

Article

COVID-19 Detection from Chest X-ray Images Using Feature Fusion and Deep Learning

Nur-A-Alam ¹, Mominul Ahsan ^{2,*} , Md. Abdul Based ³ , Julfikar Haider ²  and Marcin Kowalski ⁴ 

¹ Department of Computer Science and Engineering, Mawlana Bhashani Science and Technology University, Tangail 1902, Bangladesh; nuraalam.cse@diu-bd.net

² Department of Engineering, Manchester Metropolitan University, Chester St, Manchester M1 5GD, UK; j.haider@mmu.ac.uk

³ Department of Electrical, Electronics and Telecommunication Engineering, Dhaka International University, Dhaka 1205, Bangladesh; based@kth.se

⁴ Institute of Optoelectronics, Military University of Technology, Gen. S. Kaliskiego, 00-908 Warsaw, Poland; marcin.kowalski@wat.edu.pl

* Correspondence: md.ahsan2@mail.dcu.ie

Abstract: Currently, COVID-19 is considered to be the most dangerous and deadly disease for the human body caused by the novel coronavirus. In December 2019, the coronavirus spread rapidly around the world, thought to be originated from Wuhan in China and is responsible for a large number of deaths. Earlier detection of the COVID-19 through accurate diagnosis, particularly for the cases with no obvious symptoms, may decrease the patient's death rate. Chest X-ray images are primarily used for the diagnosis of this disease. This research has proposed a machine vision approach to detect COVID-19 from the chest X-ray images. The features extracted by the histogram-oriented gradient (HOG) and convolutional neural network (CNN) from X-ray images were fused to develop the classification model through training by CNN (VGGNet). Modified anisotropic diffusion filtering (MADF) technique was employed for better edge preservation and reduced noise from the images. A watershed segmentation algorithm was used in order to mark the significant fracture region in the input X-ray images. The testing stage considered generalized data for performance evaluation of the model. Cross-validation analysis revealed that a 5-fold strategy could successfully impair the overfitting problem. This proposed feature fusion using the deep learning technique assured a satisfactory performance in terms of identifying COVID-19 compared to the immediate, relevant works with a testing accuracy of 99.49%, specificity of 95.7% and sensitivity of 93.65%. When compared to other classification techniques, such as ANN, KNN, and SVM, the CNN technique used in this study showed better classification performance. K-fold cross-validation demonstrated that the proposed feature fusion technique (98.36%) provided higher accuracy than the individual feature extraction methods, such as HOG (87.34%) or CNN (93.64%).

Keywords: COVID-19; X-ray image; deep learning; convolutional neural network (CNN); histogram-oriented gradient (HOG); watershed segmentation



Citation: Nur-A-Alam, M.; Ahsan, M.; Based, M.A.; Haider, J.; Kowalski, M. COVID-19 Detection from Chest X-ray Images Using Feature Fusion and Deep Learning. *Sensors* **2021**, *21*, 1480. <https://doi.org/10.3390/s21041480>

Received: 8 January 2021

Accepted: 17 February 2021

Published: 20 February 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

The COVID-19 is a deadly disease caused by the newly recognized coronavirus. In December 2019, coronavirus (SARS-COV-2) infected the human body for the first time, and it can spread principally among humans through the droplets formed by the infected persons when they speak, cough or sneeze [1–6]. As the droplets are too heavy to travel far, they cannot spread person-to-person without coming in close contact [7]. Although the exact time is not yet known, a new study has estimated that the COVID-19 can be viable in the air for up to 3 hours, on copper for 4 hours and up to 72 hours on plastic and stainless steel. However, the exact answers to these questions are still not agreed upon by the general health research community and currently under investigation. COVID-19 attacks the lung

and damages the tissues of an infected person. At the early-stage, some people may not find any symptoms where most of the people had fever and cough as the core symptoms. Other secondary symptoms could be body aches, sore throat, and a headache could be all possible.

At present, COVID-19 disease is increasing daily due to the lack of quick detection methods. All over the world, a huge number of people died of this disease in 2020. The respiratory tract and lungs are the media where the virus can spread easily. As a result, inflammation occurs, and air sacs can be filled with fluid and discharge. The process is responsible for creating an obstacle in oxygen intake. Quick and accurate detection of the virus is a major challenge for doctors and health professionals around the world in order to reduce the death rate caused by this virus.

Due to the global climate changes, people have already been suffering from many other diseases, and the impact created by the COVID-19 is immeasurable. Currently, the virus has spread to almost every country in the world [8]. Recently, all over the world, America, South-East Asia, and Europe have the uppermost number of confirmed COVID-19 cases (Figure 1). On 7 January 2021, more than 85,929,428 confirmed cases of the virus and 1,876,100 deaths were reported by World Health Organization (WHO) due to the disease [8]. At present, further research on an effective screening process is required for diagnosing the virus cases and segregating the affected people. Health professionals and scientists of many countries in the world are attempting to improve their treatment plan and capacity of test through implementing multifunctional testing to stop spreading the virus and for protecting themselves from the deadly virus.

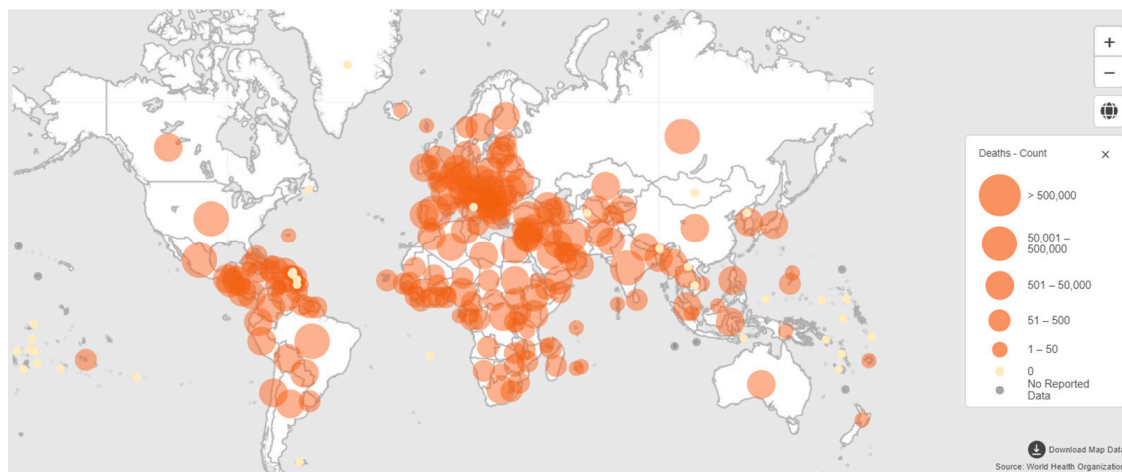


Figure 1. Map for coronavirus-related deaths across the globe reported to WHO on 7 January 2021 (source: World Health Organization, WHO).

Currently, a number of countries have developed vaccines for the COVID-19. Among them, the vaccines developed by Pfizer (USA), AstraZeneca (UK) and Moderna (USA) have been accepted and used in the USA, UK, and many countries in Europe. Based on the clinical-trial data, it has been claimed that the three popular vaccines have achieved the target of 50% efficacy and safe to use without any serious side effects [9,10]. The Pfizer vaccine is required to store at -70°C temperature. This low-temperature storage makes it challenging to transport and store all over the world, particularly in underdeveloped countries. However, the AstraZeneca vaccine requires regular fridge temperature, which will be easier for both carrying and storing worldwide. More recently, the vaccine developed by Sinovac Life Sciences in China has been approved in many countries globally, including Brazil, Indonesia, Colombia, Turkey, Chile, Uruguay and Laos [11]. Furthermore, Sputnik V was developed by Gamaleya Institute, Russia and is currently being used in Belarus, United Arab Emirates, Saudi Arabia, India and Iran [12]. Furthermore, mass vaccination worldwide still remains a huge logistical challenge [13]. Still, large-scale manufacturing is

required to produce the vaccine for covering people all over the world. Further research is required on how long the protection lasts and to find out the effectiveness of the vaccines, particularly against new variants of viruses, which are currently detected in the UK, South Africa, Brazil, and Portugal.

Deep learning with CNN also has been employed in disease diagnosis, such as cancer, via image classification. For example, Li and Shen [14] have proposed two fully convolutional residual networks to produce segmentation, feature extraction and classification result from skin lesion images. A lesion index calculation unit was used to refine the classification results. The results achieved from the deep learning frameworks showed good accuracies (0.912) in cancer diagnosis. Liao et al. [15] proposed a multitask deep learning (MTDL) method to improve diagnosis performance for twelve types of cancer, while their expression data were not adequate. Their experiments showed that the proposed method learned from the aggregation of the expression data for the twelve types of cancer to diagnose cancer accurately. However, the authors did not compare the performance with existing similar works. Yoo et al. [16] have developed an automated CNN-based method for detecting prostate cancer using diffusion-weighted magnetic resonance imaging (DWI) collected from 427 patients. First-order statistical features were extracted from slice-level probabilities, which were used to compile slice-level classification results into the patient level. The proposed method was tested on 108 patients and found good results for both slice level and patient level. However, their system could use 3D CNNs and other deep learning methods to obtain better cancer diagnosis. Esteva et al. [17] have demonstrated skin cancer classification by pre-trained Inception V3 CNN model on 129,450 clinical skin cancer images and 3374 dermatoscopic images. The CNN was trained end-to-end from the images using pixels and disease labels as inputs. The overall CNN accuracy was found as $72.1 \pm 0.9\%$ (mean \pm s.d.), whereas accuracies of 65.56% and 66.0% were achieved by two dermatologists on a subset of the validation set.

Recently, the reverse transcriptase-polymerase chain reaction (RT-PCR) diagnostic method is found to be effective in detecting the virus. However, the method has some drawbacks, including longer detection time and lower detection rate of the virus. Strict requirements in the laboratory and diverse characteristics of the testing could be attributed to the drawbacks [18,19]. Researchers are working on overcoming the limitations of RT-PCR testing to enhance diagnosing and detection of the COVID-19. According to the recommendations by WHO provided in October 2020, chest imaging examination is an effective method for the detection of clinical symptoms of people who have been affected and recovered from the virus [20]. Furthermore, other diagnostics tests are also suggested, including ultrasound, X-rays and MRI of the chest and computed tomography (CT) and needle biopsy of the lung. At present, chest X-ray is extensively used for the detection of the COVID-19 cases compared to the CT image as it takes longer for imaging, and CT scanners are not available in many underdeveloped countries. In addition, CT imaging is highly costly, and pregnant women and children may face health risks due to its high radiation [21]. On the contrary, X-ray imaging has played a great role in many medical and epidemiological cases due to its wider availability [22,23]. Chest X-ray is promising for emergency cases and treatment due to its operational speed, cost and simplicity for the radiologists. However, in prior research, some inconsistencies were observed for the chest X-ray images taken from people affected by the COVID-19 [24].

In the past, artificial intelligence (AI) techniques were employed to successfully diagnose Pneumonia either from chest X-ray images or CT [25–27]. The classification methods employed vary from Bayesian function to convolutional neural network (CNN). More recently, CNN has been found to be useful and effective in identifying COVID-19 via image classification. CNN consists of multilayer neural networks, which are highly capable of recognizing the image patterns without conducting diverse preprocessing of the images. Although several CNN models, including AlexNet, Resnet50, VGG16, VGG19, are available, VGG19 demonstrates better performance for the COVID-19 classification [18,26–28].

The proposed work here provides an intelligent machine learning architecture in order to detect COVID-19 disease using chest X-ray images. The method proposes a novel fusion of features extracted by histogram-oriented gradient (HOG) and CNN and classification by CNN. Furthermore, a modified anisotropic diffusion filtering (MADF) technique was applied to eliminate multiplicative speckle noise from the test images. The watershed segmentation technique was used to identify the fractured lung regions, which could further provide evidence for the COVID-19 attacked lungs.

The remaining sections of this research work are formulated as follows: A comprehensive study on the classification of chest X-ray images is presented in Section 2. Section 3 describes the system architecture used to identify COVID-19 from X-ray image datasets by classifying it into the COVID-19 and other classes. Section 4 presents the results with analysis achieved from this work. Then, Section 5 discusses validation, comparative performances, and limitations of the proposed system in detecting COVID-19. Finally, conclusions drawn from this research are presented in Section 6.

2. Related Literature

In recent months, researchers have investigated and analyzed chest X-ray images using deep learning algorithms to detect COVID-19. First, the images are preprocessed using the CNN technique for extracting better features, which are fed in deep learning algorithms for image classification. Ahammed et al. [29] proposed a deep neural network-based system where CNN provided high accuracy (94.03%). The authors trained the system with normal, pneumonia and COVID-19 patient's chest X-ray images. The limitation of the work was that a dataset with only 285 images was used for developing the system, and this small number of data was not perfect for training a deep learning-based system for the COVID-19 prediction.

Chowdhury et al. [30] worked with chest X-ray images to develop a novel framework named PDCOVIDNet based on parallel-dilated CNN. In the proposed method, the authors used a dilated convolution in the parallel stack that could capture and stretch necessary features for obtaining a detection accuracy of 96.58%.

Abbas et al. [31] proposed and validated a deep convolutional neural network called decompose, transfer, and compose (DeTraC) to detect COVID-19 patients from their chest X-ray images. They proposed a decomposition mechanism to check irregularities from the dataset by investigating class boundaries for obtaining a high accuracy (93.1%) and sensitivity (100%).

Azemin et al. [32] used a deep learning method based on the ResNet-101 CNN model. In their proposed method, thousands of images were used in the pre-trained phase to recognize meaningful objects and retrained to detect abnormality in the chest X-ray images. The accuracy of this method was only 71.9%.

El-Rashidy et al. [33] introduced a framework consisted of three layers: patient layer, cloud layer and hospital layer. A set of data was collected from the patient layer using some wearable sensors and a mobile app. A neural network-based deep learning model was used to detect COVID-19 using the patient X-ray images. The proposed model achieved 97.9% accuracy and 98.85% specificity.

Khan et al. [34] developed a new architecture for the diagnosis of X-ray images as the COVID-19 or normal using pre-trained deep learning models like ResNet50, VGG16, VGG19 and DensNet121, where VGG16 and VGG19 showed the best accuracies. The proposed model consisted of two phases like preprocessing and data augmentation, and transfer learning, and finally showed 99.3% accuracy.

In the proposed model by Loey et al. [35], three deep transfer models like AlexNet, GoogleNet and ResNet18 were employed on a dataset of 307 images with four different types of classes: COVID-19, normal, pneumonia bacterial and pneumonia virus. The research work was distributed into three scenarios to reduce memory consumption and execution time. At the last deep transfer model, GoogleNet achieved 100% testing accuracy and 99.9% validation accuracy.

Minaee et al. [36] reported a deep learning-based framework to detect COVID-19 from chest X-ray images using four tuning models like ResNet18, ResNet50, SqueezeNet and DensNet-121. The proposed method took advantage of data augmentation to create a transformed version of the COVID-19 images, which increased the number of samples and finally achieved 98% sensitivity and 90% specificity.

Sekeroglu et al. [37] developed a model using deep learning and machine learning classifiers where a total of 38 experiments was conducted by CNN for the detection of the COVID-19 using the chest X-ray images with high accuracy. Among them, 10 experiments were performed using 5 different machine-learning algorithms, and 14 experiments were carried out by the state-of-the-art pre-trained network for transfer learning. The system demonstrated 98.50% accuracy, 99.18% specificity and 93.84% sensitivity. They concluded that the system developed by CNN was capable of achieving COVID-19 detection from a limited number of images without any preprocessing and with minimized layers.

Wang et al. [38] developed a model using ResNet-101 and ResNet-151 with fusion effects to enhance their weight ratio dynamically. Classification of the chest X-ray images was carried out based on three classes, such as normal, COVID-19 and viral pneumonia. Performance accuracy of 96.1% was achieved during the testing phase.

Yoo et al. [39] applied chest X-ray radiography (CXR) images to classify using a deep learning-based decision-tree classifier for detecting COVID-19. This classifier compared three binary decision trees based on the PyTorch frame. The decision tree classified CXR images as normal or abnormal, where the third decision tree achieved an average accuracy of 95%.

Khalifa et al. [40] developed a classification approach for the treatment purposes of coronavirus on a single human cell-based on treatment type and treatment concentration level using deep learning and machine learning (ML) methods. Numerical features of the data sets were converted to images for building the DCNN model. The testing accuracy of treatment classification obtained by the model was as high as 98.05% compared to the other traditional ML methods, including support vector machine (SVM) and decision tree (DT). However, the proposed DCNN model showed less testing accuracy (98.2%) compared to the DT (98.5%) for the prediction of treatment concentration level. Deep transfer models (i.e., Alexnet) have not been employed in their study.

Wang et al. [41] have developed a transfer learning method (Xception model) using deep learning models for diagnosing COVID-19. The proposed method showed 96.75% diagnostics accuracy. Furthermore, Deep features and machine learning classification (Xception + SVM) were also employed to develop an efficient diagnostic method for improving the accuracy of the Xception model by 2.58%. From the result, the authors claimed that their proposed method attained higher classification accuracy and efficient diagnostic performance of the COVID-19. However, the authors have not compared their results with the existing similar works.

Sahlol et al. [42] proposed an improved hybrid classification approach using CNNs and marine predators algorithm for classifying COVID-19 images, which were obtained from international cardiothoracic radiologists. Inception architecture of CNNs was employed to extract features, and a swarm-based marine predators algorithm was used to select the most relevant features from the images. However, the research work did not consider any fusion approach to improve the classification and feature extraction of the COVID-19 images.

Most of the reported work in the literature has used chest X-ray images to diagnose COVID-19, and this highlights the importance of chest X-ray image analysis as an indispensable tool for doctors and radiographers. However, imbalance in data manipulation and lack of necessary extracted features from the images sometimes cannot provide expected accuracy in the classification result. To overcome these limitations, this work proposed fusion of features extracted by HOG and CNN and classify using CNN for improving the detection accuracy of the COVID-19.

3. Proposed Methodology

3.1. System Architecture

The proposed system considered input of the X-ray images to identify COVID-19. First of all, this system converted images from RGB to grayscale and identified the region of interest (ROI) by removing the unwanted regions. Furthermore, the system considered two feature extractors: histogram-oriented gradient (HOG) and CNN. First, the HOG technique was used to extract a feature vector from the X-ray COVID-19 dataset. Then the CNN method was used to extract another feature vector from the same images. These two features were fused and used as the input to train the classification model. The number of features extracted by one technique was not large enough to accurately identify COVID-19. However, the fusion approach of extracting features by two different techniques could provide a large number of features for accurate identification. Fusion was considered as a concatenation between the two individual vectors in this context.

Speckle-affected and low-quality X-ray images along with good quality images were used in our experiment for conducting tests. If training and testing are performed with only selected good quality X-ray images in an ideal situation, the output accuracy may be found higher. However, this does not represent a real-life scenario, where the image database would be a mix of both good- and poor-quality images. Therefore, this approach of using different quality images would test how well the system can react to such real-life situations.

A modified anisotropic diffusion filtering technique was employed to remove multiplicative speckle noise from the test images. The application of these techniques could effectively overcome the limitations in input image quality. Next, the feature extraction was carried out on the test images. Finally, the CNN classifier performed a classification of X-ray images to identify whether it was COVID-19 or not. Figure 2 shows the basic steps of the proposed system architecture, which is also represented by Algorithm 1.

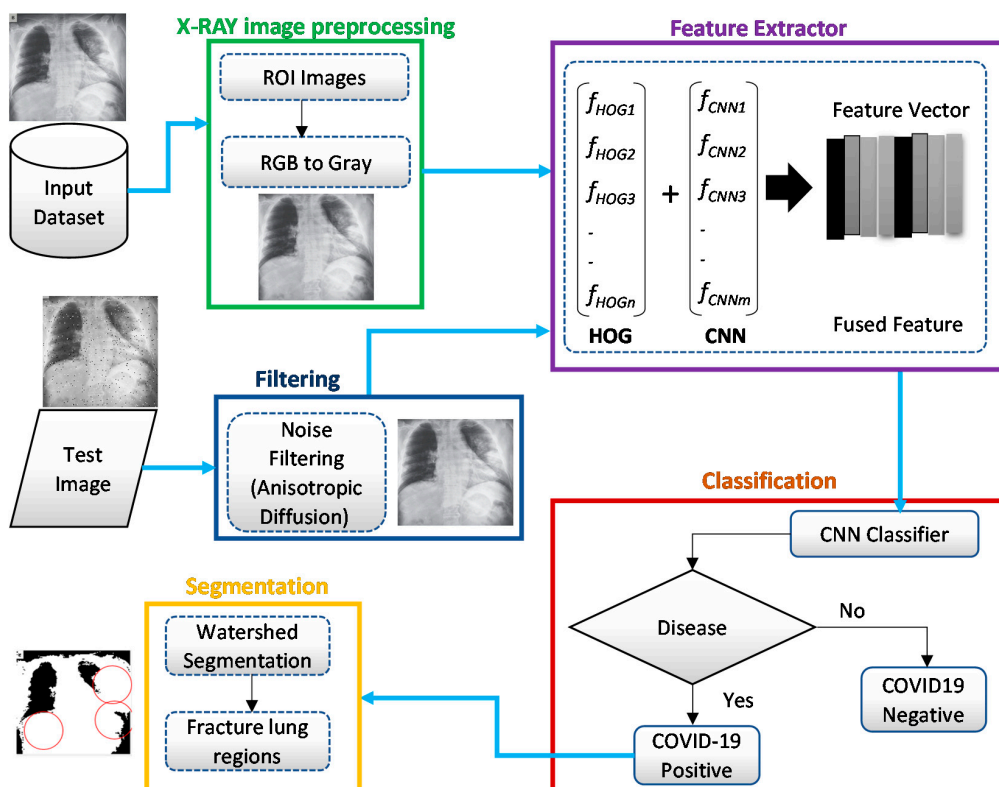


Figure 2. Overview of the proposed intelligent system architecture for identifying COVID-19 from chest X-ray images.

3.2. Dataset Used

The chest X-ray images of the patients were acquired and stored in a commonplace. The images were categorized as either COVID-19-positive or negative as a reference to evaluate the performance of the intelligent system. In this work, three standard datasets were employed to validate the system's performance.

(1) The benchmark data set [43] used in our experimental evaluation consisted of two main categories with 819 COVID-19-positive and 1341 normal chest X-ray images;

(2) Cohen's data set [44] contained a total of 660 images with 390 positive COVID-19 X-ray images;

(3) Another publicly available [45] dataset was used with 770 images of the COVID-19 and 1500 normal images.

The databases contained various sizes of images ranging from 512×512 pixels to 657×657 pixels. The acquired images were in both grayscale and RGB formats, and the RGB images were converted to grayscale images. Any feature extraction method can easily detect features from grayscale images compared to the images in other formats. To convert RGB to grayscale image, Equation (1) is used for calculating the grayscale value (I) by forming a weighted sum of the monochrome colors, red (R), green (G), and blue (B).

$$I = (W_r \cdot R) + (W_g \cdot G) + (W_b \cdot B) \quad (1)$$

W_r , W_g , and W_b are the weights of red, green, and blue colors, respectively, with a value of 0.30, 0.59, and 0.11, summing to a total equal to 1.

Furthermore, the data formats of the images included png and jpeg with bit depths of 8-bit (grayscale) and 24-bit (RGB). As the image size, format and bit depth were different in the databases, they were converted to a size of 224×224 pixels with 8-bit grayscale images and saved in png format.

Algorithm 1 Proposed Algorithm for COVID-19 Detection

Input: COVID-19 Chest X-ray image dataset (D) with resize image (M)

Extraction: Extract Feature Matrix (f).

CNN Feature Vector (F_c).

Step 1: Initialize $F_c \geq M_i, i = 1$

Step 2: Extract each image feature $D(i,1,570)$.

Step 3: $F_c(i,1) = M(x,1) + F_c(i,1)$.

Step 4: F_c = overall CNN features.

Histogram Oriented Gradient (HOG).

Step 1: Initialize. $H_0 = \text{Low pass output}, H_1 = \text{Band pass output}$.

Step 2: $HOG(i,1) = H_0(i,1) + H_1(i,1)$.

Step 3: HOG = overall Histogram Oriented Gradient

Fusion of features in Vector (V).

Training feature (V) = $[F_c, HOG]$.

test_image = $imread(img)$.

Extract test feature (T) = repeat step 1, 2 from test_image.

result (i) = $classify(training\ feature, T)$.

Output: result (i) = **COVID19 Positive or Normal**.

As the number of images available in the open repository was limited, the images from all three databases were combined to create a database for this work. The training and testing stages used both classes: COVID-19-positive and COVID-19-negative. The COVID-19-positive class contained 1979 images, and the COVID-19-negative class contained 3111 images. All data sets were divided into 0.8 portions and 0.2 portions for the training and testing, respectively. A computer having 64-bit windows, 8 GB RAM, Intel Core i5 CPU with a processing speed of 2.60 GHz was used for performing all training and testing. MATLAB 2019b was used to execute all necessary experiments. The experimentation times are computed with a GPU hardware configuration of GEFORCE RTX 2070

super. Figure 3 illustrates a comparison of the COVID-19 and normal chest X-ray images. In general, similar to pneumonia, the density of the lungs is increased in the case of the COVID-19, which causes whiteness in the lungs on radiography. An experienced radiologist can confirm the disease by the appearance of a ground-glass pattern (ground-glass opacity) due to the increased whiteness [46].

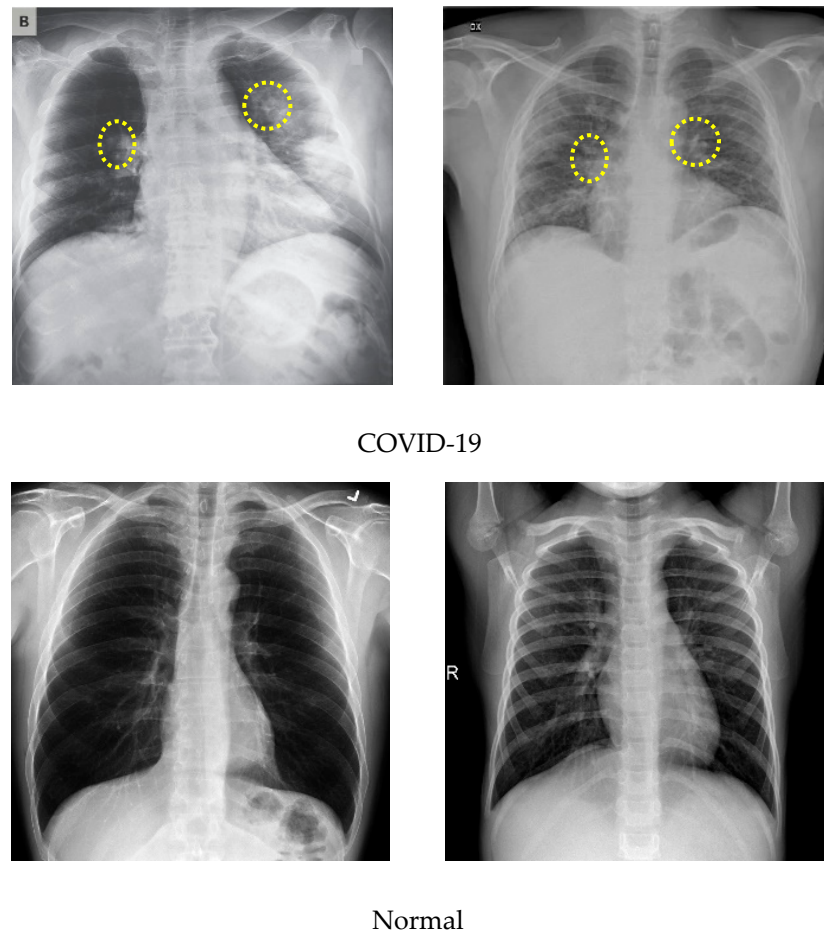


Figure 3. Comparison of the COVID-19 and normal X-ray images.

3.3. Data Preprocessing

Image processing is an important step to achieve meaningful information and accurate classification by removing noisy or deformed pixels from each image. First, the images were converted from RGB to grayscale using the MATLAB tool and resized to 224×224 pixels to be made ready as input to the system.

To eliminate superfluous text and machine annotations around images, the region of interest (ROI) was extracted for training and testing. In order to obtain meaningful information, the ROI on the chest X-ray images was defined by an area covering mainly the lung region. First, an ROI is defined by a rectangle, and a mask is created from the rectangle. Using logical indexing, the area outside the ROI was set to zero, and the extracted portion is displayed. Figure 4 illustrates example images at different preprocessing stages. For example, unnecessary symbols (tick mark in normal image) or text (B in the COVID-19 image) in the original images were removed at the ROI stage. As the images used in this study were collected from three different sources, they might differ in quality, size, or inherent noise. Therefore, the preprocessing approaches employed would normalize all the images such that they were independent of their origin, and the image size effect on the system's performance could be avoided [39].

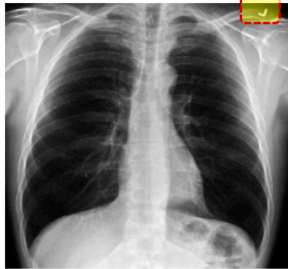




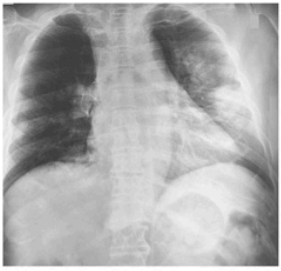
Image Category	Original Image	ROI Image	RGB to Gray
Normal			
COVID-19			

Figure 4. Proposed preprocessing stages: original image, the region of interest (ROI) image and 24-bit (RGB) to gray.

3.4. Modified Anisotropic Diffusion Filtering (MADF)

Filtering techniques preserve useful information while filtering any noise within an image. To extract the meaningful features of any noisy image, the information-preserving filtering techniques are most applicable [21]. During the testing phase, speckle-affected test images were used to assess filtering performance. Anisotropic diffusion filtering can preserve and enhance edge information while subduing any noise [24,47]. The gradient operator detects the edge information along with noise [48]. This technique finds the gradient changes in noise for strong speckle and low contrast images, which may go beyond the gradient of edge. These changes destruct the edge information than the noise, which provides less accuracy in filtering results. Similarly, speckle reducing anisotropic diffusion (SRAD) cannot fully preserve the edge information due to image over-smoothing [49]. Oriented-based non-local means (OBNLM) fails to hold detailed information and is affected by moving noise [27]. Anisotropic diffusion with memory-based speckle statistic (ADMSS) is sharper in white pixels [50].

Modified anisotropic diffusion filtering (MADF) was proposed for this work in order to preserve detailed information while reducing noise and distortion from the images. This filtering technique performs better than the other filtering methods due to its capability in eliminating multiplicative speckle noise in plane regions. The proposed method uses correlation and kurtosis values of noise to hold the useful edge information. In Equation (2), I_0 is a noisy image comprised of speckle-noise n and the original image I [27,49]. The noise part is denoted by Equation (3), where G is noise intensity and calculated from image properties in MATLAB. The mean of noise intensity is μ , which is calculated by Equation (4). Kurtosis k is calculated using Equation (5). The correlation between the image class and noise class should be minimum, which is the iteration stopping condition. This speckle suppression process continues until the noise part of the image is close to the Gaussian value. In this situation, the kurtosis value should be zero. The iteration cutoff is defined when the kurtosis value falls below 0.001 (Equation (6)), indicating a low speckle with better edge preservation. As soon as the correlation between image class and noise class is the least, the iteration will be stopped. Equation (7) calculates the correlation of image intensities (ρ_I)

and Equation (8) calculates the correlation of noise intensities (ρ_G). The proposed filtering will get the optimal result when ρ_I and ρ_G show minimum deviance.

$$I_0 = I.n \quad (2)$$

$$n = \frac{I - G}{\sqrt{G}} \quad (3)$$

$$\mu = \frac{\sum_{i=1}^N G_i}{N} \quad (4)$$

$$k = \frac{\frac{1}{N} \sum_{i=0}^N (G - \mu)^4}{\left[\frac{1}{N} \sum_{i=0}^N (G - \mu)^2 \right]^2} - 3 \quad (5)$$

$$\text{abs}(n - k) \leq 0.001 \quad (6)$$

$$\rho_I = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} i.j.p_I(i,j) - \mu I_x \mu I_y}{\frac{\sum_{i=1}^N (I_{ix} - \mu I_x)(I_{iy} - \mu I_y)}{N}} \quad (7)$$

$$\rho_G = \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} i.j.p_G(i,j) - \mu G_x \mu G_y}{\frac{\sum_{i=1}^N (G_{ix} - \mu I_x)(G_{iy} - \mu I_y)}{N}} \quad (8)$$

Figure 5 shows an example of original images and a comparison of different Anisotropic Diffusion techniques. It was very clear that the edge preservation capability of the proposed MADF technique was much better than the other techniques.



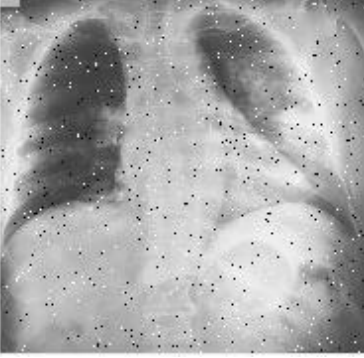

Filtering Processes applied	Image Category	
	COVID-19	Normal
Original image		
Noisy image		

Figure 5. Cont.

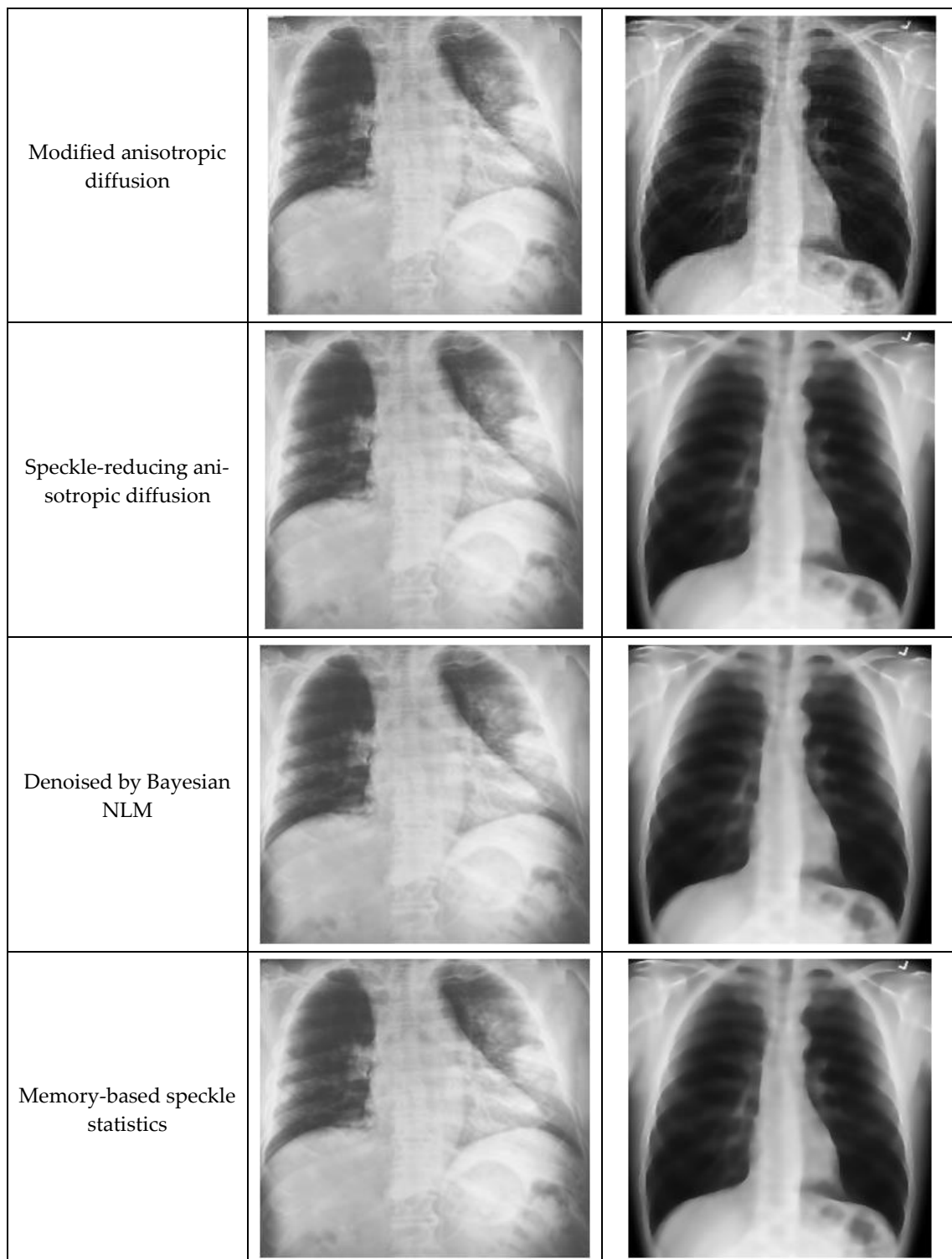


Figure 5. Illustration of images after applying different anisotropic diffusion techniques.

Upon applying MADF, the images are divided into multiple small parts of images, which is called a gradient. Then these parts are filtered one-by-one. After filtering all parts, they are merged together. In general, the anisotropic diffusion filtering method removes all noises and edge information from the images. If this method is applied to the image with many iterations, the edge information in the images is removed. If the k value is 0, then all image features or in-formation is removed. Therefore, filtering continues up to a standard k value of 0.001 in order to obtain an appropriate filtered image. In general, it is

quite difficult to extract appropriate features from the blur images, whereas the MADF method solves this issue.

3.5. Feature Extractor

3.5.1. Histogram-Oriented Gradient (HOG) Feature Extractor

Histogram-oriented gradient (HOG) system extracts features by using a selective number of histogram bins [51]. For extracting HOG features, the proposed system used a higher number of histogram bins on different regions of the images. First, the input image was scaled to 64×128 pixels and converted into a grayscale image. The gradient for every pixel in the image was calculated using Equations (9) and (10).

$$d_x = I(x + 1, y) - I(x, y) \quad (9)$$

$$d_y = I(x, y + 1) - I(x, y) \quad (10)$$

where d_x and d_y are the horizontal and vertical gradient, respectively, and $I(x, y)$ is the pixel value at (x, y) position. The gradient orientation, θ , is then calculated using Equation (11):

$$\theta(x, y) = \tan^{-1} \frac{d_y}{d_x} \quad (11)$$

The basic flow of the HOG feature extraction algorithm is shown in Figure 6.

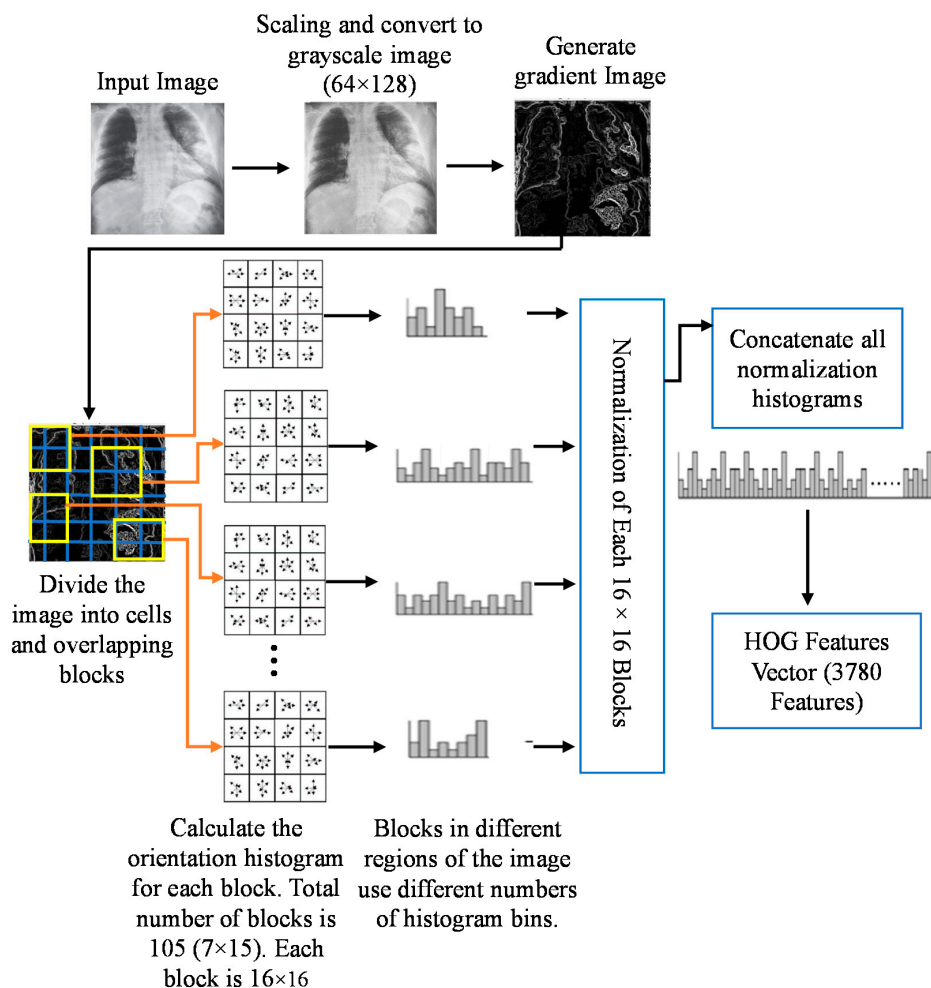


Figure 6. Basic flow of histogram-oriented gradient (HOG) feature extraction algorithm.

The gradient image was divided into cells size of 8×8 pixels to generate the histogram. As a result, a 9×1 matrix for each cell was obtained. Then the gradients were normalized by taking 16×16 blocks. Four 8×8 cells were combined to create a 16×16 block. Therefore, four 9×1 matrices or a single 36×1 matrix could be obtained. Mathematically, for a given vector V (Equation (12)), the root of the sum of squares (Equation (13)) can be calculated.

$$V = [a_1, a_2, a_3, \dots, a_{36}] \quad (12)$$

$$k = \sqrt{(a_1)^2 + (a_2)^2 + (a_3)^2 + \dots + (a_{36})^2} \quad (13)$$

After dividing all the values in vector V with the k value (Euclidian norm or the length of the vector), the resultant would be a normalized vector of size 36×1 as defined by Equation (14).

$$\text{Normalised Vector} = \left(\frac{a_1}{K}, \frac{a_2}{K}, \frac{a_3}{K}, \frac{a_4}{K}, \dots, \frac{a_{36}}{K} \right) \quad (14)$$

In this case, a total of 105 (7×15) blocks of 16×16 can be obtained. Each of these 105 blocks will have a matrix of 36×1 as features. Thus, the total number of features for the image would be $105 \times 36 \times 1 = 3780$.

3.5.2. CNN Based Feature Extractor and Classification

Image processing, particularly features extraction by employing CNN, is an important research topic in computer science [52]. An experiment was conducted using scratch and pre-trained CNN models in this proposed work. The results achieved by the scratch model were not satisfactory; however, the pre-trained model showed good performance.

VGG19 model (pre-trained) was fine-tuned to suit as a feature extractor for the experimental dataset used in this study. A VGGNet consisting of 19-layer was used to develop this network model. Experimental trial asserted that VGG19 showed better performance compared to VGG16, scratch model and other deep learning models, including ResNet50 and AlexNet. The VGG19 model was developed using sixteen convolution layers with three fully connected layers (Figure 7). A nonlinear ReLU was used in the activation function for getting the output of convolution layers, whereas the convolution part was split by five consecutive max-pooling layers. Two convolution layers were used to develop the first and second subregions, where the depth of the layers was 64 and 128. Furthermore, four consecutive convolution layers were used to build the remaining three subregions where the depth of the layers were 256, 512, and 512, respectively. Afterward, Pooling layers were employed to decrease the learnable parameter. The last layer of the proposed VGG19 model helped in obtaining the feature vector, whereas 1024 and 512 neurons existed in the two hidden layers placed before the feature collection layer. For reducing the overfitting during the implementation of the fine-tuned model, L2 regularization was employed after each fully connected layer. The CNN-based VGG19 models provide 4096 appropriate features.

3.6. Feature Fusion and Classification

Data fusion was applied in several machine learning and computer vision applications [53]. Particularly, feature fusion can combine more than one feature vector. Two feature extractors provide a feature vector of 1×4096 and 1×3780 . The feature selection process was mathematically explained by Equations (15)–(17) [54]. Equations (15) and (16) represent features extracted by HOG and CNN, respectively. The extracted feature vectors are combined by concatenation and represented by Equation (17).

$$f_{HOG\ 1 \times n} = \{ HOG_{1 \times 1}, HOG_{1 \times 2}, HOG_{1 \times 3} \dots \dots \dots HOG_{1 \times n} \} \quad (15)$$

$$f_{VGG19\ 1 \times m} = \{ VGG19_{1 \times 1}, VGG19_{1 \times 2}, VGG19_{1 \times 3} \dots \dots \dots VGG19_{1 \times m} \} \quad (16)$$

$$\text{Fused (features vector)}_{1 \times q}^{cat} = \{ f_{HOG\ 1 \times n}, f_{VGG19\ 1 \times m} \} \quad (17)$$

Then the features extracted by HOG and CNN are fused with 7876 features. 1186 score-based features were selected out of 7876 features based on maximum entropy. When the value of $i = 1$, it recalls HOG features and when $i = 2$, it recalls VGG19 features and finally adds them together. For the purpose of selecting optimal features, entropy was employed considering score values. The probability of features and entropy is defined by Equations (18) and (19). The final selected features were fed to the classifiers in order to identify COVID-19 images.

$$B_{He} = -N H e_b \sum_{i=1}^n p(f_i) \quad (18)$$

$$F_{select} = B_{He}(\max(f_i, 1186)) \quad (19)$$

where f is fused vector (1×1186), p denotes features probability, and He represents entropy. This entropy-based feature selection method selects similar or highly related features from the fused vector. Unrelated features were removed, and appropriate features were considered for classification. In this proposed system, the appropriate feature vector is 1×1186 . CNN classifier uses the selected features where the process of feature extraction and selection is shown in Figure 8. Furthermore, the feature vectors obtained by HOG and deep learning were fused to validate the proposed approach in this work.

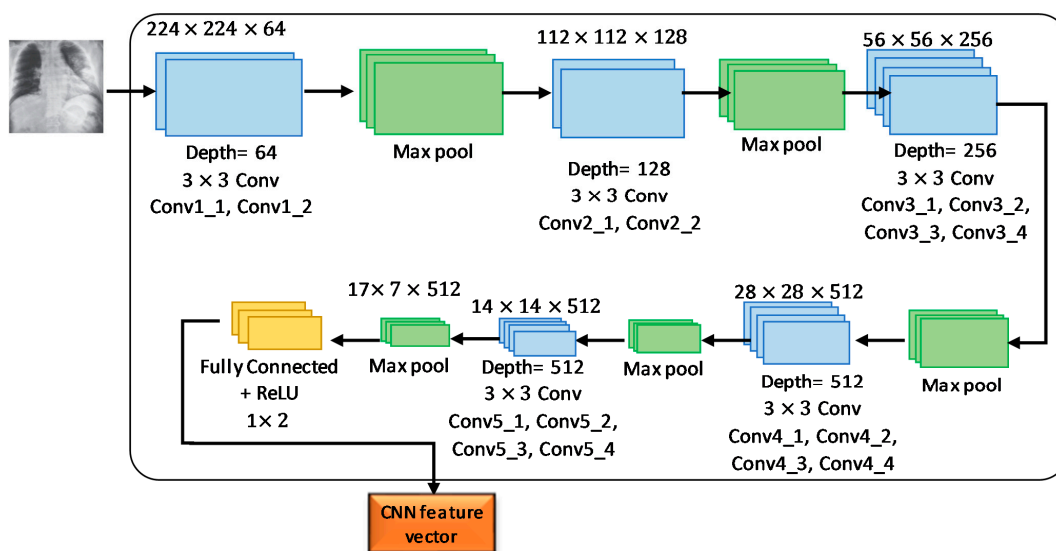


Figure 7. Illustration of image feature extraction by VGG19 pre-trained convolutional neural network (CNN) model.

The VGG19 architecture consisted of 16 layers of CNN, 3 connected layers with 1 final output layer for conducting SoftMax function. No changes in the number of connected layers and final layer are required to build the network architectures. In addition, 2×2 -pixel windows with stride 2 were used for the max-pooling layer. The first two layers and third layer of the three fully connected layers provided 4096 features and 1000 channels, respectively. The final layer represents the output layer with two neurons (COVID-19 and normal).

3.7. Segmentation of the COVID-19-Affected Region

For biomedical image segmentation, the watershed technique [55] provides better results compared to the other techniques, such as Fuzzy-C means (FCM). The conventional FCM algorithm suffers from some weaknesses in terms of initializing clusters center or determining an optimal number of clusters and sensitiveness to noise [56]. FCM segmentation method cannot detect the fracture regions from the X-ray images affected by the COVID-19. However, watershed segmentation is a fast, simple and intuitive method, which provides closed contours, requires low computational time and produces a complete division of the

image in separated regions. Segmentation was applied for the non-trivial task of separating the fracture lung regions from the X-ray images. A watershed segmentation technique was applied to segment the fracture regions of each image owing to its relatively less computational complexity and capability of providing high accuracy in segmentation. This method separated touching objects in an X-ray image and provided a complete division. Figure 9 presents different stages of image processing from filtering to segmentation of significant regions from the COVID-19-affected lung X-ray images.

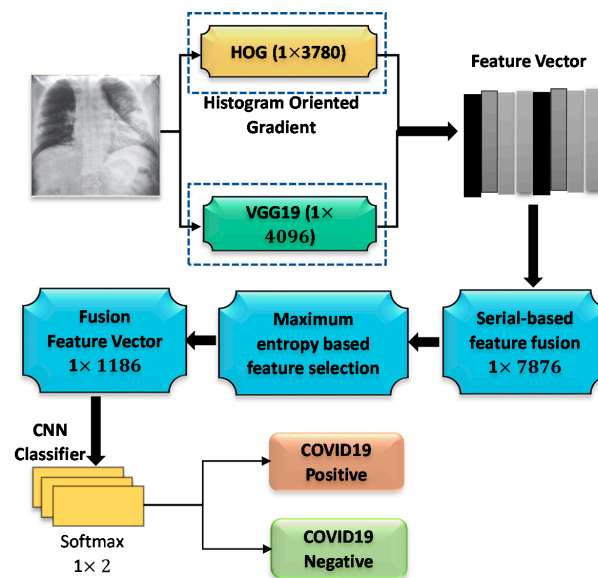


Figure 8. The proposed fusion steps for the features extracted by HOG and CNN.

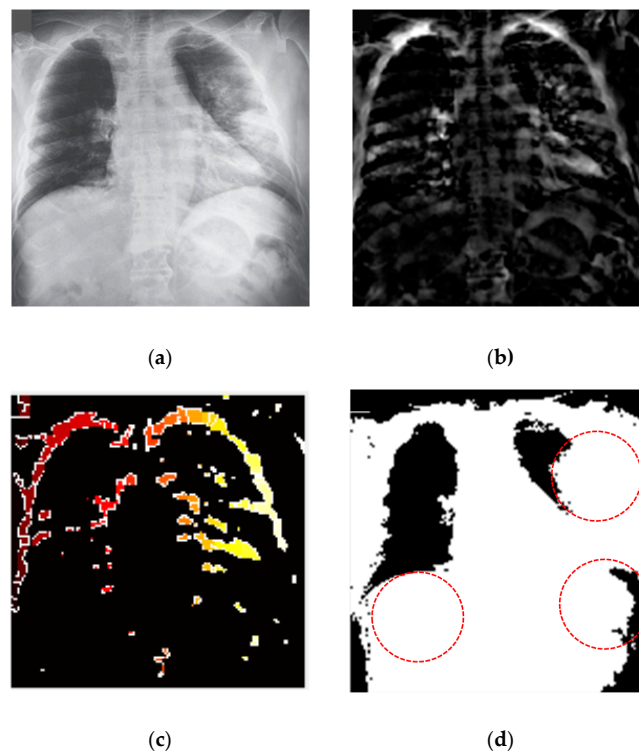


Figure 9. COVID-19 segmentation by using the watershed technique (a) applied anisotropic diffusion for filtering (b) adjusting the filtered image (c) watershed RGB image (d) fracture lung region caused by the coronavirus (COVID-19).

4. Experimental Details and Results

4.1. Datasets and Overall Performance

To validate the framework developed for intelligent COVID-19 detection, this work used a total of 5090 chest X-ray images for training, testing and validation, as shown in Table 1 without data augmentation. In this study, the distribution of the data was laminated in order to mitigate the data disequilibrium issue. The validation images were taken from the training set, but the testing set was taken before training.

Table 1. Number of images in normal and COVID 19 categories used in training, validation, and testing phases without data augmentation.

Data Sets	Number of Images		Ratio of Normal to the COVID-19 Images
	Normal	COVID-19	
Training	2489	1584	1.57
Validation	70	70	1.0
Testing	622	395	1.57

This research employed a total of 2489 normal and 1584 COVID-19-positive images for the training purpose. For the testing purpose, 622 normal images and 395 COVID-19-positive images were used. These testing images were not considered in the training dataset. This system also contained 70 validation images for both normal and COVID-19 classes. These validation images were taken from the training data set.

Three metrics, namely accuracy, specificity and sensitivity, were employed to measure the performance of the system developed for automatic COVID-19 detection from the chest X-ray images. Four different performance parameters, namely true-positive (TP), true-negative (TN), false-positive (FP) and false-negative (FN), were used to compute the metrics as defined by Equations (20)–(22).

$$\text{Accuracy (ACC)} = \frac{TP + TN}{TP + TN + FP + FN} \quad (20)$$

$$\text{Specificity (SPEC)} = \frac{TN}{TN + FP} \quad (21)$$

$$\text{Sensitivity (SEN)} = \frac{TP}{TP + FN} \quad (22)$$

Figure 10 presents the confusion matrix derived from the measured performance parameters for the overall system during the classification stage. During the course of the evaluation, the proposed method required labeled test data to validate its predicted output. The confusion matrix represents an overall system performance. The system cannot detect 27 COVID-19 positivity images and correctly detects 1952 COVID-19 positivity images out of 1979 images. The system cannot correctly detect 40 normal images and correctly detects 3071 normal images out of 3111 images.

4.2. Filtering Performance

The proposed method used COVID-19 X-ray images as test data with different speckles, noises and resolutions. To work with meaningful features, information preservation and noise reduction are the prerequisite conditions to fulfill. The current system used modified anisotropic diffusion filtering (MADF) at the image preprocessing stage. The performance measurement in MADF was assessed using three evaluation metrics, namely Signal-to-noise ratio (SNR), minimum square error (MSE) and edge preservation factors (EPF) [21]. Higher values of SNR and EPF represent more noise reduction and much edge details preservation, respectively. On the other hand, the minimum MSE value indicates less error between the input and filtered images. Classification models were run 10 times, and the highest values were reported for the performance metrics. Figure 11 presents the filtering performance using different evaluation metrics and comparison with other

techniques available in the literature. It was clear that all the existing filtering techniques produced lower MSE values indicating that the proposed technique was only slightly worse than the existing techniques. On the other hand, the SNR and EPF values were comparatively much higher in the proposed filtering technique, demonstrating its superiority over the others.

True label	Predicted label	
	COVID-19	Normal
COVID-19	TP (1952)	FN (27)
Normal	FP (40)	TN (3071)

Figure 10. Confusion matrix with overall measured performance parameters during training.

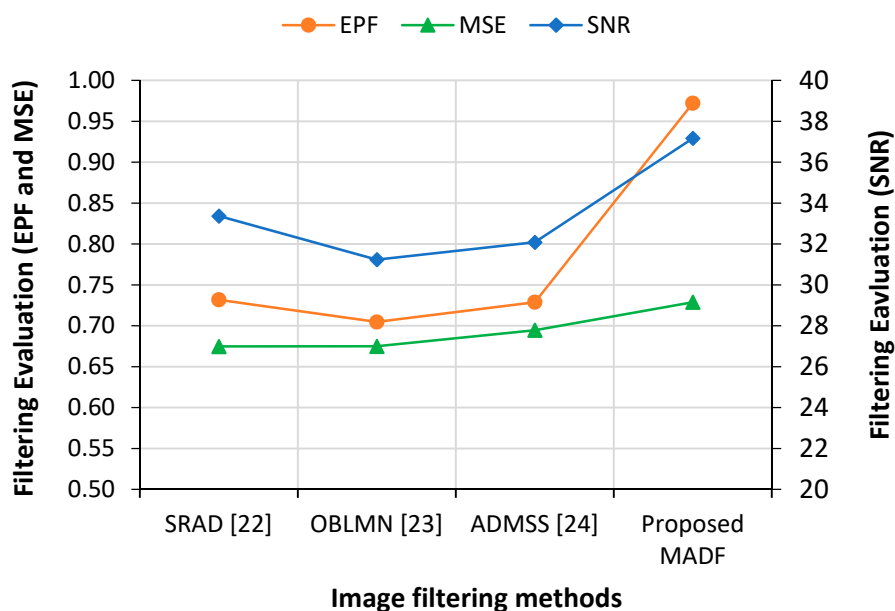


Figure 11. Performance comparison of modified anisotropic diffusion filtering (MADF) with other techniques in the literature.

4.3. Feature Extraction Performance

CNN used extracted features to train before classifying. Test features were also found from the test images using different pre-trained models to measure the performance of CNN models. Nowadays, CNN uses different pre-trained models like AlexNet, ResNet50, VGG-16, VGG-19 and ResNet50 to extract features from the training and test data sets. All these models produce similar results for standard training and testing data. Figure 12

presents a comparison of performance among different CNN models. It was apparent that VGG19, which was proposed in this work, achieved better accuracy and specificity than the other CNN models, although ResNet50 showed the best performance in terms of sensitivity.

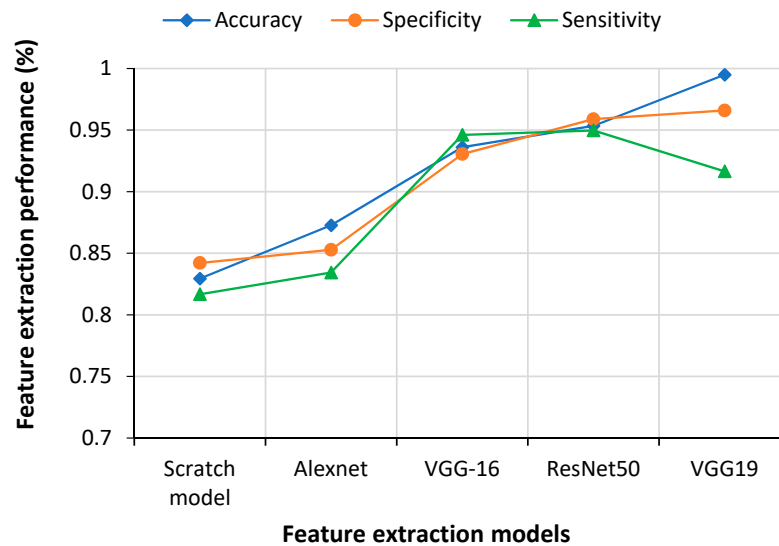


Figure 12. Performance measurement of different feature extraction models.

However, the scratch model did not show satisfactory performance compared to the CNN models. Overall, ResNet50 and VGG19 models generated much better results than the VGG16 and AlexNet models.

4.4. Classification Performance

This work proposed a fusion of feature vectors obtained by a combination of HOG and CNN techniques. This fusion vector was deliberated as the final input for the training and test datasets. Figure 13 presents a comparative study of different feature extraction approaches. The performances of different individual feature extraction techniques were less satisfactory than the fusion approach. This demonstrated that the proposed approach could classify COVID-19 cases more accurately than the single feature extraction approaches.

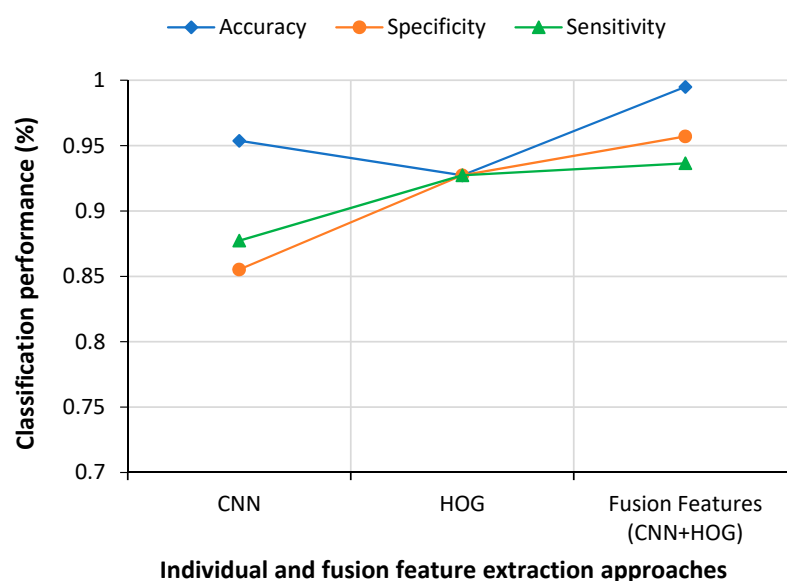


Figure 13. Comparative results of individual and fusion features.

The final classification was also performed with other popular machine-learning methods, such as artificial neural network (ANN), support vector machine (SVM) and k-nearest neighbor (KNN), in addition to CNN. The fused feature vectors were fed to during the classification to find the better classifier. CNN clearly showed the best performance, as shown in Figure 14.

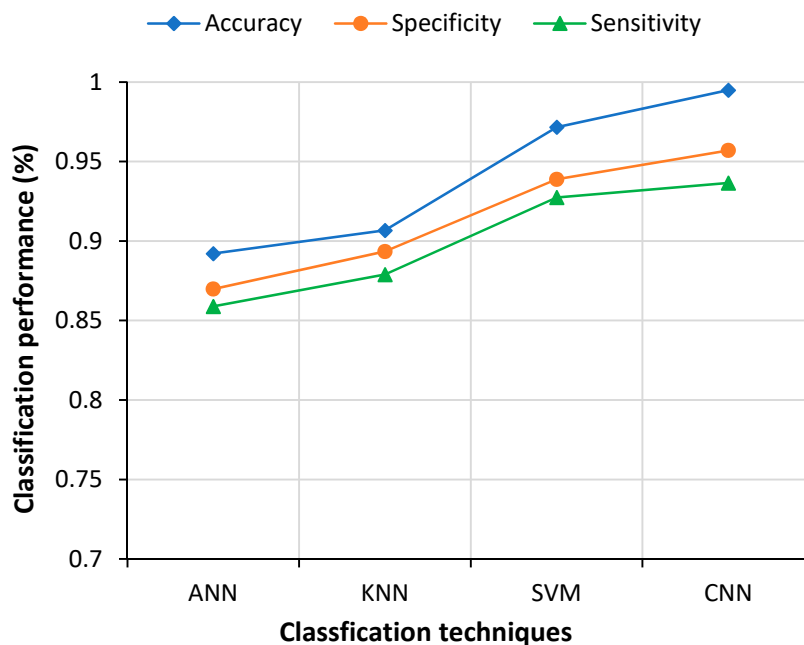
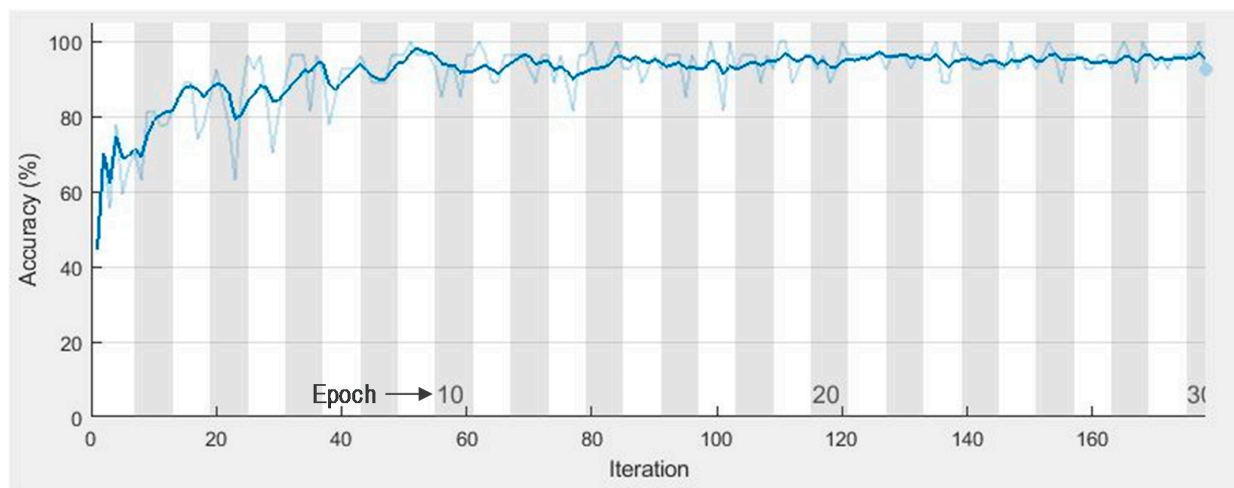


Figure 14. Comparative performance of different classifiers.

Accuracy vs. epoch curve is plotted in Figure 15a. This showed clear evidence of no overfitting situation with very close training and accuracy curves. The learning rate starts from 0.001 with mini-batch sizes of 64 and 36 epochs. The loss curve depicted in Figure 15b indicated only a little amount of lost value.



(a)

Figure 15. Cont.

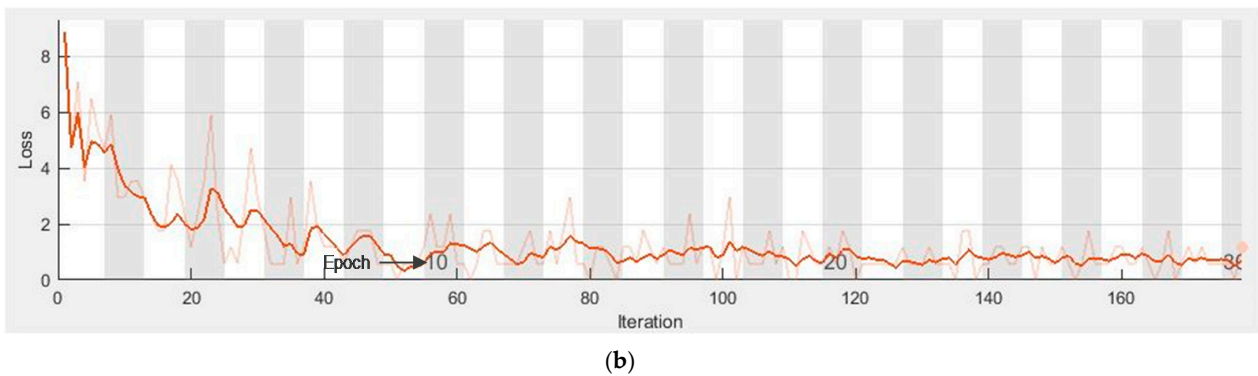


Figure 15. Learning curves (a) accuracy vs. number of epochs (b) loss vs. number of epochs.

5. Discussion

5.1. System Validation

To further validate the system's performance, generalization and k-fold validation techniques were applied.

Generalization is a term that measures a model's ability to react to new data. After being trained, this approach can digest new data and make accurate predictions. The existing work faced difficulties in the X-ray image quality, which negatively influence the generalization performance. This research used the IEEE Data Port X-ray image dataset for generalization purposes, which was completely different from the training set [57]. This dataset was previously unsighted and unknown for this model. In generalization, 3000 images were used with two classes of the COVID-19 (1100 images) and normal (1900 images). The confusion metrics of the generalization results are presented in Figure 16. Among 1100 COVID-19 images, only 95 images are miss detected. The wrong detection rate was minimum for the COVID-19 chest X-ray images, which achieved a generalization accuracy of 93.75%. This indicated the image classification reliability of the system even with a completely new set of data.

True label	Predicted label	
	COVID-19	Normal
COVID-19	TP (1005)	FN (95)
Normal	FP (244)	TN (1656)

Figure 16. Confusion matrix from generalization.

K-fold cross-validation is primarily used in applied machine learning to estimate the performance of a model to make predictions on data not used during the training of the model [58]. A k-fold cross-validation technique was performed to gain further confidence in the experimental results. CNN classifier was trained by 5-fold cross-validation after

feature extraction where feature vector was randomly distributed into 5 sub folds. Four sub folds were selected for the training dataset, and one single sub fold was selected for the testing dataset. Table 2 presents the different CNN classification results using individual and combined feature extraction methods. Single feature extraction methods provided less accuracy compared to the proposed fusion vector.

Table 2. Overall classification accuracy measured using 5-fold cross-validation.

Feature Extraction Methods	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean Accuracy
HOG	0.8732	0.8789	0.8741	0.8675	0.8730	0.8734
CNN	0.9378	0.9367	0.9387	0.9367	0.9321	0.9364
Proposed fusion (HOG+CNN)	0.9856	0.9847	0.9813	0.9827	0.9833	0.9836

5.2. Comparative Analysis

Using the relevant datasets of chest X-ray images for the COVID-19 detection is a laborious task. The researchers used different preprocessing techniques, feature extraction techniques and classification methods [29–39]. Now, it is difficult to suggest a promising technique or combination of techniques that are more effective in diagnosing COVID-19 from the chest X-ray image. In most of the cases, an accuracy above 90% was reported, and from a statistical point of view, this a very high level of accuracy. However, the goal would be increasing the accuracy level as close to 100% as misdiagnosis, even in a small number of cases, is not quite acceptable.

The comparative analyses of related work are shown in Table 3. It is fairly clear that the proposed technique produced a better classification accuracy in detecting COVID-19 compared to the other techniques proposed in the literature. However, the study by Loye et al. [35] found a higher accuracy (100%) than this study. This could be due to a much lower number of images (69 COVID-19 and 79 normal images) used in their dataset for testing the system's performance [59].

The proposed method employing feature fusion extracted by HOG and CNN (VGG19) showed better accuracy using CNN classifier. CNN provides a good result as a binary classifier for the chest X-ray dataset. However, when the classification was carried out with features extracted either by CNN or HOG individually, the accuracies were much lower compared to the reported values. HOG was the worst among the three, with an accuracy of 92.73%. Similarly, only SVM (97.16%) showed accuracy higher than 95% compared to the other two classification techniques, such as ANN (89.21%) and KNN (90.67%).

In this study, even though the image size was reduced by more than half of the original sizes, the system still demonstrated its robustness in correctly diagnosing COVID-19 cases. This could be attributed to the higher number of distinctive features obtained by the fusions of features extracted by HOG and CNN.

The proposed technique suggests a robust way in the feature extraction and noise removal phase. The proposed system used speckle-attacked and low-quality chest X-ray images in the testing phase. The MADF technique used in this work dispelled multiplicative speckle noise from the test images. It can effectively overcome the drawbacks of noisy images. Thus, the technique applied for preprocessing the input images facilitated more effective feature extraction and subsequently made a positive impact towards gaining a better classification performance. The results of the study demonstrated that the proposed feature fusion approach could diagnose COVID-19 very quickly and accurately and help the clinicians to take the necessary actions for saving the lives of COVID-19-affected people.

5.3. Limitations and Future Work

One of the limitations of this work was the imbalance of data in the datasets used for training and testing. In general, balanced data set with an equal number of normal and COVID-19 X-ray images makes the model building more comfortable, and the developed model can provide better prediction accuracy. Furthermore, the classification algorithm finds it easier to learn from a balanced dataset. Naturally, in any opensource database,

the number of normal images would be higher than the COVID-19-positive images. As the images used in this study were taken from open-source databases, the imbalance in the training and testing data sets was obvious. However, the ratio between the number of normal and COVID-19 images was maintained at 1.57 in both the training and testing data sets in order to alleviate the data imbalance problem to some extent.

Table 3. Comparison among existing methods in the COVID-19 detection.

References	Dataset	Methods	Accuracy
Ahmed et al. [29]	2971 chest X-ray images (COVID-19 = 285, normal = 1341, pneumonia = 1345)	CNN	94.03%
Chowdhury et al. [30]	2905 chest X-ray images (COVID-19 = 219, normal = 1341 and pneumonia = 1345)	Parallel-dilated CNN	96.58%
Abbas et al. [31]	196 CXR images (COVID-19 = 105, normal = 80, and SARScases = 11)	Deep CNN (DeTraC)	93.1%
Azemin et al. [32]	5982 (COVID-19 = 154 and normal = 5828)	ResNet-101 CNN	71.9%
El-Rashidy et al. [33]	750 chest X-ray images (COVID-19 = 250 and normal = 500)	CNN/ConvNet	97.95%
Khan et al. [34]	1057 X-ray images (COVID-19 = 195 and normal = 862)	VGG16+VGG19	99.3%
Loey et al. [35]	307 X-ray images (COVID-19 = 69, normal = 79, Pneumonia_bac = 79 and Pneumonia_vir = 79)	AlexNet+ Googlenet+Restnet18	100%
Minaee et al. [36]	50,184 chest X-ray images (COVID-19 = 184 and normal = 5000)	ResNet18 + ResNet50 + SqueezeNet + DenseNet-121	98%
Sekeroglu et al. [37]	6100 X-ray images (COVID-19 = 225, normal = 1583 and pneumonia = 4292)	CNN	98.50%
Wang et al. [38]	18,567 X-ray images (COVID-19 = 140, normal = 8851 and Pneumonia = 9576)	ResNet-101 + ResNet-152	96.1%
Panwar et al. [60]	284 images (COVID-19 = 142 and normal = 142)	Convolutional neural network (nCOVnet)	88.1%
Ozturk et al. [61]	625 images (COVID-19 = 125 and normal = 500)	Convolutional neural network (DarkNet)	98.08%
Khan et al. [62]	594 images (COVID-19 = 284 and normal = 310)	Convolutional neural network (CoroNet (Xception))	99%
Apostolopoulos and Mpesiana [63]	728 images (COVID-19 = 224 and normal = 504)	Transfer learning with convolutional neural networks(VGG19, MobileNet v2, Inception, Xception, InceptionResNet v2)	96.78%
Mahmud et al. [64]	610 images (COVID-19 = 305 and normal = 305)	Transfer learning with convolutional neural networks(stacked MultiResolution CovXNet)	97.4%
Benbrahim et al. [65]	320 images (COVID-19 = 160 and normal = 160)	Transfer learning with convolutional neural networks (Inceptionv3 and ResNet50)	99.01%
Martínez et al. [66]	240 images (COVID-19 = 120 and normal = 120)	Convolutional neural network (Neural Architecture Searchnetwork (NASNet))	97%
Toraman et al. [67]	1281 images (COVID-19 = 231 and normal = 1050)	Convolutional neural network (CapsNet)	97.24%
Duran-Lopez et al. [68]	6926 images (COVID-19 = 2589 and normal = 4337)	Convolutional neural network	94.43%
Proposed Method	5090 chest X-ray images (COVID-19 = 1979 and normal = 3111)	Fusion features (CNN+HOG) + VGG19 pre-train model	99.49%

Although some of the studies in the literature used balanced data sets [60,64,65] for the COVID-19 detection using CNN on X-ray images, a large number of studies employed unbalanced data [30–32,35,36]. Although, Azemin et al. [32] and Minaee et al. [36] built CNN models with hugely imbalanced data with normal to the COVID-19 ratios of 37 and 28, respectively, the models still showed high-performance matrices. In comparison, the data sets in this study with a ratio of 1.57 make it only slightly imbalanced datasets. Therefore, it can be said that the imbalanced data are not the only factor that could affect the prediction accuracy; other factors, such as data set size, filtering technique, feature extraction technique, and the machine-learning algorithm used, should also be taken into consideration.

Another limitation of this study is the small number of X-ray images. New images are continuously being made available around the world. In order to further improve the robustness and accuracy of the model, a number of images will be added to the training data set. The trained model can be stored in the cloud. A mobile app can be developed to diagnose COVID-19 within seconds in order to reduce the workload of clinicians in the hospitals. The future work will focus on the fusion of HOG and CNN features for three-dimensional volume analysis. In addition, more efficient methods can be developed by exploring other fusion approaches to improve the result.

In order to respond to the global challenge of quickly identifying COVID-19, a large number of studies were carried out in 2020 by applying deep learning techniques on lung X-ray images. The studies varied significantly in terms of deep learning architecture, feature selection technique, number of images in the training and testing data sets, testing protocols, etc. However, the majority of the studies reported high classification perfor-

mances in discriminating COVID-19 using different deep learning techniques. Concerns have been raised by Cohen et al. about the effectiveness in practical applications of the classification models with chest X-ray images [69]. Learning bias has been highlighted as one of the key challenges for the development of a robust and statistically stable deep learning model, which can confidently make the prediction for an unseen data set with similar values of the performance metrics. The bias could be due to a small number of the COVID-19 images with a publicly available large number of non-COVID-19 images [70]. Furthermore, the learning bias could be due to the source data set, meaning that the model could learn features that are related to the characteristics of the datasets (text label, scan setting, or similar age and sex) rather than the features specific to the disease impacting the overall validity of the model and its significance in clinical use [69,71]. Maguolo et al. demonstrated that, interestingly, even without containing a major part of the lungs in the X-ray images within a dataset could produce similar classification results compared to the dataset of X-ray images, including lungs [71]. Designing a fair testing protocol could be highly challenging when different datasets were merged with large differences among them. This could generate misleading results. Therefore, assessing the validity of the available testing protocol must be carefully dealt with.

In this study, despite taking some preventative measures to avoid bias, such as using relatively large data set (3111 non-COVID and 1979 COVID-19 images) without huge data imbalanced, preprocessing of images (e.g., removing text labels) to remove dataset-dependent features, extracting features by two techniques and feature fusion and testing with a set of data different from the training set, still, the trained model would not be completely free from the learning bias. Yasar et al. compiled results for the COVID-19 detection by recent studies using the X-ray images and found that seven out of sixteen (44%) studies used k-fold validation [59]. This represented that k-fold was a popular validation technique. In this study, during k-fold validation, as the images were taken from the same training and testing datasets, unlike the generalization technique where new data set were used for the validation, it provided accuracy close to the testing accuracy. This supported the argument that k-fold validation could be an unfair testing protocol. Although k-fold cross validation balanced out the predicted features' classes with the unbalanced dataset used in this study to a certain extent, still innovative and effective testing protocols must be sought to make sure the developed model would be free of any learning bias for an accurate prediction of the outcome. This could be possible in the future when the continuous effort of collecting X-ray images will create a large balanced dataset to determine if deep learning can provide a solution in the fight against the COVID-19 pandemic.

6. Conclusions

The coronavirus pandemic has stretched the healthcare systems in every country in the world to its limit as they had to deal with a large number of deaths. Early detection of the COVID-19 in a faster, easier, and cheaper way can help in saving lives and reduce the burden on healthcare professionals. Artificial intelligence can play a big role in identifying COVID-19 by applying image processing techniques to X-ray images. This work designed and developed an intelligent system for the COVID-19 identification with high accuracy and minimum complexity by combining the features extracted by histogram-oriented gradient (HOG) features and convolutional neural network (CNN). Suitable feature selection and classification are absolutely vital in the COVID-19 detection using chest X-ray images. Chest X-ray images were entered into the system in order to produce the output of the marked lung significant region, which was used to identify COVID-19. The proposed feature fusion system showed a higher classification accuracy (99.49%) than the accuracies obtained by using features obtained by individual feature extraction techniques, such as HOG and CNN. CNN produced the best classification accuracy compared to the other classification techniques, such as ANN, KNN and SVM. Furthermore, the proposed fusion technique was validated with higher accuracies using generalization and k-fold validation techniques.

Author Contributions: All authors have equal contribution to prepare and finalize the manuscript. Conceptualization, N.-A.-A., M.A., M.A.B. and J.H.; methodology, N.-A.-A., M.A., M.A.B.; software, N.-A.-A. and M.A.B.; validation, N.-A.-A., M.A., M.A.B., J.H. and M.K.; formal analysis, N.-A.-A., M.A., M.A.B., J.H., and M.K.; investigation, N.-A.-A., M.A., M.A.B., J.H., M.K.; resources, N.-A.-A. and M.A.B.; data curation, N.-A.-A. and M.A.B.; writing—original draft preparation, N.-A.-A. and M.A.B. writing—review and editing, N.-A.-A., M.A., M.A.B., J.H. and M.K.; visualization, N.-A.-A., M.A., M.A.B., J.H. and M.K.; supervision, M.A., M.A.B. and J.H.; project administration, N.-A.-A. and M.A.B.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The authors are very grateful to Farhan Matin, Associate Professor, Department of Radiology and Imaging, Uttara Adhunik Medical College and Hospital, Bangladesh, for his valuable support, suggestions, and comments.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Wu, F.; Zhao, S.; Yu, B.; Chen, Y.M.; Wang, W.; Song, Z.G.; Hu, Y.; Tao, Z.W.; Tian, J.H.; Pei, Y.Y.; et al. A new coronavirus associated with human respiratory disease in China. *Nature* **2020**, *579*, 265–269. [[CrossRef](#)]
2. Guan, W.J.; Ni, Z.Y.; Hu, Y.; Liang, W.H.; Ou, C.Q.; He, J.X.; Liu, L.; Shan, H.; Lei, C.L.; Hui, D.S.; et al. Clinical characteristics of coronavirus disease 2019 in China. *N. Engl. J. Med.* **2020**, *382*, 1708–1720. [[CrossRef](#)]
3. Chen, N.; Zhou, M.; Dong, X.; Qu, J.; Gong, F.; Han, Y.; Qiu, Y.; Wang, J.; Liu, Y.; Wei, Y.; et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: A descriptive study. *Lancet* **2020**, *395*, 507–513. [[CrossRef](#)]
4. Wang, C.; Horby, P.W.; Hayden, F.G.; Gao, G.F. A novel coronavirus outbreak of global health concern. *Lancet* **2020**, *395*, 470–473. [[CrossRef](#)]
5. Zhu, N.; Zhang, D.; Wang, W.; Li, X.; Yang, B.; Song, J.; Zhao, X.; Huang, B.; Shi, W.; Lu, R.; et al. 2020. A novel coronavirus from patients with pneumonia in China. *N. Engl. J. Med.* **2019**, *382*, 727–733. [[CrossRef](#)] [[PubMed](#)]
6. Li, Q.; Guan, X.; Wu, P.; Wang, X.; Zhou, L.; Tong, Y.; Ren, R.; Leung, K.S.; Lau, E.H.; Wong, J.Y.; et al. Early transmission dynamics in Wuhan, China, of novel coronavirus–infected pneumonia. *N. Engl. J. Med.* **2020**, *382*, 1199–1207. [[CrossRef](#)]
7. Holshue, M.L.; DeBolt, C.; Lindquist, S.; Lofy, K.H.; Wiesman, J.; Bruce, H.; Spitters, C.; Ericson, K.; Wilkerson, S.; Tural, A.; et al. First case of 2019 novel coronavirus in the United States. *N. Engl. J. Med.* **2020**, *382*, 929–936. [[CrossRef](#)] [[PubMed](#)]
8. WHO Coronavirus Disease (COVID-19) Dashboard. Available online: https://covid19.who.int/?gclid=CjwKCAjw5p_8BRBUeiwAPpJO682JEO1UwRkSSDofsqaqGeAncQYeiEeTcnMSFjd55I0lzYIHrvi4SxoCAeUQAvD_BwE (accessed on 15 October 2020).
9. Ledford, H.; Cyranoski, D.; Van, N.R. The UK has approved a COVID vaccine—here’s what scientists now want to know. *Nature* **2020**, *588*, 205–206. [[CrossRef](#)]
10. Anon. The COVID vaccine challenges that lie ahead. *Nature* **2020**, *587*, 522. [[CrossRef](#)]
11. Kim, J.H.; Marks, F.; Clemens, J.D. Looking beyond COVID-19 vaccine phase 3 trials. *Nat. Med.* **2021**, *27*, 1–7. [[CrossRef](#)]
12. Logunov, D.Y.; Dolzhikova, I.V.; Shcheblyakov, D.V.; Tukhvatulin, A.I.; Zubkova, O.V.; Dzharullaeva, A.S.; Kovyrshina, A.V.; Lubenets, N.L.; Grousova, D.M.; Erokhova, A.S.; et al. Safety and efficacy of an rAd26 and rAd5 vector-based heterologous prime-boost COVID-19 vaccine: An interim analysis of a randomised controlled phase 3 trial in Russia. *Lancet* **2021**. [[CrossRef](#)]
13. Chen, Z.; Zhang, L. Meet the Challenges of Mass Vaccination against COVID-19. *Explor. Res. Hypothesis Med.* **2021**, 1–3. [[CrossRef](#)]
14. Li, Y.; Shen, L. Skin lesion analysis towards melanoma detection using deep learning network. *Sensors* **2018**, *18*, 556. [[CrossRef](#)]
15. Liao, Q.; Ding, Y.; Jiang, Z.L.; Wang, X.; Zhang, C.; Zhang, Q. Multi-task deep convolutional neural network for cancer diagnosis. *Neurocomputing* **2019**, *348*, 66–73. [[CrossRef](#)]
16. Yoo, S.; Gujrathi, I.; Haider, M.A.; Khalvati, F. Prostate cancer detection using deep convolutional neural networks. *Sci. Rep.* **2019**, *9*, 1–10. [[CrossRef](#)]
17. Esteva, A.; Kuprel, B.; Novoa, R.A.; Ko, J.; Swetter, S.M.; Blau, H.M.; Thrun, S. Dermatologist-level classification of skin cancer with deep neural networks. *Nature* **2017**, *542*, 115–118. [[CrossRef](#)]
18. Wang, L.; Wong, A. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest X-ray images. *arXiv* **2020**, arXiv:2003.09871. [[CrossRef](#)] [[PubMed](#)]
19. Afzal, A. Molecular diagnostic technologies for COVID-19: Limitations and challenges. *J. Adv. Res.* **2020**. [[CrossRef](#)] [[PubMed](#)]
20. World Health Organization: Use of Chest Imaging in Covid-19. 2020. Available online: <https://www.who.int/publications/i/item/use-of-chest-imaging-in-covid-19> (accessed on 7 January 2021).

21. Davies, H.E.; Wathen, C.G.; Gleeson, F.V. The risks of radiation exposure related to diagnostic imaging and how to minimise them. *BMJ* **2011**, *342*. [[CrossRef](#)]
22. Cherian, T.; Mulholland, E.K.; Carlin, J.B.; Ostensen, H.; Amin, R.; Campo, M.D.; Greenberg, D.; Lagos, R.; Lucero, M.; Madhi, S.A.; et al. Standardized interpretation of paediatric chest radiographs for the diagnosis of pneumonia in epidemiological studies. *Bull. World Health Organ* **2005**, *83*, 353–359.
23. Franquet, T. Imaging of pneumonia: Trends and algorithms. *Eur. Respir. J.* **2001**, *18*, 196–208. [[CrossRef](#)] [[PubMed](#)]
24. Ng, M.Y.; Lee, E.Y.; Yang, J.; Yang, F.; Li, X.; Wang, H.; Lui, M.; Lo, C.; Leung, B.; Khong, P.; et al. Imaging profile of the covid-19 infection: Radiologic findings and literature review. *Radiol. Cardiothorac. Imaging* **2020**, *2*, e200034. [[CrossRef](#)]
25. Verma, S.G.; Prakash, S. Pneumonia classification using deep learning in healthcare. *Int. J. Innov. Technol. Explor. Eng.* **2020**, *9*, 1715–1723.
26. Xiaowei, X.; Xiangao, J.; Chunlian, M.; Peng, D.; Xukun, L.; Shuangzhi, L.; Liang, Y.; Qin, N.; Yanfei, C.; Junwei, S.; et al. A deep learning system to screen novel coronavirus disease 2019 pneumonia. *Engineering* **2020**, *6*, 1122–1129.
27. Shi, F.; Wang, J.; Shi, J.; Wu, Z.; Wang, Q.; Tang, Z.; He, K.; Shi, Y.; Shen, D. Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19. *IEEE Rev. Biomed. Eng.* **2020**. [[CrossRef](#)]
28. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.
29. Ahamed, K.; Satu, M.S.; Abedin, M.Z.; Rahaman, M.A.; Islam, S.M.S. Early Detection of Coronavirus Cases Using Chest X-ray Images Employing Machine Learning and Deep Learning Approaches. *medRxiv* **2020**. medRxiv 2020.06.07.20124594.
30. Chowdhury, N.K.; Rahman, M.M.; Kabir, M.A. PDCOVIDNet: A parallel-dilated convolutional neural network architecture for detecting COVID-19 from chest X-ray images. *Health Inf. Sci. Syst.* **2020**, *8*, 1–14. [[CrossRef](#)] [[PubMed](#)]
31. Abbas, A.; Abdelsamea, M.M.; Gaber, M.M. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. *Appl. Intell.* **2021**, *51*, 854–864. [[CrossRef](#)]
32. Che Azemin, M.Z.; Hassan, R.; Mohd Tamrin, M.I.; Md Ali, M.A. COVID-19 Deep Learning Prediction Model Using Publicly Available Radiologist-Adjudicated Chest X-Ray Images as Training Data: Preliminary Findings. *Int. J. Biomed. Imaging* **2020**, *2020*. [[CrossRef](#)]
33. El-Rashidy, N.; El-Sappagh, S.; Islam, S.M.R.; El-Bakry, H.M.; Abdelrazek, S. End-To-End Deep Learning Framework for Coronavirus (COVID-19) Detection and Monitoring. *Electronics* **2020**, *9*, 1439. [[CrossRef](#)]
34. Khan, I.U.; Aslam, N. A Deep-Learning-Based Framework for Automated Diagnosis of COVID-19 Using X-ray Images. *Information* **2020**, *11*, 419. [[CrossRef](#)]
35. Loey, M.; Smarandache, F.M.; Khalifa, N.E. Within the Lack of Chest COVID-19 X-ray Dataset: A Novel Detection Model Based on GAN and Deep Transfer Learning. *Symmetry* **2020**, *12*, 651. [[CrossRef](#)]
36. Minaee, S.; Kafieh, R.; Sonka, M.; Yazdani, S.; Soufi, G. Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning. *Med. Image Anal.* **2020**, *65*, 101794. [[CrossRef](#)]
37. Sekeroglu, B.; Ozsahin, I. Detection of COVID-19 from Chest X-Ray Images Using Convolutional Neural Networks. *SLAS Technol. Transl. Life Sci. Innov.* **2020**, *25*, 553–565. [[CrossRef](#)] [[PubMed](#)]
38. Wang, N.; Liu, H.; Xu, C. Deep Learning for The Detection of COVID-19 Using Transfer Learning and Model Integration. In Proceedings of the 2020 IEEE 10th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, 17–19 July 2020; pp. 281–284.
39. Yoo, S.H.; Geng, H.; Chiu, T.L.; Yu, S.K.; Cho, D.C.; Heo, J.; Choi, M.S.; Choi, I.H.; Van, C.C.; Nhung, N.V. Deep Learning-Based Decision-Tree Classifier for COVID-19 Diagnosis From Chest X-ray Imaging. *Front. Med.* **2020**, *7*, 427. [[CrossRef](#)] [[PubMed](#)]
40. Khalifa, N.E.M.; Taha, M.H.N.; Manogaran, G.; Loey, M. A deep learning model and machine learning methods for the classification of potential coronavirus treatments on a single human cell. *J. Nanoparticle Res.* **2020**, *22*, 1–13. [[CrossRef](#)]
41. Wang, D.; Mo, J.; Zhou, G.; Xu, L.; Liu, Y. An efficient mixture of deep and machine learning models for COVID-19 diagnosis in chest X-ray images. *PLoS ONE* **2020**, *15*, e0242535. [[CrossRef](#)]
42. Sahlol, A.T.; Yousri, D.; Ewees, A.A.; Al-Qaness, M.A.; Damasevicius, R.; Abd Elaziz, M. COVID-19 image classification using deep features and fractional-order marine predators algorithm. *Sci. Rep.* **2020**, *10*, 1–15. [[CrossRef](#)]
43. Chowdhury, M.E.; Rahman, T.; Khandakar, A.; Mazhar, R.; Kadir, M.A.; Mahbub, Z.B.; Islam, K.R.; Khan, M.S.; Iqbal, A.; Al-Emadi, N.; et al. Can ai help in screening viral and covid-19 pneumonia? *arXiv* **2020**, arXiv:2003.13145.
44. Cohen, J.P.; Morrison, P.; Dao, L. COVID-19 Image Data Collection. *arXiv* **2020**, arXiv:2003.11597v1.
45. COVID-19 X-ray Image Data Sets. Available online: <https://drive.google.com/uc?id=1coM7x3378f-Ou2l6Pg2wldOI7Dntu1a> (accessed on 10 October 2020).
46. Cleverley, J.; Piper, J.; Jones, M.M. The role of chest radiography in confirming covid-19 pneumonia. *BMJ* **2020**, *370*. [[CrossRef](#)] [[PubMed](#)]
47. Gao, M.; Kang, B.; Feng, X.; Zhang, W.; Zhang, W. Anisotropic Diffusion Based Multiplicative Speckle Noise Removal. *Sensors* **2019**, *19*, 3164. [[CrossRef](#)] [[PubMed](#)]
48. Chaoxin, Z.; Da-Wen, S. 2-Image Segmentation Techniques. In *Food Science and Technology, Computer Vision Technology for Food Quality Evaluation*; Da-Wen, S., Ed.; Academic Press: Oxford, UK, 2008; pp. 37–56.
49. Dandil, E.; Çakiroğlu, M.; Ekşi, Z.; Özkan, M.; Kurt, Ö.K.; Canan, A. Artificial neural network-based classification system for lung nodules on computed tomography scans. In Proceedings of the 2014 6th International Conference of Soft Computing and Pattern Recognition (soCPar), Tunis, Tunisia, 11–14 August 2014.

50. Ramos-Llordén, G.; Vegas-Sánchez-Ferrero, G.; Martín-Fernández, M.; Alberola-López, C.; Aja-Fernández, S. Anisotropic diffusion filter with memory based on speckle statistics for ultrasound images. *IEEE Trans. Image Process.* **2015**, *24*, 345–358. [[CrossRef](#)] [[PubMed](#)]
51. Xiang, Z.; Tan, H.; Ye, W. The excellent properties of a dense grid-based HOG feature on face recognition compared to Gabor and LBP. *IEEE Access* **2018**, *6*, 29306–29319. [[CrossRef](#)]
52. Barbhuiya, A.A.; Karsh, R.K.; Jain, R. CNN based feature extraction and classification for sign language. *Multimed. Tools Appl.* **2020**, *1–19*. [[CrossRef](#)]
53. Mostafiz, R.; Rahman, M.M.; Uddin, M.S. Gastrointestinal polyp classification through empirical mode decomposition and neural features. *SN Appl. Sci.* **2020**, *2*, 1–10. [[CrossRef](#)]
54. Saba, T.; Mohamed, A.S.; El-Affendi, M.; Amin, J.; Sharif, M. Brain tumor detection using fusion of hand crafted and deep learning features. *Cogn. Syst. Res.* **2020**, *59*, 221–230. [[CrossRef](#)]
55. Lou, S.; Pagani, L.; Zeng, W.; Jiang, X.; Scott, P.J. Watershed segmentation of topographical features on freeform surfaces and its application to additively manufactured surfaces. *Precis. Eng.* **2020**, *63*, 177–186. [[CrossRef](#)]
56. Chen, M.; Ludwig, S.A. Color image segmentation using fuzzy C-regression model. *Adv. Fuzzy Syst.* **2017**, *2017*. [[CrossRef](#)]
57. Academic Torrents. Available online: <https://academictorrents.com/> (accessed on 10 August 2020).
58. Hasan, M.M.; Islam, N.; Rahman, M.M. Gastrointestinal polyp detection through a fusion of contourlet transform and Neural features. *J. King Saud Univ. Comput. Inf. Sci* **2020**. [[CrossRef](#)]
59. Yasar, H.; Ceylan, M. A new deep learning pipeline to detect Covid-19 on chest X-ray images using local binary pattern, dual tree complex wavelet transform and convolutional neural networks. *Appl. Intell.* **2020**, *1–24*. [[CrossRef](#)]
60. Panwar, H.; Gupta, P.K.; Siddiqui, M.K.; Morales-Menendez, R.; Singh, V. Application of deep learning for fast detection of COVID-19 in X-rays using nCOVnet. *Chaos Solitons Fractals* **2020**, *138*, 109944. [[CrossRef](#)] [[PubMed](#)]
61. Ozturk, T.; Talo, M.; Yildirim, E.A.; Baloglu, U.B.; Yildirim, O.; Acharya, U.R. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Comput. Biol. Med.* **2020**, *121*, 103792. [[CrossRef](#)]
62. Khan, A.I.; Shah, J.L.; Bhat, M.M. Coronet: A deep neural network for detection and diagnosis of COVID-19 from chest X-ray images. *Comput. Methods Prog. Biomed.* **2020**, *196*, 105581. [[CrossRef](#)]
63. Apostolopoulos, I.D.; Mpesiana, T.A. Covid-19: Automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Phys. Eng. Sci. Med.* **2020**, *43*, 635–640. [[CrossRef](#)]
64. Mahmud, T.; Rahman, M.A.; Fattah, S.A. CovXNet: A multidilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. *Comput. Biol. Med.* **2020**, *122*, 103869. [[CrossRef](#)]
65. Benbrahim, H.; Hachimi, H.; Amine, A. Deep transfer learning with apache spark to detect COVID-19 in chest X-ray images. *Rom. J. Inform. Sci. Technol.* **2020**, *23*, S117–S129.
66. Martínez, F.; Martínez, F.; Jacinto, E. Performance evaluation of the NASNet convolutional network in the automatic identification of COVID-19. *Int. J. Adv. Sci. Eng. Inform. Technol.* **2020**, *10*, 662. [[CrossRef](#)]
67. Toraman, S.; Alakus, T.B.; Turkoglu, I. Convolutional capsnet: A novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks. *Chaos Solitons Fractals* **2020**, *140*, 110122. [[CrossRef](#)]
68. Duran-Lopez, L.; Dominguez-Morales, J.P.; Corral-Jaime, J.; Vicente-Diaz, S.; Linares-Barranco, A. COVID-XNet: A custom deep learning system to diagnose and locate COVID-19 in chest X-ray images. *Appl. Sci.* **2020**, *10*, 5683. [[CrossRef](#)]
69. Cohen, J.P.; Hashir, M.; Brooks, R.; Bertrand, H. On the limits of cross-domain generalization in automated X-ray prediction. In *Medical Imaging with Deep Learning*; Available online: <https://arxiv.org/abs/2002.02497> (accessed on 7 January 2021).
70. Tartaglione, E.; Barbano, C.A.; Berzovini, C.; Calandri, M.; Grangetto, M. Unveiling covid-19 from chest x-ray with deep learning: A hurdles race with small data. *Int. J. Environ. Res. Public Health* **2020**, *17*, 6933. [[CrossRef](#)] [[PubMed](#)]
71. Maguolo, G.; Nanni, L. A critic evaluation of methods for covid-19 automatic detection from x-ray images. *arXiv* **2020**, arXiv:2004.12823.