Article

COVLIAS 2.0-cXAI: Cloud-Based Explainable Deep Learning System for COVID-19 Lesion Localization in Computed Tomography Scans

Jasjit S. Suri [1,2,*], Sushant Agarwal [2,3], Gian Luca Chabert [4], Alessandro Carriero [5], Alessio Paschè [4], Pietro S. C. Danna [4], Luca Saba [4], Armin Mehmedović [6], Gavino Faa [7], Inder M. Singh [1], Monika Turk [8], Paramjit S. Chadha [1], Amer M. Johri [9], Narendra N. Khanna [10], Sophie Mavrogeni [11], John R. Laird [12], Gyan Pareek [13], Martin Miner [14], David W. Sobel [13], Antonella Balestrieri [4], Petros P. Sfikakis [15], George Tsoulfas [16], Athanasios D. Protogerou [17], Durga Prasanna Misra [18], Vikas Agarwal [18], George D. Kitas [19,20], Jagjit S. Teji [21], Mustafa Al-Maini [22], Surinder K. Dhanjil [23], Andrew Nicolaides [24], Aditya Sharma [25], Vijay Rathore [23], Mostafa Fatemi [26], Azra Alizad [27], Pudukode R. Krishnan [28], Ferenc Nagy [29], Zoltan Ruzsa [30], Mostafa M. Fouda [31], Subbaram Naidu [32], Klaudija Viskovic [6] and Mannudeep K. Kalra [33]

1   Stroke Diagnostic and Monitoring Division, AtheroPoint™, Roseville, CA 95661, USA; drindersingh1@gmail.com (I.M.S.); pomchadha@gmail.com (P.S.C.)
2   Advanced Knowledge Engineering Centre, GBTI, Roseville, CA 95661, USA; sushant.ag09@gmail.com
3   Department of Computer Science Engineering, PSIT, Kanpur 209305, India
4   Department of Radiology, Azienda Ospedaliero Universitaria (A.O.U.), 09123 Cagliari, Italy; gianchab@yahoo.com (G.L.C.); pascheale@gmail.com (A.P.); psc.dnn@gmail.com (P.S.C.D.); lucasabamd@gmail.com (L.S.); antonellabalestrieri@hotmail.com (A.B.)
5   Department of Radiology, "Maggiore della Carità" Hospital, University of Piemonte Orientale (UPO), Via Solaroli 17, 28100 Novara, Italy; profcarriero@virgilio.it
6   Department of Radiology, University Hospital for Infectious Diseases, 10000 Zagreb, Croatia; mehmedovic.armin302@gmail.com (A.M.); klaudija.viskovic@bfm.hr (K.V.)
7   Department of Pathology, Azienda Ospedaliero Universitaria (A.O.U.), 09124 Cagliari, Italy; gavinofaa@gmail.com
8   The Hanse-Wissenschaftskolleg Institute for Advanced Study, 27753 Delmenhorst, Germany; monika.turk84@gmail.com
9   Department of Medicine, Division of Cardiology, Queen's University, Kingston, ON K7L 3N6, Canada; johria@queensu.ca
10 Department of Cardiology, Indraprastha APOLLO Hospitals, New Delhi 110076, India; drnnkhanna@gmail.com
11 Cardiology Clinic, Onassis Cardiac Surgery Center, 17674 Athens, Greece; soma13@otenet.gr
12 Heart and Vascular Institute, Adventist Health St. Helena, St. Helena, CA 94574, USA; lairdjr@ah.org
13 Minimally Invasive Urology Institute, Brown University, Providence, RI 02912, USA; gyan_pareek@brown.edu (G.P.); dwsobel@gmail.com (D.W.S.)
14 Men's Health Center, Miriam Hospital, Providence, RI 02912, USA; martin_miner@brown.edu
15 Rheumatology Unit, National Kapodistrian University of Athens, 17674 Athens, Greece; psfikakis@med.uoa.gr
16 Department of Surgery, Aristoteleion University of Thessaloniki, 54124 Thessaloniki, Greece; tsoulfasg@gmail.com
17 Cardiovascular Prevention and Research Unit, Department of Pathophysiology, National & Kapodistrian University of Athens, 15772 Athens, Greece; aprotog@med.uoa.gr
18 Department of Immunology, SGPIMS, Lucknow 226014, India; durgapmisra@gmail.com (D.P.M.); vikasagr@yahoo.com (V.A.)
19 Academic Affairs, Dudley Group NHS Foundation Trust, Dudley DY1 2HQ, UK; george.kitas@nhs.net
20 Arthritis Research UK Epidemiology Unit, Manchester University, Manchester M13 9PL, UK
21 Ann and Robert H. Lurie Children's Hospital of Chicago, Chicago, IL 60611, USA; jsteji1@comcast.net
22 Allergy, Clinical Immunology and Rheumatology Institute, Toronto, ON M5G 1N8, Canada; almaini@hotmail.com
23 AtheroPoint LLC., Roseville, CA 95661, USA; surinderdhanjil@gmail.com (S.K.D.); rajvivs888@gmail.com (V.R.)
24 Vascular Screening and Diagnostic Centre, University of Nicosia Medical School, Engomi 2408, Cyprus; anicolaides1@gmail.com

25 Division of Cardiovascular Medicine, University of Virginia, Charlottesville, VA 22902, USA; as8ah@hscmail.mcc.virginia.edu

26 Department of Physiology & Biomedical Engineering, Mayo Clinic College of Medicine and Science, Rochester, MN 55905, USA; fatemi.mostafa@mayo.edu

27 Department of Radiology, Mayo Clinic College of Medicine and Science, Rochester, MN 55905, USA; alizad.azra@mayo.edu

28 Neurology Department, Fortis Hospital, Bengaluru 560076, India; prkrish12@rediffmail.com

29 Internal Medicine Department, University of Szeged, 6725 Szeged, Hungary; drnagytfer@hotmail.com

30 Invasive Cardiology Division, University of Szeged, 1122 Budapest, Hungary; zruzsa@icloud.com

31 Department of ECE, Idaho State University, Pocatello, ID 83209, USA; mfouda@isu.edu

32 Electrical Engineering Department, University of Minnesota, Duluth, MN 55812, USA; dsnaidu@d.umn.edu

33 Department of Radiology, Massachusetts General Hospital, 55 Fruit Street, Boston, MA 02114, USA; mkalra@mgh.harvard.edu

* Correspondence: jasjit.suri@atheropoint.com; Tel.: +1-(916)-749-5628

Abstract: Background: The previous COVID-19 lung diagnosis system lacks both scientific validation and the role of explainable artificial intelligence (AI) for understanding lesion localization. This study presents a cloud-based explainable AI, the "COVLIAS 2.0-cXAI" system using four kinds of class activation maps (CAM) models. Methodology: Our cohort consisted of ~6000 CT slices from two sources (Croatia, 80 COVID-19 patients and Italy, 15 control patients). COVLIAS 2.0-cXAI design consisted of three stages: (i) automated lung segmentation using hybrid deep learning ResNet-UNet model by automatic adjustment of Hounsfield units, hyperparameter optimization, and parallel and distributed training, (ii) classification using three kinds of DenseNet (DN) models (DN-121, DN-169, DN-201), and (iii) validation using four kinds of CAM visualization techniques: gradient-weighted class activation mapping (Grad-CAM), Grad-CAM++, score-weighted CAM (Score-CAM), and FasterScore-CAM. The COVLIAS 2.0-cXAI was validated by three trained senior radiologists for its stability and reliability. The Friedman test was also performed on the scores of the three radiologists. Results: The ResNet-UNet segmentation model resulted in dice similarity of 0.96, Jaccard index of 0.93, a correlation coefficient of 0.99, with a figure-of-merit of 95.99%, while the classifier accuracies for the three DN nets (DN-121, DN-169, and DN-201) were 98%, 98%, and 99% with a loss of ~0.003, ~0.0025, and ~0.002 using 50 epochs, respectively. The mean AUC for all three DN models was 0.99 (p < 0.0001). The COVLIAS 2.0-cXAI showed 80% scans for mean alignment index (MAI) between heatmaps and gold standard, a score of four out of five, establishing the system for clinical settings. Conclusions: The COVLIAS 2.0-cXAI successfully showed a cloud-based explainable AI system for lesion localization in lung CT scans.

Keywords: COVID-19 lesion; lung CT; Hounsfield units; glass ground opacities; hybrid deep learning; explainable AI; segmentation; classification; GRAD-CAM; Grad-CAM++; Score-CAM; FasterScore-CAM

1. Introduction

COVID-19, the novel coronavirus or SARS-CoV-2, the severe acute respiratory syndrome coronavirus 2, has been a rapidly spreading epidemic that was declared a global pandemic on 11 March 2020 by the World Health Organization (WHO) [1]. As of 20 May 2022, COVID-19 had infected over 521 million people worldwide and has killed nearly 6.2 million [2].

Molecular pathways [3] and imaging [4] of COVID-19 have proven to be worse in individuals with comorbidities such as coronary artery disease [5,6], diabetes [7], atherosclerosis [8], fetal programming [9], pulmonary embolism [10], and stroke [11]. Further, the evidence shows the damage to the aorta's vasa vasorum, leading to thrombosis and plaque vulnerability [12]. COVID-19 can cause severe lung damage, with abnormalities primarily in the lower region of the lung lobes [13–20]. It is challenging to distinguish COVID-19 pneumonia from interstitial pneumonia or other lung illnesses; as a result, manual classification can be skewed based on radiological expert opinion. As a result, an automated computer-aided diagnostics (CAD) system is sorely needed to categorize and characterize

the condition [21], as it delivers excellent performance due to minimal inter-and intra-observer variability.

With the advancements of artificial intelligence (AI) technology [22–24], machine learning (ML) and deep learning (DL) approaches have become increasingly popular for detection of pneumonia and its categorization. There have been several innovations in ML and DL frameworks, some of which are applied to lung parenchyma segmentation [25–27], pneumonia classification [21,25,28], symptomatic vs. asymptomatic carotid plaque classification [29–33], coronary disease risk stratification [34], cardiovascular/stroke risk stratification [35], classification of Wilson disease vs. controls [36], classification of eye diseases [37], and cancer classification in thyroid [38], liver [39], ovaries [40], prostate [41], and skin [42–44].

AI can further help in the detection of pneumonia type and can overcome the shortage of specialist personnel by assisting in investigating CT scans [45,46]. One of the key benefits of AI is its ability to emulate manually developed processes. Thus, AI speeds up the process of identifying and diagnosing diseases. On the contrary, the black-box nature of AI offers resistance to usage in clinicians' settings. Thus, there is a clear need for human readability and interpretability of deep networks, which requires identified lesions to be interpreted and quantified. We, therefore, developed an explainable AI system in a cloud framework, labeled the "COVLIAS 2.0-cXAI" system, which was our primary novelty [47–52]. The COVLIAS 2.0-cXAI design consisted of three stages (Figure 1): (i) automated lung segmentation using the hybrid deep learning ResNet-UNet model using automatic adjustment of Hounsfield units [53], hyperparameter optimization [54], and the parallel and distributed nature of design during training; (ii) classification using three kinds of DenseNet (DN) models (DN-121, DN-169, DN-201) [55–58]; and (iii) scientific validation using four kinds of class activation mapping (CAM) visualization techniques: gradient-weighted class activation mapping (Grad-CAM) [59–63], Grad-CAM++ [64–67], score-weighted CAM (Score-CAM) [68–70], and FasterScore-CAM [71,72]. The COVLIAS 2.0-cXAI was validated by a trained senior radiologist for its stability and reliability. The proposed study also considers different variations in COVID-19 lesions, such as ground-glass opacity (GGO), consolidation, and crazy paving [73–82]. The COVLIAS 2.0-cXAI design showed the reduction of model size by roughly 30% and an improvement of the online version of the AI system by two times.
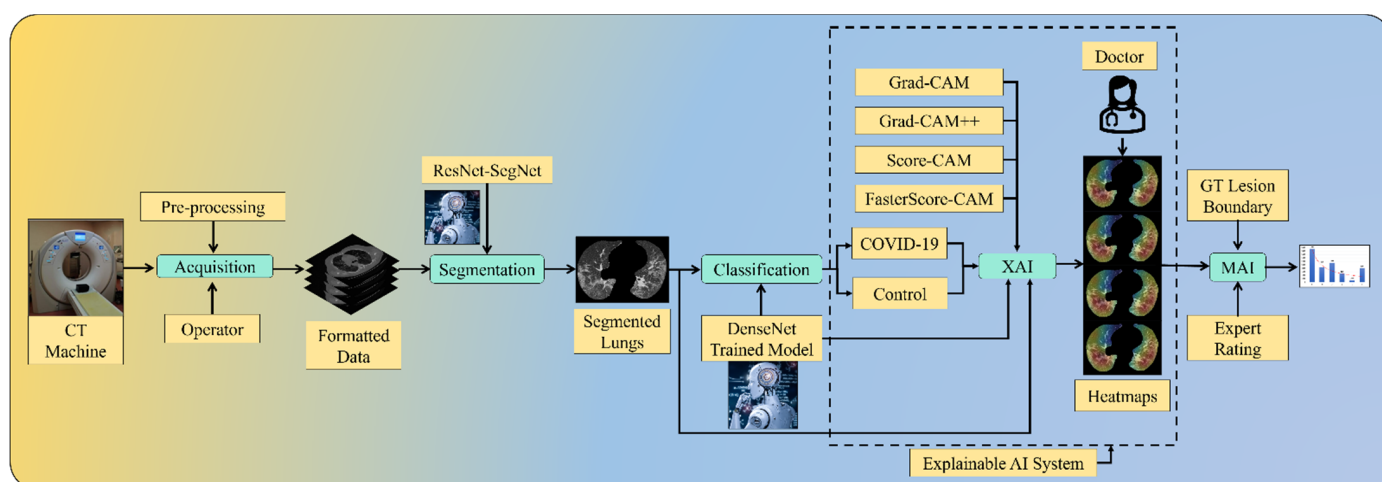

Figure 1. COVLIAS 2.0-cXAI system.

To summarize, our prime contributions in the proposed study consist of six main stages: (i) automated lung segmentation using the HDL-ResNet-UNet model; (ii) classification of COVID-19 vs. controls using three kinds of DenseNets such as DenseNet-121 [55–57,83], DenseNet-169, and DenseNet-201; the combination of segmentation and classification depicting the overall performance of the system; (iii) using explainable AI to visualize and validate the prediction of the DenseNet models using four kinds of CAM,

namely Grad-CAM, Grad-CAM++, Score-CAM, and FasterScore-CAM, for the first time. This helps us understand the AI model's learning in the input CT image [35,84–86]. (iv) Mean alignment index (MAI) between heatmaps and the gold standard score from three trained senior radiologists, a score of four out of five, establishing the system for clinical applicability. Further, a Friedman statistical test was also conducted to present the statistical significance of the scores from the three experts. (v) Application of the quantization for the trained AI model to make the system light and further ensure faster online prediction. Lastly, (vi) presents an end-to-end cloud-based CT image analysis system, including the CT lung segmentation and COVID-19 intensity map using the four CAM techniques (Figure 1).

Our study is divided into six sections. The methodology, patient demographics, image acquisition, description of the DenseNet models, and the explainable AI system used in this work are described in Section 2. Section 3 presents the background literature. In Section 4, the models' findings and their performance evaluation are presented. The discussion and benchmarking sections are in Section 4, and Section 5 presents the conclusions.

2. Methodology

2.1. Patient Demographics

Two distinct cohorts representing two different countries (Croatia and Italy) were used in the proposed study. The experimental data set included 20 Croatian COVID-19-positive individuals, 17 of whom were male, and the remainder of whom were three females. The GGO, consolidation, and crazy paving had an average value of 4. The second data set included 15 Italian control subjects, ten of whom were male, and the remainder of whom were five females. To confirm the presence of COVID-19 in the selected cohort, an RT-PCR test [87–89] was performed for both data sets.

2.2. Image Acquisition and Data Preparation

2.2.1. Croatian Data Set

A Croatian data set of 20 COVID-19-positive patients was employed in our investigation (Figure 2). This cohort was acquired between 1 March and 31 December 2020, at the University Hospital for Infectious Diseases (UHID) in Zagreb, Croatia. The patients who underwent thoracic MDCT during their hospital stay showed a positive RT-PCR test for COVID-19 and were also above the age of 18 years. These patients also had hypoxia (oxygen saturation 92%), tachypnea (respiratory rate 22 per minute), tachycardia (pulse rate > 100), and hypotension (systolic blood pressure 100 mmHg). The proposal was approved by the UHID Ethics Committee. The acquisition of the CT data was conducted using a 64-detector FCT Speedia HD scanner (Fujifilm Corporation, Tokyo, Japan, 2017).
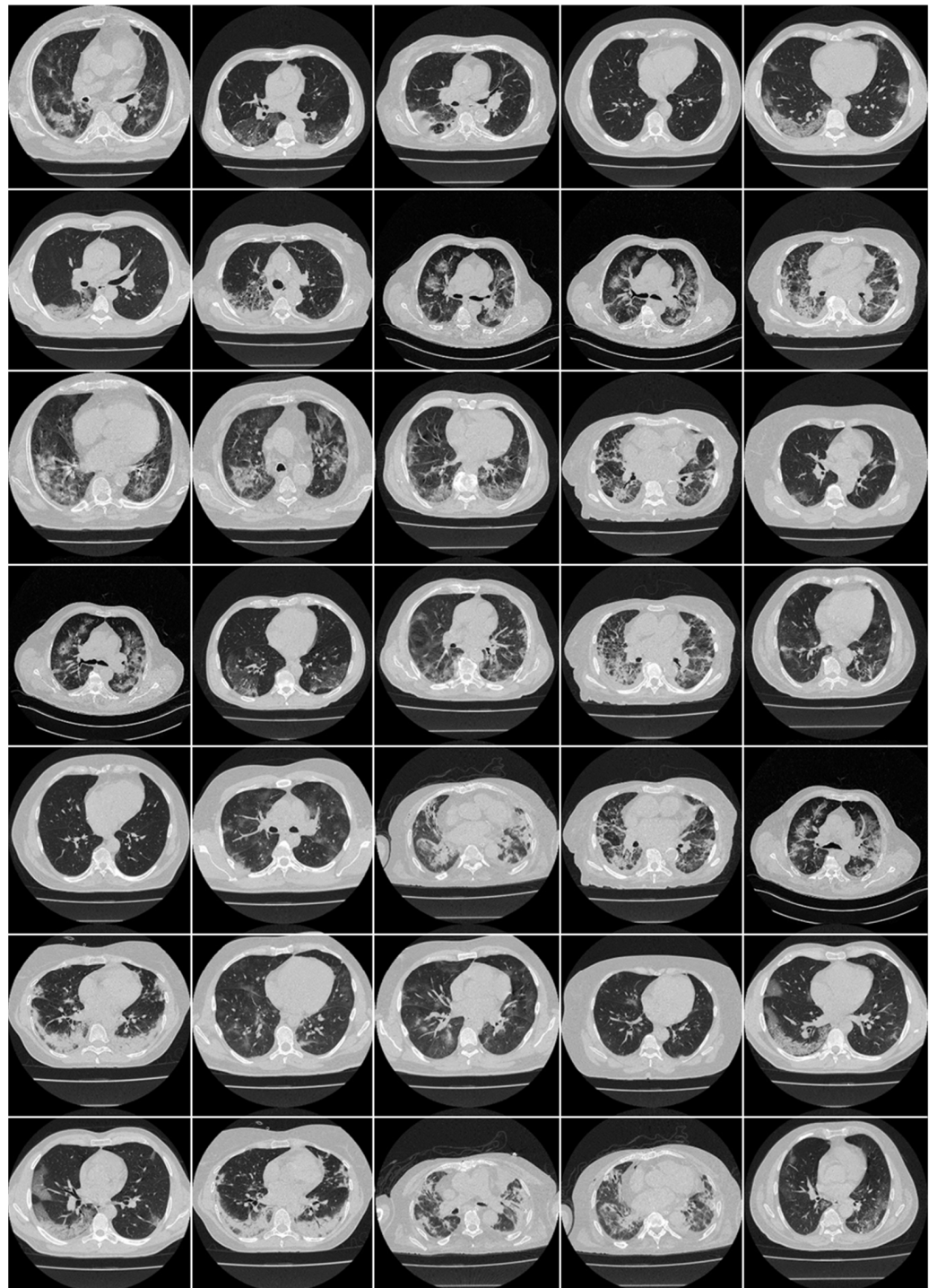
Figure 2. Raw CT slice of COVID-19 patients taken from Croatian data set.

2.2.2. Italian Data Set

The CT scans for the Italian cohort of 15 patients (Figure 3) were acquired using a 128-slice multidetector-row CT scanner (Philips Ingenuity Core, by Philips Healthcare). The breath-hold procedure was used during acquisition and no contrast agent was administered. To acquire a 1 mm thick slice, a lung kernel of a 768×768 matrix together with a soft-tissue kernel was utilized. The CT scans were carried out with a 120 kV, 226 mAs/slice detector configuration (using Philips' automated tube current modulation—Z-DOM), a spiral pitch factor of 1.08, and a 0.5 s gantry rotation time 64×0.625 detector was considered.

Figure 3. Raw control CT slice taken from Italian data set.

2.3. Artificial Intelligence Architecture

Recent deep learning developments, such as hybrid deep learning (HDL), have yielded encouraging results [26,27,90–95]. We hypothesize that HDL models are superior to SDL models (e.g., UNet [96] and SegNet [97]) due to the joint effect of the two DL models. As a result, we offer a hybrid DL (HDL) such as the ResNet-UNet model that has been trained and tested for the COVID-19-based lung segmentation database in our current study. The aim of the proposed study is directed mainly at the explainable AI (XAI) using the classification models; therefore, we have only used one HDL model.

2.3.1. ResNet-UNet Architecture

VGGNet [98–100] was highly efficient and speedy, but it had a problem with vanishing gradients. During backpropagation, it results in substantially minimal or no weight training because it is multiplied by the gradient at each epoch, and the update is very modest in the initial layers. The residual network, or ResNet [101], was created to solve this problem. Skip connections, a new link, were built into this architecture, allowing gradients to skip a specific set of layers, thus overcoming the problem of vanishing gradient. Furthermore, during the backpropagation step, the local gradient value was preserved by an identity function network. In a ResNet-UNet-based segmentation network, the encoding part of the base UNet network is substituted with ResNet architecture, thus proving a hybrid approach.

2.3.2. Dense Convolutional Network Architecture

A dense convolutional network (CNN) has an architecture that uses shorter connections across layers, thereby making them highly efficient during training [102]. DenseNet is a CNN where every layer is connected to the ones below it. The primary layer communicates with the 2nd, 3rd, 4th, and so on, whereas the secondary layer communicates with the 3rd, 4th, 5th, and so on. The key idea here was to increase the flow of information between the network layers.

To maintain the flow of the system, the input received by each layer is forwarded to all the further layers in a feature map. Unlike ResNet, it does not combine features by summarizing them; instead, it concatenates them. As a result, the "jth" layer contains J inputs and comprises feature maps from all the convolutional blocks from the subsequent "J − j" layers that receive their feature maps. Instead of only J connections, the network now has "(J(J + 1))/2" links, like standard deep learning designs. This requires fewer parameters than traditional CNN, avoiding meaningless feature maps to be learned. This paper presents three kinds of DenseNet architectures, namely, (i) DenseNet-121 (Figure 4a), (ii) DenseNet-169 (Figure 4b), and (iii) DenseNet-201 (Figure 4c). Table 1 presents the output feature map sizes of the input layer, convolution layer, dense blocks, transition layers, and fully connected layer followed by the SoftMax classification layer.

Table 1. Output feature map sizes of the three DenseNet architectures.

Layers	Output Feature Size
Input	512×512
Conv.	256×256
Max Pool	128×128
Dense Block 1	128×128
Transition Layer 1	128×128
	64×64
Dense Block 2	64×64
Transition Layer 2	64×64
	32×32
Dense Block 3	32×32
Transition Layer 3	32×32
	16×16
Dense Block 4	16×16
Classification Layer (SoftMax)	1024
	2
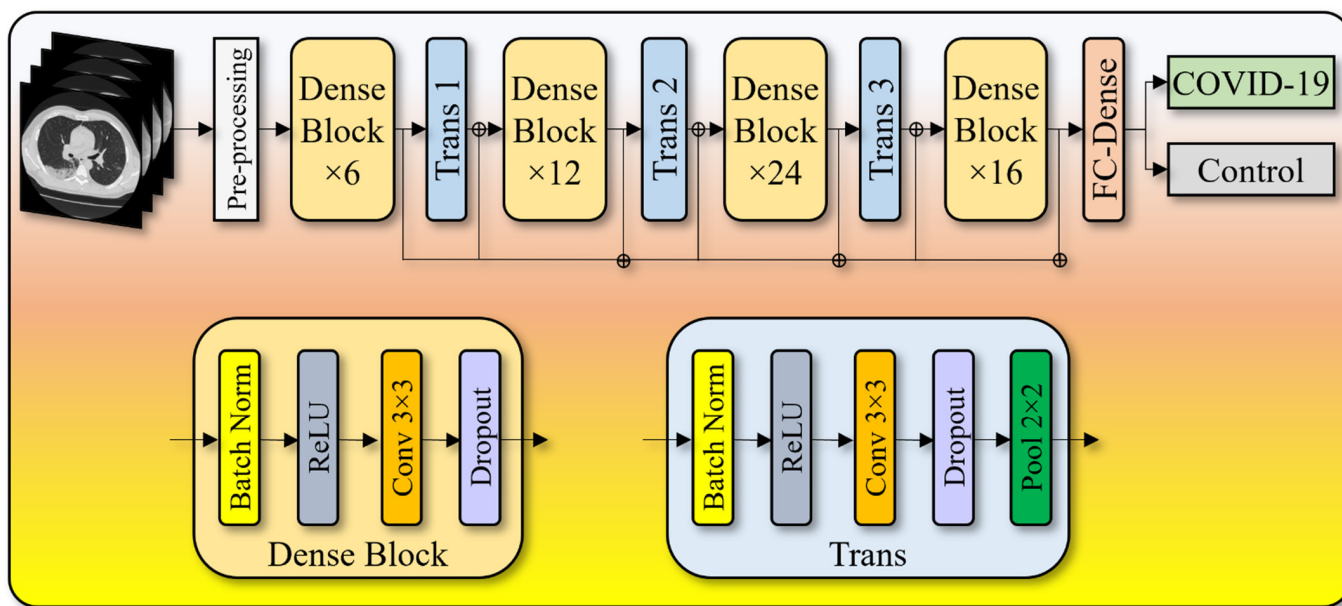
(**a**)
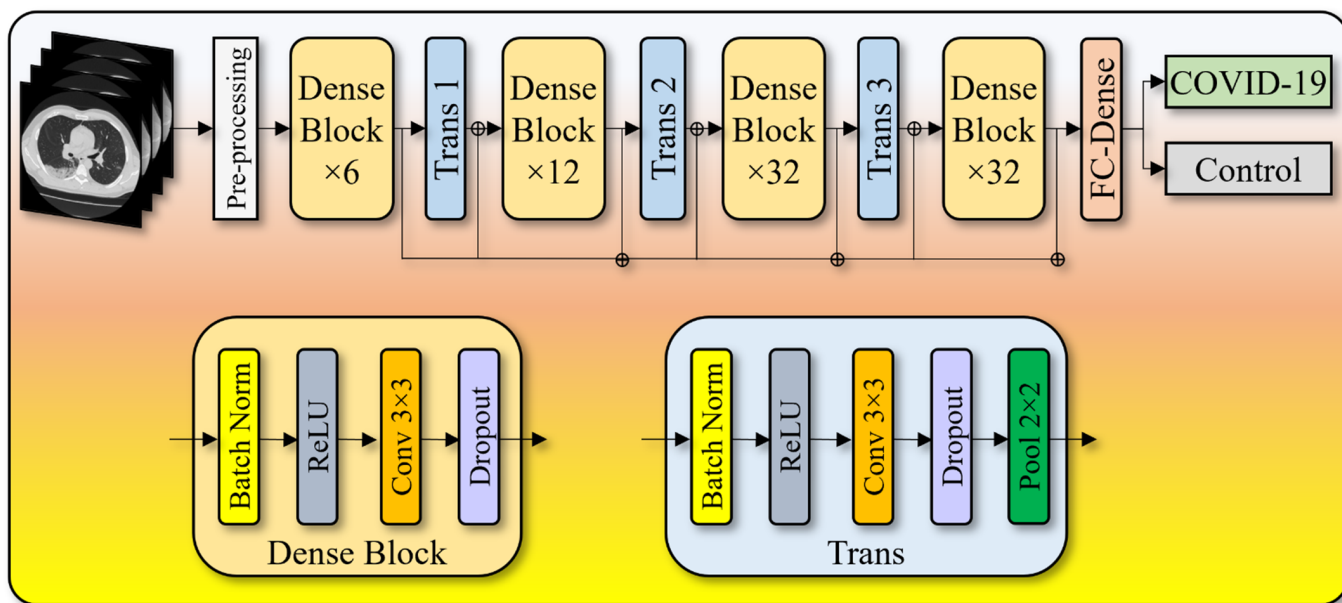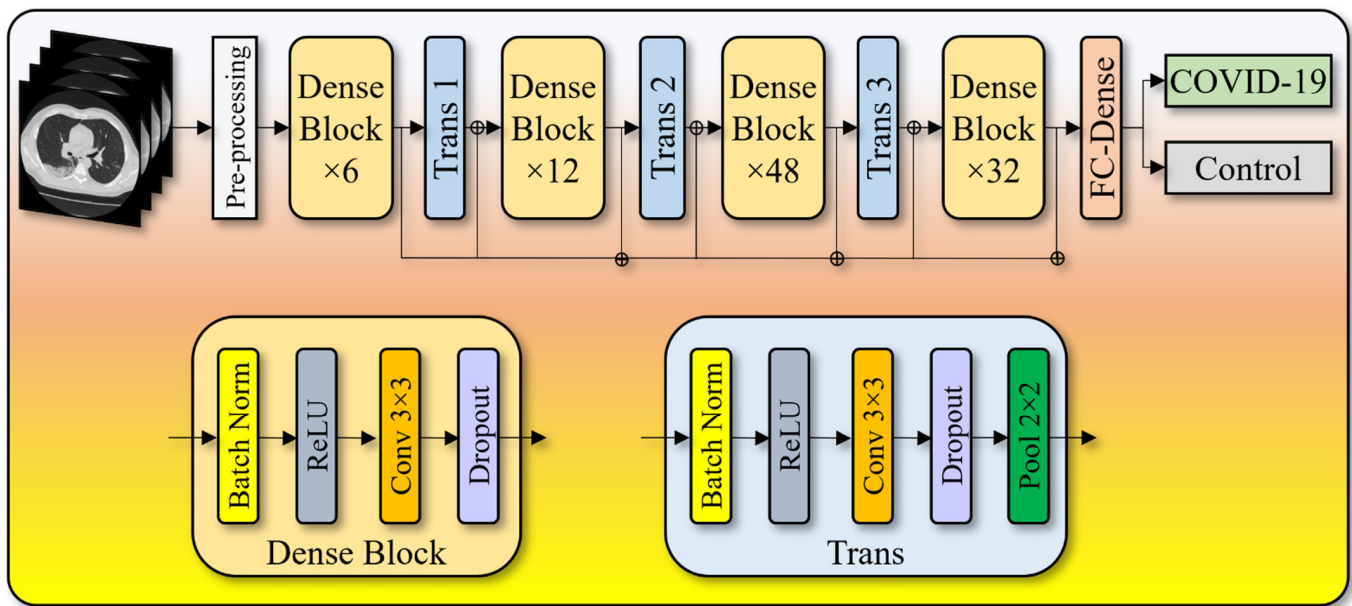

(**b**)

Figure 4. *Cont.*

(**c**)

Figure 4. (**a**) DenseNet-121 model. (**b**) DenseNet-169 model. (**c**) DenseNet-201 model.

2.4. Explainable Artificial Intelligence System for COVID-19 Lesion

We are utilizing machine learning to address more complicated problems as the technology improves and models become more accurate. As machine learning (ML) technology advances, it becomes increasingly sophisticated. This is one of the reasons to use cloud-based explainable AI (cXAI) to help understand how the ML model predicts utilizing a set of tools.

Instead of presenting individual pixels, cXAI is a new approach to displaying attributes that highlight which prominent characteristics of an image had the most significant impact on the model. The effect is seen here (image with heatmap red-yellow-blue), along with which regions contributed to our model's identification of this image as a husky. Based on the color palette, cXAI highlights the most influential areas in red, the medium influential part in yellow, and the least influential factors in blue. Understanding why a model produced the forecast it did is helpful when debugging a model's incorrect categorization or determining whether to believe its prediction. Explainability can help (i) debug the AI model, (ii) validate the results, and (iii) provide a visual explanation as to what drove the AI model to classify the image in a certain way. As part of cXAI, we present four cloud-based CAM techniques to visualize the prediction of the AI model and validate it using the color palette as described above.

Four CAM Techniques in Cloud-Based Explainable Artificial Intelligence System

Grad-CAM (Figure 5) generates a localization map that shows the critical places in the image representing the lesions by employing gradients from the target label/class settling into the final convolutional layer. The input image is fed to the model which is then transformed by the Grad-CAM heatmap (Equation (1)) to show the explainable lesions in the COVID-19 CT scans. This image then follows the typical prediction cycle, generating class probability scores before calculating the model loss. Following that, using the output from our desired model layer, we compute the gradient in terms of model loss. Finally, the gradient areas that contribute to the prediction are then preprocessed (Equation (3)), thereby overlaying the heatmap on the original grayscale scans.
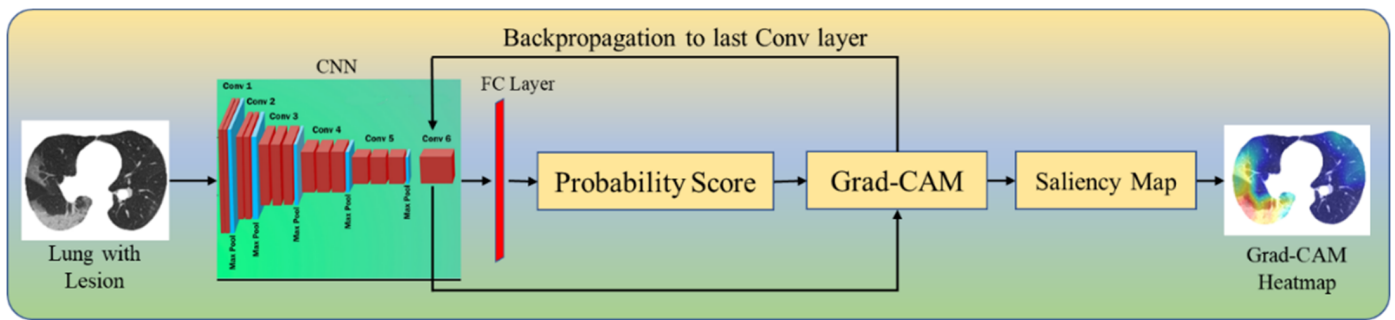
Figure 5. Grad-CAM.

Grad-CAM++ (Figure 6) is an improved version of Grad-CAM, providing a better understanding by creating an accurate localization map of the identifying object and explaining the same class objects having multiple occurrences. Grad-CAM++ generates a pictorial depiction for the class label as weights derived from the feature map of the CNN layer by considering its positive partial derivatives (Equation (2)). Then, a similar process is followed as in Grad-CAM to produce the gradient's saliency map (Equation (3)) that contributes to the prediction. This map is then overlaid with the original image.

$$w_k^c = \frac{1}{Z} \sum_i \sum_j \left(\frac{\partial Y^c}{\partial A_{ij}^k} \right) \tag{1}$$

$$w_k^c = \sum_i \sum_j a_{ij}^{kc} . relu \left(\frac{\partial Y^c}{\partial A_{ij}^k} \right) \tag{2}$$

$$where, \ Y^c = \sum_k w_k^c . \sum_i \sum_j A_{ij}^k$$

$$L_{ij}^c = \sum_k w_k^c . A_{ij}^k \tag{3}$$

where Y^c represents the final score of class c and A^k represents the global average pool of the last convolutional layer by considering its linear combination. Estimated weights for the last convolutional layer can be given by w_k^c for class c. L_{ij}^c represents a class-specific saliency map for each spatial location (i, j).
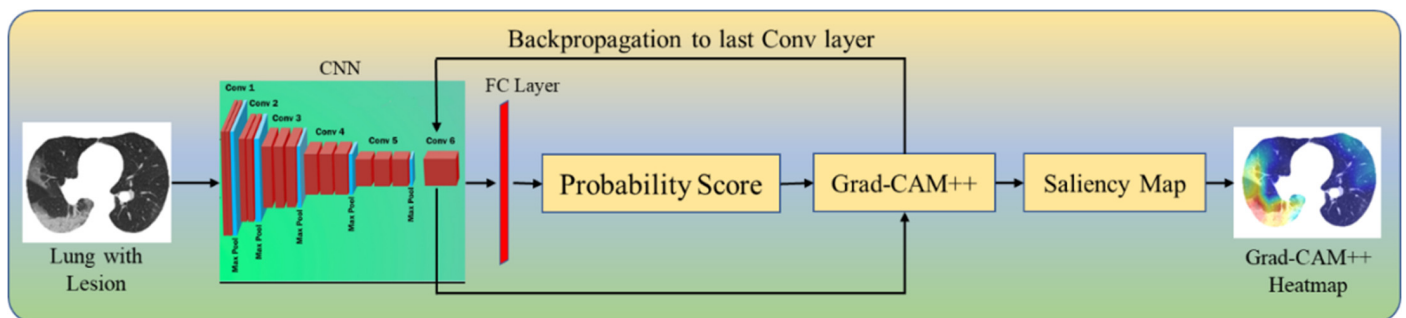

Figure 6. Grad-CAM++.

Our third CAM technique is Score-CAM (Figure 7). In this technique, the produced activation mask is used as a mask for the input image, masking sections of the image and causing the model to forecast on the partially masked image. The target class's score is then used to represent the activation map's importance. The main difference between Grad-CAM and Score-CAM is that this technique does not incorporate the use of gradients, as the propagated gradients introduce noise and are unstable. The technique is separated into the following parts to obtain the class discriminative saliency map using Score-CAM.

(i) Images are processed through the CNN model as a forward pass. The activations are taken from the network's last convolutional layer after the forward pass. (ii) Each activation map with the shape 1xmxn produced from the previous layer is sampled to the same size as the input image using bilinear interpolation. (iii) The generated activation maps are normalized with each pixel within [0, 1] to maintain the relative intensities between the pixels after upsampling. The formula given in Equation (4) is used for the normalization of the data. (iv) After the activation maps have been normalized, the highlighted areas are projected onto the input space by multiplying each normalized activation map ($1 \times X \times Y$) with the original input image ($3 \times X \times Y$) to obtain a masked image M with the shape $3 \times X \times Y$ (Equation (5)). The resulting masked images M are then fed into a CNN with SoftMax output (Equation (6)). (v) Finally, pixel-wise ReLU (Equation (7)) is applied to the final activation map generated using the sum of all the activation maps for the linear combination of the target class score and each activation map.

$$A_{i,j}^{k} = \frac{A_{i,j}^{k}}{\max A^{K} - \min A^{K}} \tag{4}$$

$$M^{k} = A^{k} \cdot I \tag{5}$$

$$S_{k} = Softmax\left(F(M^{k})\right) \tag{6}$$

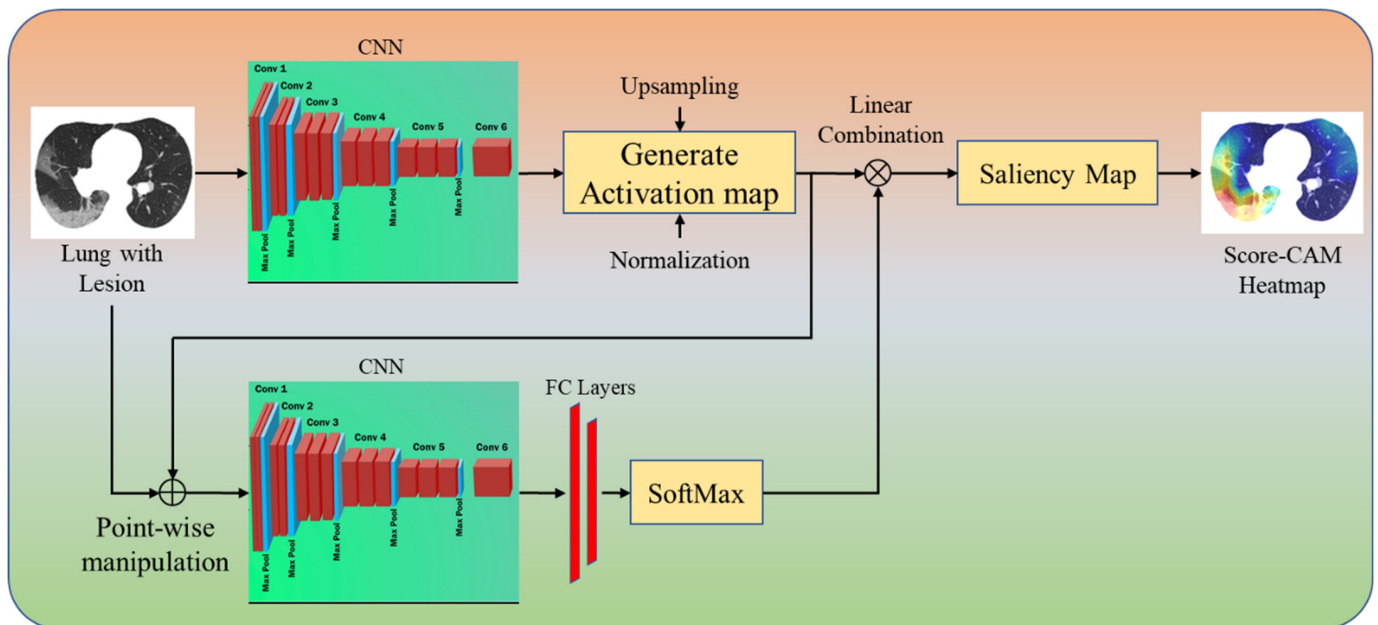$$L^{c} = ReLU\left(\sum_{k} w_{k}^{c} \cdot A^{k}\right) \tag{7}$$

Figure 7. Score-CAM++.

Finally, the fourth technique is labeled FasterScore-CAM. The main innovation of using FasterScore-CAM over the traditional Score-CAM technique is that it eliminates the channels with small variance and only utilizes the activation maps with large variance for heatmap computation and visualization. This selection of activation maps with large variance helps improve the overall speed by nearly ten-fold compared to Score-CAM.

2.5. Loss Function for Artificial-Intelligence-Based Models

During model generation, our system uses the cross-entropy (CE)-loss [103–105] function. If CE-loss can be represented by the notation α_{CE}, probability of the AI model

by p_i, gold standard label 1 and 0 by g_i and $(1 - g_i)$, respectively, then the loss function equation can be mathematically expressed as shown in Equation (8).

$$\alpha_{CE} = -[(g_i \times \log p_i) + (1 - g_i) \times \log(1 - p_i)] \tag{8}$$

2.6. Experimental Protocol

Our team has demonstrated several cross-validation (CV) protocols using the AI framework; the study uses a standardized five-fold CV technique to train the AI models [106,107]. The data consisted of 80% training data and 20% testing data. K5 CV protocol was adapted where the data were partitioned into five parts, each consisting of a unique training set and testing set and rotated cyclically for all the parts that were used independently. Note that we also used 10% of the data for validation.

The accuracy of the AI system is computed by evaluating the predicted output to the ground-truth label. The output lung mask was just black or white; these measurements were interpreted as binary (1 for white or 0 for black) values. If the symbols TP, TN, FN, and FP represent true positive, true negative, false negative, and false positive, respectively, Equation (9) may be used to evaluate the accuracy of the AI system.

$$\text{Accuracy (\%)} = \left(\frac{TP + TN}{TP + FN + TN + FP} \right) \times 100 \tag{9}$$

Precision (Equation (10)) of an AI model is given as the ratio of the correctly labeled classes by the model w.r.t total labels of the COVID-19 class including the false-positive cases. Recall (Equation (11)) of an AI model is given as the ratio of the correctly labeled COVID-19 positive class by the AI model to the total COVID-19 in the data set. F1-score (Equation (12)) is the harmonic average of the precision and recall for the given AI model [108–110].

$$\text{Precision} = \left(\frac{TP}{TP + FP} \right) \tag{10}$$

$$\text{Recall} = \left(\frac{TP}{TP + FN} \right) \tag{11}$$

$$\text{F1} - \text{Score} = 2 \times \left(\frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \right) \tag{12}$$

3. Results and Performance Evaluation

The proposed study uses the ResNet-UNet model for lung CT segmentation (see Appendix A, Figure A1) and three DenseNet models, namely, DenseNet-121, DenseNet-169, and DenseNet-201 to classify COVID-19 vs. control. The AI classification model was trained on 1400 COVID-19 and 1050 control images, giving an accuracy of 98.21% with an AUC of 0.99 ($p < 0.0001$).

A confusion matrix (CM) is a table that shows how well a classification model performs on a set of test data for which the real values are known. Table 2 presents CM for three kinds of DenseNet (DN) models (DN-121, DN-169, and DN-201). For DN-121, a total of 1382 and 1020 images were correctly classified and 18 and 30 were misclassified as COVID-19 and control. For DN-169, a total of 1386 and 1028 images were correctly classified and 14 and 22 were misclassified as COVID-19 and control. For DN-201, a total of 1388 and 1038 images were correctly classified and 12 and 12 were misclassified as COVID-19 and control.

Table 2. Confusion matrix.

DN-121	COVID	Control
COVID	99% (1382)	3% (30)
Control	1% (18)	97% (1020)
DN-169	COVID	Control
COVID	99% (1386)	2% (22)
Control	1% (14)	98% (1028)
DN-201	COVID	Control
COVID	99% (1388)	1% (12)
Control	1% (12)	99% (1038)

3.1. Results Using Explainable Artificial Intelligence

Visual Results Representing Lesion Using the Four CAM Techniques

The trained classification model from DenseNet-121, DenseNet-169, and DenseNet-201 was taken, and then cXAI was applied to it to generate the heatmap representing the lesion, thereby validating the prediction of the DenseNet models. These images which were used to train the classification models followed the pipeline described in Figure 1, where we first preprocess the CT volume with HU intensities followed by lung segmentation using the ResNet-UNet model. These segmented lung images are then fed to the classification network for the training and application of cXAI. As part of cXAI, we used four CAM techniques, namely, (i) Grad-CAM, (ii) Grad-CAM++, (iii) Score-CAM, and (iv) FasterScore-CAM to visualize the results of the classification model. Figure 8 shows the output from the cXAI, which includes the expert's lesion localization with black borders, representing the AI model's missed and correctly captured lesion.

Figures 9–14 show the visual results for the three kinds of DenseNet-based classifiers wrapped up with four types of CAM models, namely Grad-CAM (column 2), Grad-CAM++ (column 3), Score-CAM (column 4), and FasterScore-CAM (column 5) on COVID-19 vs. control segmented lung images, where the color map red shows the lesion localization using cXAI, thereby validating the prediction of the DenseNet models. Table 3 presents a comparative analysis of the three DenseNet models used in this study. The performance of the models has been compared using accuracy, loss, specificity, F1-score, recall, precision, and AUC scores. DenseNet-201 is the best-performing model when comparing the accuracy, loss, specificity, F1-score, recall, and precision. However, due to the larger model's size of 233 MB and a total number of parameters of 203 million, training the batch size of the model was kept at 4. While the batch size while training DenseNet-121 and DenseNet-169 was kept at 16 and 8 due to a smaller model size of 93 MB and 165 MB and further had a lesser number of parameters of 81 million and 143 million, respectively.
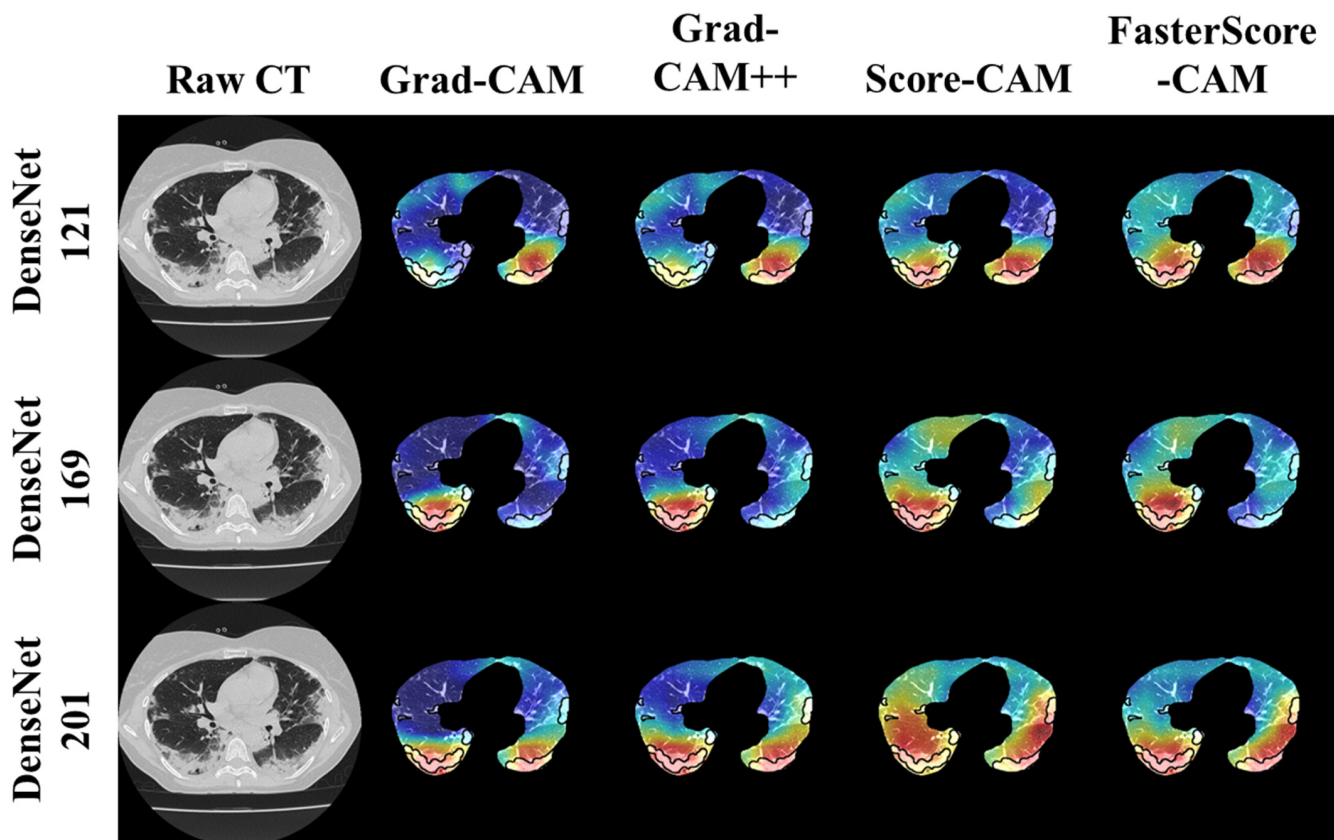
Figure 8. Heatmap using four CAM techniques using three kinds of DenseNet classifiers on COVID-19 lesion images.

Table 3. Comparative table for three kinds of DenseNet classifier models.

SN	Attributes	DN-121	DN-169	DN-201
1	# Layers	430	598	710
2	Learning Rate	0.0001	0.0001	0.0001
3	# Epochs	20	20	20
4	Loss	0.003	0.0025	**0.002**
5	ACC	98	98.5	**99**
6	SPE	0.975	0.98	**0.985**
7	F1-Score	0.96	0.97	**0.98**
8	Recall	0.96	0.97	**0.98**
9	Precision	0.96	0.97	**0.98**
10	AUC	0.99	0.99	0.99
11	Size (MB)	93	165	233
12	Batch size	16	8	4
13	Trainable Parameters	80 M	141 M	200 M
14	Total Parameters	81 M	143 M	203 M

DN-121: DenseNet-121; DN-169: DenseNet-169; DN-201: DenseNet-201; # = number of. Bold highlights the superior performance of the DenseNet-201 (DN-201) model.
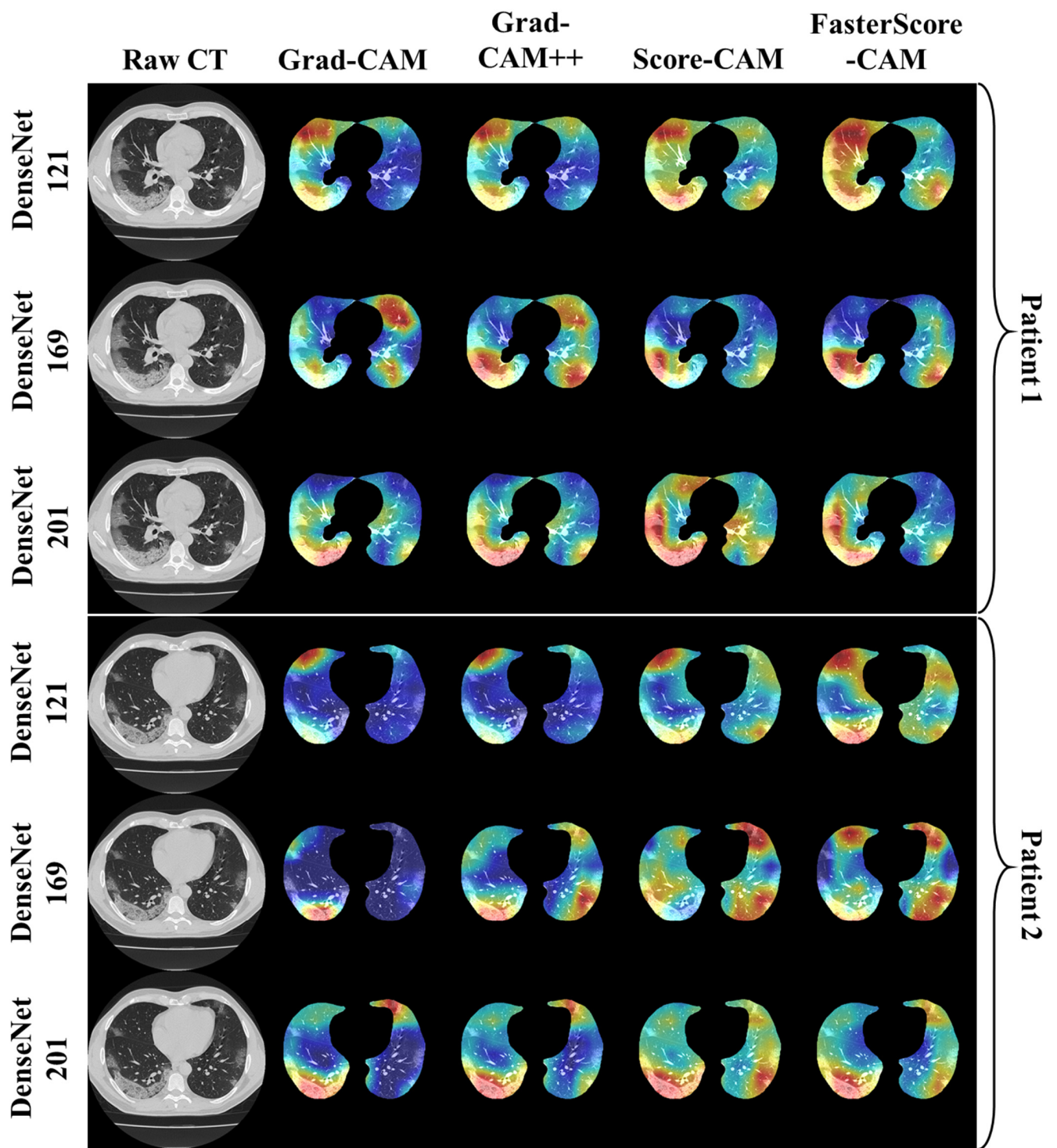
Figure 9. Heatmap using four CAM techniques and three kinds of DenseNet classifiers on COVID-19 lesion images. The top row is the CT slice for patient 1, and the bottom row is the CT slice for patient 2.

Figure 10. Heatmap using four CAM techniques using three kinds of DenseNet classifiers on COVID-19 lesion images. The top row is the CT slice for patient 1, and the bottom row is the CT slice for patient 2.
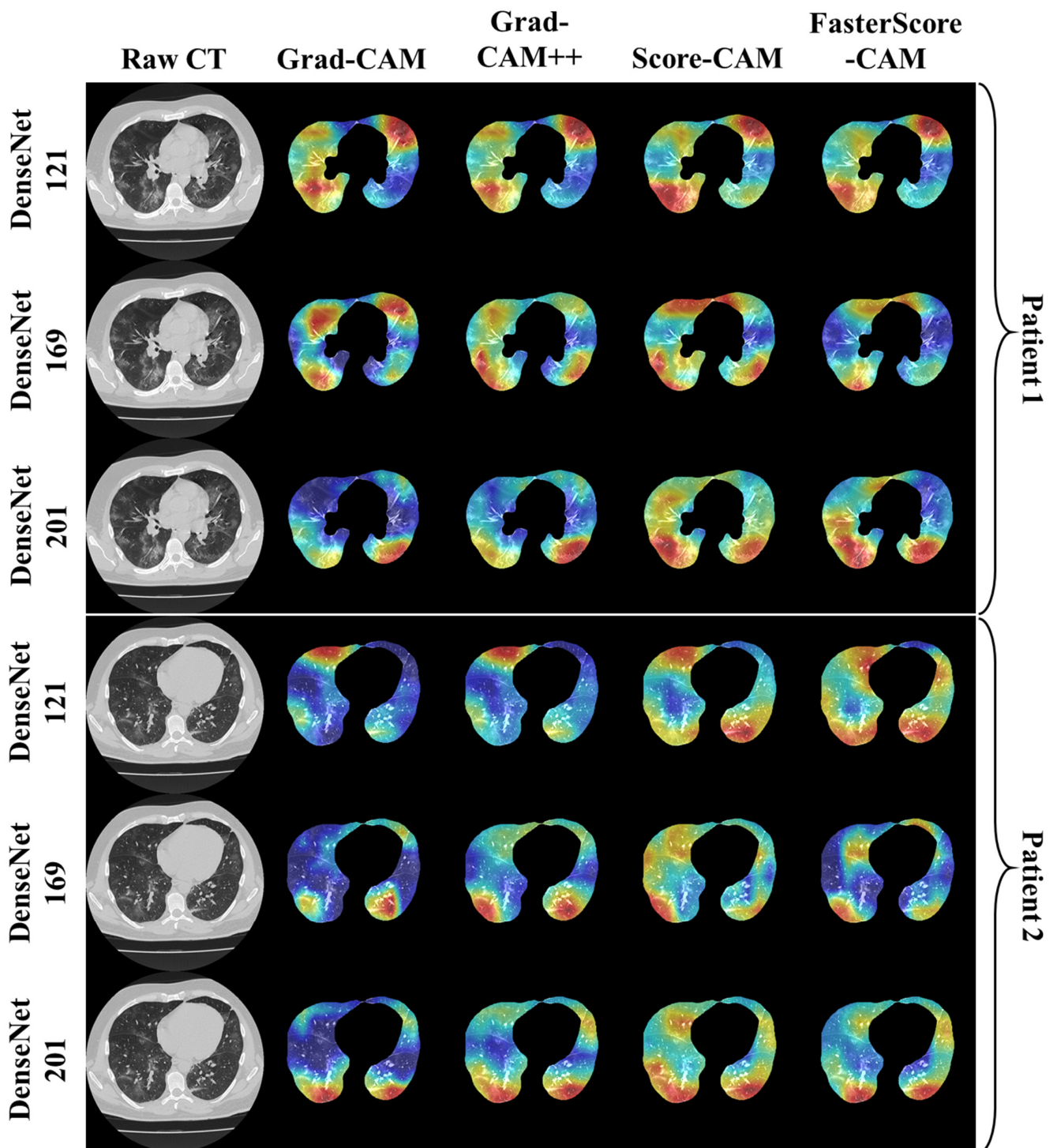
Figure 11. Heatmap using four CAM techniques using three kinds of DenseNet classifiers on COVID-19 lesion images. The top row is the CT slice for patient 1, and the bottom row is the CT slice for patient 2.
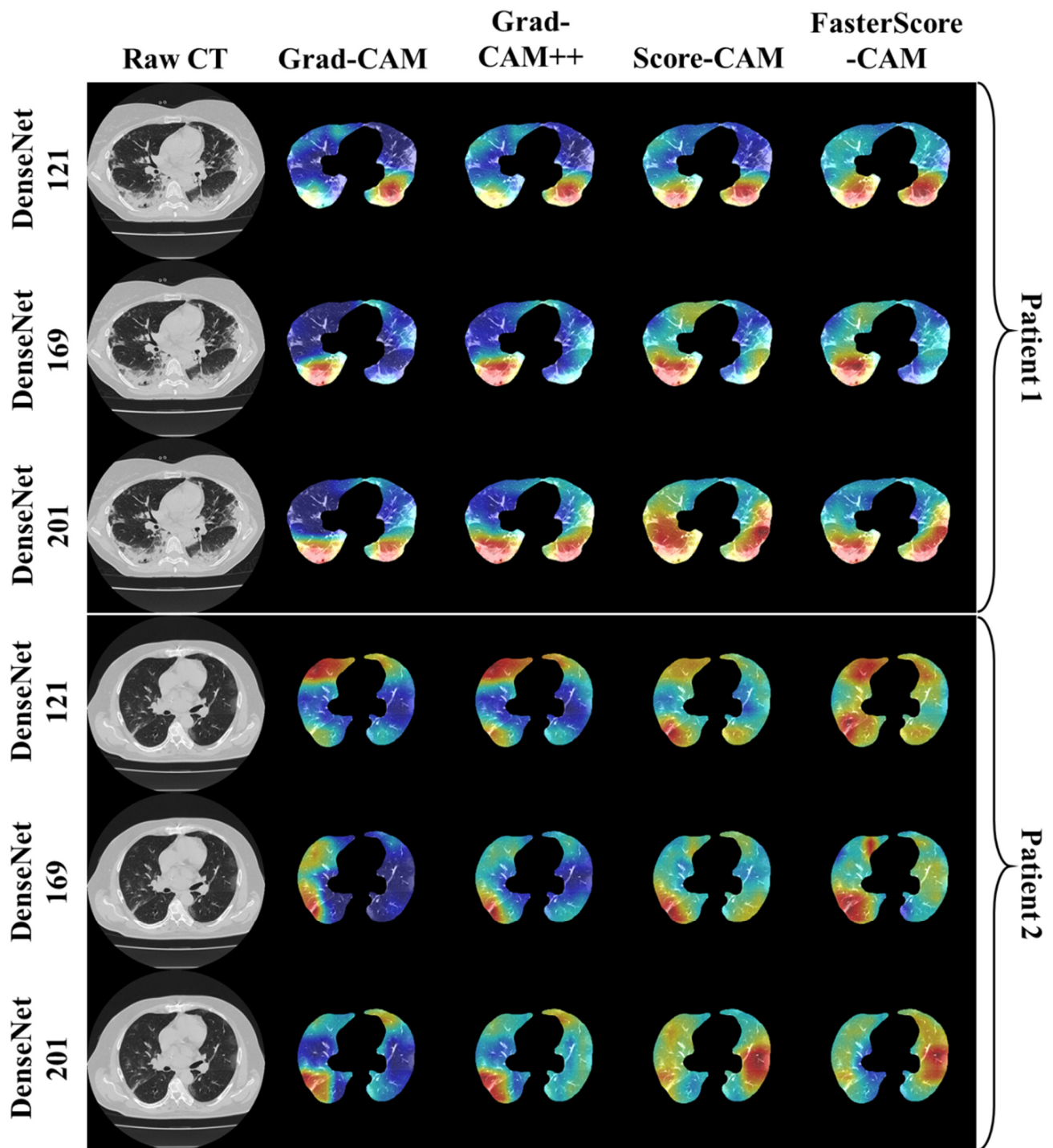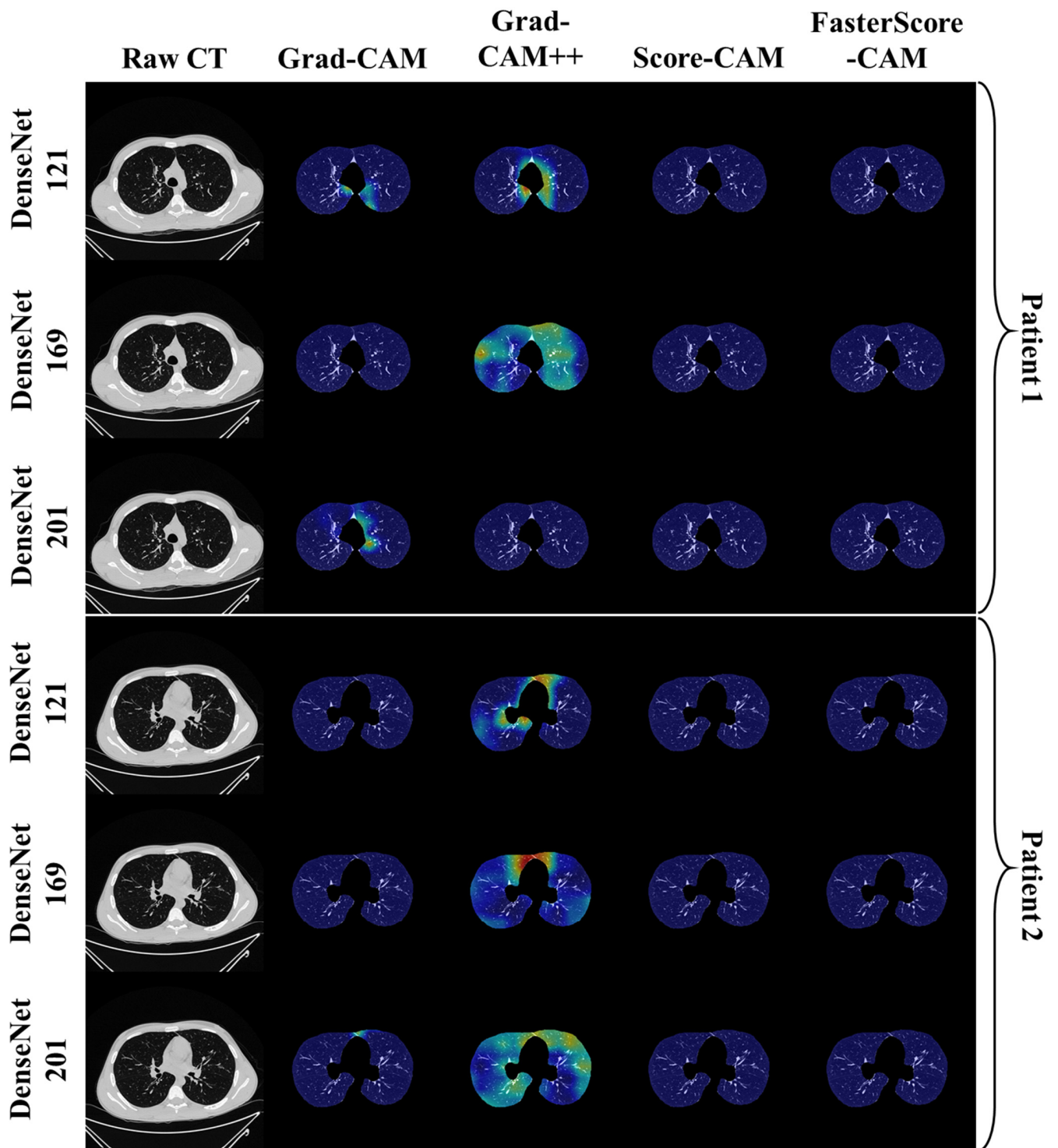
Figure 12. Heatmap using four CAM techniques using three kinds of DenseNet classifiers on control images. The top row is the CT slice for patient 1, and the bottom row is the CT slice for patient 2.

Figure 13. Heatmap using four CAM techniques using three kinds of DenseNet classifiers on control images. The top row is the CT slice for patient 1, and the bottom row is the CT slice for patient 2.

Figure 14. Heatmap using four CAM techniques using three kinds of DenseNet classifiers on control images. The top row is the CT slice for patient 1, and the bottom row is the CT slice for patient 2.
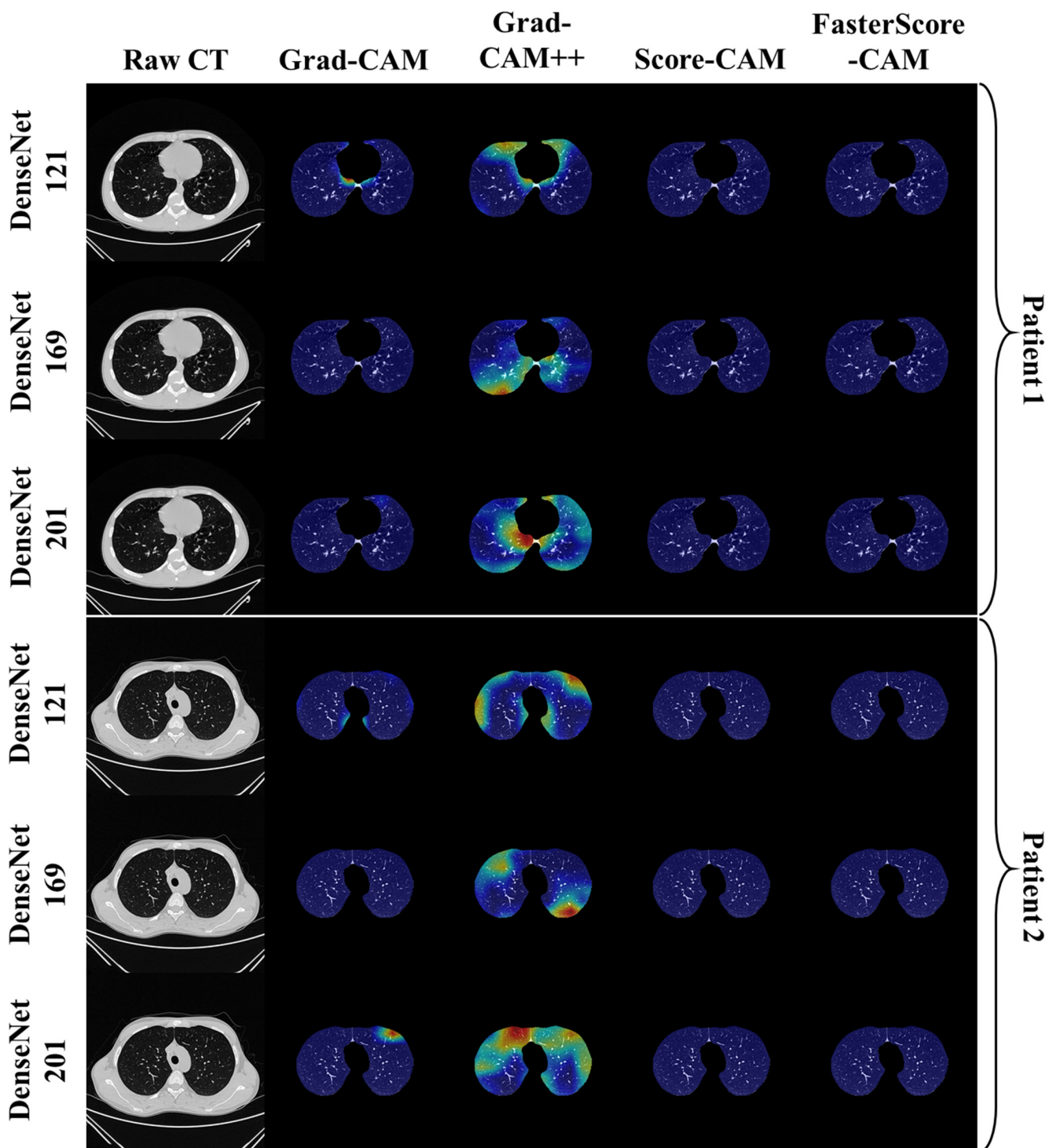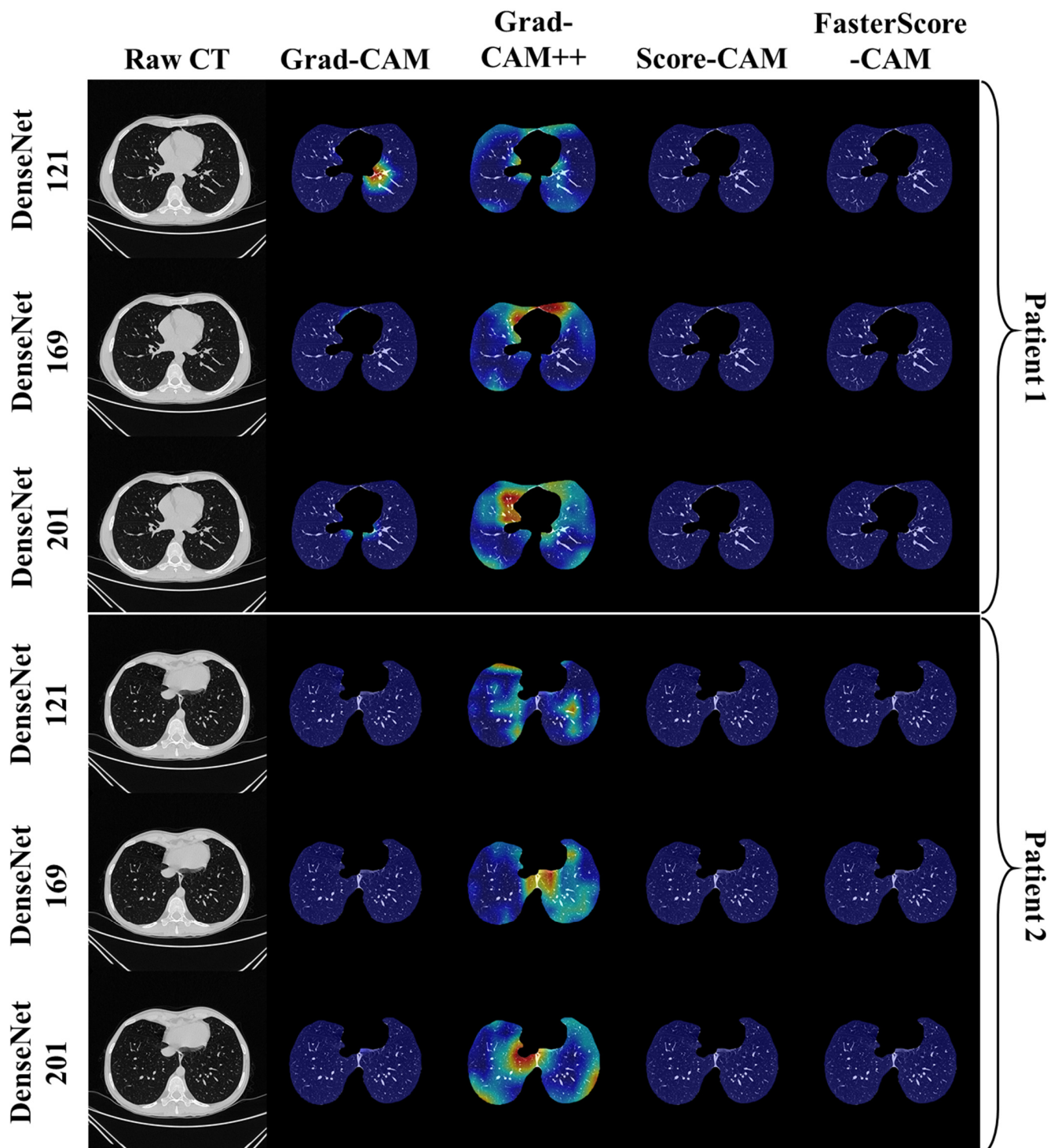
3.2. Performance Evaluation

The proposed study uses two techniques: (i) segmentation of the CT lung; and (ii) classification of the CT lung between COVID-19 vs. controls. For the segmentation part, we have presented mainly five kinds of performance evaluation metrics: (i) area error, (ii) Bland–Altman [111,112], (iii) correlation coefficient [113,114], (iv) dice similarity [115], and (v) Jaccard index. Figures 15–17 show the overlay of the ground truth lesions on heatmaps as part of the performance evaluation. The four columns represent Grad-CAM (column

2), Grad-CAM++ (column 3), Score-CAM (column 4), and FasterScore-CAM (column 5) on the segmented lung CT image. For the three DenseNet-based classification models, we introduce a new metric to evaluate the heatmap, i.e., mean alignment index (MAI). This MAI requires grading from a trained radiologist, where the radiologist rates the heatmap image between 1 and 5, with 5 being the best score. This study incorporates inter-observer analysis using three senior trained radiologists from different countries for MAI scoring on the cXAI-generated heatmap of the lesion localization on the images. The scores are then presented in the form of a bar chart (Figure 18) with grading from expert 1 (Figure 18, column 1), expert 2 (Figure 18, column 2), and expert 3 (Figure 18, column 3).

Figure 15. Overlay of ground truth annotation on heatmap using four CAM techniques on three kinds of DenseNet classifiers for COVID-19 lesion images as part of the performance evaluation.

Figure 16. Overlay of ground truth annotation on heatmap using four CAM techniques on three kinds of DenseNet classifiers for COVID-19 lesion images as part of the performance evaluation.

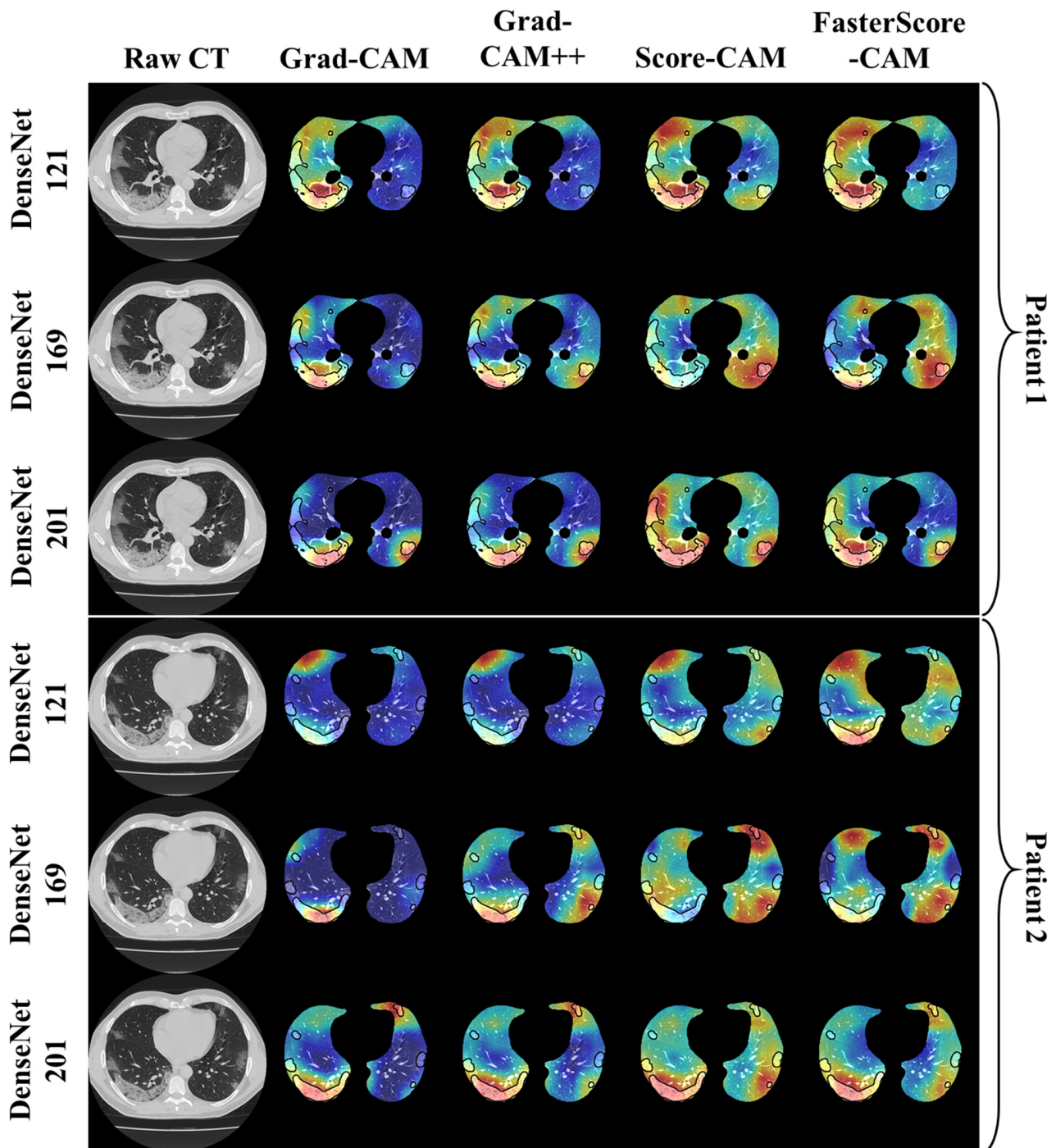**Figure 17.** Overlay of ground truth annotation on heatmap using four CAM techniques on three kinds of DenseNet classifiers for COVID-19 lesion images as part of the performance evaluation.
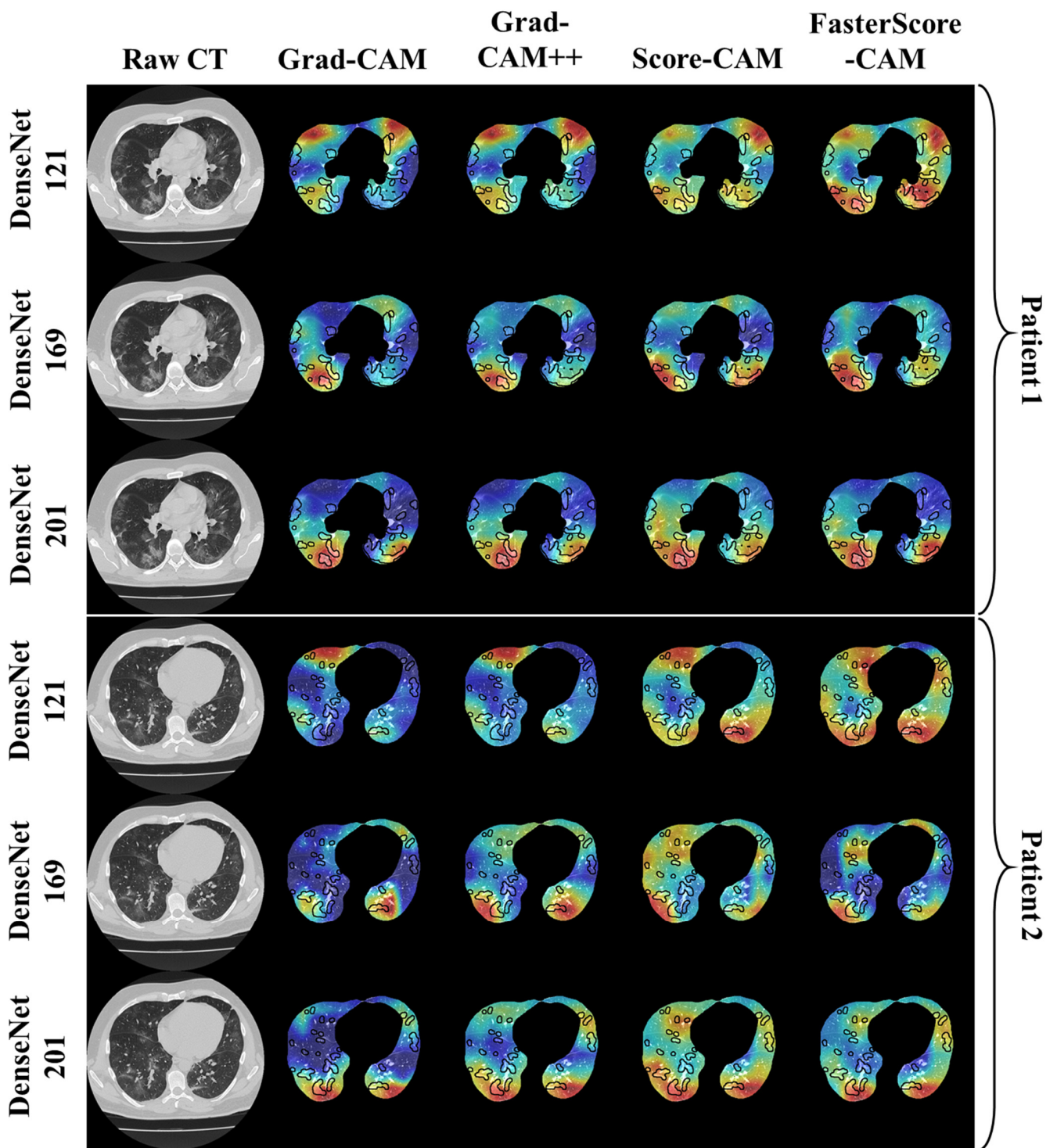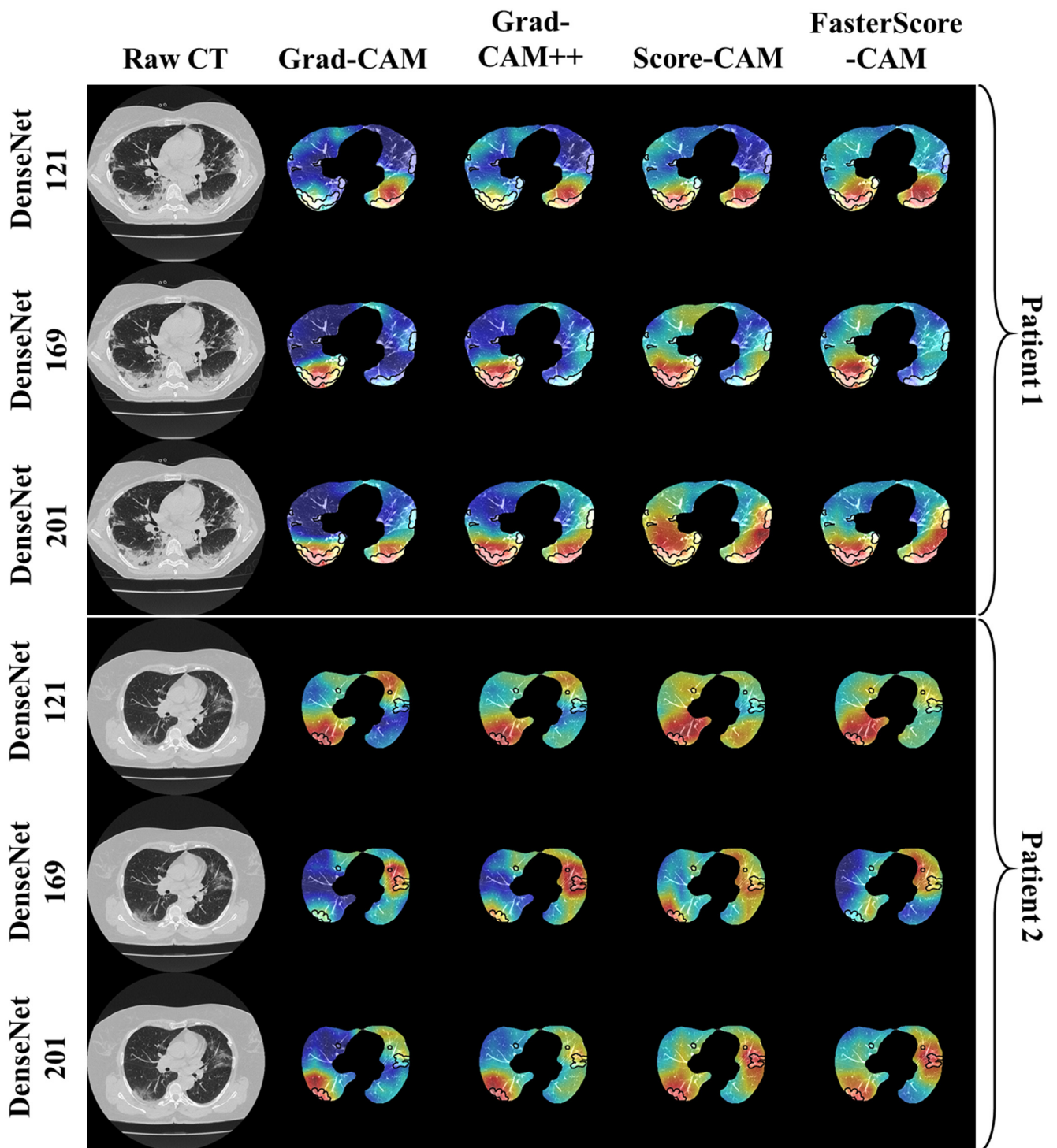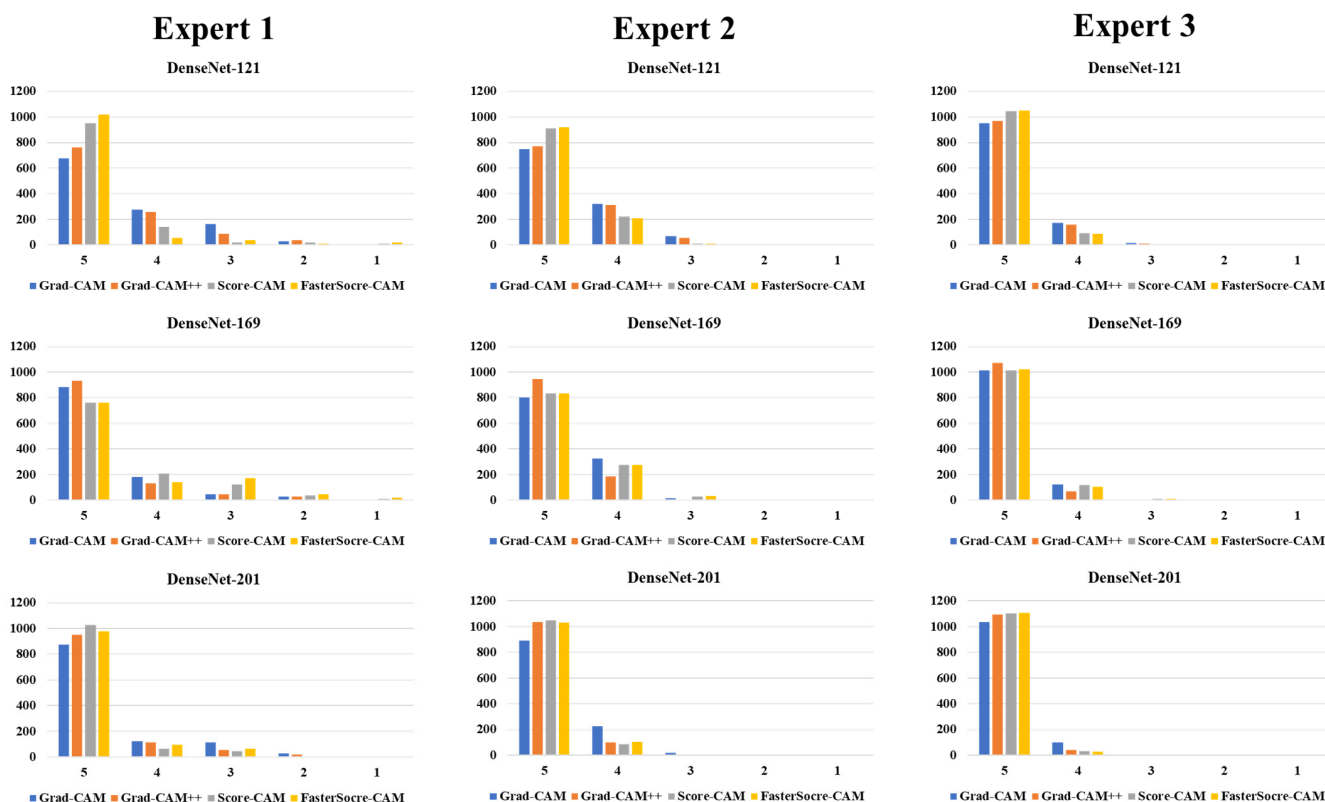
Figure 18. Bar chart representing the MAI.

3.3. Statistical Validation

This study uses the Friedman test to prove the statistically significant difference between the means of three or more groups, all of which have the same subjects [116–118]. The Friedman test's null hypothesis states that there are no differences between the sample medians. The null hypothesis will be rejected if the *p*-value calculated is less than the set significance threshold (0.05), and it can be determined that at least two of the sample medians are substantially different from each other. Further analysis of the Friedman test is presented in "Appendix A (Tables A1–A3)". It was noted that for all the MAI scores of three experts, the three classification models, namely, DenseNet-121, DenseNet-169, and DenseNet-201, and using the four CAM techniques used in XAI showed significance of $p < 0.00001$. Thus, this proves the reliability of the overall COVLIAS 2.0-cXAI system.

4. Discussion

4.1. Study Findings

To summarize, our prime contributions in the proposed study are six types of innovation in the design of COVLIAS 2.0-cXAI: (i) automated HDL lung segmentation using the ResNet-UNet model; (ii) classification of COVID-19 vs. controls using three kinds of DenseNets, namely, DenseNet-121 [55–57,83], DenseNet-169, and DenseNet-201; the combination of segmentation and classification improved the overall performance of the system; (iii) using explainable AI to visualize and validate the prediction of the DenseNet models using four kinds of CAM, namely Grad-CAM, Grad-CAM++, Score-CAM, and FasterScore-CAM, for the first time. This helps us understand the AI model's learning in the input CT image [35,84–86]. (iv) Mean alignment index (MAI) between heatmaps and the gold standard score from three trained senior radiologists, a score of four out of five, establishing the system for clinical applicability. Further, a Friedman test was also conducted to present the statistical significance of the scores from the three experts. (v) Application of the quantization to the trained AI model while making the prediction help in faster online prediction. Further, it also reduces the final trained AI model size, making the

complete system light. Lastly, (vi) presents an end-to-end cloud-based CT image analysis system, including the CT lung segmentation and COVID-19 intensity map using the four CAM techniques (Figure 1).

The proposed study presents heatmaps using four CAM techniques, namely, (i) Grad-CAM, (ii) Grad-CAM++, (iii) Score-CAM, and (iv) FasterScore-CAM. The CT lung segmentation using ResNet-UNet was adapted from our previous publication [93]. This segmented lung is then given as the input to the classification DenseNet models to train in distinguishing between COVID-19-positive and control individuals. The preprocessing involved while training the classification model consists of a Hounsfield unit (HU) adjusted to highlight the lung region (1600, −400), causing the model to train efficiently by improving the visibility of COVID-19 lesions [53]. Further, we have also designed a cloud-based AI system that takes the raw CT slice as the input and then processes this image first for lung segmentation, followed by heatmap visualization using four techniques [119–123]. Figures 19–21 represent the output from the cloud-based COVLIAS 2.0-cXAI system (Figure 22, a web-view screenshot). This COVLIAS 2.0-cXAI uses multithreading to process the four CAM techniques in a parallel manner and produces results faster than sequential processing.

While it is intuitive to examine the relationship between demographics and COVID-19 severity [22,124–126], it is not always necessarily the case that (i) there can be a relationship between demographics and COVID-19 severity, (ii) there can be data collection with all demographics parameters and COVID-19 severity, (iii) there can be data collection keeping comorbidity in mind, and/or (iv) the cohort sizes are large enough to establish the relationship between demographics and COVID-19 severity. Such conditions are prevalent in our setup and therefore no such relationship could be established; however, as part of the research, one can establish such a relationship along with survival analysis. The objective of this study was squarely not aimed at collecting demographics and relating them to COVID-19 severity; however, we have attempted this in previous studies [127].

Multilabel classification is not new [21,124,128,129]. For multilabel classification, the models are trained with multiple classes, for example, if there are two or more than two classes, then the gold standard must consist of two or more than two classes [124,129]. Note that in our study, the only two classes used were COVID-19 and controls; however, different kinds of lesions can be classified using a multiclass-based classification framework (for example, GGO vs. consolidations vs. crazy paving), which was out of the scope of the current work, but this can be part of the future study. Moreover, inclusion of unsupervised techniques can also be attempted [130].

The total data size for ResNet-UNet-based segmentation was 5000. These trained models were used for segmentation followed by classification on 2450 test CT scans consisting of 1400 COVID-19 cases and 1050 control CT scans. Three kinds of DenseNet classifiers were used for classification of COVID-19 vs. controls. Further, the COVLIAS 2.0-cXAI used the explainable AI using three kinds of Grad-CAM for heatmap generation. Thus, overall, the system used 7450 CT images, which is relatively large. Due to the radiologists' time and cost reasons, the test data set was nearly 33% of the total data set of the system, which is considered reasonable.

Figure 19. COVLIAS 2.0 cloud-based display of the lesion images using four CAM models.

COVLIAS 2.0: AI-Based COVID-19 Lesion Detector

Results

Raw CT Input Slice

CT Lung Segmentation

Lesion using Grad-CAM

Lesion using Grad-CAM++

Lesion using Score-CAM

Lesion using FasterScore-CAM
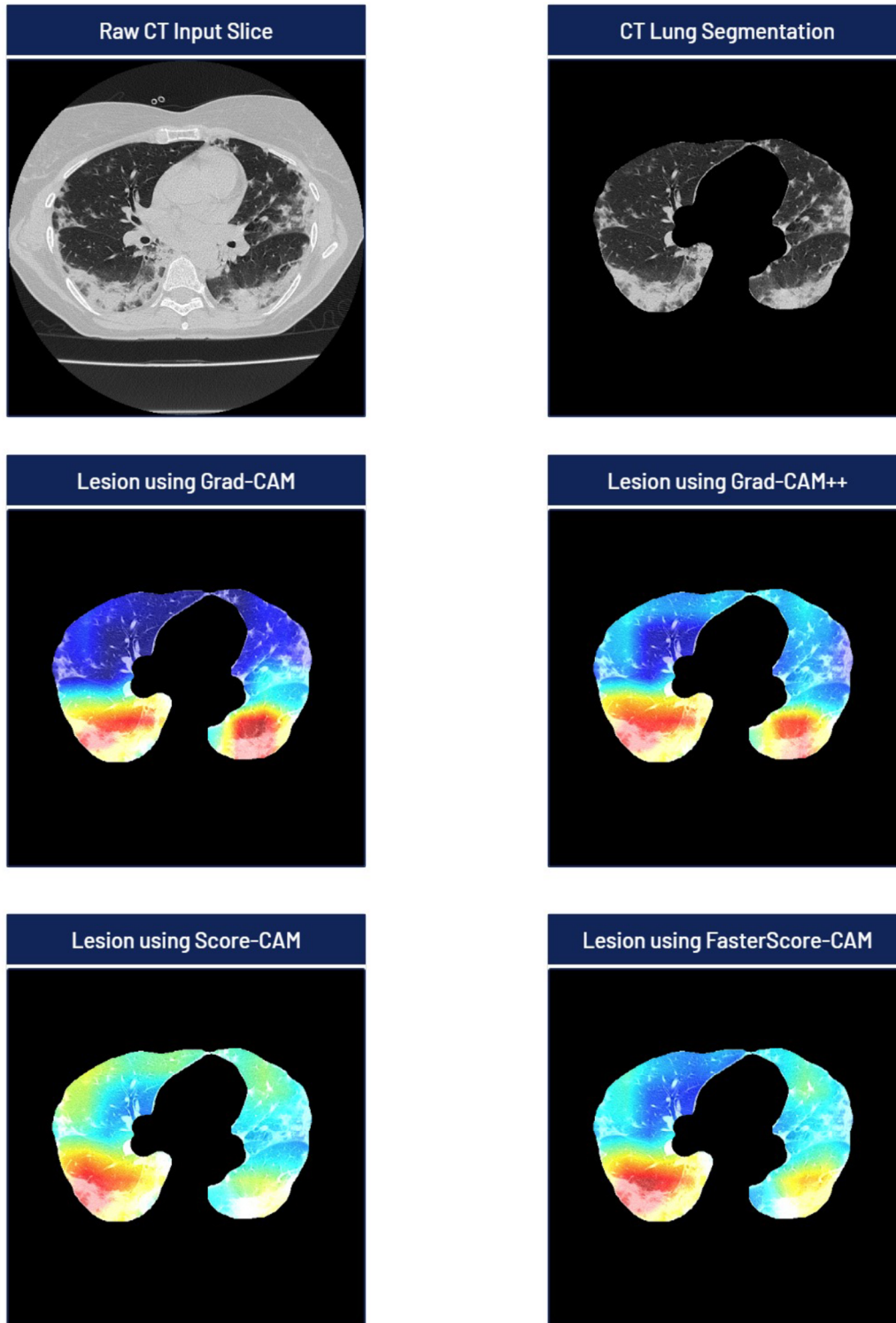
Figure 20. COVLIAS 2.0 cloud-based display of the lesion images using four CAM models.
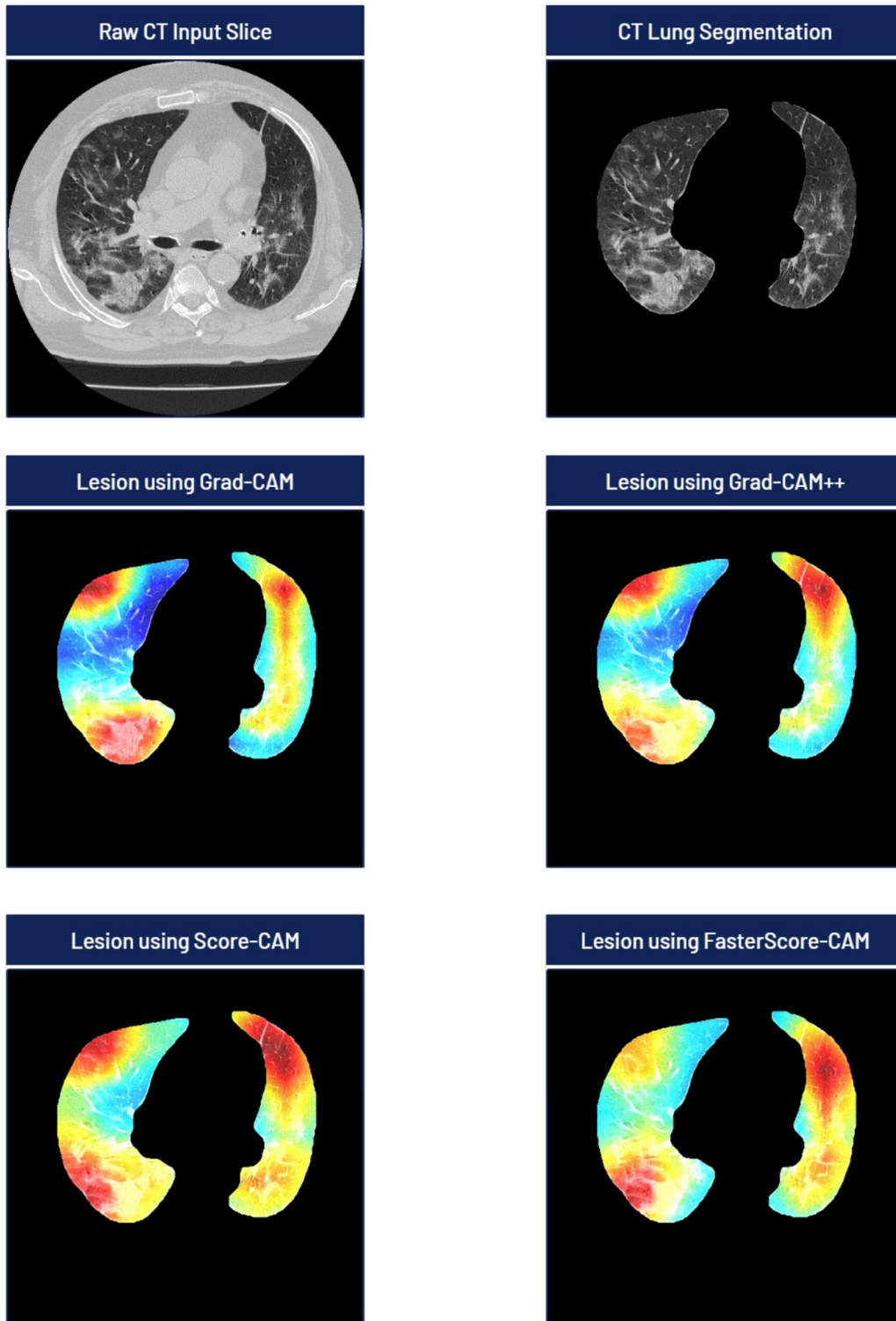
Figure 21. COVLIAS 2.0 cloud-based display of the lesion images using four CAM models.
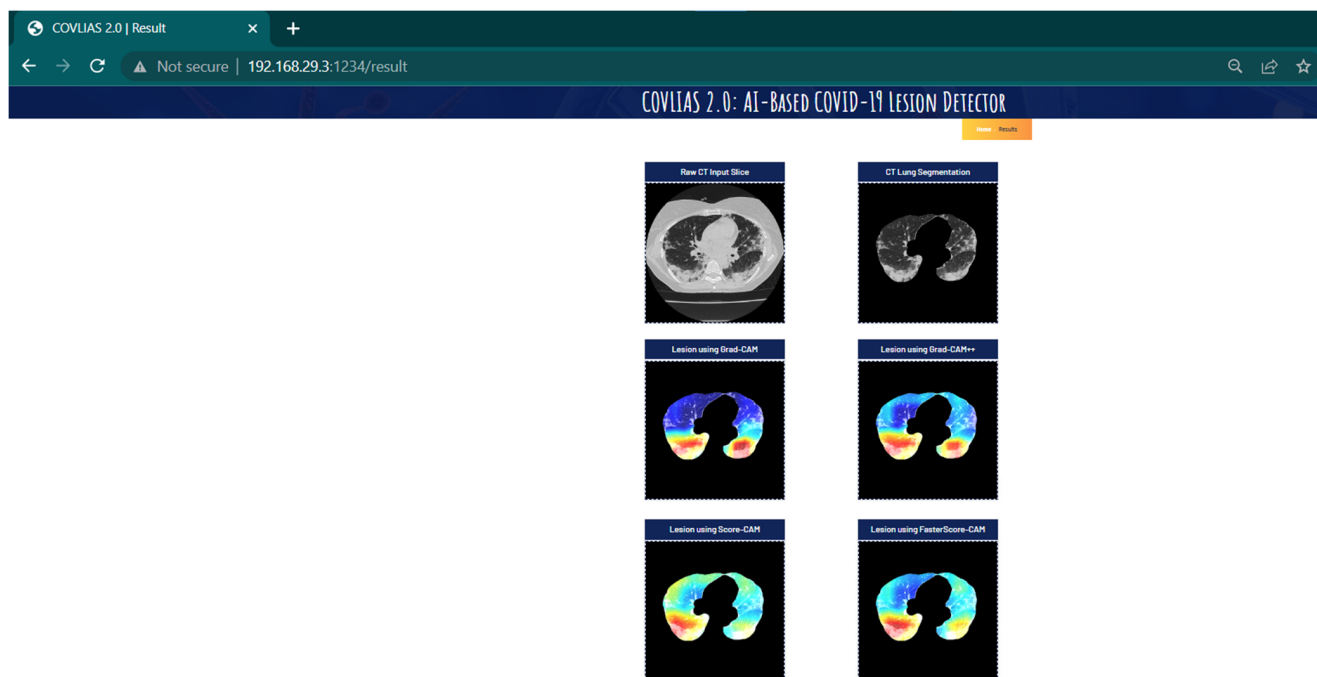
Figure 22. A web-view screenshot.

4.2. Memorization vs. Generalization for Longitudinal Studies

Generalization is the process where the AI model does not purely depend upon the data sample size for best performance [34,131]. Since the models were trained using K5 cross-validation (CV) protocol (80:20), and the accuracy was predicted on the test data set, which was not part of the training data sets, the process of memorization was thus less likely to happen. Note that for every CV protocol, the "memorization vs. generalization" needs to be evaluated independently, especially keeping the treatment paradigm for longitudinal data sets, which was out of scope for the current settings. From our past experiences, the effect of generalization can be retained in the deep learning framework to a certain degree. In our recent experiments, where we had applied "unseen test data" on our trained AI models, it resulted in encouraging accuracy [27,132], which justifies "superior generalization" in deep learning frameworks, unlike in machine learning frameworks. Since COVLIAS 2.0-cXAI is a deep learning framework, we thus conclude that the cloud-based "COVLIAS 2.0-cXAI" can be adopted for the longitudinal data sets during the monitoring phase.

4.3. A Special Note on Training Data Set

We trained the segmentation model using ResNet-UNet on 5000 COVID-19 images. An unseen data set of 2450 (1400 COVID-19 and 1050 control images) was used for testing. Since the training data set was quite large, we did not use augmentation during training protocol. Note that the unseen data (2450) was also not augmented. While several studies have been published that used the augmentation protocol [36,90,94,133–135] during classification, our DenseNet models for classification were never modified and never underwent change in rotation, tilt, or orientation. Further, note that we used the DICOM image directly, which contains orientation information. This information was used to solve the problem of rotation, tilting, or any abnormal orientation. This orientation information in the DICOM tag was used to fix the orientation of the image so that the lung is always vertically straight in the image.

4.4. A Special Note on Four CAM Models

While DL has demonstrated accuracy in image classification, object recognition, and image segmentation, model interpretability, a key component in model explainability, comprehension, and debugging, is one of the most significant issues. That poses an intriguing question: how can you trust a model's decisions if you cannot fully justify how it got there? There has been the latest trend in the growth of XAI for a better understanding of the AI black boxes [49,136–139]. Grad-CAM or Grad-CAM++ produces a coarse clustering map showing the key regions in the picture for predicting any target idea (say, "COVID-19" in a classification network) by using the gradients of any target concept (say, "COVID-19" in a classification network) in the final convolutional layer. In contrast, Score-CAM is built on the idea of perturbation-based approaches that mask portions of the original input and measure the change in target score. The produced activation mask is handled as a mask for the input image, masking sections of the input image and causing the model to predict the partially masked image. The target class score is then used to reflect the significance of the class activation map. While Score-CAM is an excellent method, it, however, takes more time to process compared to other CAM methods. FasterScore-CAM makes Score-CAM more efficient. This was achieved using only the dominating channels with significant variances as the mask image. Thus, a CAM version that is ten times faster than Score-CAM is produced.

4.5. Benchmarking the Proposed Model against Previous Strategies

We present the benchmarking strategy in Table 4, and this includes studies that utilized the CAM technique for COVID-19-based lesion localization. Lu et al. [140] presented CGENet, a deep graph model for COVID-19 detection on CT images. First, they established the appropriate backbone network for the CGENet adaptively. The authors then devised a novel graph-embedding mechanism to merge the spatial relationship into the feature vectors. Finally, to improve classification performance, they picked the extreme learning machine (ELM) [24] as the classifier for the proposed CGENet. Based on five-fold cross-validation, the suggested CGENet obtained an average accuracy of 97.78% on a large publicly available COVID-19 data set with ~2400 CT slices. They also compared the performance of CGENet against five existing methods. In addition, based on COVID-19 samples, the Grad-CAM maps were used to offer a visual explanation of CGENet. The authors did not report the AUC values and did not compare the other CAM methods such as Grad-CAM++, Score-CAM, and FasterScore-CAM.

At Tlemcen Hospital in Algeria, Lahsaini et al. [141] first gathered a data set of 4986 COVID and non-COVID images validated by RT-PCR assays. Then, to conduct a comparative analysis, they performed transfer learning on DL models that received the highest results on the ImageNet data set, such as DenseNet-121, DenseNet-201, VGG16, VGG19, Inception Resnet-V2, and Xception [142]. Finally, they proposed an explainable model for detecting COVID-19 in chest CT images and explaining the output decision based on the DenseNet-201 architecture. According to the results of the tests, the proposed design has a 98.8% accuracy rate. It also uses Grad-CAM to provide a visual explanation. The authors did not compare them with other CAM methods such as Grad-CAM++, Score-CAM, and FasterScore-CAM.

Zhang et al. [143] investigated whether combining chest CT and chest X-ray data can assist AI to diagnose more accurately. Approximately 5500 CT slices were collected from 86 participants for this study. The convolutional block attention module was used to create an end-to-end multiple-input deep convolutional attention network (MIDCAN) (CBAM). One of our model's inputs received a CT picture, while the other received an X-ray image. Grad-CAM was also used to create an explainable heatmap. The suggested MIDCAN had accuracy of 98.02%, sensitivity of 98.1%, and specificity of 97.95%. The authors did not compare the other CAM methods such as Grad-CAM++, Score-CAM, and FasterScore-CAM.

Monta et al. [144] presented the Fused-DenseNet-Tiny, a lightweight DCNN model based on a truncated and concatenated DenseNet. Transfer learning, partial layer freezing, and feature fusion were used to train the model to learn CXR features utilizing 9208 CXR. The proposed model was shown to be 97.99% accurate during testing. Despite its lightweight construction, the Fused-DenseNet-Tiny cannot outperform its larger cousin due to its limited extraction capabilities. The authors also used Grad-CAM to explain the trained AI model visually. The authors did not report the AUC values and did not compare the other CAM methods such as Grad-CAM++, Score-CAM, and FasterScore-CAM.

Table 4. Benchmarking table.

C0	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12	C13	C14	C15	C16
SN	Author	Year	TP	TS	IS^2	TM	DL Model	Modality	XAI	Heatmap Models	AUC	SEN	SPE	PRE	F1	ACC
1	Lu et al. [140]	2021		2482	100 to 500	5	CGENet	CT	✗	Grad-CAM	✗	97.9	97.7	97.7	97.8	97.8
2	Lahsaini et al. [141]	2021	177	4968	✗	6	Transferred DenseNet201	CT	✗	Grad-CAM	0.988	99.5	98.2	97.8	98	98.2
3	Zhang et al. [143]	2021	86	5504	1024(CT) 2048(X-Ray)	8	MIDCAN	CT, X-ray	✗	Grad-CAM	0.98	98.1	98	97.9	98	98
4	Monta et al. [144]	2021		9208	299	7	Fused-DenseNet-Tiny	X-ray	✗	Grad-CAM	✗	✗	✗	98.4	98.3	98
5	Proposed Suri et al.	2022	80	5000	512	3	DenseNet-121 DenseNet-169 DenseNet-201	CT	✓	Grad-CAM Grad-CAM++ Score-CAM FasterScore-CAM	0.99 0.99 0.99	0.96 0.97 0.98	0.975 0.98 0.985	0.96 0.97 0.98	0.96 0.97 0.98	98 98.5 99

TP: total patients; TS: total slice; IS: image size; TM: total models; AUC: area under the curve; SEN (%): sensitivity (or recall); SPE (%): specificity; PRE (%): precision; ACC (%): accuracy.

4.6. Strengths, Weakness, and Extensions

The study presented COVLIAS 2.0-cXAI, a cloud-based XAI system for COVID-19 lesion detection and visualization. The cXAI system presented a comparison of four heatmap techniques, (i) Grad-CAM, (ii) Grad-CAM++, (iii) Score-CAM, and (iv) FasterScore-CAM for the first time using three DenseNet models, namely, DenseNet-121, DenseNet-169, and DenseNet-201 for COVID-19 lung CT images. To improve the prediction of the three DenseNet models, we first segment the CT lung using a hybrid DL model ResNet-UNet and then pass it to the classification network. Applying quantization to the three trained AI models, namely, DenseNet-121, DenseNet-169, and DenseNet-201, while making the prediction, helps in faster online prediction. Further, it also reduces the final trained AI model size, making the complete system light. The overall cXAI system incorporates validation of the lesion localization using expert grading, thereby generating an MAI score. Lastly, the study presents an end-to-end cloud-based CT image analysis system (COVLIAS 2.0-cXAI), including the CT lung segmentation (ResNet-UNet) and COVID-19 lesion intensity map using cXAI techniques. This study uses inter-observer variability similar to other variability measurements [145] to score the MAI for lesion localization, which was further validated using the Friedman test.

Even though the three AI models, DenseNet-121, DenseNet-169, and DenseNet-201, produced promising results on a data set from a single location, the study was limited to one observer due to cost, time, and radiologists' availability. Several kinds of DenseNet systems have been developed which can be tried and the current DenseNet can be replaced by [146–148]; as part of the extension to this study, more AI models can be explored and can incorporate the use of the HDL model for binary or multiclass-based classification [128] framework. Explainable AI is an emerging field and many new strategies can be incorporated [47,50,149–157]. New techniques have evolved such as SHAP [52,158] and UMAP [159]. Heatmaps produced by Grad-CAM have been used for XAI in several applications [64], where the generated heatmaps are the threshold to compute the lesions which are then compared against the gold standard [49]. Choi et al. [48] used SHAP to

demonstrate the high-risk factors responsible for higher phosphate. Further, to improve the speed of the AI model, model optimization techniques such as weight clustering and AI model pruning [160–164] can be applied [115,165–169]. Techniques such as storage reduction are necessary when dealing with AI solutions [51,54,170–172]. Fusion of conventional image processing can be used with AI to improve the performance of the system [173,174]. These AI technologies are likely to benefit long-COVID [175].

5. Conclusions

The proposed study is the first pilot study that integrates a cloud-based explainable artificial intelligence system using four techniques, namely, (i) Grad-CAM, (ii) Grad-CAM++, (iii) Score-CAM, and (iv) FasterScore-CAM-based lesion localization using three DenseNet models, namely, DenseNet-121, DenseNet-169, and DenseNet-201. Thus, it compares the methods and explainability of the four different CAM strategies for COVID-19-based CT lung lesion localization. DenseNet-121, DenseNet-169, and DenseNet-201 demonstrated an accuracy performance of 98%, 98.5%, and 99%, respectively. The study incorporated a hybrid DL (ResNet-UNet) for COVID-19-based CT lung segmentation using independent cross-validation and performance evaluation schemes. To validate the lesion, three trained senior radiologists scored the lesion localization on the CT lung data set and then compared it against the heatmap generated by cXAI, resulting in the MAI score. Overall, ~80% of CT scans were above an MAI score of four out of five, demonstrating matching lesion locations using cXAI vs. gold standard, thus proving the clinical applicability. Further, the Friedman test was also performed on the MAI scores by comparing the three radiologists. The online cloud-based COVLIAS 2.0-cXAI achieves (i) CT lung image segmentation and (ii) generation of four CAM techniques in less than 10 s for one CT slice. The COVLIAS 2.0-cXAI demonstrated reliability, high accuracy, and clinical stability.

Appendix A

Figure A1. ResNet-UNet architecture.

Table A1. Friedman test using DenseNet-121 model on the MAI score from three experts.

	XAI	Experts	Min.	25th Percentile	Med	75th Percentile	Max	DF-1	DF-2	*p* Value	F
DenseNet-121	Grad-CAM	Expert 1	2	4	5	5	5	2	2278	<0.00001	171.81
		Expert 2	3	4	5	5	5				
		Expert 3	2.7	4.2	4.6	4.8	5				
	Grad-CAM++	Expert 1	2	4	5	5	5	2	2278	<0.00001	244.9
		Expert 2	3	4	5	5	5				
		Expert 3	2.8	4.3	4.6	4.8	5				
	Score-CAM	Expert 1	1	5	5	5	5	2	2278	<0.00001	740.1
		Expert 2	3	5	5	5	5				
		Expert 3	2	4.5	4.7	4.9	5				
	FasterScore-CAM	Expert 1	1	5	5	5	5	2	2278	<0.00001	1072.54
		Expert 2	3	5	5	5	5				
		Expert 3	2.8	4.5	4.7	4.8	5				

Min: minimum; Med: median; Max: maximum; F: Friedman statistics.

Table A2. Friedman test using DenseNet-169 model on the MAI score from three experts.

	XAI	Experts	Min.	25th Percentile	Med	75th Percentile	Max	DF-1	DF-2	*p* Value	F
DenseNet-169	Grad-CAM	Expert 1	2	5	5	5	5	2	2278	<0.00001	432.84
		Expert 2	3	4	5	5	5				
		Expert 3	2.7	4.4	4.6	4.8	5				
	Grad-CAM++	Expert 1	2	5	5	5	5	2	2278	<0.00001	689.05
		Expert 2	3	5	5	5	5				
		Expert 3	3.2	4.5	4.7	4.8	5				
	Score-CAM	Expert 1	1	4	5	5	5	2	2278	<0.00001	282.56
		Expert 2	3	4	5	5	5				
		Expert 3	2.8	4.5	4.7	4.8	5				
	FasterScore-CAM	Expert 1	1	4	5	5	5	2	2278	<0.00001	253.15
		Expert 2	3	4	5	5	5				
		Expert 3	2.7	4.4	4.4	4.8	5				

Min: minimum; Med: median; Max: maximum; F: Friedman statistics.

Table A3. Friedman test using DenseNet-201 model on the MAI score from three experts.

	XAI	Experts	Min.	25th Percentile	Med	75th Percentile	Max	DF-1	DF-2	*p* Value	F
DenseNet-201	Grad-CAM	Expert 1	2	5	5	5	5	2	2278	<0.00001	499.3
		Expert 2	3	5	5	5	5				
		Expert 3	2.8	4.5	4.7	4.9	5				
	Grad-CAM++	Expert 1	2	5	5	5	5	2	2278	<0.00001	1151.78
		Expert 2	3	5	5	5	5				
		Expert 3	2.7	4.6	4.7	4.9	5				
	Score-CAM	Expert 1	3	5	5	5	5	2	2278	<0.00001	1719.93
		Expert 2	3	5	5	5	5				
		Expert 3	3	4.6	4.7	4.9	5				
	FasterScore-CAM	Expert 1	3	5	5	5	5	2	2278	<0.00001	1239.82
		Expert 2	3	5	5	5	5				
		Expert 3	2.9	4.6	4.7	4.9	5				

Min: minimum; Med: median; Max: maximum; F: Friedman statistics.

References

1. Cucinotta, D.; Vanelli, M. WHO Declares COVID-19 a Pandemic. *Acta Biomed.* **2020**, *91*, 157–160. [CrossRef] [PubMed]
2. WHO Coronavirus (COVID-19) Dashboard. Available online: https://covid19.who.int/ (accessed on 24 January 2022).
3. Suri, J.S.; Puvvula, A.; Biswas, M.; Majhail, M.; Saba, L.; Faa, G.; Singh, I.M.; Oberleitner, R.; Turk, M.; Chadha, P.S.; et al. COVID-19 pathways for brain and heart injury in comorbidity patients: A role of medical imaging and artificial intelligence-based COVID severity classification: A review. *Comput. Biol. Med.* **2020**, *124*, 103960. [CrossRef] [PubMed]
4. Cau, R.; Mantini, C.; Monti, L.; Mannelli, L.; Di Dedda, E.; Mahammedi, A.; Nicola, R.; Roubil, J.; Suri, J.S.; Cerrone, G.; et al. Role of imaging in rare COVID-19 vaccine multiorgan complications. *Insights Imaging* **2022**, *13*, 44. [CrossRef] [PubMed]
5. Saba, L.; Gerosa, C.; Fanni, D.; Marongiu, F.; La Nasa, G.; Caocci, G.; Barcellona, D.; Balestrieri, A.; Coghe, F.; Orru, G.; et al. Molecular pathways triggered by COVID-19 in different organs: ACE2 receptor-expressing cells under attack? A review. *Eur. Rev. Med. Pharm. Sci.* **2020**, *24*, 12609–12622. [CrossRef]
6. Onnis, C.; Muscogiuri, G.; Paolo Bassareo, P.; Cau, R.; Mannelli, L.; Cadeddu, C.; Suri, J.S.; Cerrone, G.; Gerosa, C.; Sironi, S.; et al. Non-invasive coronary imaging in patients with COVID-19: A narrative review. *Eur. J. Radiol.* **2022**, *149*, 110188. [CrossRef] [PubMed]
7. Viswanathan, V.; Puvvula, A.; Jamthikar, A.D.; Saba, L.; Johri, A.M.; Kotsis, V.; Khanna, N.N.; Dhanjil, S.K.; Majhail, M.; Misra, D.P. Bidirectional link between diabetes mellitus and coronavirus disease 2019 leading to cardiovascular disease: A narrative review. *World J. Diabetes* **2021**, *12*, 215. [CrossRef]
8. Fanni, D.; Saba, L.; Demontis, R.; Gerosa, C.; Chighine, A.; Nioi, M.; Suri, J.S.; Ravarino, A.; Cau, F.; Barcellona, D.; et al. Vaccine-induced severe thrombotic thrombocytopenia following COVID-19 vaccination: A report of an autoptic case and review of the literature. *Eur. Rev. Med. Pharm. Sci.* **2021**, *25*, 5063–5069. [CrossRef]
9. Gerosa, C.; Faa, G.; Fanni, D.; Manchia, M.; Suri, J.; Ravarino, A.; Barcellona, D.; Pichiri, G.; Coni, P.; Congiu, T. Fetal programming of COVID-19: May the barker hypothesis explain the susceptibility of a subset of young adults to develop severe disease? *Eur. Rev. Med. Pharmacol. Sci.* **2021**, *25*, 5876–5884.
10. Cau, R.; Pacielli, A.; Fatemeh, H.; Vaudano, P.; Arru, C.; Crivelli, P.; Stranieri, G.; Suri, J.S.; Mannelli, L.; Conti, M.; et al. Complications in COVID-19 patients: Characteristics of pulmonary embolism. *Clin. Imaging* **2021**, *77*, 244–249. [CrossRef]
11. Kampfer, N.A.; Naldi, A.; Bragazzi, N.L.; Fassbender, K.; Lesmeister, M.; Lochner, P. Reorganizing stroke and neurological intensive care during the COVID-19 pandemic in Germany. *Acta Biomed.* **2021**, *92*, e2021266. [CrossRef]
12. Congiu, T.; Demontis, R.; Cau, F.; Piras, M.; Fanni, D.; Gerosa, C.; Botta, C.; Scano, A.; Chighine, A.; Faedda, E. Scanning electron microscopy of lung disease due to COVID-19-a case report and a review of the literature. *Eur. Rev. Med. Pharmacol. Sci.* **2021**, *25*, 7997–8003. [PubMed]
13. Pan, F.; Ye, T.; Sun, P.; Gui, S.; Liang, B.; Li, L.; Zheng, D.; Wang, J.; Hesketh, R.L.; Yang, L. Time course of lung changes on chest CT during recovery from 2019 novel coronavirus (COVID-19) pneumonia. *Radiology* **2020**, *13*, 200370.
14. Wong, H.Y.F.; Lam, H.Y.S.; Fong, A.H.; Leung, S.T.; Chin, T.W.; Lo, C.S.Y.; Lui, M.M.; Lee, J.C.Y.; Chiu, K.W.; Chung, T.W.; et al. Frequency and Distribution of Chest Radiographic Findings in Patients Positive for COVID-19. *Radiology* **2020**, *296*, E72–E78. [CrossRef] [PubMed]
15. Smith, M.J.; Hayward, S.A.; Innes, S.M.; Miller, A.S.C. Point-of-care lung ultrasound in patients with COVID-19—A narrative review. *Anaesthesia* **2020**, *75*, 1096–1104. [CrossRef] [PubMed]
16. Tian, S.; Hu, W.; Niu, L.; Liu, H.; Xu, H.; Xiao, S.Y. Pulmonary Pathology of Early-Phase 2019 Novel Coronavirus (COVID-19) Pneumonia in Two Patients With Lung Cancer. *J. Thorac. Oncol.* **2020**, *15*, 700–704. [CrossRef] [PubMed]
17. Ackermann, M.; Verleden, S.E.; Kuehnel, M.; Haverich, A.; Welte, T.; Laenger, F.; Vanstapel, A.; Werlein, C.; Stark, H.; Tzankov, A.; et al. Pulmonary Vascular Endothelialitis, Thrombosis, and Angiogenesis in COVID-19. *N. Engl. J. Med.* **2020**, *383*, 120–128. [CrossRef]

18. Aigner, C.; Dittmer, U.; Kamler, M.; Collaud, S.; Taube, C. COVID-19 in a lung transplant recipient. *J. Heart Lung Transpl.* **2020**, *39*, 610–611. [CrossRef]

19. Suri, J.S.; Rangayyan, R.M. *Recent Advances in Breast Imaging, Mammography, and Computer-Aided Diagnosis of Breast Cancer*; SPIE: Bellingham, WA, USA, 2006.

20. Peng, Q.Y.; Wang, X.T.; Zhang, L.N.; Chinese Critical Care Ultrasound Study, G. Findings of lung ultrasonography of novel corona virus pneumonia during the 2019–2020 epidemic. *Intensive Care Med.* **2020**, *46*, 849–850. [CrossRef]

21. Nillmani; Jain, P.K.; Sharma, N.; Kalra, M.K.; Viskovic, K.; Saba, L.; Suri, J.S. Four Types of Multiclass Frameworks for Pneumonia Classification and Its Validation in X-ray Scans Using Seven Types of Deep Learning Artificial Intelligence Models. *Diagnostics* **2022**, *12*, 652. [CrossRef]

22. Suri, J.S.; Agarwal, S.; Gupta, S.K.; Puvvula, A.; Biswas, M.; Saba, L.; Bit, A.; Tandel, G.S.; Agarwal, M.; Patrick, A. A narrative review on characterization of acute respiratory distress syndrome in COVID-19-infected lungs using artificial intelligence. *Comput. Biol. Med.* **2021**, *130*, 104210. [CrossRef]

23. Saba, L.; Biswas, M.; Kuppili, V.; Cuadrado Godia, E.; Suri, H.S.; Edla, D.R.; Omerzu, T.; Laird, J.R.; Khanna, N.N.; Mavrogeni, S.; et al. The present and future of deep learning in radiology. *Eur. J. Radiol.* **2019**, *114*, 14–24. [CrossRef] [PubMed]

24. Kuppili, V.; Biswas, M.; Sreekumar, A.; Suri, H.S.; Saba, L.; Edla, D.R.; Marinhoe, R.T.; Sanches, J.M.; Suri, J.S. Extreme learning machine framework for risk stratification of fatty liver disease using ultrasound tissue characterization. *J. Med. Syst.* **2017**, *41*, 152. [CrossRef] [PubMed]

25. Noor, N.M.; Than, J.C.; Rijal, O.M.; Kassim, R.M.; Yunus, A.; Zeki, A.A.; Anzidei, M.; Saba, L.; Suri, J.S. Automatic lung segmentation using control feedback system: Morphology and texture paradigm. *J. Med. Syst.* **2015**, *39*, 22. [CrossRef] [PubMed]

26. Suri, J.S.; Agarwal, S.; Pathak, R.; Ketireddy, V.; Columbu, M.; Saba, L.; Gupta, S.K.; Faa, G.; Singh, I.M.; Turk, M. COVLIAS 1.0: Lung Segmentation in COVID-19 Computed Tomography Scans Using Hybrid Deep Learning Artificial Intelligence Models. *Diagnostics* **2021**, *11*, 1405. [CrossRef]

27. Suri, J.S.; Agarwal, S.; Carriero, A.; Pasche, A.; Danna, P.S.C.; Columbu, M.; Saba, L.; Viskovic, K.; Mehmedovic, A.; Agarwal, S.; et al. COVLIAS 1.0 vs. MedSeg: Artificial Intelligence-Based Comparative Study for Automated COVID-19 Computed Tomography Lung Segmentation in Italian and Croatian Cohorts. *Diagnostics* **2021**, *11*, 2367. [CrossRef]

28. Saba, L.; Agarwal, M.; Patrick, A.; Puvvula, A.; Gupta, S.K.; Carriero, A.; Laird, J.R.; Kitas, G.D.; Johri, A.M.; Balestrieri, A.; et al. Six artificial intelligence paradigms for tissue characterisation and classification of non-COVID-19 pneumonia against COVID-19 pneumonia in computed tomography lungs. *Int. J. Comput. Assist. Radiol. Surg.* **2021**, *16*, 423–434. [CrossRef]

29. Acharya, R.U.; Faust, O.; Alvin, A.P.; Sree, S.V.; Molinari, F.; Saba, L.; Nicolaides, A.; Suri, J.S. Symptomatic vs. asymptomatic plaque classification in carotid ultrasound. *J. Med. Syst.* **2012**, *36*, 1861–1871. [CrossRef]

30. Acharya, U.R.; Faust, O.; Alvin, A.P.C.; Krishnamurthi, G.; Seabra, J.C.; Sanches, J.; Suri, J.S. Understanding symptomatology of atherosclerotic plaque by image-based tissue characterization. *Comput. Methods Programs Biomed.* **2013**, *110*, 66–75. [CrossRef]

31. Acharya, U.R.; Faust, O.; Sree, S.V.; Alvin, A.P.C.; Krishnamurthi, G.; Sanches, J.; Suri, J.S. Atheromatic™: Symptomatic vs. asymptomatic classification of carotid ultrasound plaque using a combination of HOS, DWT & texture. In Proceedings of the 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Boston, MA, USA, 30 August–3 September 2011; pp. 4489–4492.

32. Acharya, U.R.; Mookiah, M.R.; Vinitha Sree, S.; Afonso, D.; Sanches, J.; Shafique, S.; Nicolaides, A.; Pedro, L.M.; Fernandes, E.F.J.; Suri, J.S. Atherosclerotic plaque tissue characterization in 2D ultrasound longitudinal carotid scans for automated classification: A paradigm for stroke risk assessment. *Med. Biol. Eng. Comput.* **2013**, *51*, 513–523. [CrossRef]

33. Molinari, F.; Liboni, W.; Pavanelli, E.; Giustetto, P.; Badalamenti, S.; Suri, J.S. Accurate and automatic carotid plaque characterization in contrast enhanced 2-D ultrasound images. In Proceedings of the 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Lyon, France, 22–26 August 2007; pp. 335–338.

34. Banchhor, S.K.; Londhe, N.D.; Araki, T.; Saba, L.; Radeva, P.; Laird, J.R.; Suri, J.S. Wall-based measurement features provides an improved IVUS coronary artery risk assessment when fused with plaque texture-based features during machine learning paradigm. *Comput. Biol. Med.* **2017**, *91*, 198–212. [CrossRef]

35. Saba, L.; Sanagala, S.S.; Gupta, S.K.; Koppula, V.K.; Johri, A.M.; Khanna, N.N.; Mavrogeni, S.; Laird, J.R.; Pareek, G.; Miner, M.; et al. Multimodality carotid plaque tissue characterization and classification in the artificial intelligence paradigm: A narrative review for stroke application. *Ann. Transl. Med.* **2021**, *9*, 1206. [CrossRef] [PubMed]

36. Agarwal, M.; Saba, L.; Gupta, S.K.; Johri, A.M.; Khanna, N.N.; Mavrogeni, S.; Laird, J.R.; Pareek, G.; Miner, M.; Sfikakis, P.P. Wilson disease tissue classification and characterization using seven artificial intelligence models embedded with 3D optimization paradigm on a weak training brain magnetic resonance imaging datasets: A supercomputer application. *Med. Biol. Eng. Comput.* **2021**, *59*, 511–533. [CrossRef] [PubMed]

37. Acharya, U.R.; Kannathal, N.; Ng, E.; Min, L.C.; Suri, J.S. Computer-based classification of eye diseases. In Proceedings of the 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, New York, NY, USA, 30 August–3 September 2006; pp. 6121–6124.

38. Acharya, U.; Vinitha Sree, S.; Mookiah, M.; Yantri, R.; Molinari, F.; Zieleźnik, W.; Małyszek-Tumidajewicz, J.; Stępień, B.; Bardales, R.; Witkowska, A. Diagnosis of Hashimoto's thyroiditis in ultrasound using tissue characterization and pixel classification. *Proc. Inst. Mech. Eng. Part H J. Eng. Med.* **2013**, *227*, 788–798. [CrossRef] [PubMed]

39. Biswas, M.; Kuppili, V.; Edla, D.R.; Suri, H.S.; Saba, L.; Marinhoe, R.T.; Sanches, J.M.; Suri, J.S. Symtosis: A liver ultrasound tissue characterization and risk stratification in optimized deep learning paradigm. *Comput. Methods Programs Biomed.* **2018**, *155*, 165–177. [CrossRef] [PubMed]

40. Acharya, U.R.; Saba, L.; Molinari, F.; Guerriero, S.; Suri, J.S. Ovarian tumor characterization and classification: A class of GyneScan™ systems. In Proceedings of the 2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, San Diego, CA, USA, 28 August–1 September 2012; pp. 4446–4449.

41. Pareek, G.; Acharya, U.R.; Sree, S.V.; Swapna, G.; Yantri, R.; Martis, R.J.; Saba, L.; Krishnamurthi, G.; Mallarini, G.; El-Baz, A. Prostate tissue characterization/classification in 144 patient population using wavelet and higher order spectra features from transrectal ultrasound images. *Technol. Cancer Res. Treat.* **2013**, *12*, 545–557. [CrossRef]

42. Shrivastava, V.K.; Londhe, N.D.; Sonawane, R.S.; Suri, J.S. Exploring the color feature power for psoriasis risk stratification and classification: A data mining paradigm. *Comput. Biol. Med.* **2015**, *65*, 54–68. [CrossRef]

43. Shrivastava, V.K.; Londhe, N.D.; Sonawane, R.S.; Suri, J.S. A novel and robust Bayesian approach for segmentation of psoriasis lesions and its risk stratification. *Comput. Methods Programs Biomed.* **2017**, *150*, 9–22. [CrossRef] [PubMed]

44. Shrivastava, V.K.; Londhe, N.D.; Sonawane, R.S.; Suri, J.S. Reliable and accurate psoriasis disease classification in dermatology images using comprehensive feature space in machine learning paradigm. *Expert Syst. Appl.* **2015**, *42*, 6184–6195. [CrossRef]

45. Bayraktaroglu, S.; Cinkooglu, A.; Ceylan, N.; Savas, R. The novel coronavirus pneumonia (COVID-19): A pictorial review of chest CT features. *Diagn. Interv. Radiol.* **2021**, *27*, 188–194. [CrossRef]

46. Verschakelen, J.A.; De Wever, W. *Computed Tomography of the Lung*; Springer: Berlin/Heidelberg, Germany, 2007.

47. Arrieta, A.B.; Díaz-Rodríguez, N.; Del Ser, J.; Bennetot, A.; Tabik, S.; Barbado, A.; García, S.; Gil-López, S.; Molina, D.; Benjamins, R. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. *Inf. Fusion* **2020**, *58*, 82–115. [CrossRef]

48. Choi, T.Y.; Chang, M.Y.; Heo, S.; Jang, J.Y. Explainable machine learning model to predict refeeding hypophosphatemia. *Clin. Nutr. ESPEN* **2021**, *45*, 213–219. [CrossRef] [PubMed]

49. Gunashekar, D.D.; Bielak, L.; Hagele, L.; Oerther, B.; Benndorf, M.; Grosu, A.L.; Brox, T.; Zamboglou, C.; Bock, M. Explainable AI for CNN-based prostate tumor segmentation in multi-parametric MRI correlated to whole mount histopathology. *Radiat. Oncol.* **2022**, *17*, 65. [CrossRef] [PubMed]

50. Gunning, D.; Aha, D. DARPA's explainable artificial intelligence (XAI) program. *AI Mag.* **2019**, *40*, 44–58.

51. Sabih, M.; Hannig, F.; Teich, J. Utilizing explainable AI for quantization and pruning of deep neural networks. *arXiv* **2020**, arXiv:2008.09072.

52. Apley, D.W.; Zhu, J. Visualizing the effects of predictor variables in black box supervised learning models. *J. R. Stat. Soc. Ser. B (Stat. Methodol.)* **2020**, *82*, 1059–1086. [CrossRef]

53. DenOtter, T.D.; Schubert, J. Hounsfield Unit. In *StatPearls*; StatPearls Publishing LLC.: Treasure Island, FL, USA, 2022.

54. Adedigba, A.P.; Adeshina, S.A.; Aina, O.E.; Aibinu, A.M. Optimal hyperparameter selection of deep learning models for COVID-19 chest X-ray classification. *Intell.-Based Med.* **2021**, *5*, 100034. [CrossRef]

55. Chhabra, M.; Kumar, R. A Smart Healthcare System Based on Classifier DenseNet 121 Model to Detect Multiple Diseases. In *Mobile Radio Communications and 5G Networks*; Springer: Berlin/Heidelberg, Germany, 2022; pp. 297–312.

56. Hasan, N.; Bao, Y.; Shawon, A.; Huang, Y. DenseNet convolutional neural networks application for predicting COVID-19 using CT image. *SN Comput. Sci.* **2021**, *2*, 389. [CrossRef]

57. Nandhini, S.; Ashokkumar, K. An automatic plant leaf disease identification using DenseNet-121 architecture with a mutation-based henry gas solubility optimization algorithm. *Neural Comput. Appl.* **2022**, *34*, 5513–5534. [CrossRef]

58. Ruiz, J.; Mahmud, M.; Modasshir, M.; Shamim Kaiser, M.; for the Alzheimer's Disease Neuroimaging Initiative. 3D DenseNet ensemble in 4-way classification of Alzheimer's disease. In *International Conference on Brain Informatics*; Springer: Berlin/Heidelberg, Germany, 2020; pp. 85–96.

59. Jiang, H.; Xu, J.; Shi, R.; Yang, K.; Zhang, D.; Gao, M.; Ma, H.; Qian, W. A multi-label deep learning model with interpretable Grad-CAM for diabetic retinopathy classification. In Proceedings of the 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Montreal, QC, Canada, 20–24 July 2020; pp. 1560–1563.

60. Joo, H.-T.; Kim, K.-J. Visualization of deep reinforcement learning using grad-CAM: How AI plays atari games? In Proceedings of the 2019 IEEE Conference on Games (CoG), London, UK, 20–23 August 2019; pp. 1–2.

61. Panwar, H.; Gupta, P.; Siddiqui, M.K.; Morales-Menendez, R.; Bhardwaj, P.; Singh, V. A deep learning and grad-CAM based color visualization approach for fast detection of COVID-19 cases using chest X-ray and CT-Scan images. *Chaos Solitons Fractals* **2020**, *140*, 110190. [CrossRef]

62. Zhang, Y.; Hong, D.; McClement, D.; Oladosu, O.; Pridham, G.; Slaney, G. Grad-CAM helps interpret the deep learning models trained to classify multiple sclerosis types using clinical brain magnetic resonance imaging. *J. Neurosci. Methods* **2021**, *353*, 109098. [CrossRef]

63. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 618–626.

64. Chattopadhay, A.; Sarkar, A.; Howlader, P.; Balasubramanian, V.N. Grad-cam++: Generalized gradient-based visual explanations for deep convolutional networks. In Proceedings of the 2018 IEEE winter Conference on Applications of Computer Vision (WACV), Lake Tahoe, NV, USA, 12–15 March 2018; pp. 839–847.

65. Inbaraj, X.A.; Villavicencio, C.; Macrohon, J.J.; Jeng, J.-H.; Hsieh, J.-G. Object Identification and Localization Using Grad-CAM++ with Mask Regional Convolution Neural Network. *Electronics* **2021**, *10*, 1541. [CrossRef]

66. Joshua, E.S.N.; Chakkravarthy, M.; Bhattacharyya, D. Lung cancer detection using improvised grad-cam++ with 3d cnn class activation. In *Smart Technologies in Data Science and Communication*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 55–69.

67. Joshua, N.; Stephen, E.; Bhattacharyya, D.; Chakkravarthy, M.; Kim, H.-J. Lung Cancer Classification Using Squeeze and Excitation Convolutional Neural Networks with Grad Cam++ Class Activation Function. *Traitement Signal* **2021**, *38*, 1103–1112. [CrossRef]

68. Wang, H.; Wang, Z.; Du, M.; Yang, F.; Zhang, Z.; Ding, S.; Mardziel, P.; Hu, X. Score-CAM: Score-weighted visual explanations for convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Seattle, WA, USA, 14–19 June 2020; pp. 24–25.

69. Wang, H.; Du, M.; Yang, F.; Zhang, Z. Score-Cam: Improved Visual Explanations via Score-Weighted Class Activation Mapping. *arXiv* **2019**, arXiv:1910.01279.

70. Naidu, R.; Ghosh, A.; Maurya, Y.; Kundu, S.S. IS-CAM: Integrated Score-CAM for axiomatic-based explanations. *arXiv* **2020**, arXiv:2010.03023.

71. Oh, Y.; Jung, H.; Park, J.; Kim, M.S. Evet: Enhancing visual explanations of deep neural networks using image transformations. In Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, Virtual, 5–9 January 2021; pp. 3579–3587.

72. Sugeno, A.; Ishikawa, Y.; Ohshima, T.; Muramatsu, R. Simple methods for the lesion detection and severity grading of diabetic retinopathy by image processing and transfer learning. *Comput. Biol. Med.* **2021**, *137*, 104795. [CrossRef]

73. Cozzi, D.; Cavigli, E.; Moroni, C.; Smorchkova, O.; Zantonelli, G.; Pradella, S.; Miele, V. Ground-glass opacity (GGO): A review of the differential diagnosis in the era of COVID-19. *Jpn. J. Radiol.* **2021**, *39*, 721–732. [CrossRef]

74. Wu, J.; Pan, J.; Teng, D.; Xu, X.; Feng, J.; Chen, Y.C. Interpretation of CT signs of 2019 novel coronavirus (COVID-19) pneumonia. *Eur. Radiol.* **2020**, *30*, 5455–5462. [CrossRef]

75. De Wever, W.; Meersschaert, J.; Coolen, J.; Verbeken, E.; Verschakelen, J.A. The crazy-paving pattern: A radiological-pathological correlation. *Insights Imaging* **2011**, *2*, 117–132. [CrossRef]

76. Niu, R.; Ye, S.; Li, Y.; Ma, H.; Xie, X.; Hu, S.; Huang, X.; Ou, Y.; Chen, J. Chest CT features associated with the clinical characteristics of patients with COVID-19 pneumonia. *Ann. Med.* **2021**, *53*, 169–180. [CrossRef]

77. Salehi, S.; Abedi, A.; Balakrishnan, S.; Gholamrezanezhad, A. Coronavirus Disease 2019 (COVID-19): A Systematic Review of Imaging Findings in 919 Patients. *AJR Am. J. Roentgenol.* **2020**, *215*, 87–93. [CrossRef]

78. Xie, X.; Zhong, Z.; Zhao, W.; Zheng, C.; Wang, F.; Liu, J. Chest CT for Typical Coronavirus Disease 2019 (COVID-19) Pneumonia: Relationship to Negative RT-PCR Testing. *Radiology* **2020**, *296*, E41–E45. [CrossRef] [PubMed]

79. Cau, R.; Falaschi, Z.; Pasche, A.; Danna, P.; Arioli, R.; Arru, C.D.; Zagaria, D.; Tricca, S.; Suri, J.S.; Karla, M.K.; et al. Computed tomography findings of COVID-19 pneumonia in Intensive Care Unit-patients. *J. Public Health Res.* **2021**, *10*, 2270. [CrossRef] [PubMed]

80. Yang, X.; He, X.; Zhao, J.; Zhang, Y.; Zhang, S.; Xie, P. COVID-CT-dataset: A CT scan dataset about COVID-19. *arXiv* **2020**, arXiv:2003.13865.

81. Shalbaf, A.; Vafaeezadeh, M. Automated detection of COVID-19 using ensemble of transfer learning with deep convolutional neural network based on CT scans. *Int. J. Comput. Assist. Radiol. Surg.* **2021**, *16*, 115–123.

82. Gozes, O.; Frid-Adar, M.; Greenspan, H.; Browning, P.D.; Zhang, H.; Ji, W.; Bernheim, A.; Siegel, E. Rapid ai development cycle for the coronavirus (COVID-19) pandemic: Initial results for automated detection & patient monitoring using deep learning ct image analysis. *arXiv* **2020**, arXiv:2003.05037.

83. Solano-Rojas, B.; Villalón-Fonseca, R.; Marín-Raventós, G. Alzheimer's disease early detection using a low cost three-dimensional densenet-121 architecture. In Proceedings of the International Conference on Smart Homes and Health Telematics, Hammamet, Tunisia, 24–26 June 2020; pp. 3–15.

84. Saba, L.; Suri, J.S. *Multi-Detector CT Imaging: Principles, Head, Neck, and Vascular Systems*; CRC Press: Boca Raton, FL, USA, 2013; Volume 1.

85. Murgia, A.; Erta, M.; Suri, J.S.; Gupta, A.; Wintermark, M.; Saba, L. CT imaging features of carotid artery plaque vulnerability. *Ann. Transl. Med.* **2020**, *8*, 1261. [CrossRef]

86. Saba, L.; di Martino, M.; Siotto, P.; Anzidei, M.; Argiolas, G.M.; Porcu, M.; Suri, J.S.; Wintermark, M. Radiation dose and image quality of computed tomography of the supra-aortic arteries: A comparison between single-source and dual-source CT Scanners. *J. Neuroradiol.* **2018**, *45*, 136–141. [CrossRef]

87. Bustin, S.A.; Benes, V.; Nolan, T.; Pfaffl, M.W. Quantitative real-time RT-PCR—A perspective. *J. Mol. Endocrinol.* **2005**, *34*, 597–601. [CrossRef]

88. Dramé, M.; Teguo, M.T.; Proye, E.; Hequet, F.; Hentzien, M.; Kanagaratnam, L.; Godaert, L. Should RT-PCR be considered a gold standard in the diagnosis of COVID-19? *J. Med. Virol.* **2020**, *92*, 2312–2313. [CrossRef]

89. Gibson, U.E.; Heid, C.A.; Williams, P.M. A novel method for real time quantitative RT-PCR. *Genome Res.* **1996**, *6*, 995–1001. [CrossRef]

90. Jain, P.K.; Sharma, N.; Giannopoulos, A.A.; Saba, L.; Nicolaides, A.; Suri, J.S. Hybrid deep learning segmentation models for atherosclerotic plaque in internal carotid artery B-mode ultrasound. *Comput. Biol. Med.* **2021**, *136*, 104721. [CrossRef] [PubMed]

91. Jena, B.; Saxena, S.; Nayak, G.K.; Saba, L.; Sharma, N.; Suri, J.S. Artificial intelligence-based hybrid deep learning models for image classification: The first narrative review. *Comput. Biol. Med.* **2021**, *137*, 104803. [CrossRef] [PubMed]

92. Jain, P.K.; Sharma, N.; Saba, L.; Paraskevas, K.I.; Kalra, M.K.; Johri, A.; Nicolaides, A.N.; Suri, J.S. Automated deep learning-based paradigm for high-risk plaque detection in B-mode common carotid ultrasound scans: An asymptomatic Japanese cohort study. *Int. Angiol.* **2022**, *41*, 9–23. [CrossRef] [PubMed]

93. Suri, J.S.; Agarwal, S.; Elavarthi, P.; Pathak, R.; Ketireddy, V.; Columbu, M.; Saba, L.; Gupta, S.K.; Faa, G.; Singh, I.M.; et al. Inter-Variability Study of COVLIAS 1.0: Hybrid Deep Learning Models for COVID-19 Lung Segmentation in Computed Tomography. *Diagnostics* **2021**, *11*, 2025. [CrossRef] [PubMed]

94. Skandha, S.S.; Nicolaides, A.; Gupta, S.K.; Koppula, V.K.; Saba, L.; Johri, A.M.; Kalra, M.S.; Suri, J.S. A hybrid deep learning paradigm for carotid plaque tissue characterization and its validation in multicenter cohorts using a supercomputer framework. *Comput. Biol. Med.* **2022**, *141*, 105131. [CrossRef]

95. Gupta, N.; Gupta, S.K.; Pathak, R.K.; Jain, V.; Rashidi, P.; Suri, J.S. Human activity recognition in artificial intelligence framework: A narrative review. *Artif. Intell. Rev.* **2022**, 1–54. [CrossRef]

96. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Munich, Germany, 5–9 October 2015; pp. 234–241.

97. Badrinarayanan, V.; Kendall, A.; Cipolla, R. SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.* **2017**, *39*, 2481–2495. [CrossRef]

98. Mateen, M.; Wen, J.; Song, S.; Huang, Z. Fundus image classification using VGG-19 architecture with PCA and SVD. *Symmetry* **2018**, *11*, 1. [CrossRef]

99. Wen, L.; Li, X.; Li, X.; Gao, L. A new transfer learning based on VGG-19 network for fault diagnosis. In Proceedings of the 2019 IEEE 23rd International Conference on Computer Supported Cooperative Work in Design (CSCWD), Porto, Portugal, 6–8 May 2019; pp. 205–209.

100. Xiao, J.; Wang, J.; Cao, S.; Li, B. Application of a novel and improved VGG-19 network in the detection of workers wearing masks. *J. Phys. Conf. Ser.* **2020**, *1518*, 012041. [CrossRef]

101. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

102. Zhou, T.; Ye, X.; Lu, H.; Zheng, X.; Qiu, S.; Liu, Y. Dense Convolutional Network and Its Application in Medical Image Analysis. *Biomed. Res. Int.* **2022**, *2022*, 2384830. [CrossRef]

103. De Boer, P.-T.; Kroese, D.P.; Mannor, S.; Rubinstein, R.Y. A tutorial on the cross-entropy method. *Ann. Oper. Res.* **2005**, *134*, 19–67. [CrossRef]

104. Jamin, A.; Humeau-Heurtier, A. (Multiscale) Cross-Entropy Methods: A Review. *Entropy* **2019**, *22*, 45. [CrossRef] [PubMed]

105. Shore, J.; Johnson, R. Properties of cross-entropy minimization. *IEEE Trans. Inf. Theory* **1981**, *27*, 472–482. [CrossRef]

106. Browne, M.W. Cross-Validation Methods. *J. Math. Psychol.* **2000**, *44*, 108–132. [CrossRef] [PubMed]

107. Refaeilzadeh, P.; Tang, L.; Liu, H. Cross-validation. *Encycl. Database Syst.* **2009**, *5*, 532–538.

108. Powers, D.M. Evaluation: From precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv* **2020**, arXiv:2010.16061.

109. Juba, B.; Le, H.S. Precision-recall versus accuracy and the role of large data sets. In Proceedings of the AAAI Conference on Artificial Intelligence, Honolulu, HI, USA, 27 January–1 February 2019; pp. 4039–4048.

110. Yacouby, R.; Axman, D. Probabilistic extension of precision, recall, and f1 score for more thorough evaluation of classification models. In Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems, Punta Cana, Dominican Republic, 16 March 2020; pp. 79–91.

111. Dewitte, K.; Fierens, C.; Stockl, D.; Thienpont, L.M. Application of the Bland-Altman plot for interpretation of method-comparison studies: A critical investigation of its practice. *Clin. Chem.* **2002**, *48*, 799–801, author reply 801–792. [CrossRef]

112. Giavarina, D. Understanding Bland Altman analysis. *Biochem. Med.* **2015**, *25*, 141–151. [CrossRef]

113. Asuero, A.G.; Sayago, A.; Gonzalez, A. The correlation coefficient: An overview. *Crit. Rev. Anal. Chem.* **2006**, *36*, 41–59. [CrossRef]

114. Taylor, R. Interpretation of the correlation coefficient: A basic review. *J. Diagn. Med. Sonogr.* **1990**, *6*, 35–39. [CrossRef]

115. Eelbode, T.; Bertels, J.; Berman, M.; Vandermeulen, D.; Maes, F.; Bisschops, R.; Blaschko, M.B. Optimization for medical image segmentation: Theory and practice when evaluating with dice score or jaccard index. *IEEE Trans. Med. Imaging* **2020**, *39*, 3679–3690. [CrossRef] [PubMed]

116. Zimmerman, D.W.; Zumbo, B.D. Relative power of the Wilcoxon test, the Friedman test, and repeated-measures ANOVA on ranks. *J. Exp. Educ.* **1993**, *62*, 75–86. [CrossRef]

117. Schemper, M. A generalized Friedman test for data defined by intervals. *Biom. J.* **1984**, *26*, 305–308. [CrossRef]

118. Ishitaki, T.; Oda, T.; Barolli, L. A neural network based user identification for Tor networks: Data analysis using Friedman test. In Proceedings of the 2016 30th International Conference on Advanced Information Networking and Applications Workshops (WAINA), Crans-Montana, Switzerland, 23–25 March 2016; pp. 7–13.

119. Hayes, B. Cloud computing. *Commun. ACM* **2008**, *51*, 9–11. [CrossRef]

120. Saiyeda, A.; Mir, M.A. Cloud computing for deep learning analytics: A survey of current trends and challenges. *Int. J. Adv. Res. Comput. Sci.* **2017**, *8*, 68–72.

121. Karar, M.E.; Alsunaydi, F.; Albusaymi, S.; Alotaibi, S. A new mobile application of agricultural pests recognition using deep learning in cloud computing system. *Alex. Eng. J.* **2021**, *60*, 4423–4432. [CrossRef]

122. Singh, R.P.; Haleem, A.; Javaid, M.; Kataria, R.; Singhal, S. Cloud computing in solving problems of COVID-19 pandemic. *J. Ind. Integr. Manag.* **2021**, *6*, 209–219. [CrossRef]

123. Cresswell, K.; Williams, R.; Sheikh, A. Using cloud technology in health care during the COVID-19 pandemic. *Lancet Digit. Health* **2021**, *3*, e4–e5. [CrossRef]

124. Jamthikar, A.D.; Gupta, D.; Mantella, L.E.; Saba, L.; Laird, J.R.; Johri, A.M.; Suri, J.S. Multiclass machine learning vs. conventional calculators for stroke/CVD risk assessment using carotid plaque predictors with coronary angiography scores as gold standard: A 500 participants study. *Int. J. Cardiovasc. Imaging* **2021**, *37*, 1171–1187. [CrossRef]

125. Jamthikar, A.; Gupta, D.; Khanna, N.N.; Saba, L.; Araki, T.; Viskovic, K.; Suri, H.S.; Gupta, A.; Mavrogeni, S.; Turk, M.; et al. A low-cost machine learning-based cardiovascular/stroke risk assessment system: Integration of conventional factors with image phenotypes. *Cardiovasc. Diagn.* **2019**, *9*, 420–430. [CrossRef]

126. Jamthikar, A.; Gupta, D.; Saba, L.; Khanna, N.N.; Araki, T.; Viskovic, K.; Mavrogeni, S.; Laird, J.R.; Pareek, G.; Miner, M. Cardiovascular/stroke risk predictive calculators: A comparison between statistical and machine learning models. *Cardiovasc. Diagn. Ther.* **2020**, *10*, 919. [CrossRef] [PubMed]

127. Agarwal, M.; Saba, L.; Gupta, S.K.; Carriero, A.; Falaschi, Z.; Paschè, A.; Danna, P.; El-Baz, A.; Naidu, S.; Suri, J.S. A novel block imaging technique using nine artificial intelligence models for COVID-19 disease classification, characterization and severity measurement in lung computed tomography scans on an Italian cohort. *J. Med. Syst.* **2021**, *45*, 28. [CrossRef] [PubMed]

128. Tandel, G.S.; Balestrieri, A.; Jujaray, T.; Khanna, N.N.; Saba, L.; Suri, J.S. Multiclass magnetic resonance imaging brain tumor classification using artificial intelligence paradigm. *Comput. Biol. Med.* **2020**, *122*, 103804. [CrossRef] [PubMed]

129. Suri, J.S.; Bhagawati, M.; Paul, S.; Protogerou, A.D.; Sfikakis, P.P.; Kitas, G.D.; Khanna, N.N.; Ruzsa, Z.; Sharma, A.M.; Saxena, S.; et al. A Powerful Paradigm for Cardiovascular Risk Stratification Using Multiclass, Multi-Label, and Ensemble-Based Machine Learning Paradigms: A Narrative Review. *Diagnostics* **2022**, *12*, 722. [CrossRef]

130. Karlafti, E.; Anagnostis, A.; Kotzakioulafi, E.; Vittoraki, M.C.; Eufraimidou, A.; Kasarjyan, K.; Eufraimidou, K.; Dimitriadou, G.; Kakanis, C.; Anthopoulos, M. Does COVID-19 Clinical Status Associate with Outcome Severity? An Unsupervised Machine Learning Approach for Knowledge Extraction. *J. Pers. Med.* **2021**, *11*, 1380. [CrossRef]

131. Srivastava, S.K.; Singh, S.K.; Suri, J.S. Effect of incremental feature enrichment on healthcare text classification system: A machine learning paradigm. *Comput. Methods Programs Biomed.* **2019**, *172*, 35–51. [CrossRef]

132. Jain, P.K.; Sharma, N.; Saba, L.; Paraskevas, K.I.; Kalra, M.K.; Johri, A.; Laird, J.R.; Nicolaides, A.N.; Suri, J.S. Unseen Artificial Intelligence—Deep Learning Paradigm for Segmentation of Low Atherosclerotic Plaque in Carotid Ultrasound: A Multicenter Cardiovascular Study. *Diagnostics* **2021**, *11*, 2257. [CrossRef]

133. Skandha, S.S.; Gupta, S.K.; Saba, L.; Koppula, V.K.; Johri, A.M.; Khanna, N.N.; Mavrogeni, S.; Laird, J.R.; Pareek, G.; Miner, M. 3-D optimized classification and characterization artificial intelligence paradigm for cardiovascular/stroke risk stratification using carotid ultrasound-based delineated plaque: Atheromatic™ 2.0. *Comput. Biol. Med.* **2020**, *125*, 103958. [CrossRef]

134. Saba, L.; Biswas, M.; Suri, H.S.; Viskovic, K.; Laird, J.R.; Cuadrado-Godia, E.; Nicolaides, A.; Khanna, N.N.; Viswanathan, V.; Suri, J.S. Ultrasound-based carotid stenosis measurement and risk stratification in diabetic cohort: A deep learning paradigm. *Cardiovasc. Diagn.* **2019**, *9*, 439–461. [CrossRef]

135. Sanagala, S.S.; Nicolaides, A.; Gupta, S.K.; Koppula, V.K.; Saba, L.; Agarwal, S.; Johri, A.M.; Kalra, M.S.; Suri, J.S. Ten Fast Transfer Learning Models for Carotid Ultrasound Plaque Tissue Characterization in Augmentation Framework Embedded with Heatmaps for Stroke Risk Stratification. *Diagnostics* **2021**, *11*, 2109. [CrossRef]

136. Alle, S.; Kanakan, A.; Siddiqui, S.; Garg, A.; Karthikeyan, A.; Mehta, P.; Mishra, N.; Chattopadhyay, P.; Devi, P.; Waghdhare, S.; et al. COVID-19 Risk Stratification and Mortality Prediction in Hospitalized Indian Patients: Harnessing clinical data for public health benefits. *PLoS ONE* **2022**, *17*, e0264785. [CrossRef] [PubMed]

137. Sakai, A.; Komatsu, M.; Komatsu, R.; Matsuoka, R.; Yasutomi, S.; Dozen, A.; Shozu, K.; Arakaki, T.; Machino, H.; Asada, K. Medical professional enhancement using explainable artificial intelligence in fetal cardiac ultrasound screening. *Biomedicines* **2022**, *10*, 551. [CrossRef] [PubMed]

138. Gu, R.; Wang, G.; Song, T.; Huang, R.; Aertsen, M.; Deprest, J.; Ourselin, S.; Vercauteren, T.; Zhang, S. CA-Net: Comprehensive attention convolutional neural networks for explainable medical image segmentation. *IEEE Trans. Med. Imaging* **2020**, *40*, 699–711. [CrossRef] [PubMed]

139. González-Nóvoa, J.A.; Busto, L.; Rodríguez-Andina, J.J.; Fariña, J.; Segura, M.; Gómez, V.; Vila, D.; Veiga, C. Using Explainable Machine Learning to Improve Intensive Care Unit Alarm Systems. *Sensors* **2021**, *21*, 7125. [CrossRef] [PubMed]

140. Lu, S.Y.; Zhang, Z.; Zhang, Y.D.; Wang, S.H. CGENet: A Deep Graph Model for COVID-19 Detection Based on Chest CT. *Biology* **2021**, *11*, 33. [CrossRef]

141. Lahsaini, I.; Daho, M.E.H.; Chikh, M.A. Deep transfer learning based classification model for COVID-19 using chest CT-scans. *Pattern Recognit. Lett.* **2021**, *152*, 122–128. [CrossRef]

142. Chollet, F. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 1251–1258.

143. Zhang, Y.D.; Zhang, Z.; Zhang, X.; Wang, S.H. MIDCAN: A multiple input deep convolutional attention network for COVID-19 diagnosis based on chest CT and chest X-ray. *Pattern Recognit. Lett.* **2021**, *150*, 8–16. [CrossRef]

144. Montalbo, F.J.P. Diagnosing COVID-19 chest X-rays with a lightweight truncated DenseNet with partial layer freezing and feature fusion. *Biomed. Signal Process. Control* **2021**, *68*, 102583. [CrossRef]

145. Acharya, U.R.; Joseph, K.P.; Kannathal, N.; Min, L.C.; Suri, J.S. Heart rate variability. *Adv. Card. Signal Process.* **2007**, 121–165. [CrossRef]

146. Li, F.; Liu, M.; Alzheimer's Disease Neuroimaging, I. Alzheimer's disease diagnosis based on multiple cluster dense convolutional networks. *Comput. Med. Imaging Graph.* **2018**, *70*, 101–110. [CrossRef]

147. Hai, J.; Qiao, K.; Chen, J.; Tan, H.; Xu, J.; Zeng, L.; Shi, D.; Yan, B. Fully Convolutional DenseNet with Multiscale Context for Automated Breast Tumor Segmentation. *J. Healthc. Eng.* **2019**, *2019*, 8415485. [CrossRef] [PubMed]

148. Huang, G.; Liu, Z.; Pleiss, G.; van der Maaten, L.; Weinberger, K. Convolutional Networks with Dense Connectivity. *arXiv* **2020**, arXiv:2001.02394. [CrossRef] [PubMed]

149. Aghamohammadi, M.; Madan, M.; Hong, J.K.; Watson, I. Predicting heart attack through explainable artificial intelligence. In Proceedings of the International Conference on Computational Science, Faro, Portugal, 12–14 June 2019; pp. 633–645.

150. Tjoa, E.; Guan, C. A survey on explainable artificial intelligence (xai): Toward medical xai. *IEEE Trans. Neural Netw. Learn. Syst.* **2020**, *32*, 4793–4813. [CrossRef] [PubMed]

151. Lipton, Z.C. The Mythos of Model Interpretability: In machine learning, the concept of interpretability is both important and slippery. *Queue* **2018**, *16*, 31–57. [CrossRef]

152. Chalkiadakis, I. *A Brief Survey of Visualization Methods for Deep Learning Models from the Perspective of Explainable AI*. 2018. Available online: https://xueshu.baidu.com/usercenter/paper/show?paperid=19130tr0dy6h0tt06w3e0e30xa687535 (accessed on 22 May 2022).

153. Duell, J.; Fan, X.; Burnett, B.; Aarts, G.; Zhou, S.-M. A comparison of explanations given by explainable artificial intelligence methods on analysing electronic health records. In Proceedings of the 2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), Virtual, 27–30 July 2021; pp. 1–4.

154. Gilpin, L.H.; Bau, D.; Yuan, B.Z.; Bajwa, A.; Specter, M.; Kagal, L. Explaining explanations: An overview of interpretability of machine learning. In Proceedings of the 2018 IEEE 5th International Conference on Data Science and Advanced Analytics (DSAA), Turin, Italy, 1–3 October 2018; pp. 80–89.

155. Hossain, M.S.; Muhammad, G.; Guizani, N. Explainable AI and mass surveillance system-based healthcare framework to combat COVID-19 like pandemics. *IEEE Netw.* **2020**, *34*, 126–132. [CrossRef]

156. Yang, G.; Ye, Q.; Xia, J. Unbox the black-box for the medical explainable ai via multi-modal and multi-centre data fusion: A mini-review, two showcases and beyond. *Inf. Fusion* **2022**, *77*, 29–52. [CrossRef]

157. Zhang, Y.; Weng, Y.; Lund, J. Applications of Explainable Artificial Intelligence in Diagnosis and Surgery. *Diagnostics* **2022**, *12*, 237. [CrossRef]

158. Lundberg, S.M.; Lee, S.-I. A unified approach to interpreting model predictions. In Proceedings of the 31st International Conference on Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; Volume 30.

159. McInnes, L.; Healy, J.; Melville, J. Umap: Uniform manifold approximation and projection for dimension reduction. *arXiv* **2018**, arXiv:1802.03426.

160. Acharya, U.R.; Mookiah, M.R.; Vinitha Sree, S.; Yanti, R.; Martis, R.J.; Saba, L.; Molinari, F.; Guerriero, S.; Suri, J.S. Evolutionary algorithm-based classifier parameter tuning for automatic ovarian cancer tissue characterization and classification. *Ultraschall Med.* **2014**, *35*, 237–245. [CrossRef]

161. Hu, H.; Peng, R.; Tai, Y.-W.; Tang, C.-K. Network trimming: A data-driven neuron pruning approach towards efficient deep architectures. *arXiv* **2016**, arXiv:1607.03250.

162. Chowdhury, A.; Santamaria-Pang, A.; Kubricht, J.R.; Qiu, J.; Tu, P. Symbolic Semantic Segmentation and Interpretation of COVID-19 Lung Infections in Chest CT volumes based on Emergent Languages. *arXiv* **2020**, arXiv:2008.09866.

163. Jeczmionek, E.; Kowalski, P.A. Flattening Layer Pruning in Convolutional Neural Networks. *Symmetry* **2021**, *13*, 1147. [CrossRef]

164. Jang, Y.; Lee, S.; Kim, J. Compressing Convolutional Neural Networks by Pruning Density Peak Filters. *IEEE Access* **2021**, *9*, 8278–8285. [CrossRef]

165. Brodzicki, A.; Piekarski, M.; Jaworek-Korjakowska, J. The Whale Optimization Algorithm Approach for Deep Neural Networks. *Sensors* **2021**, *21*, 8003. [CrossRef] [PubMed]

166. Kennedy, J.; Eberhart, R. Particle swarm optimization. In Proceedings of the ICNN'95-International Conference on Neural Networks, Perth, WA, Australia, 27 November–1 December 1995; pp. 1942–1948.

167. Mirjalili, S.; Lewis, A. The whale optimization algorithm. *Adv. Eng. Softw.* **2016**, *95*, 51–67. [CrossRef]

168. Price, K.V. Differential evolution. In *Handbook of Optimization*; Springer: Berlin/Heidelberg, Germany, 2013; pp. 187–214.

169. Rana, N.; Latiff, M.S.A.; Abdulhamid, S.i.M.; Chiroma, H. Whale optimization algorithm: A systematic review of contemporary applications, modifications and developments. *Neural Comput. Appl.* **2020**, *32*, 16245–16277. [CrossRef]

170. Zhuang, Z.; Tan, M.; Zhuang, B.; Liu, J.; Guo, Y.; Wu, Q.; Huang, J.; Zhu, J. Discrimination-aware channel pruning for deep neural networks. In Proceedings of the Annual Conference on Neural Information Processing Systems 2018, NeurIPS 2018, Montreal, QC, Canada, 3–8 December 2018; Volume 31.

171. Wu, T.; Li, X.; Zhou, D.; Li, N.; Shi, J. Differential Evolution Based Layer-Wise Weight Pruning for Compressing Deep Neural Networks. *Sensors* **2021**, *21*, 880. [CrossRef]

172. Tung, F.; Mori, G. Deep Neural Network Compression by In-Parallel Pruning-Quantization. *IEEE Trans. Pattern Anal. Mach. Intell.* **2020**, *42*, 568–579. [CrossRef]

173. El-Baz, A.; Gimel'farb, G.; Suri, J.S. *Stochastic Modeling for Medical Image Analysis*, 1st ed.; CRC Press: Boca Raton, FL, USA, 2015.

174. El-Baz, A.S.; Acharya, R.; Mirmehdi, M.; Suri, J.S. *Multi Modality State-of-the-Art Medical Image Segmentation and Registration Methodologies*; Springer Science & Business Media: Boca Raton, FL, USA, 2011; Volume 2.

175. Cau, R.; Faa, G.; Nardi, V.; Balestrieri, A.; Puig, J.; Suri, J.S.; SanFilippo, R.; Saba, L. Long-COVID diagnosis: From diagnostic to advanced AI-driven models. *Eur. J. Radiol.* **2022**, *148*, 110164. [CrossRef]