to test their experimental perspectives for identifying malware using color images. An active DL-based IIoT malware detection technique has been reported using P.S.E., sparse-autoencoder, and LSTM to train active learners (Khowaja and Khuwaja, 2021). The fusion framework achieved 95.1% and 86.9% accuracy on detection and adversarial malware detection, respectively. Moreover, the DL-based Bidirectional-Gated Recurrent-Unit-CNN technique has been reported to detect IoT malware and achieved 98% accuracy (Chaganti et al., 2022). Additionally, an industrial-based IoT malware is analyzed using a DCNN model that extracts dynamic patterns. The model's performance is compared to state-of-the-art models, and it achieves an accuracy of 97.81% (Naeem et al., 2020). Furthermore, the effectiveness of the AMD model in detecting IIOT malware was assessed using an imbalanced benchmark dataset. Multiple experimental scenarios were conducted using the Leopard dataset and resulted in a reported accuracy of 98.05% (M. and Sethuraman, 2023). Recently, the iMDA (IoT Malware Detection Architecture) has been utilized to detect new IoT malware and its performance was compared with well-known deep CNN models such as SqueezeNet, ShuffleNet, Inceptionv3, GoogleNet, DesneNet, etc . The iMDA technique achieved highly accurate results with an accuracy of 97.93%, F1-Score of 93.94%, precision of 98.64%, MCC of 87.96%, and recall of 88.73% (Asam et al., 2022). All the above-reported work is measured and analyzed in terms of Accuracy and Precision, although the datasets selected are imbalanced. This research work is examined under the benchmark IoT dataset publicly available on Kaggle, and performance evaluation metrics are selected as F1-Score, MCC, and Recall, along with Accuracy and Precision.

3. Deep squeezed-boosted and ensemble learning (DSBEL) framework

The proposed novel approach comprised of developed a new deep CNN named the Squeezed-Boosted Boundary-Region Split-Transform-Merge SB-BR-STM and ensemble classifiers. The proposed IoT malware detection scheme is comprised of three arrangement schemes: (1) the proposed SB-BR-STM CNN and (2) the DSBEL framework, and (3)

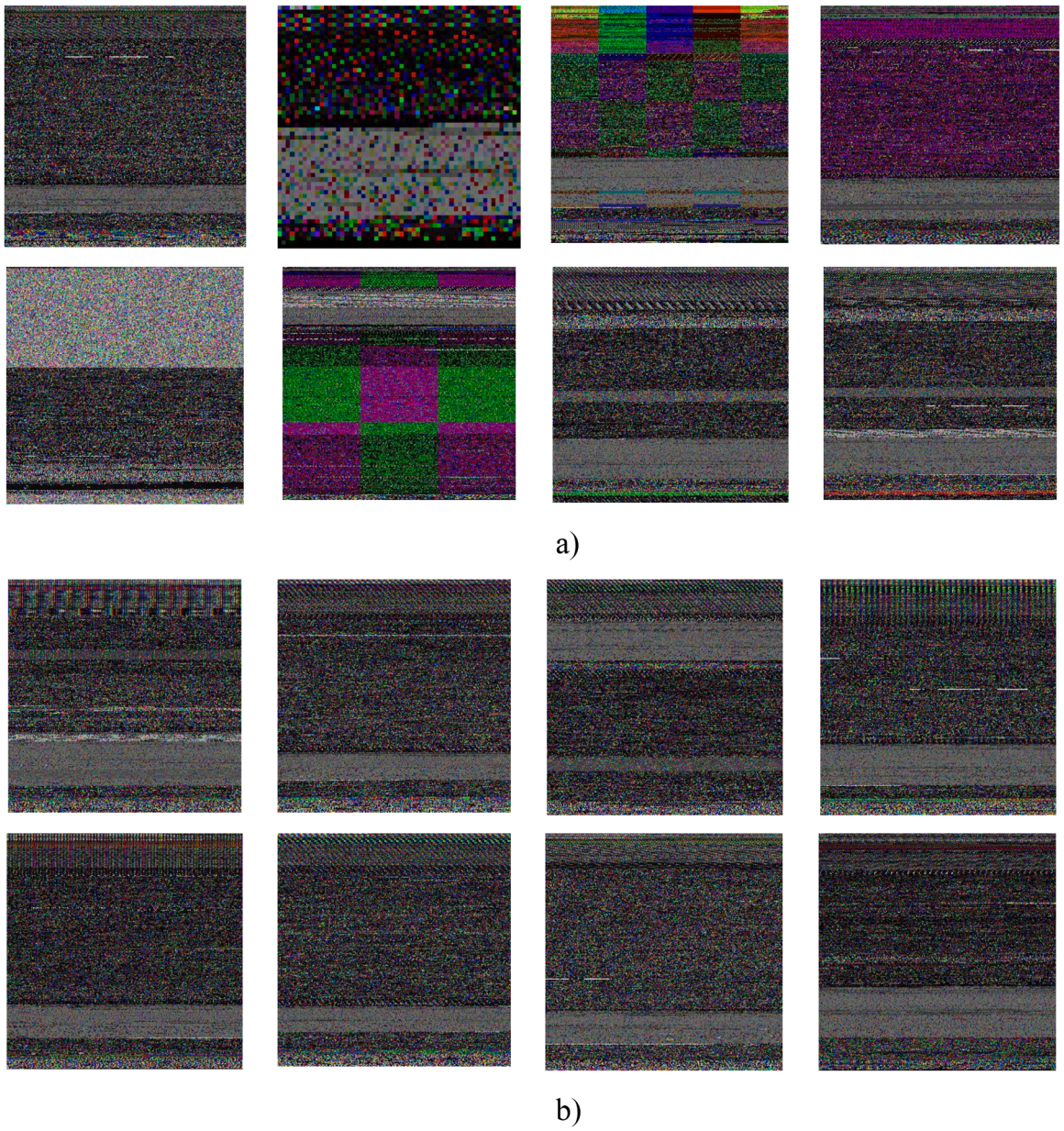

Fig. 4. IOT_Malware visual representation (a) malware (b) benign ware images.

Table 2
Hyperparameters for training CNNs.

Hyper Parameters	Values
Learning	10^{-3}
Epochs	20
Optimizer	S.G.D.
Batch Dim.	16
Momentum	0.950

Table 3
Details of the assessment metrics.

Metric Symbol	Description
Accuracy	Count accuracy in the percentage of the infected points
Precision	Count how precise the model is, which is the ratio of predicted infected points to the total infected points
Recall / Sensitivity	Count recall, the proportion of the correctly identified infected points and benign points
MCC	Mathews Correlation Coefficient
T.P.	Predicted Correctly Infected Points
TN	Predicted Correctly Benign Points
F.P.	Predicted Incorrectly Infected Points
F.N.	Mispredicted Benign Points

evaluating the current CNNs. These current customized CNN is used as both learned from scratch and as fine-tuned T.L. on IOT_Malware dataset. Moreover, data augmentation has been performed to improve learning and generalization. Fig. 2 is the graphical view of the overall framework.

3.1. Data augmentation

The models of CNN perform better for a large number of labeled data and perform better in generalization. Sometimes, the data points are different from the network requirements. Data augmentation is the process through which the data points are arranged according to the network requirement by image transformations (Shorten and Khoshgoftaar, 2019), which includes image rotation (0-360 degrees), image

Table 4
Performance measurements of the proposed SB-BR-STM. With the existing models using trained from scratch.

Models	Accuracy %	F1-Score	MCC	Recall	Precision
SqueezeNet	93.47	86.76	79.14	78.26	97.34
ShuffleNet	94.72	89.30	80.90	82.68	97.08
Inceptionv3	94.89	89.11	80.91	82.28	97.17
VGG-16	95.38	90.14	81.29	84.29	96.86
ResNet-50	95.62	90.84	81.79	85.70	96.64
GoogleNet	95.93	91.72	82.21	87.93	95.85
DenseNet201	96.17	91.90	82.56	87.93	96.25
Proposed SB-BR-STM	**97.18**	**94.48**	**84.57**	**91.15**	**98.07**

Table 5
Performance measurements of SB-BR-STM with the existing models training using TL.

Models	Accuracy %	F1-Score	MCC	Recall	Precision
SqueezeNet	96.14	87.20	80.88	80.26	95.46
ShuffleNet	96.33	91.45	81.35	87.72	95.52
Inceptionv3	96.47	91.71	81.95	86.80	97.20
VGG-16	96.58	91.31	82.51	86.23	97.02
ResNet-50	96.58	91.41	83.11	86.04	97.49
GoogleNet	96.60	91.94	83.44	86.82	97.71
DenseNet201	96.66	92.50	83.71	88.26	97.17

scaling (0.5 -1), shearing (-0.5, +0.5), image transformation (grayscale to RGB and vice versa), and reflection (in the right and left direction). Making the data set more robust and generalizing it to a network can be done with the help of the augmentation process.

3.2. Proposed SB-BR-STM CNN

In this manuscript, a new deep SB-BR-STM CNN is developed to detect IoT Malware images. The channels are initially systematically split and employ Region-Edge and dilated convolutional operations. Consequently, the channels are squeezed and merged and further fed into fully connected layers. The channels are split into four multi-paths and squeezed to preserve the reduced maps, then boosted after merging for getting diverse feature-maps. The novel channel S.B. approach incorporated at STM block in a newly modified fashion for capturing minor texture and contrast variation of malicious patterns. The idea of S. B. is employed on the channel at the abstract, mid, and final levels.

Three STM-based blocks are implemented systematically and have the same topology in the proposed SB-BR-STM. Four dilated convolutional blocks constitute the architecture of the SB-BR-STM, as presented by Eq. (1). Each block applies the average-pooling and max-pooling operations methodically to preserve the region and boundary pattern (Khan et al., 2022), as organized by Eqs. (2)–(3). These operations help efficiently assess region homogeneity inside the infected region and determine boundaries, edges, and textural variations.

Table 6
Performance analysis of the proposed hybrid frameworks.

Classifier	Accuracy %	F1-Score	Precision	MCC	Recall	AUC
SVM	97.71	91.87	99.80	85.88	85.11	98.83
MLP	97.79	92.72	99.61	87.11	86.72	99.18
AdaboostM1	97.91	99.46	99.14	89.73	90.14	99.46
Ensemble (SVM-MLP)	98.13	95.71	99.09	93.89	92.56	99.48
DSBEL (SVM-MLP- AdaboostM1)	**98.50**	**97.12**	**98.42**	**91.91**	**95.97**	**99.51**
Comparative Analysis with the reported work						
AlexNet (Asam et al., 2022)	92.86	68.07	99.60	58.74	51.71	0.9685
VGG19 (Asam et al., 2022)	95.38	83.53	99.02	74.29	72.23	0.9739
Xception (Asam et al., 2022)	96.57	93.42	97.37	86.51	90.74	0.9882
iMDA (Asam et al., 2022)	97.93	93.94	98.64	87.96	88.73	0.9938
ResNet101 (Almomani et al., 2022)	97.98	98.02	96.49	—	99.58	0.9942
ShuffleNet (Almomani et al., 2022)	98.09	98.11	97.23	—	98.99	0.9985
MD-IIOT (Naeem et al., 2020)	97.46	94.80	95.39	—	94.22	—

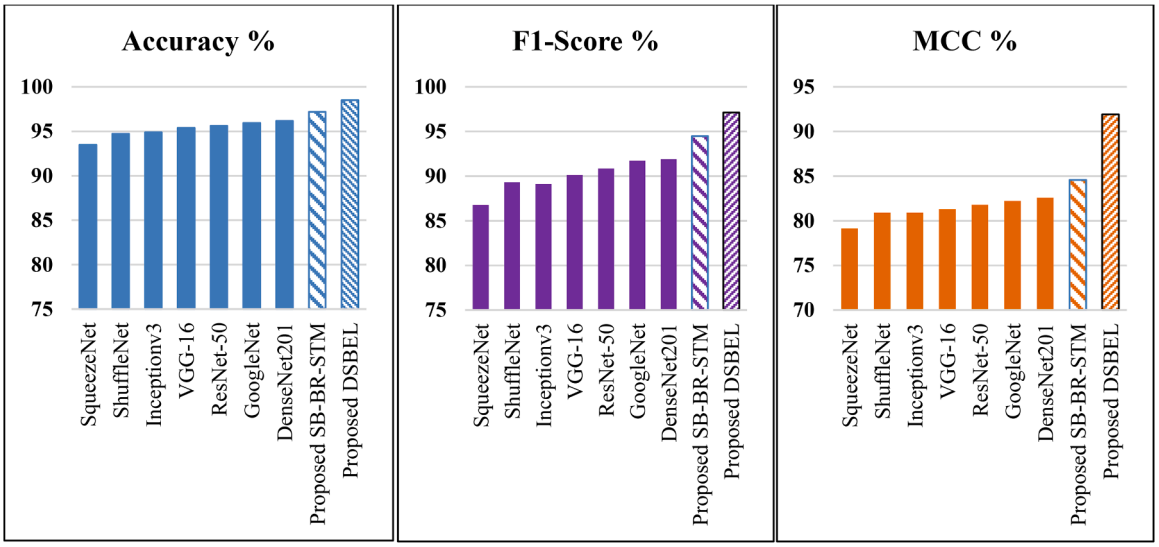

Fig. 5. Comparative analysis of existing CNNs VS proposed SB-BR-STM and DSBEL.

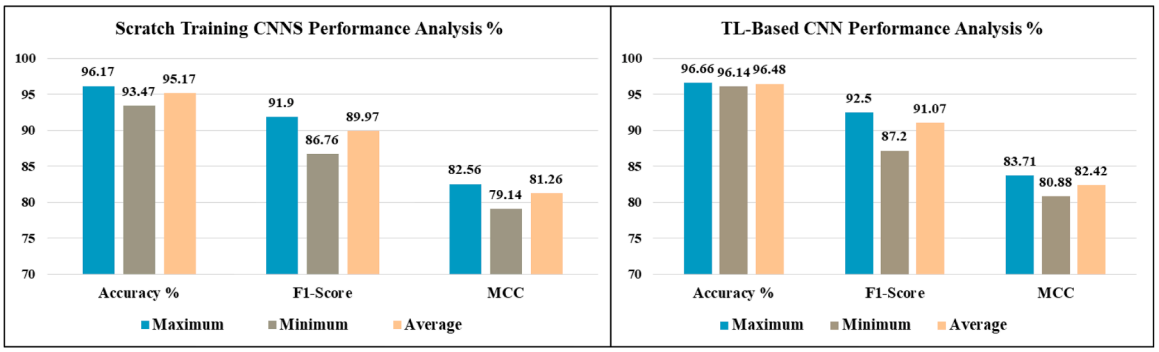

Fig. 6. Performance analysis of the existing CNNs.

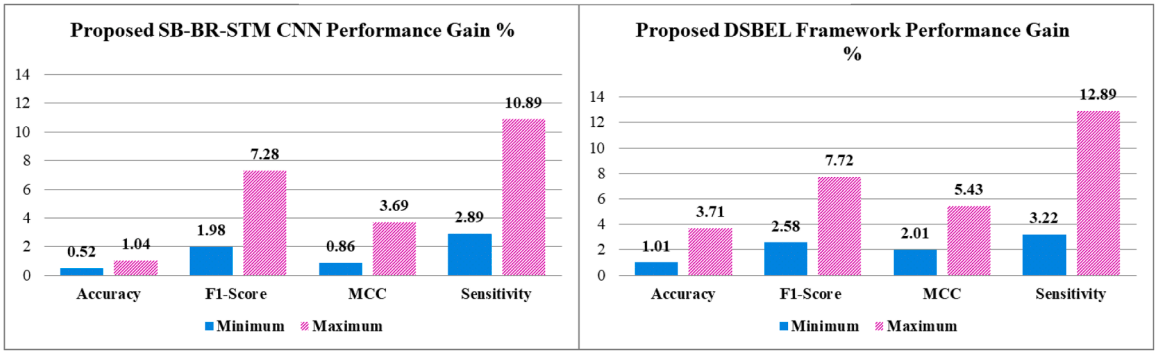

Fig. 7. Performance improvements of the proposed SB-BR-STM CNN and DSBEL framework.

$$W_{s,t} = \sum_{x=1}^{j} \sum_{y=1}^{k} w_{s+x-1,t+y-1} * u_{x,y} \tag{1}$$

$$W_{s,t}^{avg} = \frac{1}{m^2} \sum_{x=1}^{m} \sum_{y=1}^{m} w_{s+x-1,t+y-1} \tag{2}$$

$$W_{s,t}^{max} = \max_{x=1,\ldots,m,y=1,\ldots,m} w_{s+x-1,t+y-1} \tag{3}$$

In Eq. (1), w represents the input feature map having a dimension of s, t, and u represents the filter having size x, y. the acquired feature vector sorts from the lower level (1) to the upper level (s+x-1) and (t+y-1). The average and max pooling having m size window are represented in Eqs. (2)–(3). As depicted in Eq. (4), channel S.B. operation is improved at every convolutional block (B, C, and D, E) for learning diverse infected feature sets. For attaining diverse feature maps, B & C blocks are produced by TL On the other hand, training from scratch Blocks is D & E. Dimensions of each channel of S.B. convolutional STM block are 32-128, 64-256, and 128-512, respectively [56]. The main application of T.L. is solving the problem of the target domain by learning the network from the source domain to achieve better performance. Moreover, homogenous operations control distorted regions and outliers in the acquisition of input images (Khan et al., 2022). And for achieving optimal features and reducing connection intensity, the boosted channel is processed in block F.

Wd and We are the original blocks of D and E channels, respectively, as shown in Eq. (4). Similarly, Wb and Wc are depicted as auxiliary block A and C channels generated using T.L., respectively. Then, a(.) operation will concatenate the original and auxiliary channels. Additionally, dropout layer is used to reduce overfitting and achieve target-specific features. Z_p represents neuron in Eq. (5). The proposed CNN is represented diagrammatically in Fig. 3.

$$W_{Boosted} = a(w_b||w_c||w_d||w_e) \tag{4}$$

$$W = \sum_{P}^{p} \sum_{Q}^{q} z_p W_{Boosted} \tag{5}$$

$$\sigma(w) = \frac{e^{x_i}}{\sum_{i=1}^{c} e^{x_c}} \tag{6}$$

3.3. Significance of using an auxiliary channel

Improvements in the representative capacity of the developed CNN are possible by adding auxiliary channels. These channels are generated using TL-based deep CNN, and for getting prominent feature maps, squeezed at each STM block and merged at different initial, mid, and high levels. The prominent and diverse channels are learned from different CNN, as shown in the blocks, and concatenated at the next level, which helps in improving image information locally and globally. S.B.-based deep CNN learns complex malicious patterns.

3.4. Ensemble ML classifiers

In the proposed DSBEL framework, Deep-boosted feature space based on the developed SB-BR-STM is fed to ensemble classifiers to detect malware images. The DSBEL framework is applied to acquire a Deep-boosted feature vector and fed into a voting-based ensemble ML classifier. The ensemble ML classifier will take the feature space and detect the malware images by the majority-voting-based method.

The boosted feature vectors are extracted from the proposed SB-BR-STM of diverse channels and fed into the ensemble classifier, as defined in Eq. (6). Mainly three ML classifiers are used, SVM, M.L.P. (Gardner and Dorling, 1998), and AdaBooSTM1 (Schapire, 2013), having the activation functions $f_{SVM}(.)$, $f_{MLP}(.)$, and $f_{AdaBoost}(.)$ as shown in Eqs. (7)–(10). The proposed CNN is trained with carefully selected hyper-parameters, which not only minimize training errors but also minimize empirical risk [61]. Moreover, ML classifiers aim to reduce test errors of

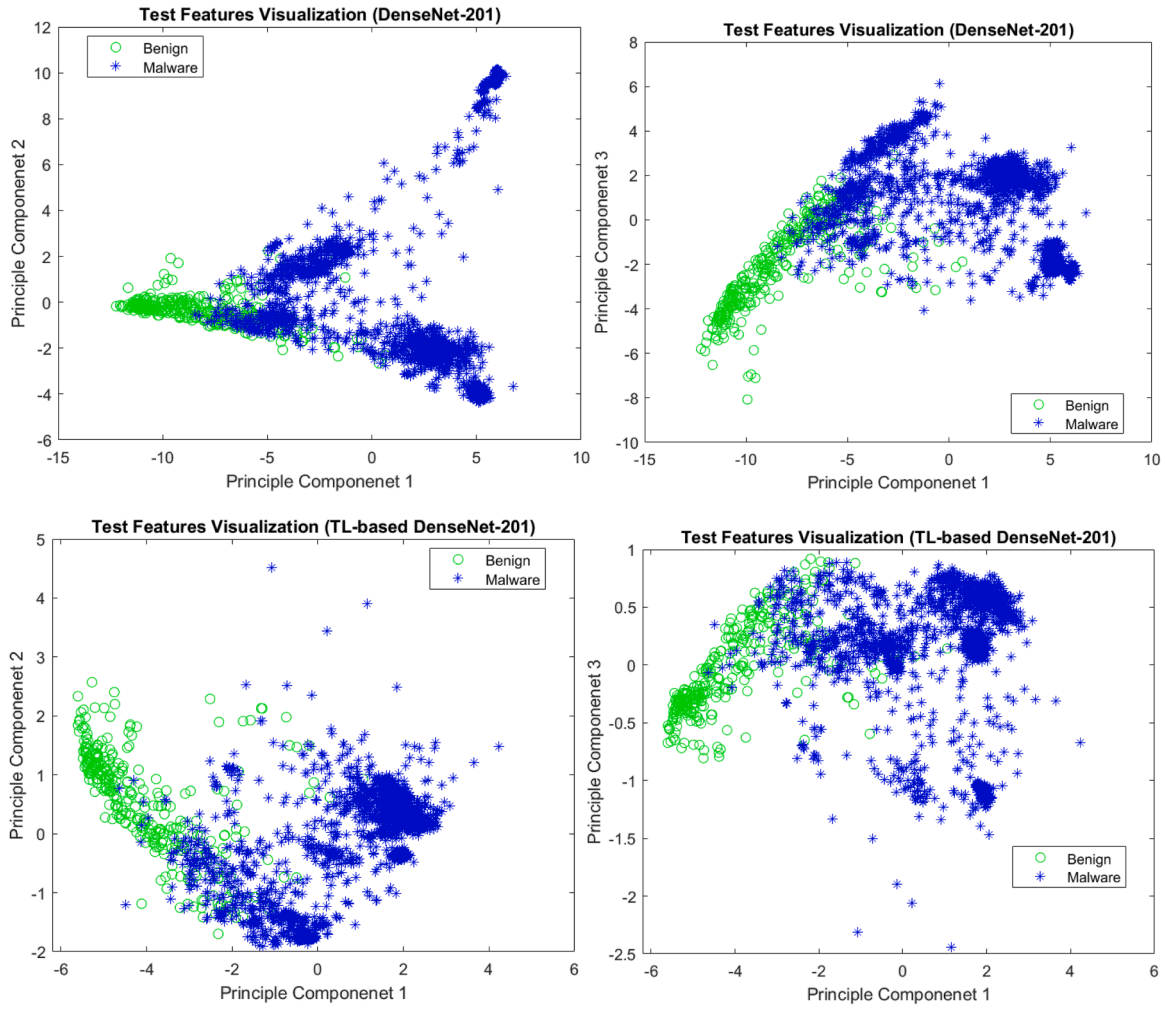

Fig. 8. Feature visualization PC1, PC2, and PC3 on the best performing in existing CNN (DenseNet-201).

training set by fixed distribution, exploiting the minimal operational principle and providing generalization.

$$h_{MLP} = f_{MLP}(w) \tag{7}$$

$$h_{SVM} = f_{SVM}(w) \tag{8}$$

$$h_{AdaBoost} = f_{AdaBoost}(w) \tag{9}$$

$$h_{final} = f_{Ensemble}(f_{MLP}(w), f_{SVM}(w), f_{AdaBoost}(w)) \tag{10}$$

Encouragement towards combining multiple feature vectors into a single information feature vector and performance improvement aspires from ensemble learning. Also, the unsatisfactory performance risk is avoided by using the extracted feature vectors from a single model. Depending upon the fusion level, applicability can be on classifier ensemble and feature boosting. Boosted feature sets are implicated by the featured ensemble and will feed to the ML classifier to acquire final vectors. On the other hand, the integration of the decision from multiple ML classifiers by the ensemble classifier voting strategy is shown in Eq. (10). Both features boosting and ensemble classifier techniques have been used in the proposed DSBEL framework.

3.5. Customized CNN utilization

Many of the current CNN are adopted for detecting malware in IoT and android-based systems (Mahmood et al., 2014; Vidas et al., 2014; Khan et al., 2020). The additional layers modify the abstract and final layers of the existing CNN for the requirements of the input and target-class dimension in the dataset. In the train-from-scratch phenomenon, the model learns by back propagating the weights and updating it concerning errors as initial weights are assigned randomly. And on the other hand, in the T.L. techniques, models use pre-trained weights for the convolution layers for convergence improvement. Additionally, TL has been utilized to derive the parameters of the modified prior CNN. The CNN is designed to capture the characteristics of the target domain, particularly for the malware dataset, using pre-trained filters from ImageNet.

4. Experimental setup

4.1. Dataset

For detecting malware in IoT using deep CNN, all the packets entering the network are represented in the image form because CNN networks are processing the images effectively. Image representation provides an efficient means of processing images and applying them to the trained CNN model for the detection of malware intrusion. In this study, the model was trained using the IoT_Malware dataset (Asam et al., 2022), which is based on the byte sequences of executable files. The IoT-Malware dataset is based on the byte sequences found in executable files (Wan et al., 2020). Within the IOT_Malware dataset, there are two main directories: one containing 2486 benign images and the other

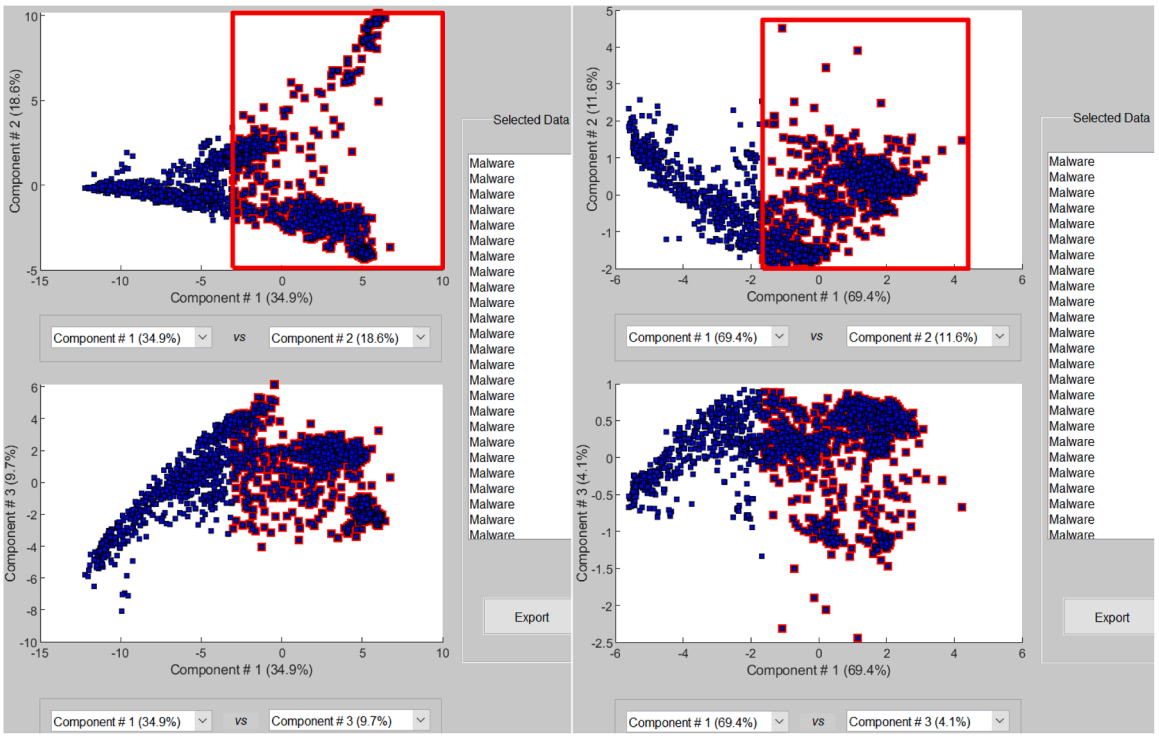

Fig. 8. (*continued*).

comprising 1473 images of malware. The directory distribution is mentioned in Table 1, and the imagery is represented in Fig. 4.

4.2. Implementation detail

The ratio was 70:30% of the data set for training and testing the proposed architecture. Moreover, on the training set, hold-out-validating is performed during the model training at 80:20%. To examine the robustness of the model, this hold-out cross-validation was conducted. All the optimized parameters are selected for validating the model. Table 2 shows the details selected for the model training. MATLAB 2022b using as a tool for developing customized CNNs. NVIDIA GPU GeForce GTX-T dell computer has 32 G.B. of RAM and was used for experimentations. Roundabout 2-4 hours ~ 1-each CNN took 10 minutes/epoch during the training.

4.3. Performance evaluation metrics

Standard performance metrics were used to evaluate the customized CNNs and proposed SB-BR-STM. These metrics are depicted in Table 3, along with the mathematical explanation. In the classification metrics, Accuracy, F1-Score, and Precision or Sensitivity are included along with True Positive (T.P.), True Negative (T.N.), False Positive (F.P.), and False Negative (F.N.). Eqs. (7)–(11) represent Accuracy, Presicion, Sensitivity, MCC, and F1-Score, respectively. We used Accuracy, Precision, and F1-Score as optimizing metrics for classification in the experimental setup.

$$Accuracy = \frac{Predicted\ Infected\ points\ + Predicted\ Benign\ Points}{Total\ Points}\ X\ 100 \tag{7}$$

$$Precision = \frac{Predicted\ Infected\ Points}{Predicted\ Infected + Incorrectly\ Predicted\ Infected}\ X\ 100 \tag{8}$$

$$Sensitivity = \frac{Predicted\ Infected\ Points}{Total\ Infected\ Points}\ X\ 100 \tag{9}$$

$$MCC = \frac{(TP\ X\ TN) - (FP\ X\ FN)}{\sqrt{(TP + FP)\ X\ (FP + FN)\ X\ (TN + FP)\ X\ (TN + FN)}} \tag{10}$$

$$F - Score = 2\ X \frac{Pre\ X\ Rec}{Pre + Rec} \tag{11}$$

5. Results and discussion

This research presents a novel DSBEL framework and SB-BR-STM CNN for discriminating malware-infected images and benign images in IoT networks. IOT_Malware data set is used to train and validate the proposed network model. This research also evaluates the existing customized state-of-the-art networks and compares the performance with the novel DSBEL framework. Customizing of the existing CNN is incorporated in a modified fashion and implemented using both trains from scratch and T.L. basis. The effectiveness of the proposed malware detection framework is assessed using standard performance metrics.

Tables 4 and 5 present the results obtained from training the model from scratch and using T.L., respectively.

Table 6 shows the results of the machine learning classifiers and ensemble classifiers.

5.1. Performance analysis of the proposed SB-BR-STM

Standard imbalance IOT_Malware dataset is used to evaluate the proposed DSBEL performance using standard metrics, Accuracy, F1-score, and MCC. A data augmentation technique was applied to images to increase the learning of the models, and ensemble ML classifiers (SVM, MLP, AdaBooSTM1) are used to detect malicious files, which helped improve the trained model's robustness and generalization. Prediction of the malware in the infected network using the proposed SB-BR-STM showing significant improvements over the traditional networks. In the malware images, textural variation is better explored using the SB-BR-STM by systematically using information related to edge and boundary. Extracting features with different granularity is done using the splitting channel S.B. and merge technique. The TL concepts and

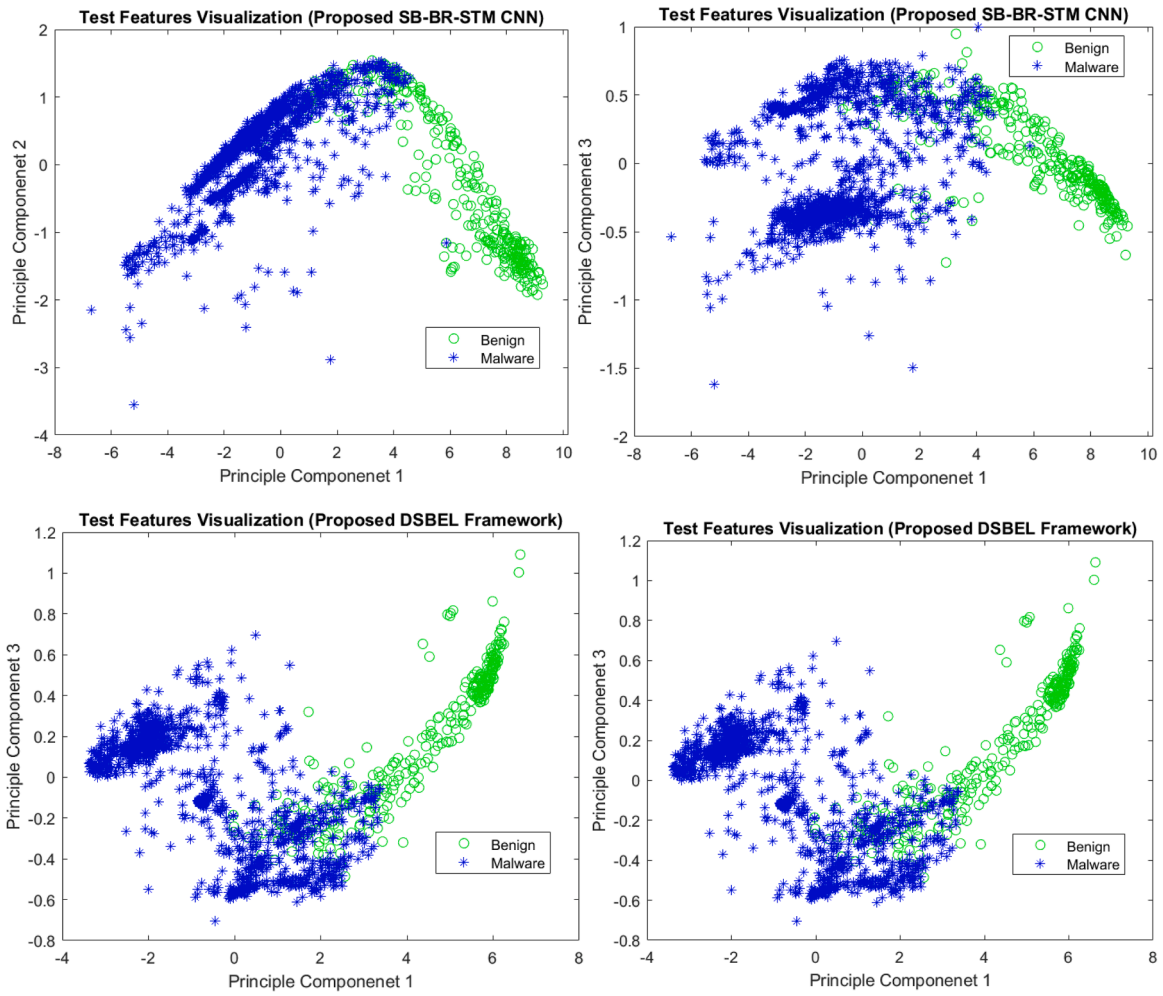

Fig. 9. Feature visualization PC1, PC2, and PC3 for the proposed DSBEL and SB-BR-STM CNN.

STM incorporated improved the performance of the developed SB-BR-STM compared to the existing CNNs. Significance performance is quantified using the MCC, F1-score, AUC-ROC, Accuracy, Precision, and Recall reported in the current study.

5.2. Customized CNN

The existing CNN is customized and compared to its performance with the proposed SB-BR-STM, as shown in Table 4. Training the customized existing models using both training from scratch and Transfer Learning is shown in Tables 4 and 5. From the tables, it can be better analyzed that the models trained using T.L. perform better than training from scratch. The performance gain of our proposed model and the existing networks using T.L. are (0.52-1.04) % Accuracy, (1.98-7.28) % F1-Score, (0.86-3.69) % MCC, (2.89-10.89) % Sensitivity and (0.36-2.61) % of Precision.

5.3. Proposed boosted and ensemble learning framework

The employed framework is a hybrid learning-based strategy in which the proposed CNN extracts features with the addition of strong ML classifiers. We extracted a deep feature vector from the proposed boosted deep CNN at the end layer and fed it into the ensemble ML competitive classifiers: SVM, MLP, and AdaBoostM1. A diverse decision feature space is obtained using the three classifiers and boosted deep feature spaces. Consequently, the boosted features are generated by integrating all these deep features, which maximizes the diversity of feature space, and the discrimination ability of the ML classifier enhances by an ensemble of ML classifiers.

Feature maps are effectively obtained from our proposed model and further fed into the ensemble classifiers (SVM, MLP, and AdabooSTM1) to detect malware in the network packets.

Table 6 and Fig. 7 shows that selected ML classifiers apply one by one and observe the model performance. After that, a majority of voting-based Ensemble ML classifiers are applied, which shows better performance due to the ensemble technique. The architecture's performance assessment measures Accuracy, Recall, Precision, F-score, and MCC. SB-STM effectiveness is evaluated in the last of the experiment for the proposed SB-BR-STM. Performance gain as (1.01-3.71) % Accuracy, (2.58-7.72) % F1-Score, (2.01-5.43) % MCC, (3.22-12.89) % Sensitivity, and (0.73-2.22) % of precision is showing the significant improvements of the proposed SB-BR-STM in the malware detection systems, as depicts in Table 4 and Fig. 7.

5.4. Detection analysis

Detection and precision rate are the main assessment metrics of the effectiveness of a malware detection architecture. Fig. 5 shows the detection performance of existing models and compares them with the model proposed in this research using Accuracy, F1-score, and MCC. The comparison illustrated good Precision by some existing customized models with low Recall. Fig. 6 shows the performance gain ranging from

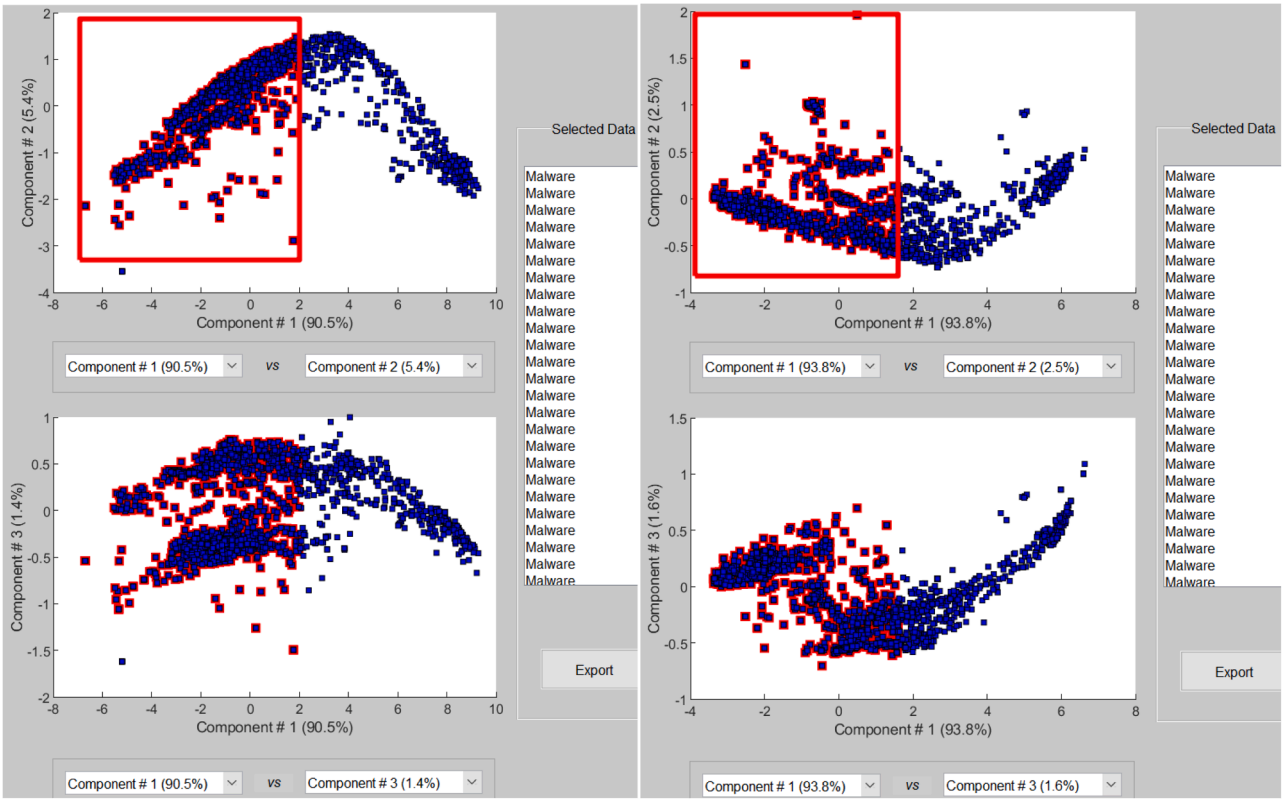

Fig. 9. (*continued*).

minimum to maximum and a comparison with the existing standard CNN architectures. Tables 4 and 6 depict the summarized results of the proposed SB-BR-STM and DSBEL.

5.5. Feature space-based analysis

The decision-making of the image is benign or malicious in the proposed model and can be better analyzed through feature space visualization. Prominent visual features and better discrimination factors of the models are associated, which helps to lower the variance and improves the learning rate of the model. Figs. 8 and 9 show the proposed SB-BR-STM and DSBEL principal components of feature space visualization. IoT malware features are extracted at different levels by channel squeezed and boosted techniques in STM blocks. Moreover, by incorporating channel concatenation, STM boosted the reduced prominent features. An upgrade in identifying distinct and diverse features is shown by visualization for the developed DSBEL and improved detection rate of the IoT malware files.

5.6. ROC and PR-curve analysis

These are the graphical representation of the classifier's capability of discriminating at all the possible values dimension. The optimum detection cut-off of a classifier can significantly access by the ROC curve (Hajian-Tilaki, 2013). Fig. 10 shows the visualized results using the ROC curve with malware images as positive class detection organized by AUC. High values of AUC show the low false positive of the proposed framework with greater sensitivity and considerable performance in the filtration of the malware-infected network.

6. Conclusion

The timely identification of malicious activity through IoT malware detection can provide valuable insights for early detection and mitigation strategies in the future. In this regard, we have introduced a deep DSBEL framework that assembles the developed SB-BR-STM's boosted features and ensemble classifiers to detect malware-attacked packets in the IoT network. The SB-BR-STM comprises the STM blocks that employ TL-based SB ideas and global-boundary and local-regional operations to preserve diverse and boosted features. Moreover, ensemble learning is used to detect ssmalware patterns based on the obtained features from SB-BR-STM for better discrimination and generalization of the DSBEL framework. The proposed novel hybrid framework is empirically assessed and shows prominent performance with an Accuracy of 98.50%, F1-Score of 97.12%, Recall of 95.97 %, and Precision of 98.42 %. The proposed DSBEL framework may be proficient enough to find attacks of cross platforms malware and stringent environments. The malware includes certain similarities in either forum, and these similar features can help in their identification and detection. In the future, online and android malware detection can be performed with the help of a robust DSBEL framework for real-time realization. Morover, the proposed boosted, hybrid and ensemble framework may be robust against zero-day attacks and cloud IoT server attacks. Furthermore, the robust framework can be used in smart homes, healthcare systems, and industrial control systems to detect malware attacks.

CRediT authorship contribution statement

Conceptualization, writing— review and editing, Supervision, methodology, resources, Software, investigation, Formal analysis, visualization, S.H.K. and W.A., writing— original draft preparation, review and editing, W.U. and H.K.A; Formal analysis, Validation, data curation, J.I. and A.R.; Project administration, Validation, funding acquisition, A.O.A. and T.J.A. All authors have read and agreed to the published version of the manuscript.

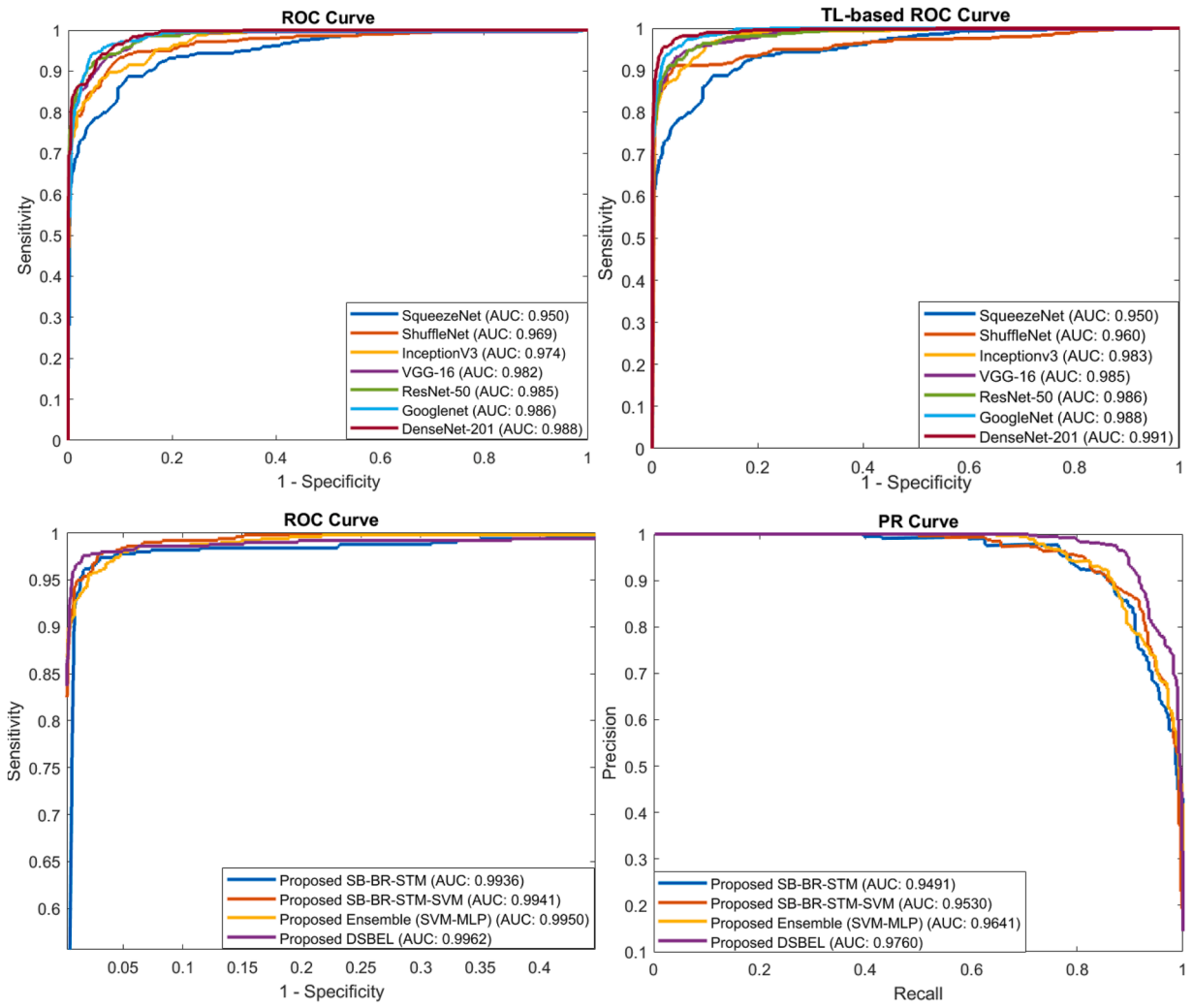

Fig. 10. ROC and PR curves for the developed SB-BR-STM CNN and DSBEL framework contrasts existing CNNs.

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.

Acknowledgments and Funding

The funding of the Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R384), Princess Nourah bint Abdulrahman University, P.O. Box 84428, Riyadh 11671, Saudi Arabia. Moreover, this research work was also funded by Institutional Fund Projects under grant no. IFPIP: 812-611-1443. The authors gratefully acknowledge the technical and financial support provided by the Ministry of Education and King Abdulaziz University, DSR, Jeddah, Saudi Arabia. Furthermore, we thank the Department of Computer Systems Engineering, the University of Engineering and Applied Sciences (UEAS), Swat, Pakistan, for providing the necessary computational resources and a healthy research environment.

References

Madakam, S, Ramaswamy, R, Tripathi, S., 2015. Internet of Things (IoT): a literature review. J. Comput. Commun. 03, 164–173. https://doi.org/10.4236/jcc.2015.35021.

Vuran, MC, Salam, A, Wong, R, Irmak, S., 2018. Internet of underground things in precision agriculture: architecture and technology aspects. Ad Hoc Netw. 81, 160–173. https://doi.org/10.1016/j.adhoc.2018.07.017.

Zafar, MM, Rauf, Z, Sohail, A, Khan, AR, Obaidullah, M, Khan, SH, et al., 2022. Detection of tumour infiltrating lymphocytes in CD3 and CD8 stained histopathological images using a two-phase deep CNN. Photodiagnosis Photodyn. Ther. 37, 102676 https://doi.org/10.1016/j.pdpdt.2021.102676.

SM, Riazul Islam, Kwak, Daehan, M, Humaun Kabir, Hossain, M, Kwak, Kyung-Sup, 2015. The Internet of Things for health care: a comprehensive survey. IEEE Access 3, 678–708. https://doi.org/10.1109/ACCESS.2015.2437951.

Zahoor, MM, Qureshi, SA, Bibi, S, Khan, SH, Khan, A, Ghafoor, U, et al., 2022. A new deep hybrid boosted and ensemble learning-based brain tumor analysis using MRI. Sensors 22, 2726. https://doi.org/10.3390/s22072726.

Khan SH. COVID-19 Detection and Analysis From Lung CT Images using Novel Channel Boosted CNNs 2022. 2209.10963.

Khan, A, Khan, SH, Saif, M, Batool, A, Sohail, A, Waleed Khan, M, 2023. A survey of deep learning techniques for the analysis of COVID-19 and their usability for detecting omicron. J. Exp. Theor. Artif. Intell. 1–43. https://doi.org/10.1080/0952813X.2023.2165724.

Rauf, Z, Sohail, A, Khan, SH, Khan, A, Gwak, J, Maqbool, M., 2023. Attention-guided multi-scale deep object detection framework for lymphocyte analysis in IHC histological images. Microscopy 72, 27–42. https://doi.org/10.1093/jmicro/dfac051.

Iyer, B, Patil, N., 2018. IoT enabled tracking and monitoring sensor for military applications. Int. J. Syst. Assur. Eng. Manag. 9, 1294–1301. https://doi.org/10.1007/s13198-018-0727-8.

Qamar, S, Khan, SH, Arshad, MA, Qamar, M, Gwak, J, Khan, A., 2022. Autonomous drone swarm navigation and multitarget tracking with island policy-based

optimization framework. IEEE Access 10, 91073–91091. https://doi.org/10.1109/ACCESS.2022.3202208.
Arshad, MA, Khan, SH, Qamar, S, Khan, MW, Murtza, I, Gwak, J, et al., 2022. Drone navigation using region and edge exploitation-based deep CNN. IEEE Access 10, 95441–95450. https://doi.org/10.1109/ACCESS.2022.3204876.
Zahoor MM, Khan SH. Brain tumor MRI classification using a novel deep residual and regional CNN 2022. 2211.16571v2.
Mikhalevich, IF, Trapeznikov, VA., 2019. Critical infrastructure security: alignment of views. In: 2019 Systems of Signals Generating and Processing in the Field of on Board Communications, IEEE, pp. 1–5. https://doi.org/10.1109/SOSG.2019.8706821.
Vignau, B, Khoury, R, Hallé, S, Hamou-Lhadj, A., 2021. The evolution of IoT Malwares, from 2008 to 2019: survey, taxonomy, process simulator and perspectives. J. Syst. Archit. 116, 102143 https://doi.org/10.1016/j.sysarc.2021.102143.
Chaganti, R, Ravi, V, Pham, TD., 2022. Deep learning based cross architecture internet of things malware detection and classification. Comput. Secur. 120, 102779 https://doi.org/10.1016/j.cose.2022.102779.
Zahoora, U, Khan, A, Rajarajan, M, Khan, SH, Asam, M, Jamal, T., 2022. Ransomware detection using deep learning based unsupervised feature extraction and a cost sensitive Pareto Ensemble classifier. Sci. Rep. 12, 15647. https://doi.org/10.1038/s41598-022-19443-7.
Ngo, Q-D, Nguyen, H-T, Le, V-H, Nguyen, D-H., 2020. A survey of IoT malware and detection methods based on static features. ICT Express 6, 280–286. https://doi.org/10.1016/j.icte.2020.04.005.
Asam, M, Hussain, SJ, Mohatram, M, Khan, SH, Jamal, T, Zafar, A, et al., 2021. Detection of exceptional malware variants using deep boosted feature spaces and machine learning. Appl. Sci. 11 https://doi.org/10.3390/app112110464.
Asam, M, Khan, SH, Akbar, A, Bibi, S, Jamal, T, Khan, A, et al., 2022. IoT malware detection architecture using a novel channel boosted and squeezed CNN. Sci. Rep. 12, 15498. https://doi.org/10.1038/s41598-022-18936-9.
R, V, DeepMalNet, K.P.S., 2018. Evaluating shallow and deep networks for static PE malware detection. ICT Express 4, 255–258. https://doi.org/10.1016/j.icte.2018.10.006.
Vinayakumar, R, Alazab, M, Soman, KP, Poornachandran, P, Venkatraman, S., 2019. Robust intelligent malware detection using deep learning. IEEE Access 7, 46717–46738. https://doi.org/10.1109/ACCESS.2019.2906934.
Shalaginov A, Øverlier L. A novel study on multinomial classification of x86/x64 Linux ELF malware types and families through deep neural networks. Malware Anal. Using Artif. Intell. Deep Learn., Cham: Springer International Publishing; 2021, p. 437–53. 10.1007/978-3-030-62582-5_17.
Bendiab, G, Shiaeles, S, Alruban, A, Kolokotronis, N., 2020. IoT malware network traffic classification using visual representation and deep learning. In: 2020 6th IEEE Conference on. Network Softwarization, Volume 1, IEEE, pp. 444–449. https://doi.org/10.1109/NetSoft48620.2020.9165381.
Muzaffar, A, Ragab Hassen, H, Lones, MA, Zantout, H, 2022. An in-depth review of machine learning based Android malware detection. Comput. Secur. 121, 102833 https://doi.org/10.1016/j.cose.2022.102833.
Deng, H, Guo, C, Shen, G, Cui, Y, Ping, Y., 2023. MCTVD: a malware classification method based on three-channel visualization and deep learning. Comput. Secur. 126, 103084 https://doi.org/10.1016/j.cose.2022.103084.
Su, J, Danilo Vasconcellos, V, Prasad, S, Daniele, S, Feng, Y, Sakurai, K, 2018. Lightweight classification of IoT malware based on image recognition. In: 2018 IEEE 42nd Annual Computer Software and Applications Conference, IEEE, pp. 664–669. https://doi.org/10.1109/COMPSAC.2018.10315.
Ren, Z, Wu, H, Ning, Q, Hussain, I, Chen, B., 2020. End-to-end malware detection for android IoT devices using deep learning. Ad Hoc Netw. 101, 102098 https://doi.org/10.1016/j.adhoc.2020.102098.
Hussain SJ, Ahmed U, Liaquat H, Mir S, Jhanjhi N, Humayun M. IMIAD: intelligent malware identification for android platform. 2019 Int. Conf. Comput. Inf. Sci., IEEE; 2019, p. 1–6. 10.1109/ICCISci.2019.8716471.
Shafiq, M, Tian, Z, Bashir, AK, Du, X, Guizani, M., 2021. CorrAUC: a malicious bot-IoT traffic detection method in IoT network using machine-learning techniques. IEEE Internet Things J 8, 3242–3254. https://doi.org/10.1109/JIOT.2020.3002255.
Zhang Y, Yang Y, Wang X. A novel android malware detection approach based on convolutional neural network. Proc. 2nd Int. Conf. Cryptogr. Secur. Priv., New York, NY, USA: ACM; 2018, p. 144–9. 10.1145/3199478.3199492.
Xu, K, Li, Y, Deng, RH, Chen, K, 2018. DeepRefiner: multi-layer android malware detection system applying deep neural networks. In: 2018 IEEE European Symposium on Security and Privacy, IEEE, pp. 473–487. https://doi.org/10.1109/EuroSP.2018.00040.
Alzaylaee, MK, Yerima, SY, DL-Droid, Sezer S., 2020. Deep learning based android malware detection using real devices. Comput. Secur. 89, 101663 https://doi.org/10.1016/j.cose.2019.101663.
Ye, Y, Li, T, Adjeroh, D, Iyengar, SS., 2018. A survey on malware detection using data mining techniques. ACM Comput. Surv. 50, 1–40. https://doi.org/10.1145/3073559.
Cozzi, E, Graziano, M, Fratantonio, Y, Balzarotti, D., 2018. Understanding Linux malware. In: 2018 IEEE Symposium on Security and Privacy, IEEE, pp. 161–175. https://doi.org/10.1109/SP.2018.00054.
Palla, TG, Tayeb, S., 2021. Intelligent mirai malware detection in IoT devices. In: 2021 IEEE World AI IoT Congress, IEEE, pp. 0420–0426. https://doi.org/10.1109/AIIoT52608.2021.9454215.
Pa, YMP, Suzuki, S, Yoshioka, K, Matsumoto, T, Kasama, T, Rossow, C., 2016. IoTPOT: a novel honeypot for revealing current IoT threats. J. Inf. Process. 24, 522–533. https://doi.org/10.2197/ipsjjip.24.522.
Khowaja, SA, Khuwaja, P., 2021. Q-learning and LSTM based deep active learning strategy for malware defense in industrial IoT applications. Multimed. Tools Appl. 80, 14637–14663. https://doi.org/10.1007/s11042-020-10371-0.
Naeem, H, Ullah, F, Naeem, MR, Khalid, S, Vasan, D, Jabbar, S, et al., 2020. Malware detection in industrial internet of things based on hybrid image visualization and deep learning model. Ad Hoc Netw. 105, 102154 https://doi.org/10.1016/j.adhoc.2020.102154.
M, G, Sethuraman, SC, 2023. A comprehensive survey on deep learning based malware detection techniques. Comput. Sci. Rev. 47, 100529 https://doi.org/10.1016/j.cosrev.2022.100529.
Shorten, C, Khoshgoftaar, TM., 2019. A survey on image data augmentation for deep learning. J. Big Data 6, 1–48. https://doi.org/10.1186/S40537-019-0197-0/FIGURES/33.
Khan, SH, Shah, NS, Nuzhat, R, Majid, A, Alquhayz, H, Khan, A., 2022. Malaria parasite classification framework using a novel channel squeezed and boosted CNN. Microscopy 71, 271–282. https://doi.org/10.1093/jmicro/dfac027.
Khan, SH, Khan, A, Lee, YS, Hassan, M, Jeong, WK., 2022. Segmentation of shoulder muscle MRI using a new region and edge based deep auto-encoder. Multimed. Tools Appl. https://doi.org/10.1007/s11042-022-14061-x.
Gardner, M., Dorling, S., 1998. Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences. Atmos. Environ. 32, 2627–2636. https://doi.org/10.1016/S1352-2310(97)00447-0.
Schapire RE. Explaining adaboost. Empir Inference Festschrift Honor Vladimir N Vapnik 2013:37–52. 10.1007/978-3-642-41136-6_5/COVER.
Mahmood, R, Mirzaei, N, Malek, S., 2014. EvoDroid: segmented evolutionary testing of Android apps. In: Proceedings of the 22nd ACM SIGSOFT International Symposium on Foundations of Software Engineering, New York, NY, USA: ACM, pp. 599–609. https://doi.org/10.1145/2635868.2635896.
Vidas T, Tan J, Nahata J, Tan CL, Christin N, Tague P. A5 automated analysis of adversarial android applications. Proc. 4th ACM Work. Secur. Priv. Smartphones Mob. Devices, New York, NY, USA: ACM; 2014, p. 39–50. 10.1145/2666620.2666630.
Khan, A, Sohail, A, Zahoora, U, Qureshi, AS., 2020. A survey of the recent architectures of deep convolutional neural networks. Artif. Intell. Rev. 1–68. https://doi.org/10.1007/s10462-020-09825-6.
Wan, T-L, Ban, T, Lee, Y-T, Cheng, S-M, Isawa, R, Takahashi, T, et al., 2020. IoT-malware detection based on byte sequences of executable files. In: 2020 15th Asia Joint Conference on Information Security, IEEE, pp. 143–150. 10.1109/AsiaJCIS50894.2020.00033.
Almomani, I, Alkhayer, A, El-Shafai, W., 2022. An automated vision-based deep learning model for efficient detection of android malware attacks. IEEE Access 10, 2700–2720. https://doi.org/10.1109/ACCESS.2022.3140341.
Hajian-Tilaki, K., 2013. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Casp. J. Intern. Med. 4, 627–635.

Dr. Saddam Hussain Khan (Gold Medalist) is an Assistant Professor at the Department of Computer System Engineering, the University of Engineering and Applied Science (UEAS) in Swat, Pakistan. He received the bachelor's degree from the University of Engineering and Technology (UET) Peshawar, the master's degree from the University of Engineering and Technology (UET) Taxila, and the Ph.D. degree from the Pakistan Institute of Engineering and Applied Sciences (PIEAS), Islamabad, Pakistan. His research interests include computer vision, deep neural networks, machine learning, medical image analysis, biomedical and bioinformatics, deep learning in cyber security, Block Chain and IoT, Federated learning, and deep reinforcement learning. E-Mail: saddamhkhan@ueas.edu.pk ORCID: https://orcid.org/0000-0002-6681-1987 Google Scholar: https://scholar.google.com/citations?user=j_ImpdYAAAAJ&hl=en

Dr.Tahani Jaser Alahmadi received a Ph.D. degree from the Faculty of Information Technology, Griffith University in Australia, in 2019. She has B.S. in Computer Science and M.S. in Information Technology (Data Management). She is currently working as an Assistant Professor at the Faculty of Computer and Information Sciences, IS Department, Princess Nourah bint Abdulrahman University, Saudi Arabia. She is a Member of the Golden Key Society and Media Access Australia. She received multiple awards such as Google Doctoral Consortium Award, and the Institute for Integrated and Intelligent Systems (IIIS) award for Quality and Impact Research. Her research interests are innovative research methods in data analysis and mining, machine learning, pattern recognition, image processing, Accessibility, Usability and Sentiment Analysis. Tahani Jaser Alahmadi, tjalahmadi@pnu.edu.sa https://orcid.org/0000-0002-0067-692X Tahani Alahmadi - Google Scholar

Engr. Wasi Ullah is Lab Engineer at the Department of Computer Systems Engineering, University of Engineering and Applied Science (UEAS), Swat, Pakistan. Previously, He worked as Lecturer at Department of Electrical Engineering, Abasyn University, Peshawar. He received the Bachelor's Degree from the University of Engineering and Technology (UET) Peshawar, and Master's Degree from the Abasyn University Peshawar. His research interests include computer vision, deep neural networks, machine learning, deep learning in cyber security, IoT, Computer Networks and Mobile Adhoc Networks. **E-Mail:** wasi.ullah@ueas.edu.pk

Dr. Javed Iqbal (SM'18) is currently working as Associate Professor and Chair in the Department of Computer Systems Engineering at University of Engineering and Applied Sciences, Swat Pakistan. Prior to his recent appointment, he worked as Associate Professor and Team Lead 'Communication and Network Research Group' in Electrical Engineering Department of Sarhad University of Science and IT, Peshawar, Pakistan. He completed his Ph.D. from Politecnico Di Torino and MS from BTH, Sweden in 2015 and 2009 respectively. He was awarded scholarship for Ph.D studies from Higher Education Commission of Pakistan (HEC), Pakistan. Dr. Javed is involved in reviewing for various reputed impact factor journals and conferences including IEEE. He has published several papers in reputed journals and conferences and two book chapters. He is an active Senior Member of IEEE and also serving as as Associate Editor, IEEE Access. His recent research interests are the application of Machine Learning and Data Sciences for smart systems especially in the Healthcare and Transportation domain through Knowledge management, predictive analysis, and visualization. **E-Mail:** javed@ueas.edu.pk **Google Scholar:** https://scholar.google.com/citations?user=BWFXhNIAAAAJ&hl=en&oi=ao

Dr. Azizur Rahim (Member, IEEE) received an M.S. degree in Electrical Engineering from COMSATS, Islamabad, Pakistan, and a Ph.D. Degree in Computer Application Technologies from The Alpha Laboratory, School of Software, Dalian University of Technology, Dalian, China. Currently, he is working as Assistant Professor with the Department of Computer Systems Engineering, University of Engineering and Applied Sciences, Swat, Pakistan. Previously, he worked as an Assistant Professor at the School of Electrical Engineering and Computer Science (SEECS), National University of Science and Technology (NUST), Islamabad, Pakistan. His research interests include Mobile and Social Computing, Mobile Social Networks, Vehicular Social Networks Artificial Intelligence, Big Data Analytics, Data Communication Networks. E-Mail: aziz.rahim@ieee.org Google Scholar: https://scholar.google.com/citations?user=AmUvey0AAAAJ&hl=en

Dr. Hend Khalid Alkahtani is an assistant professor at Princess Nourah Bint Abdulrahman University, College of Computer and Information Sciences IS Department. She has a PhD in information security, Department of computer science, Loughborough University (2018), Master of Science with Concentration in Information Management, Department of Engineering Management, The George Washington University (1993) and Bachelor of science, Computer Science, School of Engineering and Applied Science, The George Washington University (1988-1992). She has 23 years of Work experience as a lecturer, worked as a computer center president and as a statistic center president. She is a member in IEEE. She received an award from SIDF Academy: Leading Creative Transformation in Critical Time Program, Stanford University, Center for Professional Development. Hend Khalid Alkahtani, hkalqahtani@pnu.edu.sa https://orcid.org/0000-0001-7507-5267

Dr. Wajdi Alghamdi is PhD holder specialized in Computer Science (Data mining) from the Department of Computing, Goldsmiths College, University of London, London, United Kingdom. His is currently an assistant professor at Information Technology Department, Computing and Information Technology College, King Abdul-Aziz University, Jeddah, Saudi Arabia. He is mostly interested in Knowledge Discovery in Databases, Data Mining and Statistical Computing. His main research is focusing on applying Machine Learning and Statistical Learning methods to genotype, phenotype and clinical data in-order to discover patterns of interest, including the identification of clinical and genetic predictors with respect to diseases.

Dr. Alaa Omran Almagrabi received B.Sc (Computer science) degree from Jeddah teaching College in 2003 and the Master degree in Information Technology in 2009 and Ph.D (Computer Science in 2014) from La Trobe University in Melbourne, Australia. In 2014, he was appointed as Assistant Professor with the Department of Information System in the Computer Science and Information Technology at the University of King Abdulaziz in Jeddah, Saudi Arabia. In 2019, he was promoted to Associate Professor rank in the department of Information System. In 2023, he was promoted to full Professor rank in the department of Information System. My research area includes Pervasive Computing, Networking, Data mining, System analysis and design, and ontology domains.