

STRENGTHENING DEEP LEARNING MODEL FOR ROBUST SCREENING OF VOLUMETRIC CHEST RADIOGRAPHIC SCANS

Muhammad Owais¹, Taimur Hassan², Neha Gour¹, Iyyakutti Iyappan Ganapathi¹, and Naoufel Werghi¹

¹ C2PS and KUCARS, Department of Electrical Engineering and Computer Science, Khalifa University, UAE

² Department of Electrical, Computer and Biomedical Engineering, Abu Dhabi University, UAE

ABSTRACT

The emerging deep learning algorithms have shown significant potential in the development of efficient computer-aided diagnosis tools for automated detection of lung infections using chest radiographs. However, many existing methods are slice-based and require manual selection of appropriate slices from the entire CT scan, which is tedious and requires expert radiologists. To overcome these limitations, we propose a recurrent 3D Inception network (R3DI-Net) that sequentially exploits spatial and 3D structural features of the entire CT scan, ultimately leading to improved diagnostic performance. Additionally, the proposed method flexibly handles input CT scans with a variable number of slices without incurring performance degradation. A quantitative evaluation of R3DI-Net was made using a combined collection of three publicly accessible datasets containing a sufficient number of data samples. Our method outperforms various existing methods by achieving remarkable performances of 98.39%, 98.36%, 98.1%, and 98.64% in terms of accuracy, F1-score, sensitivity, and average precision, respectively.

Index Terms—R3DI-Net, computer-aided diagnosis, lung infection, radiology.

1. INTRODUCTION

The recent breakthrough in the field of artificial intelligence (AI) methods has had remarkable success in developing efficient computer-aided diagnosis (CAD) tools in the medical field [1–4]. Such CAD methods employed a new class of AI methods, known as “deep learning” (DL) algorithms such as neural networks to aid in the diagnosis of different medical conditions using medical imaging data [5,6]. Deep learning algorithms have shown a great potential in this area by learning to extract important features from medical imaging data and making predictions based on those features [7]. Despite its success, deep learning in medical image analysis is still a rapidly evolving field with ongoing research and development intended to improve its performance and generalizability.

Numerous CAD solutions have been proposed for lung infection analysis to improve the accuracy of diagnosis and treatment planning using chest radiographs like computed

tomography (CT) and X-ray scans [8]. These methods can recognize patterns in the CT scan indicative of specific diseases, such as changes in lung tissue density or fluid in the lungs. However, an accurate diagnostic decision for these methods depends on selecting appropriate slices from the whole CT volume resulting in manual effort and time. In addition, slice-based CAD methods have some limitations in lung CT data analysis: 1) Limited spatial context: Only a single slice of the CT scan is explored at a time, which can result in a surplus of spatial context. This can make it difficult for the model to precisely recognize lesions or structures spread across multiple slices [9]. 2) Lack of robustness: They are sensitive to the choice of slice orientation and thickness and may cause performance degradation when the slices are not representative of the original structures [10]. 3) Annotation bias: These methods may be biased towards the annotations provided during training, as the model only sees the slices that were annotated. This can affect the generalizability of the model to new cases where the annotated slices may not be representative of the complete CT scan [11]. In contrast, volumetric data analysis methods consider the whole scan and have the advantage of providing a detailed view of the anatomy and can be more robust and accurate in detecting lesion regions [12].

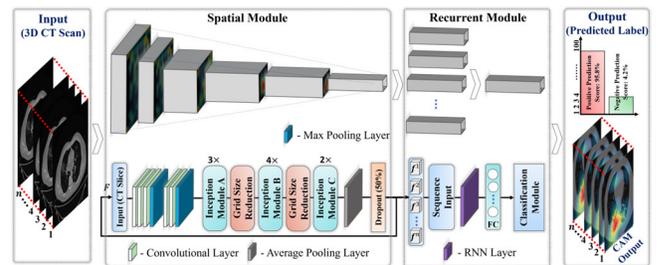


Fig. 1. Overall architecture of the proposed recurrent 3D Inception network (R3DI-Net), comprising two main building blocks: spatial module (SM) and recurrent module (RM).

To overcome the limitations of traditional slice-based methods, a new sequence-based CAD model (Fig. 1) is proposed for effective screening of lung infections utilizing the complete CT volume of varying lengths. This study presents the following key contributions:

- A sequence-based recurrent 3D Inception network (R3DI-Net) is proposed that leverages the power of the inception

module, followed by the recurrent module, to extract multi-scale spatial and 3D structural features, respectively.

- The network design can handle a variable-length CT volume as input and employs sequential training to achieve optimal convergence, utilizing transfer learning to analyze volumetric data.
- The proposed R3DI-Net will be made publicly accessible for further research, development, and educational purposes.

2. PROPOSED METHOD

2.1 Network Structure

The proposed R3DI-Net consists of two sub-networks, referred to as the spatial module and the recurrent module. The following subsections provide a comprehensive overview of each sub-network in detail.

2.1.1. Spatial Module: The spatial module takes advantage of the abilities of inception blocks [13] to analyze both the spatial and channel-wise information at various scales contained in each input CT slice. The design mainly comprises three key building blocks known as inception modules A, B, and C, as well as some additional convolutional, average pooling, and max pooling layers, as illustrated in Fig. 1. Module A is used to extract information from the input image at multiple scales. It consists of multiple parallel branches, each with a different filter size, that extract information from different parts of the image. The outputs from these branches are then concatenated together to form a multi-scale representation of the input image. Module B is similar to Inception A but also includes a pooling layer to reduce the spatial resolution of the input tensor. This block helps the network extract more abstract and high-level features while retaining the necessary information for accurate object recognition. Module C is used to get the final prediction and typically consists of multiple parallel branches, each with a different filter size, that extract information from different parts of the image. The outputs from these branches are then concatenated and passed through a few fully connected layers to make the final prediction. Additionally, there is a grid size reduction block to decrease the spatial dimension of the input tensor from the previous stage. The purpose of reducing the grid size is to reduce the computational complexity of the network, as well as to provide a form of spatial invariance, which means that the network can activate the key regions (i.e., lesion patterns) regardless of their position in the image. Mathematically, the input tensor \mathbf{S}_i with dimensions $w_i \times h_i \times d_i$ is subjected to the following layer-wise transformations after passing through these building blocks:

$$\begin{bmatrix} \Psi_A(\mathbf{S}_i) \\ \Psi_B(\mathbf{S}_i) \\ \Psi_C(\mathbf{S}_i) \\ \Psi_{GS}(\mathbf{S}_i) \end{bmatrix} = \begin{bmatrix} h_{1,3,3}(\mathbf{S}_i) \cdot h_{1,5}(\mathbf{S}_i) \cdot h_1(\mathbf{S}_i) \cdot h_1(A_3(\mathbf{S}_i)) \\ h_{1,7,7}(\mathbf{S}_i) \cdot h_{1,7}(\mathbf{S}_i) \cdot h_1(\mathbf{S}_i) \cdot h_1(A_3(\mathbf{S}_i)) \\ h_{1,3,3'}(\mathbf{S}_i) \cdot h_{1,3'}(\mathbf{S}_i) \cdot h_1(\mathbf{S}_i) \cdot h_1(A_3(\mathbf{S}_i)) \\ h_{1,3,3}(\mathbf{S}_i) \cdot h_3(\mathbf{S}_i) \cdot M_3(\mathbf{S}_i) \end{bmatrix} \quad (1)$$

The functions $\Psi_A(\cdot)$, $\Psi_B(\cdot)$, $\Psi_C(\cdot)$, and $\Psi_{GS}(\cdot)$ correspond to the inception modules A, B, C, and grid size reduction block,

respectively, represented as transfer functions. The operation h_{x_1, \dots, x_n} denotes the convolutional layers applied sequentially with filter sizes ranging from x_1 to x_n . The * and ' symbols indicate sequential and parallel separable convolution operations, respectively. The functions $A_x(\cdot)$ and $M_x(\cdot)$ denote average-pooling and max-pooling operations with a filter size of x , respectively. Finally, the dot operation (\cdot) indicates depth-wise feature concatenation.

2.1.2. Recurrent Module: The recurrent module employs a modified recurrent neural network (RNN) termed the long short-term memory (LSTM) model to explore deeper into the 3D structural information [9,14]. The LSTM model consists of a series of cascaded LSTM cells [14], including a memory cell unit and a group of three gate units: input, forget, and output. It is suitable for processing sequences of 2D CT slices of both fixed and variable lengths and is designed to capture 3D structural dependencies among the slices. Integrating a cascade of spatial and recurrent modules enables the implementation of transfer learning for volumetric data without affecting the total number of training parameters. The advantages of LSTM were incorporated to develop the R3DI-Net for efficient screening of volumetric CT data.

2.2 Network Workflow

The proposed network comprises three convolutional layers followed by a max pooling layer which yields a low-level semantic feature map of size $73 \times 73 \times 64$ from each CT slice of size 299×299 . Two more convolutional layers are added, followed by another max pooling layer to refine the input tensor further, resulting in a new feature space of size $35 \times 35 \times 192$. The subsequent stack of inception modules, which includes three A modules, four B modules, two C modules, and two grid size reduction blocks (as depicted in Fig. 1), extracts multilevel semantic information from the output tensor of the previous layer in succession. This ultimately results in a high-level feature map of size $8 \times 8 \times 2048$. Finally, this high-level feature map is further reduced to a feature vector \mathbf{f} of size $1 \times 1 \times 2048$ by the 8×8 average-pooling layer. The final output feature vector is a multi-scale semantic representation of the input slice.

The proposed spatial module processes each of the n slices ($\mathbf{S}^1, \mathbf{S}^2, \dots, \mathbf{S}^n$) successively, and outputs a set of n feature vectors ($\mathbf{f}^1, \mathbf{f}^2, \dots, \mathbf{f}^n$) of size $1 \times 1 \times 2048 \times n$. These feature vectors are then accumulated and processed by the recurrent module to extract 3D structural features and make a prediction. The accumulated set of n feature vectors ($\mathbf{f}^1, \mathbf{f}^2, \dots, \mathbf{f}^n$) is passed to the LSTM layer through a sequence input layer of the recurrent module, which exploits the 3D structural dependencies among these feature vectors and generates a single feature vector \mathbf{h}_n of size $1 \times 1 \times 600$. Subsequently, a fully connected layer (FC) further analyzes \mathbf{h}_n by mapping it into a low-dimensional feature vector of size $1 \times 1 \times 128$. Finally, based on the highest probability score, the classification module, including FC and softmax activation function, yields predicted class labels for the CT volume.

2.3 Training Loss

To optimize the convergence of the R3DI-Net, a sequential training approach was utilized that leveraged the power of transfer learning in volumetric data analysis. In the first step, the first inception module was trained using the training dataset $\langle [S_{Tr}]_{i=1}^p, [l_{Tr}]_{i=1}^p \rangle$ in order to identify and learn the spatial features. Then, each CT slice was processed using the trained inception module, resulting in the transformation of all training data samples $[S_{Tr}]_{i=1}^p$ into feature vectors $[f_{Tr}]_{i=1}^p$. This generated a new training dataset $\langle [f_{Tr}]_{i=1}^p, [l_{Tr}]_{i=1}^p \rangle$ in the feature domain. The recurrent module was trained to recognize the 3D structural dependencies from the feature-level training data samples in the second stage. The complete two-step loss function of the R3DI-Net can be formulated as:

$$Loss = \begin{cases} \arg \min_{w_{SM}} \mathcal{L}_{SM} \left(\chi_{w_{SM}} \left(\langle [S_{Tr}]_{i=1}^p, [l_{Tr}]_{i=1}^p \rangle \right), \langle [S_{Va}]_{i=1}^q, [l_{Va}]_{i=1}^q \rangle \right) \\ \arg \min_{w_{RM}} \mathcal{L}_{RM} \left(\chi_{w_{RM}} \left(\langle [f_{Tr}]_{i=1}^p, [l_{Tr}]_{i=1}^p \rangle \right), \langle [f_{Va}]_{i=1}^q, [l_{Va}]_{i=1}^q \rangle \right) \end{cases} \quad (2)$$

where χ_{SM} and χ_{RM} represent spatial and recurrent modules as transfer functions, respectively. The symbols $\mathcal{L}_{SM}(\cdot)$ and $\mathcal{L}_{RM}(\cdot)$ signify cross-entropy loss functions [15] for spatial and recurrent modules, respectively. The notations $[S_{Tr}]_{i=1}^p$ and $[S_{Va}]_{i=1}^q$ present training and validation data samples, respectively. The training and validation data samples in feature space are represented as $[f_{Tr}]_{i=1}^p$ and $[f_{Va}]_{i=1}^q$, respectively. The actual class label for training and validation samples are indicated by $[l_{Tr}]_{i=1}^p$ and $[l_{Va}]_{i=1}^q$, respectively.

3. RESULTS AND ANALYSIS

3.1 Dataset and Experimental Setup

A large-scale database was created to evaluate the proposed method by merging three chest CT datasets previously used in a study [12]. The combined database contains 5471 CT samples from 2789 patients. The database was categorized into two groups according to ground-truth labels. The simulations were carried out using the deep-learning toolbox on MATLAB (R2019a) framework on a desktop PC. Hyperparameter initialization is done using the default parametric settings in the deep-learning toolbox. The experimental setup includes five-fold cross-validation, with 70% of the database for training, 10% for validation, and 20% for testing. Finally, the performance of the proposed method and other methods were evaluated based on standard performance metrics, including average accuracy (ACC), F1-score (F1), sensitivity (SEN), and average precision (AP) [15].

3.2 Testing Results (Ablation Studies)

The R3DI-Net leverages both spatial and 3D structural information from a CT scan volume to diagnose whether or not the given case is infected. We initially evaluated the validation performance of the proposed R3DI-Net using various window sizes to determine the optimal slice count.

The window size (i.e., length of consecutive slices) is a crucial factor in the system's performance. A small window size leads to a drop in performance due to the loss of structural information, while a large window size increases processing time without significantly improving performance. Hence, we evaluated the validation performance of our R3DI-Net for 20 different window sizes (ranging from 1 to 20), as depicted in Fig. 2. The red box in Fig. 2 shows the highest validation performance for all performance metrics (ACC, F1, SEN, and AP) for a window size of 15 ($w = 15$).

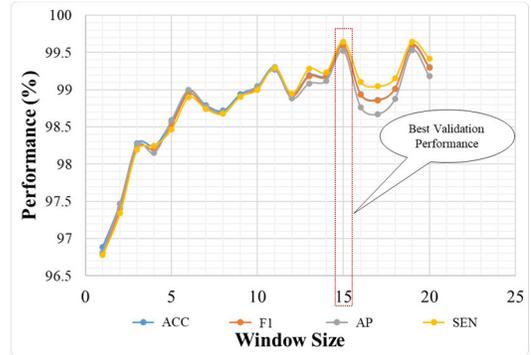


Fig. 2. Validation results of the proposed R3DI-Net with respect to different sizes of windows to find the optimal window size.

Table 1. Quantitative performance comparison of R3DI-Net versus the baseline Inception-Net, with average results emphasized in bold. (Note: “Std: standard deviation” and “unit: %”).

Model	ACC (Std)	F1 (Std)	SEN (Std)	AP (Std)
Inception-Net [13] (slice-based prediction)	94.57 (3.09)	94.41 (3.15)	93.94 (3.8)	94.89 (2.51)
R3DI-Net (slice-based prediction)	95.03 (3.91)	94.86 (4.03)	94.42 (4.62)	95.31 (3.41)
R3DI-Net (scan-based prediction)	98.39 (1.88)	98.36 (1.9)	98.1 (2.36)	98.64 (1.45)

Subsequently, the effectiveness of the model was assessed based on the testing dataset, as highlighted in Table 1. This comparison demonstrates the added benefits of the second-stage recurrent module in improving the results by incorporating structural context across multiple slices. The proposed R3DI-Net (which includes a recurrent module) outperforms the baseline Inception-Net model with an average improvement of 3.82%, 3.95%, 4.16%, and 3.75% in ACC, F1, SEN, and AP, respectively. In addition, the R3DI-Net shows an improvement of 3.36%, 3.5%, 3.68%, and 3.33% in ACC, F1, SEN, and AP, respectively, for scan-based prediction compared to slice-based prediction. A t-test analysis revealed that the R3DI-Net (scan-based prediction) achieved an average p-value of 0.0001 ($p < 0.01$) compared to the R3DI-Net (slice-based prediction) and the Inception-Net (slice-based prediction), respectively. The lower p-values ($p < 0.01$) indicate that the R3DI-Net (scan-based prediction) outperforms them significantly with a 99% confidence level. The quantitative comparison between the proposed R3DI-Net model and the Inception-Net baseline

model for the slice-based prediction is summarized in Table 1. The proposed model still demonstrates improved performance with a slight increase compared to the baseline model.

Table 2. Quantitative performance comparison of R3DI-Net versus the baseline DS-Net and DSS-Net, with average results emphasized in bold. (Note: “Std: standard deviation” and “unit: %”).

Model	ACC (Std)	F1 (Std)	SEN (Std)	AP (Std)
DS-Net [12] (slice-based prediction)	93.31 (2.94)	93.1 (3.02)	92.49 (3.60)	93.72 (2.44)
R3DI-Net (slice-based prediction)	95.03 (3.91)	94.86 (4.03)	94.42 (4.62)	95.31 (3.41)
DSS-Net [12] (scan-based prediction)	96.58 (2.79)	96.53 (2.77)	96.01 (3.54)	97.07 (2.00)
R3DI-Net (scan-based prediction)	98.39 (1.88)	98.36 (1.9)	98.1 (2.36)	98.64 (1.45)

We further compared the performance of our R3DI-Net model to an existing baseline method [12] using the same experimental protocol (as shown in Table 2). The baseline method [12] consists of two networks: the Dilated Shuffle Subnetwork (DS-Net) and the Dilated Shuffle Sequential Network (DSS-Net). DS-Net is designed for slice-based predictions, while DSS-Net is responsible for scan-based predictions. Our R3DI-Net outperformed the DS-Net [12], with average improvement of 1.72%, 1.76%, 1.93%, and 1.59% in ACC, F1, SEN, and AP, respectively, for slice-based prediction. And in terms of scan-based prediction, our final R3DI-Net showed improvement of 1.81%, 1.83%, 2.09%, and 1.57% in ACC, F1, SEN, and AP compared to the DSS-Net [12].

3.3 Comparison

Table 3 presents a comparison of the performance of our R3DI-Net with different state-of-the-art CAD diagnostic methods [16–26]. The R3DI-Net outperforms in terms of qualitative performance parameters comparison to all competitive models (as shown in Table 3). Additionally, the DenseNet201-based method by Jaiswal *et al.* [26] is second best among the other methods [16–25]. However, the R3DI-Net outperforms DenseNet201 [27] convolutional network (used by Jaiswal *et al.* [26]) with average gains of 4.22%, 4.33%, 4.64%, and 4.01% in ACC, F1, SEN, and AP, respectively. A t-test analysis confirms that the R3DI-Net shows significant improvement over the method of Jaiswal *et al.* [26] at a 99% confidence score, with an average p-value of 0.00003 ($p < 0.01$). In another study, Martínez *et al.* [25] used the existing pretrained NASNet [28] convolutional network to diagnose COVID-19 infections using chest CT scans automatically. The NASNet-based method by Martínez *et al.* [25] is ranked third among the other methods [16–24]. However, our R3DI-Net shows significantly higher quantitative results than the method by Martínez *et al.* [25], with an improvement of 4.71%, 4.87%, 5.28%, and 4.45% in ACC, F1, SEN, and AP, respectively.

Table 3. Performance comparison of R3DI-Net with state-of-the-art techniques with average results emphasized in bold (“unit: %”).

Study	ACC	F1	SEN	AP
Brunese <i>et al.</i> [16]	89.66	89.54	87.81	91.43
Farooq <i>et al.</i> [17]	90.30	90.22	88.53	92.17
Minaee <i>et al.</i> [18]	89.84	89.48	89.06	89.91
Khan <i>et al.</i> [19]	91.54	91.33	90.47	92.26
Alsharman <i>et al.</i> [20]	89.73	89.53	88.73	90.4
Misra <i>et al.</i> [21]	92.96	92.76	92.14	93.41
Hu <i>et al.</i> [22]	91.65	91.52	90.44	92.69
Ardakani <i>et al.</i> [23]	90.30	90.26	88.64	92.17
Apostolopoulos <i>et al.</i> [24]	92.95	92.85	91.94	93.81
Martínez <i>et al.</i> [25]	93.68	93.49	92.82	94.19
Jaiswal <i>et al.</i> [26]	94.17	94.03	93.46	94.63
Proposed (R3DI-Net)	98.39	98.36	98.1	98.64

4. DISCUSSION AND CONCLUSION

2D-CNNs typically consider only spatial features from each slice to make a diagnostic decision, disregarding the 3D structural information resulting in performance issues. In contrast, 3D-CNNs consider additional 3D structural features from the entire CT scan for a diagnostic decision. The approaches with 3D-CNNs have a higher number of training parameters that require significant computational resources. This work addresses these challenges with a proposed sequence-based 3D model for accurate analysis of CT volumetric data. The proposed network incorporates multi-scale spatial features (using the inception module) and 3D structural features (using the recurrent module) to achieve state-of-the-art results. This network design employs transfer learning with minimal increase in training parameters and is equipped to handle variable-length sequences for volumetric data analysis. Fig. 3 displays examples of correctly classified (true-positives and true-negatives) and misclassified (false-positives and false-negatives) cases, along with prediction scores and class activation maps (taken from the second-last layer of the inception module). Incorrect predictions may also occur due to the presence of small lesion patterns or insufficient data annotation. In our future work, we will further continue to investigate the effectiveness of transformers in analyzing volumetric data.

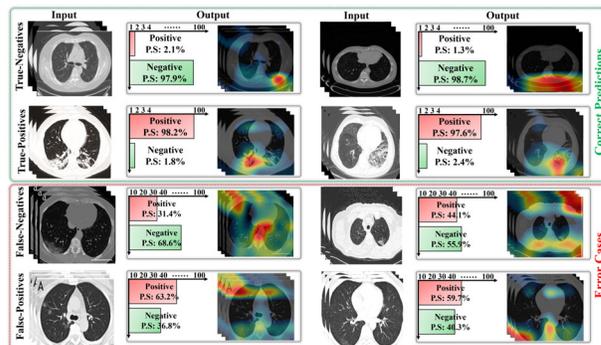


Fig. 3. Visualization of correctly classified (true-positives and true-negatives) and misclassified (false-positives and false-negatives) cases made by the proposed framework (“PS: Prediction score”).

REFERENCES

- [1] A. Castiglione, P. Vijayakumar, M. Nappi, S. Sadiq, and M. Umer, "Covid-19: Automatic detection of the novel coronavirus disease from CT images using an optimized convolutional neural network," *IEEE Trans. Ind. Inform.*, vol. 17, no. 9, pp. 6480–6488, 2021.
- [2] M. Zhang, R. Chu, C. Dong, J. Wei, W. Lu, and N. Xiong, "Residual learning diagnosis detection: An advanced residual learning diagnosis detection system for COVID-19 in industrial internet of things," *IEEE Trans. Ind. Inform.*, vol. 17, no. 9, pp. 6510–6518, 2021.
- [3] T. Lan, Z. Cai, and B. Ye, "A novel spline algorithm applied to COVID-19 computed tomography image reconstruction," *IEEE Trans. Ind. Inform.*, 2022.
- [4] Y. Fang *et al.*, "Sensitivity of chest CT for COVID-19: Comparison to RTPCR," *Radiology*, vol. 296, no. 2, pp. E115–E117, 2020, Art. no. 200432.
- [5] B. Hassan *et al.*, "Joint segmentation and quantification of chorioretinal biomarkers in optical coherence tomography scans: A deep learning approach," *IEEE Trans. Instrum.*, vol. 70, pp.1–17, 2021.
- [6] B. Taha, N. Werghi, and J. Dias, "Automatic polyp detection in endoscopy videos: A survey," In 2017 13th IASTED International Conference on Biomedical Engineering (BioMed), pp. 233–240, 2017.
- [7] S. Javed *et al.*, "Multiplex cellular communities in multi-gigapixel colorectal cancer histology images for tissue phenotyping," *IEEE Trans. Image Process.*, vol. 29, pp. 9204–9219, 2020.
- [8] M. Owais *et al.*, "Light-weighted ensemble network with multilevel activation visualization for robust diagnosis of COVID19 pneumonia from large-scale chest radiographic database," *Appl. Soft Comput.*, vol. 108, 2021, Art. no 107490.
- [9] M.M.N. Abid, T. Zia, M. Ghaffoor, and D. Windridge, "Multi-view convolutional recurrent neural networks for lung cancer nodule identification," *Neurocomputing*, vol. 453, pp. 299–311, 2021.
- [10] Ö. Çiçek *et al.*, "3D U-Net: learning dense volumetric segmentation from sparse annotation," *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 2016, pp. 424–432.
- [11] L. Zhang, "Generalizing deep learning for medical image segmentation to unseen domains via deep stacked transformation," *IEEE Trans. Med. Imaging*, vol. 39, pp. 2531–2540, 2020.
- [12] M. Owais, H. Sultan, N.R. Baek, Y.W. Lee, M. Usman, D.T. Nguyen, G. Batchuluun, and K.R. Park, "Deep 3D Volumetric Model Genesis for Efficient Screening of Lung Infection Using Chest CT Scans," *Mathematics*, vol. 10, no. 21, 2022, Art. no. 4160.
- [13] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2818–2826.
- [14] M. Schuster and K.K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Trans. Signal Process.*, vol. 45, pp. 2673–2681, 1997.
- [15] J. Heaton, "Artificial intelligence for humans," *Neural Netw. Deep Learn.*, vol. 3, 2015.
- [16] L. Brunese, F. Mercaldo, A. Reginelli, and A. Santone, "Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays," *Comput. Meth. Programs Biomed.*, vol. 196, 2020, Art. no. 105608.
- [17] M. Farooq and A. Hafeez, "COVID-ResNet: A deep learning framework for screening of COVID19 from radiographs," 2020, *arXiv:2003.14395*.
- [18] S. Minaee *et al.*, "Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning," *Med. Image Anal.*, vol. 65, 2020, Art. no. 101794.
- [19] I.U. Khan and N. Aslam, "A deep-learning-based framework for automated diagnosis of COVID-19 using X-ray images," *Information*, vol. 11, no. 9, 2020, Art. no. 419.
- [20] N. Alsharman and I. Jawarneh, "GoogleNet CNN neural network towards chest CT-coronavirus medical image classification," *J. Comput. Sci.*, vol. 16, no. 5, pp. 620–625, 2020.
- [21] S. Misra *et al.*, "Multi-channel transfer learning of chest X-ray images for screening of COVID-19," *Electronics*, vol. 9, no. 9, 2020, Art. no. 1388.
- [22] R. Hu, G. Ruan, S. Xiang, M. Huang, Q. Liang, and J. Li, "Automated diagnosis of COVID-19 using deep learning and data augmentation on chest CT," *medRxiv*, 2020.
- [23] A.A. Ardakani *et al.*, "Application of deep learning technique to manage COVID-19 in routine clinical practice using CT images: Results of 10 convolutional neural networks," *Comput. Biol. Med.*, vol. 121, 2020, Art. no. 103795.
- [24] I.D. Apostolopoulos, and T.A. Mpesiana, "Covid-19: Automatic detection from X-ray images utilizing transfer learning with convolutional neural networks," *Australas. Phys. Eng. Sci. Med.*, vol. 43, pp. 635–640, 2020.
- [25] F. Martinez, F. Martinez, and E. Jacinto, "Performance evaluation of the NASNet convolutional network in the automatic identification of COVID19," *Int. J. Adv. Sci. Eng. Inf. Techn.*, vol. 10, no. 2, pp. 662–667, 2020, Art. no. 662.
- [26] A. Jaiswal *et al.*, "Classification of the COVID-19 infected patients using densenet201 based deep transfer learning," *J. Biomol. Struct. Dyn.*, vol. 39, no. 15, pp. 5682–5689, 2021.
- [27] G. Huang, Z. Liu, L. van der Maaten, and K.Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4700–4708.
- [28] B. Zoph, V. Vasudevan, J. Shlens, and Q.V. Le, "Learning transferable architectures for scalable image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2018, pp. 8697–8710.