MAUN Fen Bil. Dergi., 13, 1, 15-25 Araştırma Makalesi/ Research Article
MAUN J. of Sci., 13, 1, 15-25 DOI : 10.18586/msufbd.1625377.

# Effect of Hybrid Activation and Loss Functions for Pneumonia Classification in Chest X-Ray Images

Yasin ÖKKAN[1], Sibel BARIN ÖZKAN[2]
[1] Department of Computer Technologies, Zonguldak Bulent Ecevit University, Zonguldak, Turkey
[2] Department of Computer Technologies, Zonguldak Bulent Ecevit University, Zonguldak, Turkey

✉: yasin.ozkan@beun.edu.tr [1]0000-0002-2029-0856 [2]0000-0001-8302-7441

Received (Geliş): 22.01.2025 Revision (Düzeltme):07.04.2025 (Kabul): 22.04.2025

## ABSTRACT

Pneumonia is a major infectious disease that causes significant deaths worldwide and early diagnosis is crucial. Chest X-ray images, widely used in diagnosis, require detailed analysis for accurate results. In this study, deep learning-based models are used to classify chest X-ray images as normal or pneumonia. A publicly available chest X-ray dataset was used and five models (MobileNetV2, ResNet50, VGG19, Xception and ViT) were compared. Among them, VGG19 achieved the highest accuracy of 88.14%. Moreover, a proposed hybrid activation function integrated into the VGG19 model improved the classification performance, reaching an accuracy of 91.67%. Moreover, performance evaluations with different loss functions showed that the proposed hybrid loss function gave the best result with an accuracy of 94.71%. Unlike previous studies, this research provides a new perspective for pneumonia detection based on deep learning models by presenting a novel combination of hybrid activation and loss functions.

**Keywords**: Chest X-Ray, Deep learning, Hybrid activation function, Loss function, Pneumonia

# Chest X-Ray Görüntülerinde Pnömoni Sınıflandırması İçin Hibrit Aktivasyon ve Kayıp Fonksiyonlarının Etkisi

## ÖZ

Pnömoni, dünya çapında önemli ölümlere neden olan önemli bir enfeksiyon hastalığıdır ve erken tanı çok önemlidir. Teşhiste yaygın olarak kullanılan göğüs röntgeni görüntüleri, doğru sonuçlar için ayrıntılı analiz gerektirir. Bu çalışmada, göğüs röntgeni görüntülerini normal veya pnömoni olarak sınıflandırmak için derin öğrenme tabanlı modeller kullanılmıştır. Halka açık bir göğüs röntgeni veri kümesi kullanılmış ve beş model (MobileNetV2, ResNet50, VGG19, Xception ve ViT) karşılaştırılmıştır. Bunlar arasında VGG19 %88,14 ile en yüksek doğruluğu elde etmiştir. Ayrıca, VGG19 modeline entegre edilen önerilen bir hibrit aktivasyon fonksiyonu, sınıflandırma performansını artırarak %91,67'lik bir doğruluğa ulaşmıştır. Ayrıca farklı kayıp fonksiyonlarıyla yapılan performans değerlendirmeleri, önerilen hibrit kayıp fonksiyonunun %94,71 doğrulukla en iyi sonucu verdiğini göstermiştir. Önceki çalışmalardan farklı olarak bu araştırma, hibrit aktivasyon ve kayıp fonksiyonlarının yeni bir kombinasyonunu sunarak derin öğrenme modellerine dayalı pnömoni tespiti için yeni bir bakış açısı sunmaktadır.

**Anahtar Kelimeler**: Göğüs röntgeni, Derin öğrenme, Hibrit aktivasyon fonksiyonu, Kayıp fonksiyonu, Zatürre

## INTRODUCTİON

Chest X-Ray (CXR) is a cornerstone diagnostic tool in medical imaging, extensively utilized for detecting a wide range of pulmonary and cardiovascular conditions. It plays a pivotal role in the identification of lung diseases, including pneumonia, which, if left undiagnosed or untreated, can result in severe complications [1]. In this regard, CXR are particularly effective for diagnosing conditions like pneumonia, as they enable the detection of infiltrates and opacities within the lungs [2]. Despite their effectiveness, the interpretation of traditional CXR images can be labor-intensive and prone to human error, even for seasoned radiologists, particularly when dealing with large volumes of data. To mitigate these challenges and enhance diagnostic efficiency, the adoption of digital imaging technologies has become increasingly prevalent, offering a means to accelerate the diagnostic process while minimizing the risk of oversight.

Digital CXR offers many advantages over traditional film-based methods. Digital systems allow high-resolution images to be obtained, while at the same time facilitating the process of storing and sharing images digitally. This is particularly advantageous for healthcare facilities in remote areas, as digital images can be quickly transferred over the internet and reviewed by remote specialists [3]. With the availability of digital images, image analysis processes have become automated, saving time and labor, especially when working with large patient data.

At this point, artificial intelligence techniques have an important place in the analysis of CXR images. In this field, especially deep learning methods and CNNs, which stand out among these methods, are used effectively. Thanks to their ability to learn and classify complex patterns in images, CNNs have achieved highly successful results in medical imaging [4]. In early diagnosis of diseases such as pneumonia, CNNs can ease the workload of radiologists and improve accuracy. However, training deep learning models requires large amounts of labeled data. In this study, the “Chest X-Ray dataset” containing a total of 5,856 labeled images classified as normal and pneumonia was used. The labeling process was carried out by expert radiologists. This is where transfer learning comes into play as an approach that allows training models with high accuracy even with limited labeled data. Transfer learning is the process of adapting a model that has been trained on larger datasets to smaller datasets, which provides a significant advantage in the field of medical imaging [5]. Although various deep learning models have been applied to pneumonia detection in previous studies, there is still a gap in the literature regarding the combined effect of specially designed activation and loss functions on classification performance.

This paper provides an in-depth study of the effectiveness of deep learning methods, specifically CNN and transfer learning approaches, for analyzing CXR images. In this context, the performance of different architectures such as MobileNetV2, ResNet50, VGG19, Xception and ViT is evaluated. The aim of the study is to optimize these methods to accurately classify diseases such as pneumonia. The choice of activation and loss functions, which play a critical role in improving model performance, significantly affects the overall performance of the model. Accordingly, the evaluation of how different combinations of activation functions and loss functions can improve the classification accuracy of CXR images is one of the main focuses of this study. In conclusion, this study aims to demonstrate how different deep learning techniques and functional structures contribute to more accurate and efficient diagnostic processes in healthcare.

The remaining sections of this paper are presented in a systematic structure. In the second section, a literature review is given, comprehensively reviewing the previous research that forms the basis of the study. In the third section, the dataset used in the study, model development, transfer learning, ViT, the proposed activation function and the proposed loss functions are explained in detail and this content is presented under the heading materials and methods. In the fourth section, the obtained findings are evaluated, their relation with the literature is discussed and the contributions of the study are discussed under the title of results and discussion. Finally, in the fifth section, general conclusions are summarized, the scientific outputs of the study are highlighted, and suggestions for future work are presented under the conclusion heading.

## LITERATURE REVIEW

Pneumonia is a disease caused by inflammation of the lungs and early diagnosis is critical. CXR are widely used to accurately diagnose pneumonia. These images play an important role in the early detection of diseases such as pneumonia by providing rapid information about the condition of the lungs. Nowadays, the use of artificial intelligence and deep learning techniques has led to a significant improvement in the analysis of CXR images. In the literature, it is emphasized that these techniques have improved the success in pneumonia classification and new methods should be explored for more effective solutions in the future.

Khan et al. [6] proposed a technique based on deep learning to distinguish COVID-19 infections from other infections. Three distinct pre-trained architectures EfficientNetB1, NasNetMobile, and MobileNetV2 were utilized for the classification of COVID-19. The training phase was conducted using an expanded dataset, and two separate learning strategies were implemented for classification. To enhance performance, not only were the models optimized, but also the hyperparameters were meticulously adjusted. Additionally, refining the classification layer further contributed to improved accuracy. The proposed approach effectively distinguished four categories—COVID-19, viral pneumonia, lung opacity, and normal—achieving an impressive accuracy rate of 96.13%.

Shelke et al. [7] proposed a classification model that analyzes CXR images for accurate diagnosis of COVID-19. The model classifies X-rays into four classes: normal, pneumonia, tuberculosis (TB), and COVID-19, and classifies COVID-19 images as mild, moderate, and severe based on severity. The VGG-16 model was used for pneumonia, TB and normal classification and 95.9% accuracy was achieved. For normal, pneumonia and COVID-19 discrimination, 98.9% accuracy was achieved using DenseNet-161, and for severity classification, ResNet-18 model performed the best and achieved 76% accuracy.

Ibrahim et al. [8] used an AlexNet-based deep learning model to classify COVID-19, bacterial pneumonia, non-COVID-19 viral pneumonia and normal CXR images. The model was evaluated under two-class, three-class, and four-class classification setups, yielding high accuracy rates. Specifically, it attained 94.43% accuracy in distinguishing non-COVID-19 viral pneumonia from normal cases, 91.43% for bacterial pneumonia versus normal, 99.16% for COVID-19 against normal, and 99.62% for COVID-19 versus bacterial pneumonia. Furthermore, the model achieved 94.00% accuracy in the three-class scenario and 93.42% in the four-class classification task.

Abbas et al. [9] proposed a deep CNN model called DeTraC to classify CXR images for the diagnosis of COVID-19. DeTraC incorporates a class decomposition mechanism that effectively addresses irregularities in image datasets by analyzing class boundaries. Experimental findings confirmed its capability to identify COVID-19 cases using an extensive dataset compiled from multiple hospitals globally. The model achieved an accuracy of 93.1% in distinguishing COVID-19 X-ray images from normal and severe acute respiratory syndrome cases, along with a precision rate of 100%.

Gielczyk et al. [10] proposed a machine learning-based method for classifying CXR images for COVID-19 diagnosis. In addition, some preprocessing methods such as thresholding, blurring and histogram equalization were also examined. As a result of the analysis, 97%, 96% and 99% F1-scores were obtained for three classes (healthy, COVID-19 and pneumonia) respectively.

Alshmrani et al. [11] designed a deep learning architecture integrating VGG19 and CNN for multi-class classification of COVID-19, pneumonia, lung cancer, tuberculosis (TB) and lung opacity. The model was trained on a comprehensive dataset containing 3,615 COVID-19, 6,012 lung opacity, 5,870 pneumonia, 20,000 lung cancer, 1,400 tuberculosis, and 10,192 normal CXR images. The results of the experiments showed exceptional performance, with an accuracy of 96.48%, 93.75% recall, 97.56% precision, 95.62% F1-score, and an AUC of 99.82%.

Asif et al. [12] proposed an Inception V3-based Deep CNN (DCNN) model for automatic detection of COVID-19 pneumonia from CXR images. The dataset used for training included 864 COVID-19, 1,345 viral pneumonia, and 1,341 normal X-ray images. Using transfer learning, the model achieved a classification accuracy of more than 98%, with 97% training accuracy and 93% validation accuracy, demonstrating its effectiveness in detecting COVID-19 pneumonia.

Zhao et al. [13] proposed the AM_DenseNet model to classify 14 different chest diseases using X-ray images. By incorporating dense connections and attention mechanisms, the model enhanced feature extraction while mitigating class imbalance through the Focal Loss function. Experimental evaluations demonstrated its effectiveness, achieving an AUC of 0.8537 for multi-class classification.

Okolo et al. [14] developed a Transformer-based deep learning model for the classification of CXR images. Initially, the ViT model was examined, and subsequently, an improved version called Input Enhanced ViT (IEViT) was proposed to boost performance. Experimental results across four datasets showed that the IEViT model surpassed ViT, yielding up to 5.82% improvement in F1-score, a 3% increase in recall, and a 6.41% rise in precision.

Kolonne et al. [15] developed a deep learning framework based on CXR analysis to distinguish between normal cases, pneumonia and COVID-19 pneumonia. By augmenting the MobileNetV2 architecture with additional layers, the model's capacity was enhanced. The system's performance, assessed through 5-fold cross-validation, demonstrated high reliability, achieving 98.65% accuracy and 98.15% recall.

Souid et al. [16] used a modified version of the MobileNetV2 model to detect lung abnormalities from CXR images. The model was trained on the NIH Chest X-Ray 14 dataset, where it delivered an average AUC of 0.811 and an accuracy exceeding 90%, confirming its effectiveness in medical image classification.

These studies highlight the advancements in deep learning-driven diagnostic tools for lung diseases, pneumonia, COVID-19, and other infections. The integration of sophisticated architectures not only supports fast and precise diagnoses but also facilitates early disease detection, ultimately accelerating treatment procedures.

## MATERIALS AND METHODS

### Dataset

CXR is an essential, widely accessible imaging technique used in diagnosing various lung conditions. It is particularly crucial for detecting respiratory infections, including pneumonia and COVID-19, enabling healthcare professionals to diagnose patients early and accurately. In this study, we utilized the publicly available CXR dataset [17], which consists of 5,840 CXR images representing both healthy individuals and those with pneumonia. This dataset was obtained from the Kaggle platform and is openly accessible for research purposes. The dataset is organized into two primary folders: train and test, each containing two subfolders labeled normal and pneumonia. All images in the dataset were pre-processed and resized to a fixed resolution of 600×600 pixels. Images are in grayscale and JPEG format. Sample images from these classes are shown in Figure 1. A detailed breakdown of the training and testing datasets can be found in Table 1.

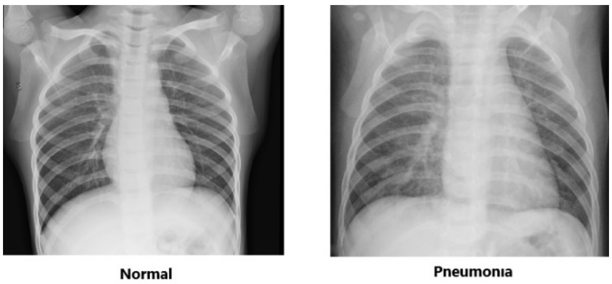


**Figure 1.** Example images of normal and pneumonia classes in the CXR dataset [17].

**Table 1.** Summary of the CXR dataset.

| | Normal | Pneumonıa | Total |
|---|---|---|---|
| Test | 234 | 390 | 624 |
| Train | 1341 | 3875 | 5216 |
| Total | 1575 | 4265 | 5840 |

Table 1 shows the distribution of test and training data in the CXR dataset. There are 5840 images in total, of which 1575 are normal and 4265 are pneumonia cases. The training data contains more samples than the test data.

**Model Development**

Deep learning has emerged as an important tool in medical image analysis. This method automatically extracts meaningful information from complex image data, enabling fast and accurate diagnosis of diseases. At the core of deep learning lies CNN, which play an important role especially in visual recognition tasks. CNNs are highly effective in automatically classifying patterns in medical images by using convolutional layers and pooling operations to extract hierarchical features from images. In many studies, CNNs have been shown to successfully detect diseases such as pneumonia, tuberculosis and COVID-19 from CXR images and provide important support to healthcare professionals in the diagnosis process [4]. This demonstrates the benefits of deep learning-based systems, especially in emergency medicine and similar fields where rapid diagnosis is critical.

However, transfer learning plays an important role to further improve the performance of deep learning models and make them more efficient. Transfer learning allows for faster and more accurate training on smaller data sets by using pre-trained models that have been trained on large and labeled data sets (e.g. ImageNet) [5]. This increases the ability of the model to generalize to smaller data sets. In recent years, the ViT architecture has also gained an important place in the deep learning community. Unlike traditional CNNs, ViT has a transformer-based structure and excels in modeling long-range dependencies, especially on large datasets [18]. In this study, ViT-based models are used for the classification of CXR images, and the performance of the model is further improved by transfer learning and the aim is to accurately detect diseases. In this context, the combination of ViT and transfer learning offers a significant advantage over previously used CNN methods.

In this study, experiments were conducted on various deep learning architectures to classify pneumonia diseases. In the first stage, a comprehensive performance evaluation was performed on the CXR dataset using deep learning architectures such as MobileNetV2, ResNet50, VGG19, Xception and ViT. As a result of the experiments, it was observed that the VGG19 model achieved the highest accuracy. This finding led to the decision to use the VGG19 model for the rest of the study. In the second stage, the performances of ReLU, Softmax, Sigmoid, LeakyReLU activation functions as well as new activation functions developed on VGG19 were analyzed. Each function is analyzed in terms of model accuracy and learning capacity and the best performing activation function is identified. In the third section, the performances of different loss functions (MSE, MAE, Binary Cross-Entropy and the proposed loss function) are compared in the output layer of VGG19. These steps aim to increase the applicability and efficiency of deep learning-based models in the classification of pneumonia diseases. Figure 2 presents a diagram showing the general workflow of the study. The parts marked in red indicate the methods with the most successful classification results. We continue with the models and combinations with the highest performance in the workflow.

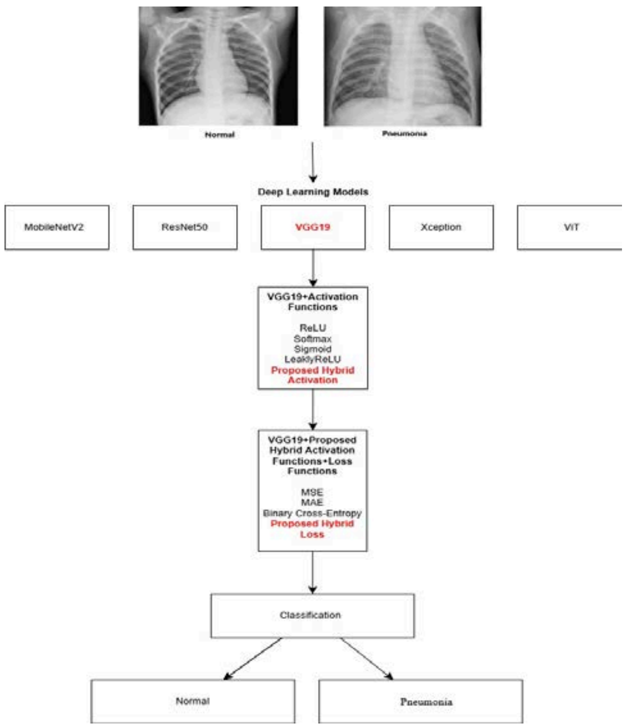


**Figure 2.** Workflow diagram of the presented architecture

During the training of the deep learning-based models used in this study, the underlying hyperparameters were carefully selected. In order to maximize the performance of the models during the training process, the same hyperparameters were used for each model. The hyperparameters used in the study are presented in Table 2.

**Table 2.** Hyperparameters used in training models.

| Hyperparameters | Values |
|---|---|
| Epochs | 200 |
| Learning Rate | 0.001 |
| Batch Size | 32 |

| | |
|---|---|
| Optimization Method | Adam |
| Early Stopping | Yes, early stop on loss of validation |
| Dropout | 0.3 |

**Transfer Learning Models**

Transfer learning enables a model to use previously learned knowledge in a new task, helping to achieve strong performances with limited data. In this study, we investigate the performance of four different models based on transfer learning: MobileNetV2, ResNet50, VGG19 and Xception.

MobileNetV2 is a lightweight model designed to work efficiently on mobile devices with low resource requirements. By using depthwise separable convolutions, it keeps performance high while reducing the number of parameters [19]. ResNet50 enables training deep networks with the residual learning principle, making it ideal for deeper and more complex networks [20]. VGG19 performs powerful feature extraction with successive convolution layers and is widely used for transfer learning [21]. Xception provides more efficient feature extraction using decomposed convolutions and provides high accuracy on large datasets [22].

**Vision Transformer (ViT)**

ViT is a model based on the transformers architecture, different from the traditional CNN for image processing tasks. First introduced by Dosovitskiy et al., this method processes images into fixed-size patches and converts each patch into a vector. These vectors are then passed to a transformer model, which learns local and global features from the image [18]. The success of ViT is especially evident when optimized with large data sets and powerful computational resources. Unlike traditional CNNs, the self-attention mechanism underlying ViT can model the relationships between different parts in images much more efficiently [23].

Recently, ViT has demonstrated high performance in image classification tasks, outperforming CNNs. This is mainly because transformer-based structures learn more comprehensive and flexible features based on global information. The use of ViT in areas such as medical imaging has made a huge impact, especially in applications that are considered to have limited data. For example, the ViT model has been shown to provide successful results in the diagnosis of diseases such as COVID-19 [24]. However, the challenges of ViT, such as large data requirements and training times, may require the effective use of transfer learning methods on smaller data sets. In this context, the successful applications of ViT provide a significant advancement in the field of medical imaging.

**Proposed Activation Function**

The activation functions used in neural networks significantly affect the learning capacity and overall performance of the network. Activation functions increase the network's ability to model nonlinear relationships, allowing for better results on complex data [25]. ReLU (Rectified Linear Unit) is a function commonly used in deep networks that allows the network to learn fast by converting sub-zero inputs to zero [26]. Sigmoid and tanh (hyperbolic tangent) functions keep the output of the model within a certain range by limiting its output; however, these functions can cause problems such as gradient disappearance [27]. The Softplus function, on the other hand, provides a linear growth similar to ReLU, but alleviates the problem of gradient fading by offering a smoother transition [28]. Each of these activation functions offers various advantages and limitations according to the different learning processes and needs of the network.

The proposed hybrid activation function is a combination that exhibits different behavior in positive and negative regions. For positive values, the Softplus function is preferred because Softplus provides a smooth transition for zero and small positive inputs and increases slowly without increasing the gradients. This feature helps to make learning stable and robust when the model has negative gradients. The mathematical representation of the Softplus activation function is presented in Equation 1.

$$Softplus(x) = \ln\,(1 + e^{x}) \tag{1}$$

For negative values, the Tanh function is used, since Tanh provides a symmetric output between -1 and 1, giving a wider range of activation for negative inputs. In this way, the model follows a stronger and more balanced learning process in negative regions. The mathematical representation of the Tanh activation function is presented in Equation 2.

$$Tanh(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{2}$$

This combination of Softplus and Tanh offers advantages in both positive and negative regions, allowing the model to learn in a more flexible and balanced way. Promoting stable learning for both small and large values, this hybrid function offers an ideal solution for adapting to larger data ranges and complex problems. The mathematical representation of the hybrid function combining Softplus and Tanh activations is presented in Equation 3.

$$f(x) = \begin{cases} \ln(1 + e^{x}) & if \quad x \geq 0 \\ \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} & if \quad x < 0 \end{cases} \tag{3}$$

**Proposed Loss Function**

In image classification tasks, one of the most important elements that guide the learning process of the model is the loss functions. Loss functions calculate the model's errors by measuring the difference between the model's predicted values and the actual labels and ensure that these errors are minimized [29]. A loss function plays a critical role in the optimization process so that the model can classify correctly. Traditionally, among the loss functions used in classification problems, Cross-Entropy loss (CE) is widely preferred to improve the model's ability to predict the correct class [30], while methods such as contrastive loss allow the model to become more generalizable by learning the similarities and differences between images [31]. The mathematical representations of the cross-entropy loss and contrastive loss functions are presented in Equations (4) and (5) respectively.

$$L_{CE}(y, \ddot{y}) = -\sum_{i=1}^{C} y_i \log(\ddot{y}_i) \quad (4)$$

$$L_{contrastive} = \frac{1}{2}[y.D^2(x_1, x_2) + (1 - y).\max(0, m - D(x_1, x_2))^2] \quad (5)$$

In this context, Hybrid Contrastive-CrossEntropy Loss (HCCEL) is proposed as a new loss function. HCCEL aims to optimize both class accuracy and visual similarities by combining cross-entropy loss and contrastive loss. HCCEL includes two components: First, CE improves the model's ability to predict the correct class [30]. This component aims for the model to predict the correct probability for each class. Secondly, contrastive loss (CL) helps the model learn visual similarities between images by making the model zoom in on similar images and zoom out on images belonging to different classes [31]. By combining these two components, HCCEL optimizes class accuracy and similarity-driven learning at the same time. The mathematical representation of the new hybrid activation function formed by the combination of the two loss functions is presented in Equation (6).

$$L_{HCCEL} = L_{CE} + \lambda.L_{contrastive} \quad (6)$$

The parameter λ is a hyperparameter that determines the weight of the Contrastive Loss component. In the HCCEL function, this parameter balances between class accuracy and visual similarity-based learning. The value of λ should be set so that the model both makes accurate class predictions and learns visual similarities. A high λ value increases the effect of contrastive loss, allowing the model to learn more similarities, which often helps to learn better visual representations. However, too high a λ can lead the model to neglect class accuracy. A low λ can only optimize class accuracy, with less emphasis on learning visual similarities. Therefore, the correct setting of λ is critical for the model to successfully learn both objectives.

The λ value was determined by Grid Search, a systematic hyperparameter search method. Grid Search involves trying different values of λ with fixed steps in a given range (e.g. between 0.1 and 1.0) and evaluating the model performance for each value. As a result of the comparisons, the highest accuracy was obtained at a value of 0.5.

## RESULT AND DISCUSSION

CXR images are an important imaging modality often used for medical diagnosis and early detection of diseases. These images are used to visually assess the condition of the lungs, heart and other chest organs. CXR play a critical role in the diagnosis of conditions such as lung infections, cancer, heart disease and pneumonia. However, these images often contain complex structures and are difficult to interpret accurately as many diseases with similar symptoms can be visually very close to each other. Therefore, the use of artificial intelligence and deep learning methods offers great potential for detecting subtle differences in CXR images and making accurate diagnoses.

In this study, different deep learning architectures were used to classify normal and pneumonia labeled data in CXR images. These architectures include MobileNetV2, ResNet50, VGG19, Xception and ViT. In the first part of the study, the classification performance of these five different models on the dataset is comprehensively evaluated and presented in Table 3.

The model's performance is evaluated using critical metrics such as True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). TP refers to the number of correctly identified positive instances, TN represents the correctly classified negative cases, FP denotes instances that were incorrectly predicted as positive, and FN indicates positive samples that were incorrectly classified as negative.

**Table 3.** Performance metrics on the CXR dataset.

| Models | True Positive | False Positive | False Negative | True Negative | Accuracy (%) | Recall (%) | Precision(%) | F1-score(%) |
|---|---|---|---|---|---|---|---|---|
| MobileNetV2 | 161 | 73 | 10 | 380 | 86.70 | 94.15 | 68.80 | 79.51 |
| ResNet50 | 113 | 121 | 26 | 364 | 76.44 | 81.29 | 48.29 | 60.59 |
| VGG19 | 183 | 51 | 23 | 367 | 88.14 | 88.83 | 78.21 | 83.18 |
| Xception | 75 | 159 | 146 | 244 | 51.12 | 33.94 | 32.05 | 32.97 |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Vit | 123 | 111 | 0 | 390 | 82.21 | 100.00 | 52.56 | 58.91 |

These values are essential for calculating key performance metrics, including accuracy, precision, recall, and the F1-score, which provide a more nuanced evaluation of the model's performance. Accuracy represents the proportion of correctly classified instances out of all predictions. Precision measures the proportion of predicted positive cases that are truly positive, while recall assesses the model's ability to identify actual positive instances. The F1-score combines both precision and recall to give a balanced view of the model's overall predictive performance. These metrics are especially valuable when working with imbalanced datasets, ensuring a more reliable evaluation of the model's effectiveness in classification tasks [32]. The mathematical formulas for these metrics are outlined in Equations (7), (8), (9), and (10).

$$Accuracy = \frac{(\mathrm{TP+TN})}{(\mathrm{TP+TN+FP+FN})} \tag{7}$$

$$\mathrm{Precision} = \frac{\mathrm{TP}}{(\mathrm{TP + FP})} \tag{8}$$

$$\mathrm{Recall} = \frac{\mathrm{TP}}{(\mathrm{TP + FN})} \tag{9}$$

$$\mathrm{F1 - Score} = 2 \times \frac{\mathrm{Precision \times Recall}}{(\mathrm{Precision + Recall})} \tag{10}$$

In the evaluations performed on the models in Table 4, Sigmoid was used as the activation function and binary_crossentropy as the loss function. With this combination, the classification performances of the model were evaluated and performance metrics were calculated according to the results obtained. According to the accuracy values, VGG19 achieved the highest accuracy rate with 88.14%. This model was very successful in pneumonia detection. MobileNetV2 ranked second with 86.70% accuracy, while ViT performed well with 82.21% accuracy. ResNet50 had an average performance with 76.44% accuracy, while Xception was the least successful with 51.12% accuracy. These results show that VGG19 and MobileNetV2 have the best classification performance, while Xception performs poorly with very low accuracy. The high accuracy and balanced performance of VGG19 demonstrates its effective classification ability on CXR data. Therefore, after comparisons with other models, the VGG19 architecture was chosen for more in-depth analysis and improvements.

In the second part of the study, the proposed hybrid activation function was integrated into the VGG19 model in addition to the ReLU, Softmax, Sigmoid, LeakyReLU activation functions. With this integration, the classification performance of the model is tested with different activation functions. The results obtained are presented in Table 4.

**Table 4.** Performance metrics of various activation functions on the VGG19 model.

| Activation Functions | True Positive | False Positive | False Negative | True Negative | Accuracy (%) | Recall (%) | Precision(%) | F1-score(%) |
|---|---|---|---|---|---|---|---|---|
| ReLU | 161 | 73 | 52 | 338 | 79.97 | 75.59 | 68.80 | 72.04 |
| Softmax | 177 | 57 | 32 | 358 | 85.74 | 84.69 | 75.64 | 79.91 |
| Sigmoid | 183 | 51 | 23 | 367 | 88.14 | 88.83 | 78.21 | 83.18 |
| LeakyReLU | 152 | 82 | 68 | 322 | 75.96 | 69.09 | 64.69 | 66.96 |
| Proposed Hybrid | 197 | 37 | 15 | 375 | 91.67 | 92.92 | 84.19 | 88.34 |

Table 4 clearly shows the differences between the classification performance of the different activation functions used on the VGG19 model. The sigmoid activation function performs the best with strong metrics such as 88.14% accuracy, 88.83% recall and 78.21% precision, indicating that the model shows high success in true positive predictions. Softmax activation ranked second with an accuracy of 85.74% and an F1 score of 79.91%, again providing a successful classification performance. The other activation functions ReLU and LeakyReLU achieved lower results with 79.97% and 75.96% accuracy respectively. These results indicate that both functions lead to more errors in classification performance and false positive predictions. The proposed hybrid activation function provides the highest performance with 91.67% accuracy, allowing the model to outperform in all metrics. These findings reveal that the hybrid activation function offers a significant advantage, especially in terms of true positive classifications and improving the overall performance of the model. The study will continue with the hybrid activation function integrated into the VGG19 model, which yields the most successful results.

In the third part of the study, in addition to the hybrid activation function integrated into the VGG19 model, performance evaluations were performed using MSE, MAE, Binary Cross-Entropy and the proposed loss function. The performance metrics obtained with each loss function are presented in Table 5.

**Table 5.** Performance metrics of various loss functions on the model obtained with VGG19+proposed hybrid loss function.

| Loss Functions | True Positive | False Positive | False Negative | True Negative | Accuracy (%) | Recall(%) | Precision(%) | F1-score(%) |
|---|---|---|---|---|---|---|---|---|
| MSE | 189 | 45 | 28 | 362 | 88.30 | 87.10 | 80.77 | 83.81 |
| MAE | 172 | 62 | 38 | 352 | 83.97 | 81.90 | 73.50 | 77.48 |
| Binary Cross-Entropy | 197 | 37 | 15 | 375 | 91.67 | 92.92 | 84.19 | 88.34 |
| Proposed Hybrid loss | 210 | 24 | 9 | 381 | 94.71 | 95.89 | 89.74 | 92.72 |

When Table 5 is analyzed, the Proposed Hybrid loss function provides the most successful results with 94.71% accuracy. The Binary Cross-Entropy loss function ranks second with 91.67% accuracy and shows a very strong performance. The MSE loss function has a lower accuracy with 88.30% accuracy, while the MAE loss function shows the lowest accuracy with 83.97% accuracy. These results show that the Proposed Hybrid loss function is more effective than the other loss functions and provides a better overall performance in the classification task.
In conclusion, this study highlights the effectiveness of the developed hybrid activation function and the Proposed Hybrid loss function. Both functions improve the overall classification performance of the model, leading to significant improvements in key performance metrics such as accuracy. These findings suggest that hybrid structures offer an effective approach for performance improvements in deep learning models.

The contributions of the study to the literature enrich the field by providing a new perspective to existing research. In Table 6, the findings of previous studies in this field are examined comparatively with a systematic approach and presented in detail.

**Table 6.** Summary of studies with CXR in the literature.

| Study | Year | Dataset | Objective | Method | Accuracy (%) |
|---|---|---|---|---|---|
| Khan et al. [6] | 2022 | COVID-19 Radiography Database | Classifying viral pneumonia, lung opacity and normal images. | Three pre-trained models (EfficientNetB1, NasNetMobile, MobileNetV2) were used with fine-tuning on an augmented dataset. | EfficientNetB1:96.13 NasNetMobile:94.81 MobileNetV2:93.96 |
| Shelke et al. [7] | 2021 | Dataset of 2271 images from Clinico Diagnostic Lab in India. | Classifying normal, pneumonia, tuberculosis and COVID-19 images. | DenseNet-161 | 98.9 |
| Ibrahim et al. [33] | 2024 | Novel Corona Virus 2019 Dataset | To classify CXR images into four categories COVID-19, non-COVID-19 viral pneumonia, bacterial pneumonia, and normal | AlexNet | 93.42 |
| Saraiva et al. [34] | 2019 | CXR Dataset | Classify pneumonia and normal images | CNN model is proposed. | 94.40 |
| Gulgun and Erol [35] | 2020 | CXR Dataset | Classify pneumonia and normal I mages | The performance of three different models is compared: CNN model, | CNN Model: 80.4 CNN Model with Data |

| | | | | | |
|---|---|---|---|---|---|
| | | | | CNN Model Applying Data Augmentation Technique, VGG16 | Augmentation Technique: 93.4 VGG16: 85.6 |
| Proposed Model | 2025 | CXR Dataset | Classify pneumonia and normal images | The proposed VGG19 transfer learning architecture integrating activation and loss function is implemented on the original dataset. | 94.71 |

Table 6 provides a summary of various studies that highlight the effectiveness of deep learning models in diagnosing COVID-19 and other pulmonary diseases. These studies illustrate significant advancements in CXR image classification, leveraging diverse datasets and model architectures. The findings emphasize the continuous progress in developing robust and accurate deep learning approaches for medical image analysis.

Khan et al. [6] examined EfficientNetB1, NasNetMobile and MobileNetV2 models on COVID-19 Radiography Database and reported that EfficientNetB1 model provided 96.13% accuracy rate. Shelke et al. [7], the highest success was achieved with 98.9% accuracy using the DenseNet-161 model on a dataset containing 2271 images. Ibrahim et al. [33], 93.42% accuracy was achieved using the AlexNet model.

Saraiva et al. [34], Gulgun and Erol [35] and the proposed work were performed on the same CXR dataset. Saraiva et al. [34], 94.40% accuracy was achieved with the proposed CNN model. Gulgun and Erol [35], the performances of the standard CNN model, the CNN model with data augmentation, and the VGG16 model were compared, and 80.4%, 93.4%, and 85.6% accuracy rates were obtained, respectively. In the proposed study, the VGG19 transfer learning architecture was applied with a special activation and loss function and achieved an accuracy of 94.71%. In particular, the proposed model outperformed the other methods and demonstrated that transfer learning as well as customized functions can improve the classification performance. This achievement is due to the significant impact of the combination of loss function and activation function on improving the model performance. In the proposed work, the customized activation and loss functions used in combination with the VGG19 transfer learning architecture improved the accuracy of the model to 94.71%. This combination provides an effective strategy to overcome the local minimum problems often encountered in deep learning models and enables the model to gain stronger generalization capability.

The customized loss function minimized the classification errors, while the activation functions helped the model to make more precise and accurate decisions during the learning process. This interaction enhanced the model's capacity to provide better discrimination and accuracy and played a critical role in improving the results. In conclusion, these findings show that the combination of loss function and activation function is an important tool for optimizing the performance of deep learning models and clearly demonstrate why the proposed model is more successful than other methods.

## CONCLUSION

Artificial intelligence and deep learning techniques play a crucial role in medical imaging, particularly in the early diagnosis and classification of lung diseases. This study evaluates the effectiveness of various deep learning architectures in distinguishing between normal and pneumonia-labeled CXR images. Performance analyses were conducted on multiple transfer learning models, including MobileNetV2, ResNet50, VGG19, Xception, and ViT, to compare their classification capabilities. The findings indicate that VGG19 achieved the highest accuracy (88.14%), making it the most suitable model for further exploration. In the next phase, different activation functions ReLU, Softmax, Sigmoid, and LeakyReLU along with a proposed hybrid activation function were integrated into the VGG19 model. The hybrid activation function outperformed others, enhancing the model’s accuracy to 91.67%, demonstrating its effectiveness in improving classification performance. Further, various loss functions, including MSE, MAE, Binary Cross-Entropy, and the proposed hybrid loss function, were incorporated into the VGG19 model alongside the hybrid activation function. The proposed hybrid loss function yielded the best results, achieving 94.71% accuracy. These findings suggest that hybrid loss functions are more effective in classification tasks, significantly enhancing the model’s overall performance and true positive detection rates.

This study shows that deep learning architectures and hybrid structures offer an effective approach to improve classification accuracy in medical imaging applications. The developed hybrid activation function and loss function combinations can be a powerful tool for accurate diagnosis and early detection by improving the classification accuracy of the model. Future work should test these hybrid structures on larger datasets and evaluate them more extensively in real-world applications. Furthermore, combinations of different deep learning architectures and loss functions can be explored to further improve the accuracy and generalization ability of the model. However, the study also has certain limitations. The models and combinations used in the study were only evaluated on a single dataset. In addition, the models were only tested on static images, and the performance in real-time clinical settings has not yet been validated. These limitations should be addressed in future studies to better validate the robustness and clinical applicability of the proposed methods.

## REFERENCES


[1] Ruuskanen, O., Lahti, E., Jennings, L. C., & Murdoch, D. R. (2011). Viral pneumonia. The Lancet, 377(9773), 1264-1275.

[2] Hammoudi, K., Benhabiles, H., Melkemi, M., Dornaika, F., Arganda-Carreras, I., Collard, D., & Scherpereel, A. (2021). Deep learning on chest X-ray images to detect and evaluate pneumonia cases at the era of COVID-19. Journal of medical systems, 45(7), 75.

[3] MacMahon, H. (2003). Digital chest radiography: practical issues. Journal of thoracic imaging, 18(3), 138-147.

[4] Rajpurkar, P. (2017). CheXNet: Radiologist-Level Pneumonia Detection on Chest X-Rays with Deep Learning. ArXiv abs/1711, 5225.

[5] Pan, S. J., & Yang, Q. (2009). A survey on transfer learning. IEEE Transactions on knowledge and data engineering, 22(10), 1345-1359.

[6] Khan, E., Rehman, M. Z. U., Ahmed, F., Alfouzan, F. A., Alzahrani, N. M., & Ahmad, J. (2022). Chest X-ray classification for the detection of COVID-19 using deep learning techniques. Sensors, 22(3), 1211.

[7] Shelke, A., Inamdar, M., Shah, V., Tiwari, A., Hussain, A., Chafekar, T., & Mehendale, N. (2021). Chest X-ray classification using deep learning for automated COVID-19 screening. SN computer science, 2(4), 300.

[8] Ibrahim, A. U., Ozsoz, M., Serte, S., Al-Turjman, F., & Yakoi, P. S. (2024). Pneumonia classification using deep learning from chest X-ray images during COVID-19. Cognitive computation, 16(4), 1589-1601.

[9] Abbas, A., Abdelsamea, M. M., & Gaber, M. M. (2021). Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. Applied Intelligence, 51, 854-864.

[10] Giełczyk, A., Marciniak, A., Tarczewska, M., & Lutowski, Z. (2022). Pre-processing methods in chest X-ray image classification. Plos one, 17(4), e0265949.

[11] Alshmrani, G. M. M., Ni, Q., Jiang, R., Pervaiz, H., & Elshennawy, N. M. (2023). A deep learning architecture for multi-class lung diseases classification using chest X-ray (CXR) images. Alexandria Engineering Journal, 64, 923-935.

[12] Asif, S., Wenhui, Y., Jin, H., & Jinhai, S. (2020, December). Classification of COVID-19 from chest X-ray images using deep convolutional neural network. In 2020 IEEE 6th international conference on computer and communications (ICCC) (pp. 426-433). IEEE.

[13] Zhao, J., Li, M., Shi, W., Miao, Y., Jiang, Z., & Ji, B. (2021, April). A deep learning method for classification of chest X-ray images. In Journal of Physics: Conference Series (Vol. 1848, No. 1, p. 012030). IOP Publishing.

[14] Okolo, G. I., Katsigiannis, S., & Ramzan, N. (2022). IEViT: An enhanced vision transformer architecture for chest X-ray image classification. Computer Methods and Programs in Biomedicine, 226, 107141.

[15] Kolonne, S., Fernando, C., Kumarasinghe, H., & Meedeniya, D. (2021, December). MobileNetV2 based chest X-rays classification. In 2021 International Conference on Decision Aid Sciences and Application (DASA) (pp. 57-61). IEEE.

[16] Souid, A., Sakli, N., & Sakli, H. (2021). Classification and predictions of lung diseases from chest x-rays using mobilenet v2. Applied Sciences, 11(6), 2751.

[17] Kaggle, Available at: https://www.kaggle.com/datasets/alifrahman/chestxraydataset/data, Accessed: 02 December 2024.

[18] Dosovitskiy, A., Springenberg, J. T., Riedmiller, M., & Brox, T. (2014). Discriminative unsupervised feature learning with convolutional neural networks. Advances in neural information processing systems, 27.

[19] Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4510-4520).

[20] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

[21] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[22] Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1251-1258).

[23] Vaswani, A. (2017). Attention is all you need. Advances in Neural Information Processing Systems.

[24] Touvron, H., Cord, M., Douze, M., Massa, F., Sablayrolles, A., & Jégou, H. (2021, July). Training data-efficient image transformers & distillation through attention. In International conference on machine learning (pp. 10347-10357). PMLR.

[25] Akalın, F. (2024). Survival Classification in Heart Failure Patients by Neural Network-Based Crocodile and Egyptian Plover (CEP) Optimization Algorithm. Arabian Journal for Science and Engineering, 49(3), 3897-3914.

[26] Nair, V., & Hinton, G. E. (2010). Rectified linear units improve restricted boltzmann machines. In Proceedings of the 27th international conference on machine learning (ICML-10) (pp. 807-814).

[27] Hochreiter, S. (1998). The vanishing gradient problem during learning recurrent neural nets and problem solutions. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems, 6(02), 107-116.

[28] Sun, K., Yu, J., Zhang, L., & Dong, Z. (2020). A convolutional neural network model based on improved softplus activation function. In International Conference on Applications and Techniques in Cyber Intelligence ATCI 2019: Applications and Techniques in Cyber Intelligence 7 (pp. 1326-1335). Springer International Publishing.

[29] Goodfellow, I., Bengio, Y., & Courville, A. (2016). Deep Learning. MIT Press.

[30] Mao, A., Mohri, M., & Zhong, Y. (2023, July). Cross-entropy loss functions: Theoretical analysis and applications. In International conference on Machine learning (pp. 23803-23828). PMLR.

[31] Hadsell, R., Chopra, S., & LeCun, Y. (2006, June). Dimensionality reduction by learning an invariant mapping. In 2006 IEEE computer society conference on computer vision and pattern recognition (CVPR'06) (Vol. 2, pp. 1735-1742). IEEE.

[32] Akalın, F. (2024). Değiştirilmiş Yapay Arı Kolonisi Optimizasyon Algoritması ve İstatistiksel Modelleme ile Sentetik Veri Üretimi. Journal of the Institute of Science & Technology/Iğdır Üniversitesi Fen Bilimleri Enstitüsü Dergisi, 14(4).

[33] Ibrahim, A. U., Ozsoz, M., Serte, S., Al-Turjman, F., & Yakoi, P. S. (2024). Pneumonia classification using deep

learning from chest X-ray images during COVID-19. *Cognitive computation*, *16*(4), 1589-1601.

[34] Saraiva, A. A., Santos, D. B. S., Costa, N. J. C., Sousa, J. V. M., Ferreira, N. M. F., Valente, A., & Soares, S. (2019, February). Models of learning to classify X-ray images for the detection of pneumonia using neural networks. In *Bioimaging* (pp. 76-83).

[35] Gülgün, O. D., & Erol, H. (2020). Classification performance comparisons of deep learning models in pneumonia diagnosis using chest x-ray images. *Turkish Journal of Engineering*, *4*(3), 129-141.