International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

An Enhancement to Fused and Cascaded Squeeze and Excitation Network for Pneumonia Prediction with Deep Learning Techniques

Ramitha M A
Research Scholar, Department of Computer Science and Engineering,
Karpagam Academy of Higher Education,
Coimbatore, Tamil Nadu, India,
ramithaa@gmail.com

Dr. N Mohanasundaram
Professor, Department of Computer Science and Engineering,
Karpagam Academy of Higher Education,
Coimbatore, Tamil Nadu, India,
itismemohan@gmail.com

Dr. R. Santhosh
Professor, Department of Computer Science and Engineering,
Faculty of Engineering,
Karpagam Academy of Higher Education,
Coimbatore, Tamil Nadu, India,
santhoshrd@gmail.com

Abstract--Pneumonia is a rapidly spreading disease that poses a significant risk to the health and welfare of those affected. Accurately diagnosing pneumonia from a biomedical perspective requires the use of various diagnostic tools and the assessment of multiple clinical characteristics. However, the lack of available experts and equipment hinders the proper diagnosis of this disease. As deep learning is getting revolutionized in terms of these fields, the utilization of such advanced models brings more efficacy. This study aims to comprehensively enhance both the Inception V3-based squeeze-excitation network and the Densenet 121 with squeeze-excitation network and analyze the performance of both in addition to state-of-the-art models, for accurate pneumonia detection from medical images. This paper aims to develop a robust deep-learning model capable of effectively identifying Pneumonia by analyzing the characteristics derived from chest X-ray images. The study seeks to identify the most efficient model for pneumonia detection by comparing and assessing the performance of these models. The study consists of three main phases. a) Data collection from Kermany and RSNA Pneumonia challenge datasets, which consist of X-ray images, b) Preprocessing the data to remove noises and anomalies using a Gaussian filter to improve image quality, and c) Conducting the classification task by utilizing the Densenet 121-based squeeze-excitation network and the Inception V3-based squeeze-excitation network. The Inception V3 SENet model demonstrated superior performance. With an accuracy of 0.9539%, recall of 0.95%, and f1-score of 0.95%, it surpassed the Densenet model's performance in terms of accuracy (0.74), recall (0.81), and f1-score (0.69), as well as outperformed other state-of-the-art models.

Keywords: Classification, Deep learning, Inception v3, Densenet 121, Lung X-ray, Pneumonia, Squeeze-excitation

I. INTRODUCTION

The COVID-19 pandemic began in 2019 with the initial identification of the coronavirus disease in Wuhan, China. From there, it rapidly propagated to various nations, causing a significant escalation in both its scale and severity. Consequently, the World Health Organization categorized it as a worldwide pandemic. The impact of this viral infection varied across different nations, with the outbreak peaking at different times and resulting in a global crisis. As of 2022, thousands of new cases and hundreds of deaths are recorded each day. As per [1], the global count of COVID-19 cases has surpassed 26.4 million, leading to 5.25 million fatalities. The symptoms and medical complications experienced by an individual are influenced by their

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

underlying health conditions or co-existing ailments. Nevertheless, the current diagnostic methods like RT-PCR, antigen, and antibody tests have drawbacks as they are expensive, intricate, and need specialized technology that may not be accessible in remote regions [2]. Although RT-PCR is considered the standard, its accuracy is only 70%.

Apart from other complexities linked to COVID-19 exposure, there is a possibility of contracting pneumonia in the lungs, where fluid accumulates in the air sacs. To verify the presence of COVID-19 infection-induced pneumonia, it is crucial to perform a chest X-ray examination[3]. The likelihood of misdiagnosis and false positives is higher due to the similarity of symptoms observed in COVID-19 and other lung inflammatory infections. This paper introduces an efficient deep-learning model that provides a detailed and informative classification of pneumonia, even in the face of current obstacles arising from the absence of specific expertise and tools essential for precise diagnosis and accurate results. The various steps involved in the classification process are explained thoroughly in different sections of the paper.

A. Key Highlights

The purpose of this research is to present a successful combination strategy for categorizing pneumonia, along with the following goals:

- Develop an effective hybrid method for classifying pneumonia using inception v3 squeeze-excitation networks.
- Most state-of-the-art models are developed as standalone solutions, which could potentially lead to issues with efficiency and performance.
- To demonstrate the superiority and effectiveness of integrating the SE network into the models, we will analyze and compare their performance metrics with other state-of-the-art works.
- The proposed system underwent experimental analysis utilizing state-of-the-art models, revealing exceptional performance across multiple metrics.

II. LITERATURE REVIEW

Shah & Shah (2022) [4] conducted a thorough investigation of various models utilized in the detection of pneumonia by analyzing chest X-ray images extensively. The effectiveness of various models with similar objectives is assessed in the paper through extensive quantitative analysis. While the paper also discusses other models such as an adversarial training-based approach, MUXConv, NSGANetV1, P-Net, Quantum Neural Network, GAN, and Ensemble Feature Scheme (EFS), the results highlight the exceptional performance of CheXNet. The authors aimed to create a valuable resource for both novices and established researchers in the field by compiling and analyzing a vast amount of research information, encompassing datasets, model designs, and outcomes. It is important to note that the limitations of quantifying certain trade-offs solely through qualitative analysis are acknowledged in the paper. Nonetheless, the insights provided in the study contribute significantly to the pneumonia detection community, with CheXNet emerging as a particularly noteworthy model.

Jain et al. (2022) [5]. The utilization of state-of-the-art deep-learning AI technology is suggested for the automatic identification and classification of pneumonia in chest X-ray images. This method is both cost-efficient and convenient, given the widespread availability of these images. The study employed a set of seven pre-trained Convolutional Neural Networks (CNNs) known as VGG16, VGG19, DenseNet201, Xception, InceptionV3, NasnetMobile, and ResNet152. These CNNs were utilized to accurately categorize pneumonia into up to five distinct classes. The database used for the research comprised a total of 18,603 scans, encompassing cases involving two, three, and five classes of pneumonia. The most accurate classification output was obtained by DenseNet201, VGG16, and VGG19, accordingly.

Ravi et al. (2022) [6] Suggested a method for classifying pediatric pneumonia using X-ray images by utilizing a deep feature fusion technique that incorporates transfer learning and a stacked ensemble meta-classifier. The emphasis is on large-scale learning that takes into account cost sensitivity. Sophisticated features are integrated into transfer learning models such as Xception, InceptionResNetV2, DenseNet201, and NASNetMobile to assess the significance of various pneumonia types, while also considering the issue of class imbalance when performing the backpropagation process. The penultimate layer, which is the global average pooling, is utilized to extract features. Afterwards, the aforementioned characteristics undergo Kernel Principal Component Analysis (KPCA) in order to decrease their dimensionality. The condensed attributes are merged and input into a meta-classifier that is stacked with multiple layers. This classifier then evaluates if the CXR image exhibits indications of pneumonia or if it's considered normal.

Hu et al. (2022) [7]. A design proposal was presented, recommending the integration of radionics to enhance the design using a collection of 812 chest X-ray images. The research employed three sets of images, consisting of 262 cases of COVID-19, 288 cases of non-COVID-19 pneumonia, and 262 healthy subjects. These sets were identified as the training, validation, and independent test groups. The models were trained through 50 iterations, with the training and validation data randomly distributed within the training-validation set at a ratio of 7:1. The deep neural network structures were evaluated using sensitivity, specificity, accuracy, and ROC curves, which included AUC measurements for each set. The findings indicate that among the models considered, VGG-19 demonstrated superior performance, surpassing the other discussed models

Bharati, et al [8]. Presenting a revolutionary method for identifying pneumonia by analyzing chest X-ray images. This innovative strategy combines deep residual neural networks (ResNets) to extract crucial characteristics, along with the utilization of Support Vector Machines (SVM) for precise classification. By leveraging the hierarchical features extracted by ResNets and the classification capabilities of SVM, the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

introduced scheme achieves superior precision in pneumonia classification compared to other concurrent approaches. This highlights the effectiveness of combining ResNets with SVM as a promising solution for accurate and efficient pneumonia diagnosis through chest radiography.

III. METHODOLOGY

The proposed framework's general structure is depicted in Figures 1 and 2, which entails providing ample input data to train the deep learning model, irrespective of the specific deep learning framework being used. So, for that data collection of lung X-ray images in order to feed the further stages. Once the data are collected, these are undergone preprocessing for the removal of anomalies and noises in raw files. Finally, the results obtained from both models, namely the Inception V3-based squeeze-excitation network hybrid and the Densenet 121-based squeeze-excitation network, are fed into the classifier.

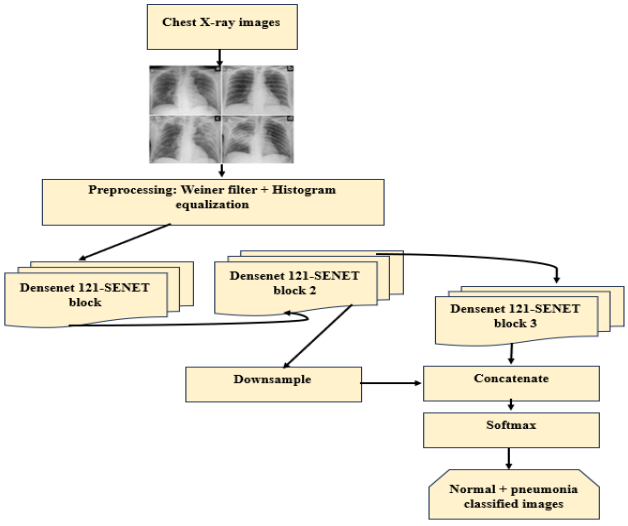


Figure. 1. The overall architecture of the Densenet 121-SE

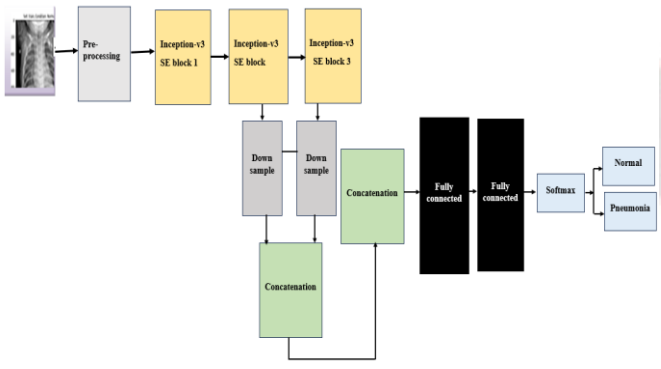


Figure. 2. Proposed Inception-V3 -SE architecture diagram

A. *Data Collection*

The dataset was bought to execute the recommended processes. The model was prepared on the Kermany dataset and the RSNA Pneumonia Detection Challenge dataset in this review [9]. This technique will bring about a stage-by-stage detachment. There is a subcategory for each photograph class (Pneumonia/Typical) in every one of the 3 organizers (train, test, and Val). At the Guangzhou Ladies and Youngsters' Clinical Center in Guangzhou, images of the chest (front/back) were taken. The patients' normal clinical treatment included chest X-beam screening. The lucidity of all chest radiographs was then checked, and the low quality. The images were assessed by two experts before being allowed to be utilized in tutoring the computer-based intelligence framework. A third individual checked the evaluation set for reviewing issues.

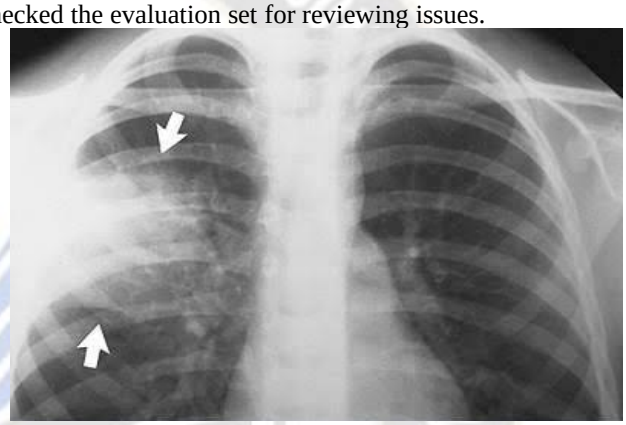


Figure 3. Pneumonia indicating chest X-ray images

B. *Preprocessing*

All chest radiographs were checked and pre-processed to exclude scans of low quality or images that were unreadable, to increase the accuracy of chest X-ray assessment.

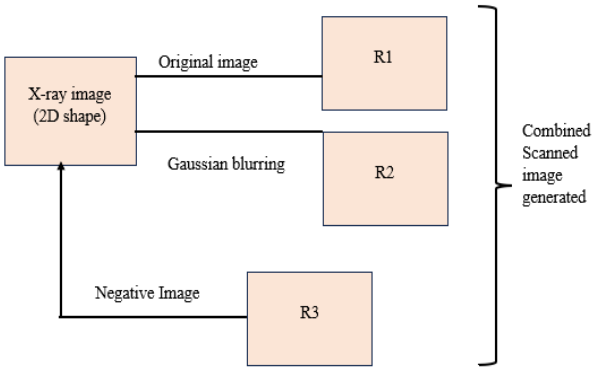


Figure 4. Pre-Processing stages

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

The augmentation technique will magnify or diminish the visuals according to the rescale operation.

For picture fading, the Gaussian filter was used, which is a two-dimensional convolution operator that employs Gaussian kernels (removing detail and noise). The main idea was to discretize the Gaussian function and utilize the Gaussian function's value at distinct places as the weight, then compute a weighted average over a specific range of neighbourhood for each pixel of the obtained gray matrix, successfully eliminating Gaussian noise. Eqs. (1) and (2) depict its one-dimensional and two-dimensional Gaussian distributions (2).

$$G(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{x^2}{2\sigma^2}} \ldots\ldots\ldots\ldots\ldots\ldots (1)$$

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}} \ldots\ldots\ldots\ldots\ldots (2)$$

The standard deviation influences how smoothly a picture is taken after Gaussian filtering. The weighted average of the domain pixels is the outcome, the higher the weight, the closer the pixel is to the centre. The image after and before Gaussian blurring is shown in Figure 5.

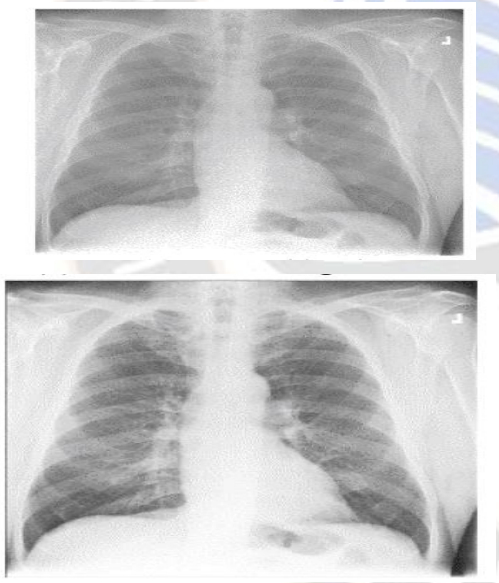


Figure 5. Comparison of the (a) image with noise and (b) Gaussian-filtered image

To create a negative image, the process entails subtracting the maximum intensity value from each pixel. In the context of an 8-bit grayscale image, where the maximum intensity is 255, the resulting image is obtained by subtracting 255 from every pixel.

The given x-ray image, the negative image of the real x-ray image and the image that blurred using a Gaussian filter are combined as shown in Figure 4 to get a 3D image. This is the input image to the proposed architecture, as given in Figure 2.

The activation patterns of various nodes within the neural network vary based on the data, as distinct inputs trigger the activation of different neurons.

C. Classification

1) Classification using Densenet121-SE model

The Dense-SEnet architecture incorporates a sequential flow of operations starting with the input image passing through a DenseNet121 model. The resulting output from DenseNet121 is then utilized as input for the SE block. The Squeeze-and-Excitation (SE) block is a computational unit that can be employed with any transformation Ftr in order to enhance its performance. X → U, where X is a real number set R H′×W′×C′ and U belongs to R H×W×C. In this scenario, Ftr is considered a convolutional operator. The collection V consists of the filter kernels that have been learned, expressed as [v1, v2, . . . , vC], with each vc representing the elements of the corresponding filter. The outcomes produced by Ftr are denoted as U, which is composed of [u1, u2, . . . , uC], where C indicates the number of output channels.

$$u_c = v_c * X = \sum v_c^g * x^g. \qquad (3)$$

The symbol ∗ represents convolution, where vc = [v1c, v2c, . . . , vC′c] and X = [x1, x2, . . . , xC′] (without bias terms for simplicity). The vsc kernel operates in 2D on a specific channel of vc and its corresponding channel of X, resulting in an output that is generated by summing across all channels. This summation considers the interdependence between channels in vc and its interaction with the filters' spatial correlation. Our objective is to enhance the network's ability to detect essential features while suppressing unimportant ones, allowing downstream transformations to exploit informative features fully. To accomplish this, we suggest a technique for modelling channel interdependencies and modifying filter responses before forwarding them to the subsequent transformation. This is achieved through a two-step process known as squeeze and excitation, which is depicted in Figure 6 using an SE building block.

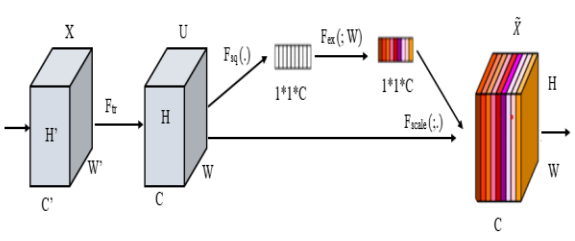


Figure 6. Squeeze - Excitation block for classification

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Squeeze: In order to tackle the issue of exploiting channel dependencies, the initial stage involves analyzing the signal transmitted to each channel within the output features[10]. The system's filters are specifically designed to function within a limited receptive field, which implies that each unit of the transformed output U is unable to utilize contextual information beyond that particular region. This poses a particular challenge, especially in the lower layers of the network where the receptive field sizes are relatively small. To overcome this challenge, our suggestion is to integrate global spatial information into a channel descriptor. This can be achieved by implementing global average pooling, which produces channel-specific statistics. In essence, the spatial dimensions of U represented as H x W, are reduced to produce a statistic z ∈ RC. The calculation of each element in z (denoted as the c-th element) is as follows:

$$z_c = F_{sq}(u_c) = \frac{1}{H*W}\sum_{i=1}^{H}\ \sum_{j=1}^{W}\ u_c(i,j). \qquad (4)$$

Excitation: The utilization of the information acquired through the squeeze operation involves a two-step process [11]. The second step is designed to capture dependencies between channels. Two conditions must be met by the function employed in the second step to accomplish this objective. Firstly, it must be flexible enough to learn non-linear interactions between channels. We aim to expand the emphasis on multiple channels rather than focusing solely on one channel by instructing the model about a non-exclusive connection. In order to accomplish this, we opted to employ a simple gating mechanism that incorporates a sigmoid activation.

$$s = F_{ex}(z,W) = \sigma\big(g(z,W)\big) = \sigma\big(W_2\delta(W_1 z)\big), \qquad (5)$$

The ReLU [30] function is denoted by δ, while W1 ∈ RCr ×C and W2 ∈ RC× Cr. In order to enhance the simplicity and flexibility of the model, we introduce a gating mechanism that utilizes two fully connected (FC) layers in conjunction with a non-linear activation function. This mechanism consists of a layer responsible for reducing dimensionality, employing parameters W1 and achieving a reduction ratio of r. This is followed by a ReLU layer, and finally, a layer that expands dimensionality, utilizing parameters W2. The resulting output of the block is obtained by scaling the transformation output U according to the activations.

$$\tilde{x}_c = F_{scale}(u_c, s_c) = s_c.u_c, \qquad (6)$$

where Xe = [xe1, xe2, . . . , xeC] and Fscale (uc, sc) represents the multiplication of a scalar s with each channel of the feature map u_c, which is a matrix of size H x W. SE-based concentration has a modest computational effort and has no effect on training time. These will be then passed to Densenet 121 and Inception-V3. This study investigates the classification performance of two advanced deep learning models, specifically Densenet-121 SE and Inception-v3 SE when applied to pre-processed images. Densenet-121 is a modern convolutional neural network structure developed specifically for recognizing visual objects. It has gained recognition for its ability to deliver cutting-edge performance while using a smaller parameter count.

DenseNet and ResNet share many similarities, but there are key differences in their architectures. DenseNet utilizes concatenation (.) as a means to fuse the output of the preceding layer with the subsequent layer, whereas ResNet employs addition (+) to combine the previous layer with the upcoming layers. The primary objective of DenseNet is to tackle this concern by densely connecting all layers. In this study, we specifically utilized DenseNet-121, which has a total of 121 layers [(5 + (6 + 12 + 24 + 16) × 2) = 121]. The structure of this architecture includes 5 convolutional and pooling layers, 3 transition layers with scales of 6, 12, and 24, 1 classification layer with a scale of 16, and 2 dense blocks that utilize 1x1 and 3x3 convolutions. In general, conventional convolutional neural networks (CNNs) compute the output of each layer (lth) by applying a non-linear transformation (Hl (.)) to the output of the preceding layer (Xl−1).

$$x_2 = m_{\ell}\big(x_{i_{-1}}\big)^2 \qquad (7)$$

The layer output functionality maps are not truly summarized by them. Instead, they are concatenated. DenseNet introduces a straightforward communication framework for improving information exchange among layers. In this framework, each layer (lth layer) incorporates inputs from the preceding layers' features. As a result, the equation undergoes another transformation, yielding:

$$X_1 = H_l\big|\big(X_0, X_1, X_{2+\cdots}, X_{l-1}\big)\big| \qquad (8)$$

The unified tensor [X0, X1, X2,..., Xl - 1] is created by combining the output maps from preceding layers. In this context, the function Hl(.) denotes a non-linear transformation that involves three essential operations: batch normalization (BN), activation (ReLU), and pooling and convolution (CONV). Figure 7 illustrates the structure of DenseNet. The growth rate, denoted as k, has a vital role in the generalization of the lth layer in the network. It is calculated as k[l] = (k[0] + k(l - 1)), where k[0] represents the initial number of channels.

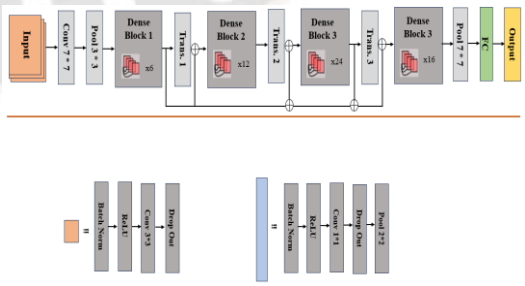


Figure 7. Schema diagram of Densenet 121

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

In the initial stage of this work, three DenseNet121-SENet blocks are sequentially interconnected. The feature map generated by the second DenseSENet block undergoes downsampling. The resulting feature map and the output from the third DenseSENet block are then concatenated and passed through a softmax classifier. This classifier generates the final output, yielding an accuracy of 74%, recall of 81%, and f1-score of 69%.

2) Classification using Inception v3- SE model

A custom-built network requires a substantial dataset for training and its architecture must undergo validation before it can be utilized in practical applications [12]. Because they have been carefully assessed by using network architectures, it makes acceptance easier and promotes dependability. We want systems that are easy to fine-tune and function better, so the training course can be completed faster than establishing a network from the roots up. InceptionV3 has been chosen over other prominent systems [13]. On the ImageNet dataset, the models were picked because they outperformed the rest. The structures of these systems were tested to evaluate which one performs better with CXR images. In comparison to their deeper variants, they have fewer parameters to train and need fewer computer resources.

Another method of network optimization is to weight characteristic channels produced by convolutions and those giving each filter a varied priority. This plan is referred to as "worldwide attention" since the value of a channel is given to all parts of it worldwide. To make pre-trained systems highly robust to datasets, Squeeze-Excitation (SE) based concentration [14] is employed. Furthermore, information help for medical judgments must be based only on well-studied procedures in a process as crucial as Covid evaluation. The SE mechanism is a very well channel-wise boosting technique that may be added to any CNN using the software. In Inception-SEnet, the input image first goes through an Inception layer before it goes through the SENet Block.

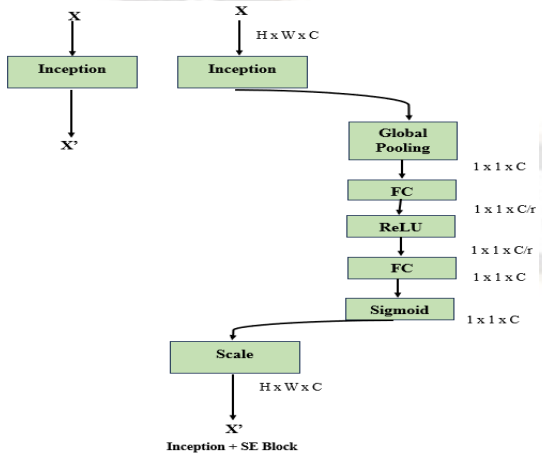


Figure 8. The original Inception block's schema (left), and the Inception block after the SE Block is added (right).

The SE-based focus introduces a low computational load without affecting the training time [15]. Figure 8 illustrates the placement strategy of the SE block for the InceptionV3 networks, showcasing its simple yet effective attention system. In this context, the SE block is exclusively utilized in conjunction with the final residual and inception blocks.

The pre-processed output serves as the input for the first Inception-SENet block in the proposed system. Three Inception-SENet blocks, followed by a max pool layer, are sequentially interconnected. Side outputs are extracted from the second and third Inception-SENet blocks. These side outputs, obtained through separate down-sampling (average pooling), are then concatenated together. The resulting feature maps are further concatenated with the output from the third SENet block. Subsequently, this composite output passes through two fully connected layers. Finally, the softmax classifier, as shown in Figure 2, categorizes the image as either normal or affected by pneumonia. The task makes use of the RMS (Root Mean Squared) Propagation Optimizer, with the evaluation of the classification model's performance carried out using the cross-entropy loss function.

The initial classification scores obtained from the InceptionV3 model pre-trained on the ImageNet classifier were adjusted by utilizing output units of size 2 to build a new model [16]. A random search of 1e-6, 1e-5, 1e-4, and 1e-3 learning rates resulted in the 1e-4 training process being picked. Using a grid search on 4, 8, 16, and 32, the batch size of 16 was found to be the most effective. Binary cross-entropy was selected as the loss. We used our database to retrain all of the network layers.

The size of the tensor produced in the 2D convolutional (conv) layer of a CNN is (H×W×C), where H×W represents the spatial dimensions of the feature map at that stage, and C represents the number of levels obtained by convolving with a filter element. To assign weights to each channel within the convolutional layer and minimize redundancy, the concentration module, utilizing the Squeeze-Excitation (SE) technique [17], can be applied to any convolutional level. Channels with low contributions will be prioritized lesser. Figure 10 depicts the general structure of the SE block.

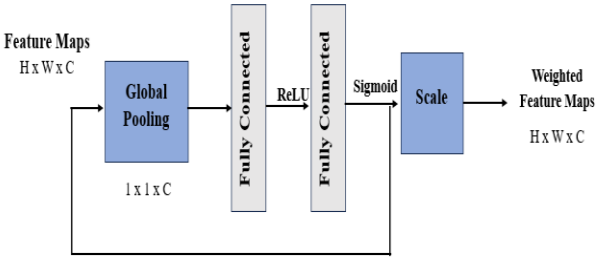


Figure 9. Squeeze-Excitation (SE) Block, H ×W - the feature map's height and width, and C - the layer's channels

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

The following points make up the SE attention block:

Squeeze Module: Applying Global Average Pooling (GAP) Eq 7, where it compresses the data from each of the previous convolution layer's extracted features into a single value per channel. Here, i and j correspond to the map's H and W proportions. From input feature maps of size (H × W × C), the tensor (1 × 1 × C) is reduced.

$$z_c = \frac{1}{H \times W} \sum_{H}^{i=1} \sum_{W}^{j=1} a_c(i,j) \qquad (9)$$

Excitation Module: During the training, it discovers the channels' adaptable scaling weights. The stimulating component is a Multi-Layer Perceptron (MLP) with a lower element. Inside this input, hidden, and output layers of the MLP, C neurons are being used. The size (1 × 1 × C) tensor is obtained by activating the network output with a sigmoid (Eq 8) that scales the weights from 0 to 1.

$$sig\ (x) = \frac{1}{1+e^{-x}} \qquad (10)$$

Scale Module: The attribute maps generated from the convolutional layer are enhanced by incorporating scaling factors through element-wise multiplication [18]. Each (H W) map is then scaled using the corresponding scaling weight.

IV.PERFORMANCE ANALYSIS

The specified hardware requirements for running the model include a Ryzen 5/6 series CPU, an NV GTX graphics card, a 1 TB HDD, and the Windows 10 operating system. The software specifications include pytorch, a Python library for creating deep learning models, and Google Collaboratory, an open-source environment for developing these models. Models are trained under 50 epochs. Experimental analysis is carried out on different state-of-the-art models (represented as L) to assess their performance by measuring accuracy, sensitivity, specificity, recall, precision, and F1-score, and examining a confusion matrix.

Table 1 provides a thorough evaluation of precision, sensitivity, and specificity, whereas Figure 10 graphically illustrates the comparative efficacy of various models. The suggested Inception-V3-SE model stands out in terms of superior training efficiency and the use of a hybrid approach, resulting in highly precise image classification.

TABLE I COMPARISON ANALYSIS OF THE ACCURACY, SENSITIVITY, SPECIFICITY

Models	Accuracy	Sensitivity	Specificity
CheXNet [6]	83	88	90
DenseNet201 [7]	87	93	95
Xception [8]	82	95	94
VGG-19 [9]	86	94	91
ResNet + SVM [10]	87	94	93
D121-SE	74	84	87
Inception-V3-SE	95.39	98	93

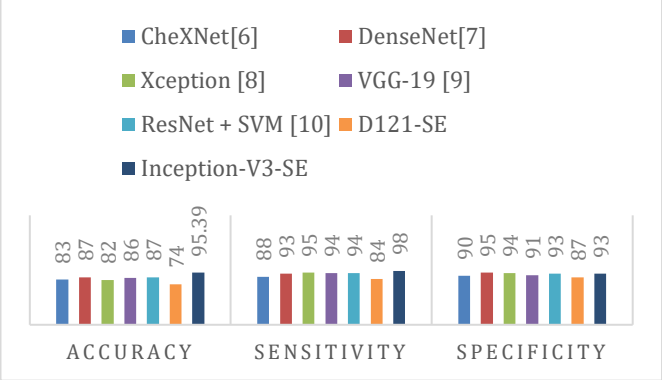


Figure 10. Models vs Accuracy, sensitivity, specificity

Table 2 displays an analysis comparing recall, precision, and F1-score, while Figure 11 illustrates various models graphically. The proposed model Inception-V3-SE performs better than other models.

TABLE II COMPARISON OF PRECISION, RECALL, FI-SCORE

Models	Precision	Recall	F1-score
CheXNet [6]	82	85	84
DenseNet201 [7]	87	81	80
Xception [8]	84	87	86
VGG-19 [9]	81	80	79
ResNet + SVM [10]	88	86	87
D121-SE	81	74	69
Inception-V3-SE	95	95	95

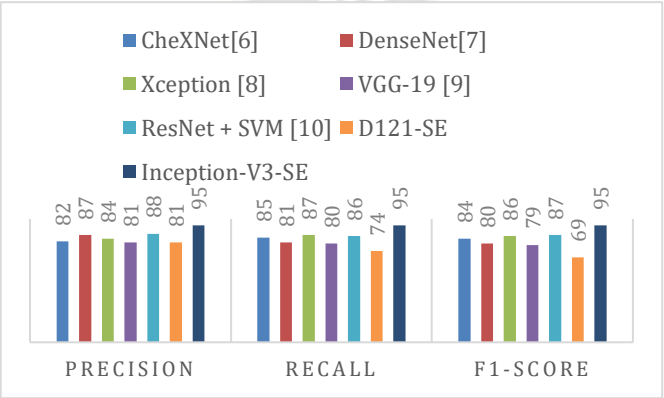


Fig. 11. Models' vs Precision, Recall, F1-SCORE

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

Based on the above tables and graphical representation of performance metrics which contains Accuracy, Sensitivity, Specificity, Precision, Recall, and F1-score, the Inception-v3-SE model demonstrates superior performance compared to other models, including the Densenet121-SE model.

From True labels and Expected Marks, the Disarray Chart makes a Diagram of the Disarray Lattice. The Normal labels return the confusion matrix as a Chart. The specific classification is addressed by the lines, though the anticipated order is addressed by the segments. Examples ordered incorrectly are delivered by off-corner to corner cells, while cases characterized wrongly are created by diagonal cells. The quantity of perceptions for every trademark property is shown in every cell of the softmax classifier's confusion matrix (Figure 12).

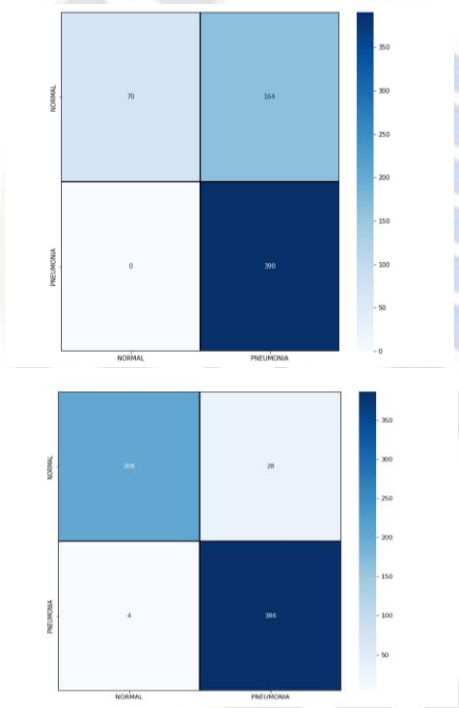
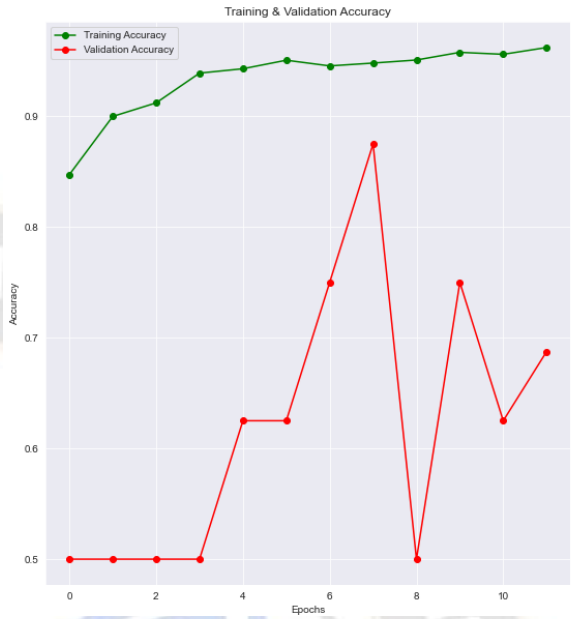
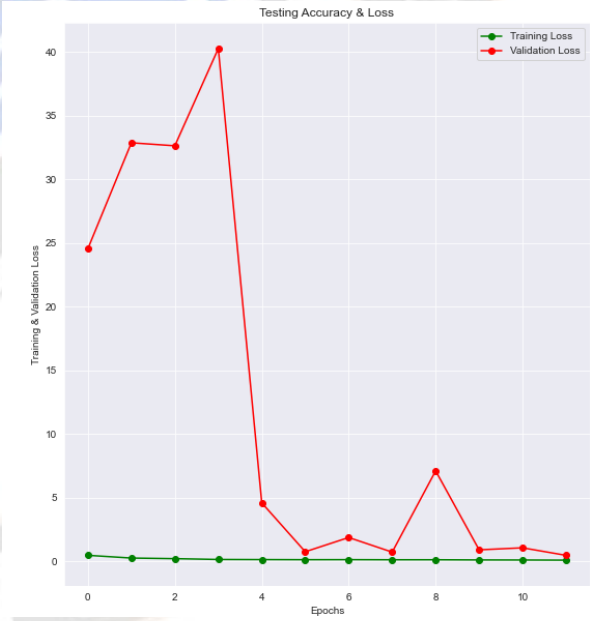


Fig. 12. Confusion matrix of a) Densenet121-SE b) Inception-V3-SE

The training and validation accuracy of the proposed Inception-V3-SE system is illustrated in Figure 13 (a & b).



(a)



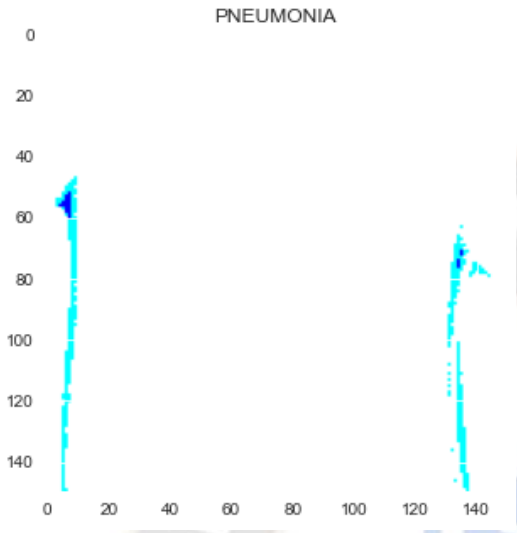
(b)

Figure 13. Training and validation (a) Accuracy, (b) Loss

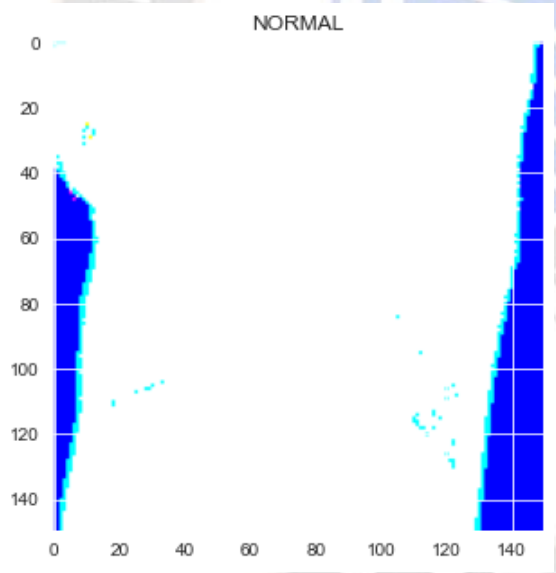
Table 3: Shows the overall output analysis of a) Densenet121-SE and b) Inception-V3-SE classifier. The performance matrices, including accuracy, precision, recall, and F1-score, are calculated from the confusion matrices obtained during the evaluation. From this analysis, it's clear that the Inception-V3-SE system demonstrates higher accuracy and

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
Article Received: *25 July 2023* ***Revised:*** *12 September 2023* ***Accepted:*** *30 September 2023*

better ability to classify pneumonia in comparison to the Densenet121-SE model.



(a)



(b)

Figure 14. Quadratic Weighted Kappa curve of Inception-V3-SE for both normal and pneumonia case

The QDK (Quadratic Weighted Kappa) is a measure that evaluates the agreement between a set of predictions and a multiclass label set. It considers both the accuracy of the matrix and the similarity between classes, going beyond the sole consideration of individual class labels. The QDK analysis of the proposed Inception-V3-SE system is illustrated in Figure 14.

Fig. 15 illustrates the ROC curve obtained via the proposed model.

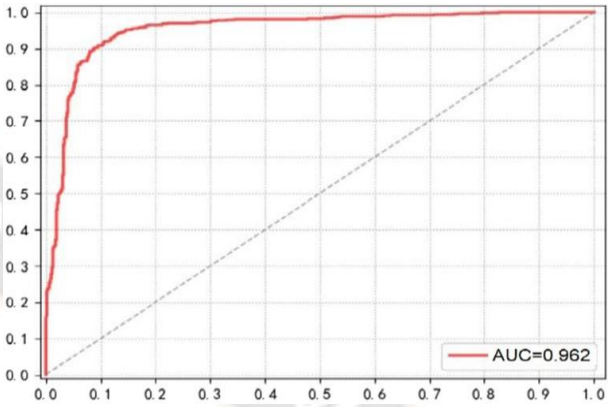


Figure.15: ROC curve of Proposed model

Additionally, Table 3 provides a comprehensive analysis of the performance of two classifiers: a) Densenet121-SE and b) Inception-V3-SE.

TABLE III DENSENET 121-SE AND b) INCEPTION V3-SE

a) Densenet121-SE

	Recall	**Precision**	**F1-score**	**Accuracy**
Normal (class 0)	1.00	0.30	0.45	.72
Pneumonia (class 1)	0.70	1.0	0.83	.71
Macro average	0.85	0.65	0.65	.73
Weighted average	0.81	0.74	0.69	.71
Proposed System	.82	.75	0.70	0.74

a) Inception-V3-SE

	Recall	**Precision**	**F1-score**	**Accuracy**
Normal (class 0)	0.88	0.98	0.93	.73
Pneumonia (class 1)	0.99	0.93	0.96	.72

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 9
***Article Received:** 25 July 2023 **Revised:** 12 September 2023 **Accepted:** 30 September 2023*

Macro average	0.94	0.96	0.94	.71
Weighted average	0.95	0.95	0.95	.70
Proposed System	.93	.92	.91	0.9539

The previous work done with the Integration of inception and SE Net gains 74% higher accuracy compared to Densenet and SE net integration. Table 4 shows a comparison of the current work (Densent-121-SE) and Inception V3 -SE. From the findings it's clear that Inception-V3-SE performs higher than the Densent-121-SE model.

TABLE CI RESEARCH WORK ANALYSIS

Research works	**Accuracy**	**Recall**	**Precision**
Densent-SENET	74	74	81
Inception-SENET	95.39	95	95

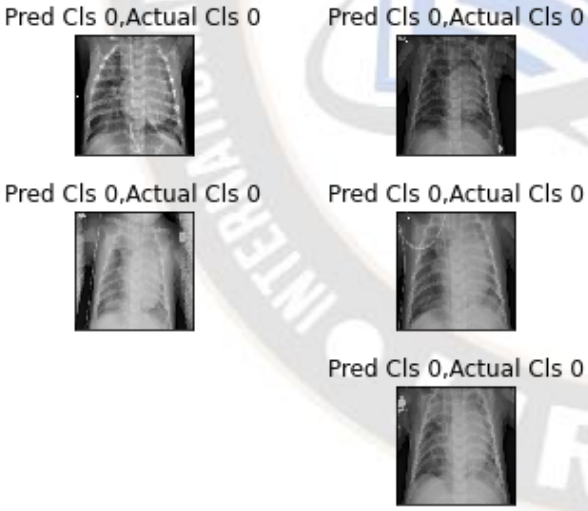


Figure 16. Actual and predicted class of proposed model

This study introduces an innovative model designed to identify pneumonia disease through analysis of a chest X-ray image dataset. In the initial phase, the model underwent training using a set of chest X-ray images. The model extracted valuable features from these images and learned to categorize them as either normal or indicative of pneumonia. After training, a set of images was presented to the model for classification, determining whether the person's chest X-ray was normal or indicated pneumonia. Comparative analysis with existing techniques revealed that our model demonstrated significantly superior performance.

V. CONCLUSION

The primary objective of this study was to develop a hybrid model that could effectively classify instances of Pneumonia. By analyzing sophisticated models such as Densenet121-SE and Inception-v3-SE, the researchers aimed to enhance the precision of pneumonia classification. The dataset used in the study comprised X-ray images that were categorized into two groups: "normal" and "pneumonia". To achieve this goal, a classifier was constructed by applying deep learning techniques to the existing Densenet-SE (Squeeze-and-Excitation) block and an Inception-v3-SE block. The experimental evaluation yielded compelling results, indicating that the proposed system utilizing the Inception-v3-SE model outperformed other state-of-the-art models across various performance measures. Notably, the proposed system attained an impressive accuracy of 0.9539, along with a Recall of 95 and Precision of 95, underscoring its superiority in pneumonia classification. This research is of great significance to the medical field, offering valuable insights into the utilization of advanced models and the efficacy of hybrid approaches. In the coming years, there is potential to delve into the utilization of transfer learning to augment the model's overall knowledge.

REFERENCES

[1] Moreno-Luna, Libertad, Rafael Robina-Ramírez, Marcelo Sánchez-Oro Sánchez, and José Castro-Serrano. "Tourism and sustainability in times of COVID-19: The case of Spain." International Journal of Environmental Research and Public Health, 18, no. 4, 2021, 1859.

[2] Martin, Julia, Noelia Tena, and Agustin G. Asuero. "Current state of diagnostic, screening and surveillance testing methods for COVID-19 from an analytical chemistry point of view." Microchemical Journal, 167, 2021, 106305.

[3] Tomar, Rohan, and Abhilasha Sharma. "Analysis & Detection of COVID-19 on Chest X-Ray Images based on Support Vector Machines." In 2022 International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), 2022, pp. 1583-1590. IEEE.

[4] Shah, Aakash, and Manan Shah. "Advancement of deep learning in pneumonia/COVID-19 classification and localization: A systematic review with qualitative and quantitative analysis." Chronic Diseases and Translational Medicine, 8, no. 03, 2022, 154-171.

[5] Nillmani, Pankaj K. Jain, Neeraj Sharma, Mannudeep K. Kalra, Klaudija Viskovic, Luca Saba, and Jasjit S. Suri. "Four types of multiclass frameworks for pneumonia classification and its

validation in X-ray scans using seven types of deep learning artificial intelligence models." Diagnostics, 12, no. 3, 2022, 652.

[6] Ravi, Vinayakumar, Harini Narasimhan, and Tuan D. Pham. "A cost-sensitive deep learning-based meta-classifier for pediatric pneumonia classification using chest X-rays." Expert Systems, 39, no. 7, 2022, e12966.

[7] Hu, Zongsheng, Zhenyu Yang, Kyle J. Lafata, Fang-Fang Yin, and Chunhao Wang. "A radiomics-boosted deep-learning model for COVID-19 and non-COVID-19 pneumonia classification using chest x-ray images." Medical physics, 49, no. 5, 2022, 3213-3222.

[8] Kumar, A. M. Prasanna, S. M. Vijaya, and G. Bharathi. "Hybrid Reverse Propagation ANN Adaptive Algorithm Based Deep Learning Image Processing for Pneumonia Detection." ACS Journal for Science and Engineering, 2, no. 2, 2022, 36-48.

[9] Khan, Wasif, Nazar Zaki, and Luqman Ali. "Intelligent pneumonia identification from chest x-rays: A systematic literature review." IEEE Access, 9, 2021, 51747-51771.

[10] Shi, Yandong, Lixiang Lian, Yuanming Shi, Zixin Wang, Yong Zhou, Liqun Fu, Lin Bai, Jun Zhang, and Wei Zhang. "Machine learning for large-scale optimization in 6g wireless networks." IEEE Communications Surveys & Tutorials, 2023.

[11] Chen, Ying, Cheng Zheng, Fei Hu, Taohui Zhou, Longfeng Feng, Guohui Xu, Zhen Yi, and Xiang Zhang. "Efficient two-step liver and tumour segmentation on abdominal CT via deep learning and a conditional random field." Computers in Biology and Medicine, 150, 2022, 106076.

[12] Zhao, Jingyuan, Xuebing Han, Minggao Ouyang, and Andrew F. Burke. "Specialized deep neural networks for battery health prognostics: Opportunities and challenges." Journal of Energy Chemistry, 2023.

[13] Karmore, Swapnili, Rushikesh Bodhe, Fadi Al-Turjman, R. Lakshmana Kumar, and Sofia K. Pillai. "IoT-based humanoid software for identification and diagnosis of COVID-19 suspects." IEEE Sensors Journal, 22, no. 18, 2020, 17490-17496.

[14] Arabian, Herag, Alberto Battistel, J. Geoffrey Chase, and Knut Moeller. "Attention-Guided Network Model for Image-Based Emotion Recognition." Applied Sciences, 13, no. 18, 2023, 10179.

[15] Albattah, Waleed, Ali Javed, Marriam Nawaz, Momina Masood, and Saleh Albahli. "Artificial intelligence-based drone system for multiclass plant disease detection using an improved efficient convolutional neural network." Frontiers in Plant Science, 13, 2022, 808380.

[16] Mujahid, Muhammad, Furqan Rustam, Roberto Álvarez, Juan Luis Vidal Mazón, Isabel de la Torre Díez, and Imran Ashraf. "Pneumonia classification from X-ray images with inception-V3 and convolutional neural network." Diagnostics, 12, no. 5, 2022, 1280.

[17] Liu, Meilin, Zidong Wang, Han Li, Peishu Wu, Fuad E. Alsaadi, and Nianyin Zeng. "AA-WGAN: Attention augmented Wasserstein generative adversarial network with application to fundus retinal vessel segmentation." Computers in Biology and Medicine, 158, 2023, 106874.

[18] Lee, Junghyup, Dohyung Kim, and Bumsub Ham. "Network quantization with element-wise gradient scaling." In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2021, pp. 6448-6457.