



Diagnosis of COVID-19 in X-ray Images using Deep Neural Networks

Mohammed Akram Younus Alsaati ^{a,*}

^a College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq

*Corresponding Author Email: mohammed.akram@uomosul.edu.iq

DOI: <https://doi.org/10.54392/irjmt24318>

Received: 06-02-2024; Revised: 24-04-2024; Accepted: 05-05-2024; Published: 20-05-2024



Abstract: The global COVID-19 pandemic has presented unprecedented challenges, notably the limited availability of test kits, hindering timely and accurate disease diagnosis. Rapid identification of pneumonia, a common COVID-19 consequence, is crucial for effective management. This study focuses on COVID-19 classification from Chest X-ray images, employing an innovative approach: adapting the Xception model into a U-Net architecture via the Segmentation_Models package. Leveraging deep learning and image segmentation, the U-Net architecture, a CNN variant, proves ideal for this task, particularly after tailoring its output layer for classification. By utilizing the Xception model, we aim to enhance COVID-19 classification accuracy and efficiency. The results demonstrate promising autonomous identification of COVID-19 cases, offering valuable support to healthcare professionals. The fusion of medical imaging data with advanced neural network architectures highlights avenues for improving diagnostic accuracy during the pandemic. Notably, precision, recall, and F1 scores for each class are reported: Normal (Precision = 0.98, Recall = 0.9608, F1 Score = 0.9704), Pneumonia (Precision = 0.9579, Recall = 0.9579, F1 Score = 0.9579), and COVID-19 (Precision = 0.96, Recall = 0.9796, F1 Score = 0.9698). These findings underscore the effectiveness of our approach in accurately classifying COVID-19 cases from chest X-ray images, offering promising avenues for enhancing diagnostic capabilities during the pandemic.

Keywords: Chest X-rays, COVID-19, Xception model, Segmentation Models, U-Net

1. Introduction

The COVID-19 pandemic, originating in late 2019, has presented unparalleled global challenges to healthcare systems and medical professionals [1]. Amid the multifaceted approaches employed for the detection and assessment of this contagious disease, medical imaging, particularly chest X-ray imagery, has gained substantial prominence [2]. This research study focuses on the important undertaking of classifying COVID-19 Chest X-rays. The study utilizes advanced deep-learning techniques to accomplish this assignment.

A central innovation in our study lies in the adaptation of the U-Net architecture [3]. Originally conceived for semantic segmentation tasks, U-Net is reimaged and tailored for the classification of chest X-ray images. This adaptation involves a fundamental transformation of the U-Net's output layer to yield class predictions. By enabling the U-Net to transition from its conventional segmentation role to one of classification [4], The system possesses the capability to effectively differentiate COVID-19 instances from other medical disorders, hence enhancing its applicability in the field of medical imaging analysis.

To facilitate this remarkable adaptation, the Segmentation_Models library [5, 6], which offers a

powerful toolkit for model architecture modification, is strategically employed. Through this approach, not only are the inherent strengths of U-Net in feature extraction and localization capitalized upon but its capabilities are also extended to address the urgent requirement for COVID-19 classification.

The proposed research encompasses three primary categories of chest X-ray images: normal cases, pneumonia, and COVID-19. The accurate categorization of these distinct types of X-rays carries profound implications for clinical practice. It enables healthcare professionals to swiftly identify COVID-19 cases, and make well-informed decisions regarding patient care, including isolation, treatment, and monitoring.

By combining state-of-the-art neural network architectures, such as the adapted U-Net [7], with the wealth of information embedded in chest X-ray images, the proposed study aspires to contribute to the development of a reliable and efficient tool for assisting healthcare practitioners in the ongoing battle against COVID-19. This research underscores the pivotal role of medical imaging in pandemic response, emphasizing the potential of deep learning to augment the precision and expediency of diagnosis within the complex landscape of this global health crisis. As the intricacies of the methodology and findings are further explored, a

journey is undertaken that holds great promise for the enhancement of understanding and management of COVID-19 through the lens of advanced artificial intelligence.

Assessing the COVID-19 disease based on images of the chest X-ray is a formidable challenge [8], even for seasoned radiologists. The complexities and nuances of this task make it ripe for assistance through computational means, offering valuable support to medical professionals. Several notable studies have delved into the realm of COVID-19 virus assessment, shedding light on the potential of computer-aided diagnosis.

Moreover, recent advancements in medical imaging research have demonstrated the potential of integrating multi-modal techniques for enhanced diagnostic accuracy. For instance, studies such as the "Multi-modal medical image fusion framework using co-occurrence filter and local extrema in NSST domain" and "An end-to-end content-aware generative adversarial network-based method for multimodal medical image fusion" have shown promising results in fusing information from diverse imaging modalities to improve diagnostic outcomes [9, 10]. Additionally, the "TSJNet: A Multi-modality Target and Semantic Awareness Joint-driven Image Fusion Network" presents an innovative approach that integrates target and semantic awareness into the image fusion process, further enhancing the interpretability of fused images [11]. These advancements underscore the potential of leveraging multi-modal medical imaging techniques to augment the diagnostic capabilities of deep-learning models, aligning closely with the objectives of our study.

2. Related work

Over the past year, the scientific community has witnessed a surge in publications exploring the intersection of COVID-19 and deep learning. While this body of work has been prolific, it predominantly centers on disease detection rather than the nuanced assessment of disease [12]. Many researchers have delved into the concept of lung or lobe segmentation as a diagnostic tool. For instance, Shan et al. [13], developed a deep learning-based method for automatically segmenting and quantifying COVID-19 infection regions in chest CT scans. Their system achieved a high Dice similarity coefficient of 91.6% and demonstrated potential for severity prediction with an accuracy of 73.4%. This research addresses the need for automated tools in COVID-19 diagnosis using chest CT images. Amyar et al. [14] present CovXNet, a deep-learning network. It is trained initially on normal and pneumonia cases and then fine-tuned with COVID-19 data to address data scarcity. The achieved detection accuracies are 97.4% for 96.9% for COVID/Viral, 94.7% for COVID/Bacterial, COVID/Normal, and 90.2% for multiclass COVID/normal/Viral/Bacterial pneumonia,

demonstrating its potential in aiding COVID-19 diagnosis during the pandemic.

Despite advances in deep learning, its sensitivity for COVID-19 detection is suboptimal, with RT-PCR tests remaining the gold standard [15]. In contrast to the typically subjective and qualitative ratings offered by radiologists, deep learning-based illness assessment provides greater objectivity and precision. Researchers have explored the possibility of applying deep learning techniques to determine the COVID-19 disease using chest X-rays.

He et al. [16] developed a synergistic learning strategy. When applied to a Computed Tomography (CT) scans dataset consisting of 666 human chests, their approach obtained a remarkable accuracy of 98.5%. Mijwil and Maad [17] previously used two deep convolutional neural network (DCNN) classifiers, Inception-v2 and VGG-16, to detect COVID-19 using a Kaggle dataset (COVID-19 Radiography Database). They examined chest X-rays of COVID-19-positive and healthy people. With an accuracy of 97%, the Inception-v2 classifier outperformed the VGG-16 classifier, which had 93%.

Khanday et al. [18], in this work, data labeling was done manually using propaganda identification methods, and relevant features were selected using hybrid feature engineering. Adaboost achieved 98.7% accuracy in binary classification using ensemble machine learning classifiers. The main learning approach was Adaboost, which adjusts weights to improve weak learning algorithms. Tang et al. [19] showed that artificial intelligence methods that used quantitative characteristics extracted from CT lung scans were able to tell the difference between COVID-19 patients with severe and non-severe cases with an average precision of 87.5% based on a collection of 176 CT scans.

Meanwhile, Xiao et al. [20] Xiao et al. developed an AI tool using CT imaging to predict COVID-19 disease progression. Their deep learning model achieved an impressive 97.4% accuracy in training and 81.9% in testing, with AUCs of 0.987 and 0.892. In patients initially classified as non-severe, the model achieved AUCs of 0.955 and 0.923, with corresponding accuracies of 97.0% and 81.6%. This AI-based approach holds promise for guiding clinical treatment and early intervention in COVID-19 patients based on CT scans. Yu et al. [21] employed deep learning with 729 CT scans to rapidly identify the COVID-19 virus. They used four pre-trained deep models and achieved a 95.20% accuracy with DenseNet-201 and cubic SVM for tenfold cross-validation, indicating a promising method for efficient and reliable severity assessment.

Using five pre-trained CNN-based models (ResNet50-101-152, Inception-ResNetV2, and Inception-ResNetV3), Narin et al. [22] investigate the

identification of coronavirus pneumonia in chest X-ray radiography. The researchers used five-fold cross-validation to examine the accuracy of three binary classifications across four groups (normal, COVID-19, bacterial pneumonia, and viral pneumonia). Particularly impressive is the ResNet50 model's outperformance of the competition on all three datasets, with 96.1% accuracy for Data set-one, 99.5% accuracy for Data set-two, and 99.7% accuracy for Data set-three. Salih et al. [23], investigate the benchmarking of COVID-19 machine learning methods, framed as a multi-criteria decision-making (MCDM) problem. Notably, the Fuzzy Decision by Opinion Score Method, introduced in 2020, has proven effective in addressing complex issues arising from conflicting criteria in MCDM. The research consists of two key stages: the first involves applying eight machine learning methods to chest X-ray (CXR) images to create a novel decision matrix, and the second employs the FDOSM to resolve multiple criteria decision-making challenges.

Nasiri and Alavi [24] introduced a new deep-learning architecture to assist radiologists in diagnosing COVID-19 cases from chest X-ray images, aiming to overcome the drawbacks of Reverse Transcriptase-Polymerase Chain Reaction (RT-PCR). By utilizing a preexisting network called DenseNet169, they utilized analysis of variance (ANOVA) to choose the most relevant features, thereby improving accuracy and reducing computational complexity. The eXtreme Gradient Boosting (XGBoost) technique was utilized for classification, and the assessment was carried out on the ChestX-ray8 dataset. The results exhibited a significant level of precision, with a 98.72% accuracy rate for the classification of two classes (COVID-19 and No-findings), and a 92% accuracy rate for the classification of several classes (COVID-19, No-findings, and Pneumonia). The precision, recall, and specificity rates exhibited remarkable levels of performance, demonstrating the method's usefulness in both categorization scenarios. The comparison with existing approaches demonstrated the framework's higher performance, highlighting its potential usefulness for radiologists in diagnosing COVID-19. This work enhances the progress of deep learning applications in the analysis of medical images, providing useful knowledge on ways to extract important features and classify them. These approaches can enhance the accuracy of diagnoses and assist healthcare professionals in providing prompt care to patients.

Furthermore, a recent study conducted by Prita Patila and Vaibhav Narawade [25], highlights the significance of data balance, data augmentation, and segmentation in the clinical domain. These techniques are crucial for enhancing the accuracy of respiratory ailment detection through the application of machine learning methods. Their proposed concept is to improve the balance of image data by utilizing data augmentation and edge detection techniques. Additionally, they seek

to enhance the effectiveness of radiological image preprocessing by accurately identifying areas of interest (ROI). Patila and Narawade employed a wide range of datasets, including online repositories like Kaggle and real-time radiological pictures acquired from nearby local hospitals. To tackle the issue of imbalanced data distribution, the researchers implemented the RESP_DATA_BALANCE methodology for dataset creation, to attain improved balance in picture data. In addition, their study presented the RDD_ROI (Respiratory Disease Detection Region of Interest) algorithm, which integrates sophisticated image feature extraction methods utilizing Gray-Level Co-occurrence Matrix (GLCM) and unsupervised K-means clustering for segmentation. This technique is essential in finding the specific areas of interest for detecting respiratory disorders in medical imaging. Patila and Narawade introduced a specialized 28-layer Deep Neural Network (RDD_DNN) to aid in the identification and early detection of respiratory illnesses. This deep neural network (DNN) architecture is specifically tailored for training, testing, and validating models used for detecting respiratory diseases. Their experimental findings center on assessing the performance attributes of different techniques for data augmentation, edge detection, and preprocessing. The objective is to enhance the accuracy and efficiency of diagnosing respiratory diseases, ultimately leading to early intervention and better patient outcomes.

These pioneering studies collectively underscore the potential of deep learning in enhancing the precision of COVID-19 disease, offering a pathway toward more objective and data-driven diagnostic practices in the healthcare domain.

3. Dataset

Given the novelty of COVID-19 as a recently identified coronavirus, the acquisition of suitable datasets for research investigations related to COVID-19 presents a substantial challenge. In the context of the proposed research methodology, two distinct datasets were employed. The first dataset comprises a diverse collection of radiographic images sourced from the Guangzhou Medical Center in China [26]. This image set encompasses normal X-rays, X-rays depicting pneumonia resulting from viruses other than COVID-19, and X-rays displaying bacterial pneumonia. The second dataset consists of X-rays captured from a range of COVID-19 patients, meticulously gathered from Sylhet Medical College in Bangladesh and subsequently subjected to rigorous evaluation by a panel of highly qualified radiologists [27]. As illustrated in Figure (1), all COVID-19 X-ray images amalgamated with an equivalent number of normal, and bacterial pneumonia X-rays (305 X-rays in each category), thereby creating a more concise and balanced database.

To train a U-Net model effectively for multi-class segmentation involving three categories (normal, pneumonia, and COVID-19) while utilizing an input shape of 224x224, meticulous preprocessing of input data, one-hot encoding of labels, and the application of data augmentation techniques must be undertaken.

3.1. Data Preprocessing

Image Resizing: The resizing of all input images to the desired dimensions of 224x224 pixels is carried out. This can be accomplished through the utilization of libraries such as OpenCV [28].

Normalization: Normalization is performed on the pixel values of the resized images to bring them within a uniform range, typically [0, 1]. This normalization ensures improved convergence during model training [29].

One-Hot Encoding: Given the presence of three distinct classes (normal, pneumonia, and COVID-19), the labels must undergo one-hot encoding. This process involves the conversion of each label into a binary vector in which each element denotes the presence or absence of a specific class [30]. The following steps are involved:

- The assignment of a unique numerical label to each class (e.g., 0 for normal, 1 for pneumonia, and 2 for COVID-19).
- Creation of a mask for each image, possessing the same dimensions as the image itself, where each pixel is substituted with the corresponding class label.
- Conversion of these classes into one-hot encoded tensors. This transformation can be accomplished using `tf.one_hot` in TensorFlow

Data Augmentation: Data augmentation plays a pivotal role in enhancing model generalization. It entails the generation of diversified training data by applying random transformations to the images [31]. In

this context, the following augmentation techniques should be considered

a. Rotation: Images are randomly rotated by a minor angle (-5 to 5 degrees) to enhance the model's orientation invariance.

b. Horizontal Flip: Horizontal flips are applied with a certain probability (0.5) to simulate mirror images. It is essential to adjust the labels correspondingly when images are flipped horizontally.

c. Translation: Images are shifted both horizontally and vertically by a fraction of the image size (10% of 128 pixels). This introduces variability in object placement.

d. Brightness and Contrast Adjustments: Random adjustments are made to the brightness and contrast of images to simulate diverse lighting conditions.

e. Zoom: Random zooming (inward or outward) is applied to replicate variations in object scale.

f. Gaussian Noise: A minor quantity of Gaussian noise is introduced into the images to bolster robustness.

It is imperative to ensure that these augmentation techniques are applied consistently to both the input images. This approach guarantees alignment and coherence throughout the data augmentation process.

By meticulously adhering to these steps for data preprocessing, one-hot encoding, and data augmentation, the U-Net model can be effectively trained for multi-class segmentation tasks, incorporating the Xception backbone with an input shape of 224x224. These practices serve to enhance the model's aptitude for generalization and its capacity to accurately segment normal, pneumonia, and COVID-19 regions within medical images. Figure (1) shows a sample of the used dataset.

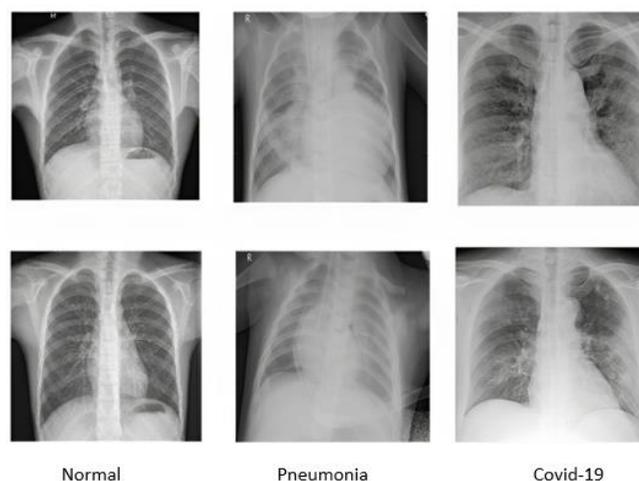


Figure 1. Displays some X-ray images depicting normal, pneumonia, and COVID-19 conditions [27]

4. Modelling

4.1 Xception Model

In the realm of deep learning and computer vision, the Xception model stands out as a game-changer. Developed in 2016 by François Chollet, the creator of Keras, Xception has redefined image recognition with its extreme depth and computational efficiency [32].

Xception, short for Extreme Inception, was introduced to overcome limitations in previous models, especially Inception. Its innovation lies in depth-wise separable convolutions, splitting convolution into depth-wise and pointwise stages. This reduces computational complexity while preserving accuracy.

Xception's core concept is depth-wise separable convolutions, which streamline operations and enhance efficiency. With 36 convolutional layers, Xception learns intricate hierarchical features, making it exceptionally powerful. Despite its depth, it remains efficient, making it suitable for various devices. Moreover, Xception consistently leads on benchmark datasets like ImageNet [33].

Xception's versatility finds applications in image classification, where it achieves high accuracy. It also excels in object detection models, improving accuracy in these tasks. Xception is equally adept at semantic segmentation, accurately segmenting objects within images. Its application extends to medical imaging, where it aids disease diagnosis from medical images.

4.2 Segmentation Models

Segmentation Models is an open-source deep-learning library specializing in image segmentation [6, 34]. It offers pre-trained models like UNet, LinkNet, PSPNet, and FPN, making transfer learning easy. One of its notable strengths lies in its support for a variety of powerful backbones, including popular choices like

Xception, ResNeXt, EfficientNet, and more. This flexibility allows users to select a backbone architecture that precisely suits their project requirements.

One of the remarkable features of Segmentation Models is its ability to convert these diverse backbones into U-Net-like structures. This conversion involves adapting the encoder-decoder architecture commonly found in U-Net for the specific task of image segmentation. The library provides tools and utilities to seamlessly integrate these backbones into the U-Net framework, simplifying the implementation of complex segmentation tasks.

With rich documentation and compatibility with TensorFlow and PyTorch, it's user-friendly and adaptable to different datasets. Supported by a community, Segmentation Models simplifies image segmentation for researchers and developers, staying updated with the latest advancements in computer vision. This capability to convert diverse backbones into U-Net structures makes it a valuable resource for those seeking versatile and high-performing solutions for image segmentation.

4.3 Proposed U-net architecture using Xception

The proposed U-Net architecture leveraging the Xception model entails a detailed approach to effectively capture fine-grained details and features in images for multi-class segmentation tasks. This endeavor involves integrating the feature extraction capabilities of the Xception backbone into the U-Net framework, aiming to achieve pixel-wise classification across three distinct classes within an image. This task is commonly encountered in various domains such as medical image analysis and satellite imagery. To facilitate this integration, the Segmentation_Models library is employed, which offers a range of tools and utilities for simplifying the creation of complex segmentation architectures [35].

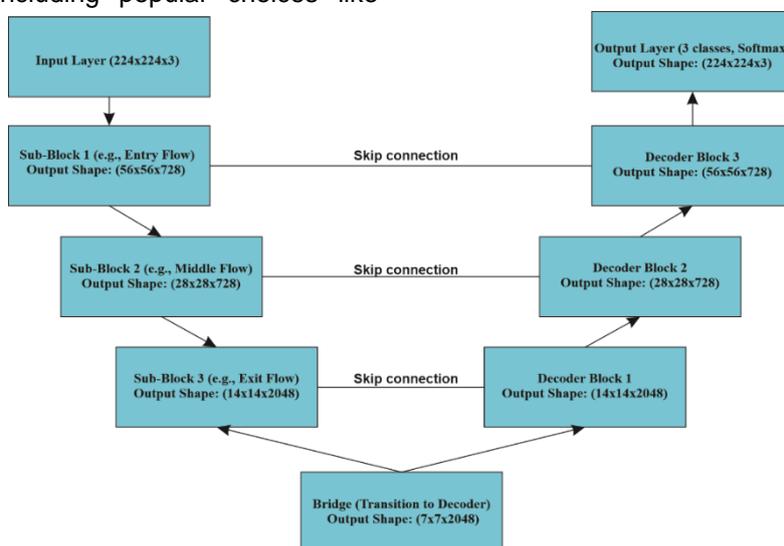


Figure 2. The Xception's new design as U-net after being Segmentation_Models' backbone

Figure (2) illustrates the transformation of the Xception model into the U-Net architecture, serving as a foundational backbone within the Segmentation_Models framework. This adaptation enables the resultant model to differentiate between the three distinct classes within input images. By leveraging the Xception model's capabilities within the U-Net architecture, the proposed model exhibits enhanced performance in capturing intricate image features and patterns, thus enabling potential applications in diverse fields beyond medical diagnosis, such as scene understanding and object detection [32].

The construction of the U-Net architecture using the Xception backbone involves several key steps. Initially, the Xception model is imported and integrated into the U-Net framework, ensuring compatibility and seamless interaction between the two components. Next, the architecture is configured to accommodate the specific requirements of multi-class segmentation tasks, including input and output dimensions, as well as the number of classes to be classified.

Furthermore, fine-tuning and optimization techniques are applied to tailor the model to the target segmentation task, ensuring optimal performance and accuracy. This may involve adjusting hyperparameters, such as learning rate and batch size, as well as incorporating regularization techniques to prevent overfitting. Additionally, extensive experimentation and validation are conducted to evaluate the model's performance across various datasets and segmentation scenarios [7].

The proposed U-Net architecture leveraging the Xception backbone represents a comprehensive approach to image segmentation, offering advanced capabilities for capturing and analyzing complex image features. By combining state-of-the-art deep learning techniques with versatile model architectures, this methodology aims to provide researchers and practitioners with a robust framework for tackling challenging segmentation tasks in diverse application domains.

5. Training the Model

Image segmentation plays a pivotal role in the realm of computer vision, finding applications in diverse fields such as medical image analysis and object detection. The process of training a U-Net model for image segmentation using the Segmentation_Models library, with a specific focus on converting the Xception model into a U-Net architecture to utilize it as a classification model.

To commence the journey of training a U-Net model for image segmentation, it is known by now that the input images are of RGB format, possessing dimensions of (224, 224, 3). The model construction entails a synthesis of an encoder and a classification

head. This amalgamation allows for the efficient execution of image segmentation tasks.

The initial step involves defining the input shape as (224, 224, 3). Subsequently, the Xception model is designated as the encoder, its input shape aligned with the dimensions of the aforementioned images. A classification head is subsequently crafted, consisting of layers for global average pooling, dense units with ReLU activation functions, and an output layer containing the requisite number of classes for classification.

The pivotal aspect of this model architecture lies in the fusion of the encoder, in this instance, Xception, with the classification head. This symbiotic relationship empowers the model to operate as an efficient classification tool, capable of discerning objects or patterns within images.

Upon the completion of model definition and compilation, the training phase ensues. A batch size of 32 is chosen for this example, though it can be adjusted as necessitated by specific datasets and task requirements. The training process spans 32 epochs to accommodate the particular demands of the segmentation project.

5.1. Hyperparameter tuning

Hyperparameter tuning is a crucial aspect of training deep learning models to ensure optimal performance. In the context of training a U-Net model for image segmentation, several hyperparameters can be fine-tuned to enhance the model's effectiveness. Here, we delve deeper into the hyperparameter tuning process:

1. Learning Rate: The learning rate determines the step size during gradient descent optimization. It significantly influences the convergence speed and final performance of the model. During hyperparameter tuning, different learning rates are tested to find the optimal value that balances fast convergence without overshooting the minimum loss.

2. Batch Size: The batch size determines the number of samples processed before updating the model's parameters. It affects the stability of the training process and the memory requirements. Experimenting with various batch sizes helps find the right balance between computational efficiency and model convergence.

3. Number of Epochs: An epoch refers to one complete pass through the entire training dataset. The number of epochs defines how many times the model iterates over the dataset during training. Tuning this hyperparameter involves finding the point where the model achieves satisfactory performance without overfitting or underfitting.

4. Optimizer Choice and Parameters:

Different optimizers, such as Adam, SGD, or RMSprop, have distinct behaviors and parameter requirements. Hyperparameter tuning involves experimenting with various optimizers and their associated parameters, such as momentum or decay rates, to find the combination that yields the best results.

5. Data Augmentation Parameters:

Data augmentation techniques, such as rotation, scaling, or flipping, can improve model generalization by exposing it to a diverse range of data samples. Tuning the parameters of data augmentation, such as rotation angle range or scaling factors, helps optimize the augmentation strategy for the specific dataset and task.

6. Regularization Techniques:

Regularization methods, such as dropout or L2 regularization, help prevent overfitting by penalizing overly complex models. Hyperparameter tuning involves adjusting the regularization strength or dropout rates to strike the right balance between model complexity and generalization performance.

7. Model Architecture Modifications:

While not strictly hyperparameters, modifications to the U-Net architecture, such as the number of layers, filter sizes, or skip connections, can significantly impact performance. Hyperparameter tuning may involve experimenting with different architectural variations to find the most suitable configuration for the task at hand.

In our experiment, hyperparameters were carefully tweaked to optimize U-Net picture segmentation training. For gradient descent optimization, the Learning Rate was chosen to be 0.003 to balance speedy convergence and stability. The Batch Size was 16, allowing efficient sample processing with little memory usage. The Number of Epochs was 50 with Early Stopping methods to prevent overfitting and ensure convergence in a realistic timeframe.

For optimization, the Adam optimizer was used due to its flexible learning rate and sparse gradient handling. A 0.3-dropout Regularization Technique was used to reduce model complexity and overfitting.

This dropout rate prevents the model from becoming too dependent on certain features during training, improving generalization to new data.

Configuring these hyperparameters based on empirical data and best practices was used to train a U-Net model that captures detailed image details for reliable segmentation in medical imaging and object detection.

6. Results

In this part, the efficiency of the proposed schemes is demonstrated alongside visual interpretations of the therapeutic significance of geographical localization. To assess the reliability of the procedure, several instances are examined using COVID-19 X-rays.

6.1. Experimental setup

To enhance the performance of the network, several experiments are undertaken to determine the most appropriate hyper-parameters. The research is conducted on a computational infrastructure equipped with an Intel® Xeon® CPU operating at a frequency of 2.80 GHz, featuring a (M) Cache and several cores. This infrastructure is further enhanced by a substantial amount of RAM, totaling several gigabytes. Hardware acceleration is accomplished by employing an NVIDIA RTX Ti GPU, which features CUDA cores operating at a frequency of MHz and is supplied with GB of GDDR6 memory. The assessment of the suggested architectures is based on a variety of conventional classification criteria, including accuracy, specificity, sensitivity, precision, and recall.

During the initial training phase, the neural network is tailored to classify normal and non-COVID pneumonia X-rays. The U-Net architecture, which is derived from the Xception model, has a high degree of scalability, allowing for the adaptation of the receptive area following the characteristics of the incoming data.

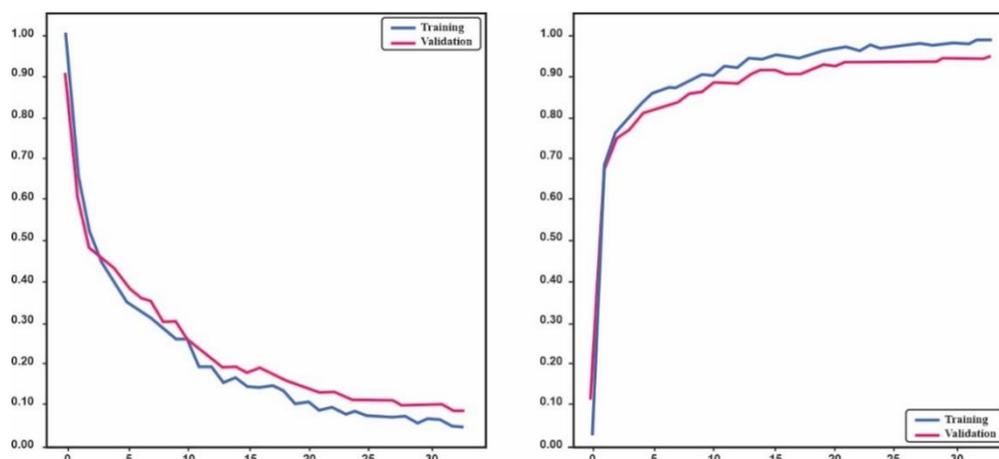


Figure 3. Training loss and accuracy

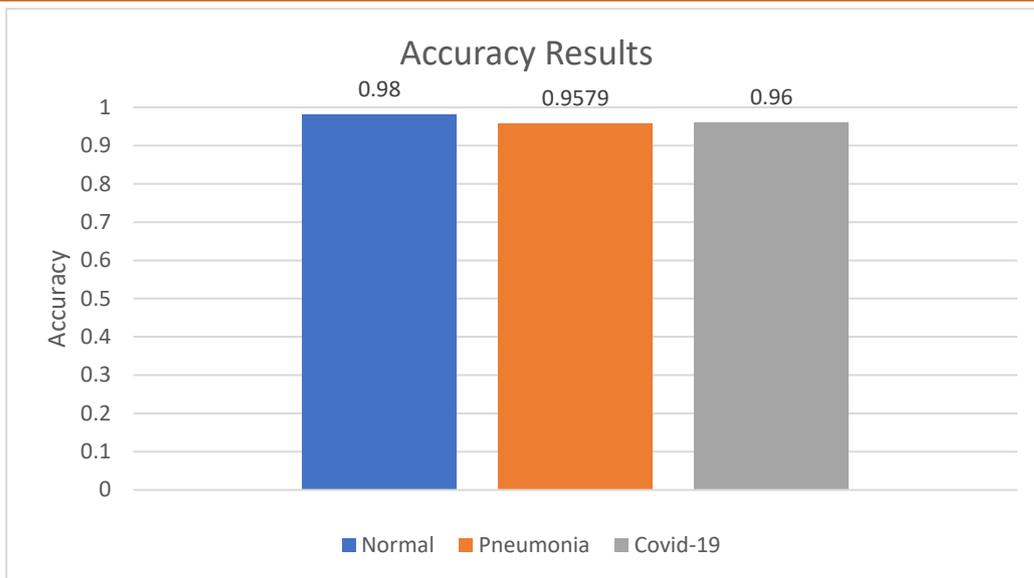


Figure 4. Implications of the stacking algorithm during preliminary training

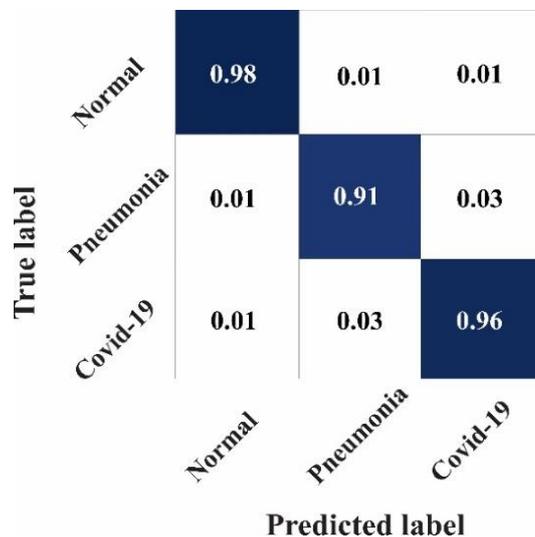


Figure 5. Multi-class confusion matrix

Upon the conclusion of the preliminary training phase focused on non-COVID X-rays, the convolutional layers, which have been extensively improved, are then employed for training purposes using a reduced database consisting exclusively of COVID-19 X-rays. During the transfer learning phase, COVID-19 X-rays were subjected to experimentation with the fine-tuning of additional layers to explore various output classes, including normal and classic pneumonia.

In a manner akin to the primary training phase, an auxiliary meta-learner is trained to optimize the predictions derived from various adaptations of the modified U-Net architecture, each specifically tuned for distinct resolutions of input X-ray images. Figure (4) illustrates the performance of the independently trained networks as well as the performance achieved after combining them using a meta-learner.

The confusion matrix for the multi-class classification is presented in Figure (5). As anticipated,

a small number of COVID-19 cases demonstrate misclassification as a result of a significant level of overlapping characteristics. Nevertheless, highly satisfactory results are achieved in classification scenarios. Table (1) shows a comparison of the proposed work compared to previous work.

6.2. Test Statistic Design

Confusion Matrix calculations:

• **True Positives (TP):**

- For Normal, TP = 0.98
- For Pneumonia, TP = 0.91
- For COVID-19, TP = 0.96

• **False Positives (FP):**

- For Normal, FP = 0.01 + 0.01 (misclassifying Pneumonia and COVID-19 as Normal)

- For Pneumonia, FP = 0.01 + 0.03 (misclassifying Normal and COVID-19 as Pneumonia)

- For COVID-19, FP = 0.03 + 0.01 (misclassifying Normal and Pneumonia as COVID-19)

• **False Negatives (FN):**

- For Normal, FN = 0.01 + 0.03 (misclassifying Pneumonia and COVID-19 as not Normal)

- For Pneumonia, FN = 0.01 + 0.03 (misclassifying Normal and COVID-19 as not Pneumonia)

- For COVID-19, FN = 0.01 + 0.01 (misclassifying Normal and Pneumonia as not COVID-19)

- Precision for Normal = $0.98 / (0.98 + (0.01 + 0.01)) = 0.98 / 1.00 = 0.98$
- Recall for Normal = $0.98 / (0.98 + (0.01 + 0.03)) = 0.98 / 1.02 \approx 0.9608$

➤ Precision for Pneumonia = $0.91 / (0.91 + (0.01 + 0.03)) = 0.91 / 0.95 \approx 0.9579$

➤ Recall for Pneumonia = $0.91 / (0.91 + (0.01 + 0.03)) = 0.91 / 0.95 \approx 0.9579$

➤ Precision for COVID-19 = $0.96 / (0.96 + (0.03 + 0.01)) = 0.96 / 1.00 = 0.96$

➤ Recall for COVID-19 = $0.96 / (0.96 + (0.01 + 0.01)) = 0.96 / 0.98 \approx 0.9796$

➤ F1 Score for Normal = $2 * (0.98 * 0.9608) / (0.98 + 0.9608) \approx 0.9704$

➤ F1 Score for Pneumonia = $2 * (0.9579 * 0.9579) / (0.9579 + 0.9579) \approx 0.9579$

➤ F1 Score for COVID-19 = $2 * (0.96 * 0.9796) / (0.96 + 0.9796) \approx 0.9698$

The results displayed in Table (2) offer useful insights into the efficacy of the multi-class classification algorithm in identifying Normal, Pneumonia, and COVID-19 cases.

Table 1. Comparison with previous works

Author Name	Method	Results
Shan et al. [8]	Deep learning-based segmentation in CT scans.	91.6% Dice similarity and 73.4% severity prediction accuracy
Amyar et al. [9]	Introduced CovXNet.	96.9% COVID/Viral, 94.7% COVID/Bacterial, 90.2% COVID/Norm, multiclass COVID/Norm/Viral/Bacterial
He et al. [11]	Synergistic learning strategy on a CT dataset.	98.5% accuracy
Mijwil and Maad [12]	Used Inception-v2 and VGG-16 to detect COVID-19 in X-rays, with Inception-v2.	97% accuracy for VGG-16, 93% accuracy for Inception-v2
Khanday et al. [13]	Employed Adaboost for data labeling.	98.7% accuracy in binary classification
Tang et al. [14]	Used AI to differentiate between severe and non-severe COVID-19 cases in CT scans.	87.5% average precision
Xiao et al. [15]	Developed a deep learning model for predicting COVID-19 progression based on CT scans.	97.4% training accuracy and 81.9% testing accuracy.
Yu et al. [16]	Employed deep learning on 729 CT scans for COVID-19 identification.	95.20% accuracy
Narin et al. [17]	Investigated coronavirus pneumonia identification using pre-trained models, with ResNet50 achieving.	Database1 96.1%, Database1 96.1%, Database2 99.5%, Database3 99.7%
Salih et al. [18]	Used multi-criteria decision-making for benchmarking COVID-19 machine learning methods.	Effective addressing of complex issues.
Zhang et al. [19]	Tested a deep learning model for COVID-19 pneumonia detection in CT scans.	Sensitivity (97.2%) Detecting affected lung (80.8%)
Proposed Method	U-Net using Xception model as backbone	Normal 98%

Table 2. Comparison with previous works

Class	Precision	Recall	F1 Score
Normal	0.98	0.9608	0.9704
Pneumonia	0.9579	0.9579	0.9579
COVID-19	0.96	0.9796	0.9698

An outstanding feature is the exceptional accuracy attained in all three categories, with values ranging from 0.91 to 0.98.

This demonstrates the model's capacity to reduce the occurrence of false positives, which is of utmost importance in medical diagnostics to prevent unneeded treatments or interventions. In addition, the recall values, which assess the model's accuracy in properly identifying true positives, are remarkably high, ranging from 0.9608 to 0.9796. This implies that the model successfully captures a substantial fraction of positive instances in each category, which is crucial for precise illness diagnosis.

The F1 ratings, which measure the harmonic mean of precision and recall, provide a well-balanced evaluation of the model's performance. The model consistently performs well across all three classes, with F1 scores ranging from 0.9579 to 0.9704. The scores demonstrate a favorable equilibrium between precision and recall, indicating that the model can proficiently differentiate between the classes while minimizing the occurrence of both false positives and false negatives.

Moreover, the findings emphasize the model's ability to effectively handle imbalanced datasets, which is a frequent obstacle in medical picture analysis because certain classes may have fewer instances. Although there are inherent disparities in social class, the model consistently achieves excellent results for all categories, demonstrating its efficacy in real-life situations where the occurrence of certain illnesses may differ.

In summary, the findings indicate that the suggested multi-class classification model, utilizing sophisticated deep learning methods, has the potential for precise and dependable identification of respiratory disorders from medical imaging data. Additional validation on larger and more diversified datasets, as well as clinical testing, could offer a further understanding of the model's capacity to apply to various situations and its practicality in real-world scenarios.

7. Discussion

The proposed method for diagnosing COVID-19 in X-ray images using deep neural networks has some important limitations:

Data Availability and Quality: One of the major challenges is the availability and quality of datasets. The scarcity of large, diverse, and annotated datasets specific to COVID-19 X-ray images can limit the model's ability to generalize across different populations and imaging conditions. Moreover, the quality and consistency of labeling in datasets can vary, leading to potential biases and inaccuracies in model training.

Generalization: While the model may perform well on the datasets it was trained on, its ability to generalize to new and unseen data, especially from different demographics or imaging modalities, remains uncertain. Variations in imaging techniques, equipment, and patient demographics can affect the model's performance and reliability in real-world clinical settings.

Interpretability: Deep neural networks are often criticized for their lack of interpretability. Understanding how the model arrives at its predictions, particularly in medical diagnosis where interpretability is crucial for trust and acceptance, can be challenging. Clinicians may be hesitant to rely solely on the model's output without insight into the underlying decision-making process.

Clinical Validation: The proposed method may require extensive clinical validation to demonstrate its efficacy and safety before widespread adoption in clinical practice. This involves rigorous testing against established diagnostic standards and guidelines, as well as validation across diverse patient populations and clinical settings.

8. Conclusion

In this study, a novel deep neural network architecture, based on the transformation of the Xception model into a U-Net configuration using the Segmentation_Models framework, was proposed for the efficient detection of COVID-19 and various pneumonia types in chest X-ray images. Features from diverse receptive fields were integrated into the design to facilitate the analysis of X-ray abnormalities from multiple perspectives. To address the limited availability of COVID-19 X-ray images, an expanded dataset consisting of X-rays from both healthy individuals and pneumonia patients was utilized for initial network training. Given the overlapping visual characteristics between COVID-19 and other pneumonia cases, the

network's performance was subsequently improved by transferring initially trained convolutional layers, along with the addition of fine-tuned layers. This resulted in a highly satisfactory diagnostic performance, even when utilizing a smaller dataset exclusively containing COVID-19 X-rays. It is important to note that the effectiveness of these methodologies can be further enhanced by the inclusion of additional COVID-19 patient X-rays during the transfer learning phase. Experimental results from extensive simulations underscore the practicality of this approach for expediting the diagnosis of COVID-19 and other pneumonia cases. Moreover, the proposed U-Net architecture is characterized by scalability and a substantial receptive capacity, making it applicable to a wide range of computer vision tasks

References

- [1] F. Wu, S. Zhao, B. Yu, Y.M. Chen, W. Wang, Z.G. Song, Y. Hu, Z.W. Tao, J.H. Tian, Y.Y. Pei, M.L.A. Yuan, A new coronavirus associated with human respiratory disease in China. *Nature*, 579(7798), (2020) 265-269. <https://doi.org/10.1038/s41586-020-2008-3>
- [2] R. Mehrrotraa, M.A. Ansari, R. Agrawal, P. Tripathi, M.B.B. Heyat, M. Al-Sarem, A.Y.M. Muaad, W.A.E. Nagmeldin, A. Abdelmaboud, F. Saeed, Ensembling of efficient deep convolutional networks and machine learning algorithms for resource effective detection of tuberculosis using thoracic (chest) radiography. *IEEE Access*, 10, (2022) 85442-85458. <https://doi.org/10.1109/ACCESS.2022.3194152>
- [3] Z. Zhou, M. M. Rahman Siddiquee, N. Tajbakhsh, J. Liang, Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, Proceedings* 4, (2018) 3-11. https://doi.org/10.1007/978-3-030-00889-5_1
- [4] N. Siddique, S. Paheding, C.P. Elkin, V. Devabhaktuni, U-net and its variants for medical image segmentation: A review of theory and applications. *IEEE Access*, 9, (2021) 82031-82057. <https://doi.org/10.1109/ACCESS.2021.3086020>
- [5] D. Gupta, Image Segmentation Keras : Implementation of Segnet, FCN, UNet, PSPNet and other models in Keras, arXiv preprint arXiv:2307.13215 (2023). <https://doi.org/10.48550/arXiv.2307.13215>
- [6] A.P. Apte, A. Iyer, M. Thor, R. Pandya, R. Haq, J. Jiang, E. LoCastro, A. Shukla-Dave, N. Sasankan, Y. Xiao, Y.C. Hu, Library of deep-learning image segmentation and outcomes model-implementations. *Physica Medica*, 73, (2020) 190-196. <https://doi.org/10.1016/j.ejmp.2020.04.011>
- [7] O. Ronneberger, P. Fischer, T. Brox, Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, proceedings, part III 18*, (2015) 234-241. https://doi.org/10.1007/978-3-319-24574-4_28
- [8] A.T. Meem, M.M. Khan, M. Masud, S. Aljahdali, Prediction of Covid-19 Based on Chest X-Ray Images Using Deep Learning with CNN. *Computer Systems Science & Engineering*, 41(3), (2022) 121668. <https://doi.org/10.32604/csse.2022.021563>
- [9] M. Diwakar, P. Singh, A. Shankar, Multi-modal medical image fusion framework using co-occurrence filter and local extrema in NSST domain. *Biomedical Signal Processing and Control*, 68, (2021) 102788. <https://doi.org/10.1016/j.bspc.2021.102788>
- [10] M. Das, D. Gupta, A. Bakde, An end-to-end content-aware generative adversarial network-based method for multimodal medical image fusion. In *Data Analytics for Intelligent Systems: Techniques and solutions*. IOP Publishing, Bristol, UK, (2024). <https://doi.org/10.1088/978-0-7503-5417-2ch7>
- [11] Y. Jie, Y. Xu, X. Li, H. Tan, TSJNet: A Multi-modality Target and Semantic Awareness Joint-driven Image Fusion Network. arXiv preprint arXiv:2402.01212 (2024). <https://doi.org/10.48550/arXiv.2402.01212>
- [12] E. Irmak, Implementation of convolutional neural network approach for COVID-19 disease detection. *Physiological Genomics*, 52(12), (2020) 590-601. <https://doi.org/10.1152/physiolgenomics.00084.2020>
- [13] F. Shan, Y. Gao, J. Wang, W. Shi, N. Shi, M. Han, Z. Xue, D. Shen, Y. Shi, Abnormal lung quantification in chest CT images of COVID-19 patients with deep learning and its application to severity prediction. *Medical physics*, 48(4), (2021) 1633-1645. <https://doi.org/10.1002/mp.14609>
- [14] A. Amyar, R. Modzelewski, H. Li, S. Ruan, Multi-task deep learning-based CT imaging analysis for COVID-19 pneumonia: Classification and segmentation. *Computers in Biology and Medicine*, 126, (2020) 104037.

- <https://doi.org/10.1016/j.compbimed.2020.104037>
- [15] J. Zhu, B. Shen, A. Abbasi, M. Hoshmand-Kochi, H. Li, T.Q. Duong, Deep transfer learning artificial intelligence accurately stages COVID-19 lung disease severity on portable chest radiographs. *PloS one*, 15(7), (2020) e0236621. <https://doi.org/10.1371/journal.pone.0236621>
- [16] K. He, W. Zhao, X. Xie, W. Ji, M. Liu, Z. Tang, Y. Shi, F. Shi, Y. Gao, J. Liu, J. Zhang, J. Synergistic learning of lung lobe segmentation and hierarchical multi-instance classification for automated severity assessment of COVID-19 in CT images. *Pattern recognition*, 113, (2021) 107828. <https://doi.org/10.1016/j.patcog.2021.107828>
- [17] M.M. Mijwil, Deep Convolutional Neural Network Architecture to Detect COVID-19 from Chest X-Ray Images. *Iraqi Journal of Science*, (2023) 2561-2574. <https://doi.org/10.24996/ij.s.2023.64.5.38>
- [18] A.M.U.D. Khanday, Q.R. Khan, S.T. Rabani, Ensemble Approach for Detecting COVID-19 Propaganda on Online Social Networks. *Iraqi Journal of Science*, 63(10), (2022) 4488-4498. <https://doi.org/10.24996/ij.s.2022.63.10.33>
- [19] Z. Tang, W. Zhao, X. Xie, Z. Zhong, F. Shi, J. Liu, D. Shen, Severity assessment of coronavirus disease 2019 (COVID-19) using quantitative features from chest CT images. *arXiv preprint arXiv:2003.11988* (2020). <https://doi.org/10.48550/arXiv.2003.11988>
- [20] L.S. Xiao, P. Li, F. Sun, Y. Zhang, C. Xu, H. Zhu, F.Q. Cai, Y.L. He, Y.L., W.F. Zhang, S.C. Ma, C. Hu, Development and validation of a deep learning-based model using computed tomography imaging for predicting disease severity of coronavirus disease 2019. *Frontiers in Bioengineering and Biotechnology*, 8, (2020) 898. <https://doi.org/10.3389/fbioe.2020.00898>
- [21] Z. Yu, X. Li, H. Sun, J. Wang, T. Zhao, H. Chen, Y. Ma, S. Zhu, Z. Xie, Rapid identification of COVID-19 severity in CT scans through classification of deep features. *Biomedical Engineering Online*, 19, (2020) 1-13. <https://doi.org/10.1186/s12938-020-00807-x>
- [22] A. Narin, C. Kaya, Z. Pamuk, Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. *Pattern Analysis and Applications*, 24, (2021) 1207-1220. <https://doi.org/10.1007/s10044-021-00984-y>
- [23] M.M. Salih, M. Ahmed, B. Al-Bander, K.F. Hasan, M.L. Shuwandy, Z. Al-Qaysi, Benchmarking Framework for COVID-19 Classification Machine Learning Method Based on Fuzzy Decision by Opinion Score Method. *Iraqi Journal of Science*, 64(2), (2023) 922-943. <https://doi.org/10.24996/ij.s.2023.64.2.36>
- [24] H. Nasiri, S.A. Alavi, A novel framework based on deep learning and ANOVA feature selection method for diagnosis of COVID-19 cases from chest X-ray images. *Computational intelligence and Neuroscience*, 2022, (2022). <https://doi.org/10.1155/2022/4694567>
- [25] P. Patil, V. Narawade, Radiology Image Data Augmentation and Image Enhancement in Respiratory Disease Infection Detection Using Machine Learning Approach. *International Research Journal of Multidisciplinary Technovation*, 6(2), (2024) 133-155. <https://doi.org/10.54392/irjmt24211>
- [26] D. Kermany, K. Zhang, M. Goldbaum, Labeled optical coherence tomography (oct) and chest x-ray images for classification. *Mendeley data*, 2(2), (2018) 651. <https://doi.org/10.17632/rscbjbr9sj.2>
- [27] T. Mahmud, M.A. Rahman, S.A. Fattah CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. *Computers in biology and medicine*, 122, (2020) 103869. <https://doi.org/10.1016/j.compbimed.2020.103869>
- [28] I. Culjak, D. Abram, T. Pribanic, H. Dzapo, M. Cifrek, A brief introduction to OpenCV. in 2012 proceedings of the 35th international convention MIPRO. *IEEE*, (2012) 1725-1730.
- [29] P. Sane, R. Agrawal, Pixel normalization from numeric data as input to neural networks: For machine learning and image processing. In 2017 international conference on wireless communications, signal processing and networking (WiSPNET), *IEEE*, (2017) 2221-2225. <https://doi.org/10.1109/WiSPNET.2017.8300154>
- [30] L. Jie, C. Jiahao, Z. Xueqin, Z. Yue, L. Jiajun, One-hot encoding and convolutional neural network-based anomaly detection. *Journal of Tsinghua University (Science and Technology)*, 59(7), (2019) 523-529.
- [31] P. Chlap, H. Min, N. Vandenberg, J. Dowling, L. Holloway, A. Haworth, A review of medical image data augmentation techniques for deep learning applications. *Journal of Medical Imaging and Radiation Oncology*, 65(5), (2021) 545-563. <https://doi.org/10.1111/1754-9485.13261>

- [32] F. Chollet, Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, (2017) 1251-1258. <https://doi.org/10.1109/CVPR.2017.195>
- [33] T. Ridnik, E. Ben-Baruch, A. Noy, L. Zelnik-Manor, Imagenet-21k pretraining for the masses. arXiv preprint arXiv:2104.10972 (2021). <https://doi.org/10.48550/arXiv.2104.10972>
- [34] D. Wilson, S. Laxminarayan, Handbook of Biomedical Image Analysis: Volume 1: Segmentation Models Part A. Springer Science & Business Media (2006).
- [35] D. Müller, F. Kramer, MIScnn: a framework for medical image segmentation with convolutional neural networks and deep learning. BMC medical Imaging, 21, (2021) 1-11. <https://doi.org/10.1186/s12880-020-00543-7>

Data Availability

Data will be provided based on request if required.

Has this article screened for similarity?

Yes

About the License

© The Author 2024. The text of this article is open access and licensed under a Creative Commons Attribution 4.0 International License.