

Article

A Novel Method for Detection of Tuberculosis in Chest Radiographs Using Artificial Ecosystem-Based Optimisation of Deep Neural Network Features

Ahmed T. Sahlol ^{1,†}, Mohamed Abd Elaziz ^{2,*,†}^(b), Amani Tariq Jamal ^{3,†}^(b), Robertas Damaševičius ^{4,5,†}^(b) and Osama Farouk Hassan ^{6,†}^(b)

- ¹ Computer Department, Faculty of Specific Education, Damietta University, Damietta 34517, Egypt; atsegypt@du.edu.eg
- ² Department of Mathematics, Faculty of Science, Zagazig University, Zagazig 44519, Egypt
- ³ Computer Science Department, King Abdulaziz University, Jeddah 21589, Saudi Arabia; Atjamal@kau.edu.sa
- ⁴ Department of Applied Informatics, Vytautas Magnus University, 44404 Kaunas, Lithuania; robertas.damasevicius@polsl.pl
- ⁵ Faculty of Applied Mathematics, Silesian University of Technology, 44-100 Gliwice, Poland
- ⁶ Department of Mathematics, Faculty of Science, Damanhour University, Damanhur 22511, Egypt; osamafarouk4@gmail.com
- * Correspondence: abd_el_aziz_m@yahoo.com
- + These authors contributed equally to this work.

Received: 1 June 2020; Accepted: 2 July 2020; Published: 8 July 2020



Abstract: Tuberculosis (TB) is is an infectious disease that generally attacks the lungs and causes death for millions of people annually. Chest radiography and deep-learning-based image segmentation techniques can be utilized for TB diagnostics. Convolutional Neural Networks (CNNs) has shown advantages in medical image recognition applications as powerful models to extract informative features from images. Here, we present a novel hybrid method for efficient classification of chest X-ray images. First, the features are extracted from chest X-ray images using MobileNet, a CNN model, which was previously trained on the ImageNet dataset. Then, to determine which of these features are the most relevant, we apply the Artificial Ecosystem-based Optimization (AEO) algorithm as a feature selector. The proposed method is applied to two public benchmark datasets (Shenzhen and Dataset 2) and allows them to achieve high performance and reduced computational time. It selected successfully only the best 25 and 19 (for Shenzhen and Dataset 2, respectively) features out of about 50,000 features extracted with MobileNet, while improving the classification accuracy (90.2% for Shenzen dataset and 94.1% for Dataset 2). The proposed approach outperforms other deep learning methods, while the results are the best compared to other recently published works on both datasets.

Keywords: Tuberculosis (TB); transfer learning; convolutional neural networks; deep learning; Artificial Ecosystem-based Optimization; image processing

1. Introduction

Tuberculosis (TB) is considered one of the world's biggest threats to humanity as it is ranked as the fifth leading cause of death globally (1.5 mln deaths annually) [1]. For this reason, the World Health Organization encourages interlacing efforts to root it out because it is rather easy to cure [2]. Although it has low specificity and difficult interpretation, posteroanterior chest radiography is one of the preferred TB screening methods. In these circumstances, an automated computer-based diagnosis system for TB could be an efficient method for widespread TB screening.



As a result, the attention of the machine learning community was attracted by this topic, and some works have tackled the problem by applying classical machine learning methods with hand-crafted feature extraction, then applying a specific classifier such as Support Vector Machine (SVM) [3], Random Forest (RF) [4], ensemble baggage tree (EBT) [5] or Optimum-Path Forest (OPF) [6], while others apply nature-inspired evolutionary algorithms [7–9] or Convolutional Neural Networks (CNN) [10]. Some works achieved brilliant results, as some of these methods achieve quite close to human performance [11]. This makes deep learning an efficient method for medical image analysis [12], breast cancer classification [13], tumor segmentation [14] and pneumonia detection [15]. For TB diagnosis, CNN has achieved performances close to the best, while providing a more simple implementation to some competing methods, which employ complex machine learning pipelines [16–19]. For example, Vajda et al. [19] use an approach that starts with several preprocessing operations on lung-like segmentation, then extracting hand-crafted features such as the eigenvalues.

operations on lung-like segmentation, then extracting hand-crafted features such as the eigenvalues of the Hessian matrix, then finally uses a classifier for diagnosis. The results were aligned with deep learning results [20], however, their multi-stage approach is more complex than a CNN based approach and needed more development work.

Previous works adopt transfer learning-based models such as pre-trained AlexNet, ResNet and GoogLeNet models, which were applied on the image classifications tasks [20–23]. These models are considered as very powerful because they have been carefully implemented, trained and optimized on about 1 million images (the Imagenet dataset [24]), and then used to distinguish between some classes. Consequently, they need a lot of memory, resources and computation, while they are more likely to suffer from over-fitting are less likely to generalize well when used for medical classification tasks with limited amounts of images available [25]. While in [26], a hybrid approach consists of two main phases was proposed; a pretrained DenseNet-121 besides four types of local features such as SIFT, GIST, LBP, and HOG, and convolutional features. The hybrid approach showed advantages over the state of the art approaches in the classification of 14 chest diseases as it reaches 84.62% of classification accuracy, compared to 80.97% accuracy of the pre-trained DenseNet-121 model.

In our previous work [27,28], a segmentation of white blood cells was done by extracting color, shape, texture and hybrid features, and then a social spider optimization algorithm was used to select the best features. We also used the hybrid method for the detection of the bad tissues in the chest radiogram images [29]. The Moth–Flame and Ant Lion search heuristic optimization methods combined with the custom neural network allowed to obtain an accuracy of 85% on Montgomery County and Shenzhen datasets.

While CNN achieves the best performance on big datasets, often the dataset is small and may not provide enough data to train a CNN from scratch leading to the "small data" problem [30]. In such cases, transfer learning may be used [31,32]. The network is first trained on a large but generic collection of images, and then used to solve a specific task [33]. The examples of pre-trained neural networks are VGGNet [34], ResNet (GoogLeNet) [35], Nasnet [36], Mobilenet [37], Inception [38] and Xception [39]. The use of CNN on chest X-Rays images has been performed in [23,40], the authors recommended using CNN because there are available images that are similar to real-life medical images. It also enables applying the computer-based systems at cheap prices. The impact of applying CNN on chest diseases has been revealed in previous works [23].

Here we suggest a new approach to optimize deep learning architectures for medical image classification purposes. Such methods might play a significant role as a computer-aided tool for image-based clinical diagnosis soon. Our specific novelty and contribution is the combination of the Artificial Ecosystem-based Optimization (AEO) algorithm [41] and the MobileNet [37] deep learning network. The AEO algorithm has not been used for the analysis of the chest X-ray images and the detection of pulmonary diseases yet. The proposed MobileNet-AEO approach is based on transfer learning and starts by extracting deep features from raw Chest X-ray images using MobileNet. Then a selection of the optimal subset of relevant features is performed using binary AEO as a feature selection technique. The motivation of this work is to apply AEO as a feature selector from deep features that

produced from CNN which is largely redundant. This can accordingly minimize the capacity and resources consumption, then improve the classification of Chest X-ray images. up to our knowledge, AEO has not applied on any real applications yet.

The other parts of the article are structured in the following order: Section 2 presents the methodology and the techniques used including the model structure and description. The proposed MobileNet-AEO approach is presented in Section 3. Datasets and evaluation measures are described in Section 4. The experimental results and comparisons with other works are presented in Section 5, while, the conclusions are presented in Section 6.

2. Material and Methods

2.1. Feature Extraction Using Convolutional Neural Networks

The main concept of transfer learning with very deep neural networks is to re-train a CNN model (on our dataset) that was previously trained on the ImageNet [24] (about 1.2 million images). Since the dataset covers a wide range of objects (1000 different categories), so the model can learn diverse types of features, then they can be reused for other classification tasks.

In this work, we used MobileNet [37] with 88 layers. The MobileNet structure is mainly based on the depth-wise separable convolutions (Conv.). Each Conv. layer is succeeded by a Batch Normalization (BN.) [42] and ReLU. Figure 1 shows that down-sampling is performed with stride-Conv. in the depth-wise Conv., and in the first layer. Then, a final average pooling is applied to reduce the spatial resolution to be 1 before exposed to the fully connected layer. Counting depth-wise and 1×1 (point-wise) Conv., shows that MobileNet contains 28 layers. It is worthwhile to mention that 95% of it's computation time is spent in 1×1 convolutions which contain 75% of MobileNet parameters. The whole MobileNet architecture consists of 17 of these blocks as seen in Figure 1. There are no pooling layers between these depth-wise blocks. Instead, there is a stride of 2 to reduce the spatial dimensions.

In our case, the last layer has only 2 channels (TB or not TB). The final layer is the soft-max layer because it is a binary classification. Hidden layers have the rectification non-linearity [20].



Figure 1. Standard MobileNet structure.

To implement transfer learning with MobileNet, we first retrieved the previously extracted bottleneck features by MobileNet, then combine them with current extracted features from our TB images by training MobileNet on existing training data. Finally, we assign another class (rather than 1000 classes as in Imagenet) at the top of the model (for classification) [43].

One of the drawbacks of CNN in general, especially, MobileNet is that it requires higher computational resources such as memory and storage capacity. In order to get over this problem, some statistical operations were applied to exclude irrelevant and correlated features, and also to make the proposed approach computationally efficient, they are listed as follows [44]:

1. Chi-square is applied to remove the features which have a high correlation values by computing the dependence between them. It is calculated between each feature for all classes, as in Equation (1):

$$\chi^2 = \sum_{k=1}^n \frac{(O_k - E_k)^2}{E_k}$$
(1)

where O_k and E_k refer to the actual and the expected feature value, respectively.

2. Tree-based classifier is used to calculate feature importance to improve the classification since it has high accuracy, good robustness, and is simple [45].

2.2. Feature Selection Using Artificial Ecosystem-Based Optimization

Recently, the Artificial Ecosystem-based Optimization (AEO) algorithm has been proposed [41]. It is inspired by the behavior in the earth system of biology of flow of the energy. Such mechanism emulates some actions of living organisms like decomposition, consumption, and production.

In general, the producers can be any type of green plants that obtained their food using the photosynthesis process. During this process (i.e., photosynthesis), the sugar (i.e., oxygen and glucose) is produced by interactions between the water and carbon dioxide subject to the absence of sunlight. Then, this obtained sugar is used by plants to make fruits, leaves, roots, and wood. So, the herbivore consumers and omnivore consumers obtained their essential food from the producers. Since the consumers are animals and they can't produce their food and they only feed only on food from either producers or other consumers. In nature, there are three types of consumers (1) carnivores, (2) herbivores, and (3) omnivores. These animals that feed only on plants (i.e., producers) named herbivores; whereas, omnivores are these animals that have the ability to eat the producers and other animals. In addition, these animals that feed only on other animals are named carnivores. Decomposers represent the organism that feeds on producers (i.e., dead plants) and consumers (i.e., animals) or on the waste from living organisms. The fungi and the most of bacteria are decomposers, where the decomposers break down the remains of organisms dies and convert them into simple molecules, for example, minerals, water, and carbon dioxide. Followed by absorbing these energy types by producers to generate sugar using photosynthesis.

These actions can be modeled in a mathematical formulation where the action of production can control the trade-off of exploitation vs. exploration during the optimization process. During the decomposition action, the consumer can control the search space which is ended by deleting the intensification. Through such a system, the plants reach food using water, carbon dioxide, the light of the sun, and bacteria and fungi which can decompose nourishment.

The update process can be summarized as follows:

1. Production Procedure: according to the followed procedure in [41], the selection of the producer position is performed in a random way and the corresponding producer is the worst. However, the best solution represented by the decomposer can be modeled as following:

$$X_1(t+1) = (1-d)X_n(t) + d \cdot X_{rand}(t),$$
(2a)

$$d = \left(1 - \frac{t}{T_{max}}\right) rand_1 \tag{2b}$$

$$X_{rand}(t) = rand_2() \cdot (ub - lb) + lb,$$
(2c)

where *t* and T_{ma} are the current iteration and total number of iterations, respectively. *ub* and *lb* represents the upper and the lower boundaries of the search space. *rand*₁ and *rand*₂ are arbitrary variables in the interval [0,1] and *d* is a the weight parameter. $X_{rand}(t)$ donates a solution that generated randomly in the search space.

l

- 2. Consumption procedure: in such a procedure, the first user feeds to the other user with a lower level of energy or on a producer. Each set of users known as omnivores, vegetarian or herbivores, and carnivores has its mechanism in modernizing its position as follows:
 - (a) The herbivores locations can be modernized just with respect to the producers:

$$X_i(t+1) = X_i(t) + K \cdot (X_i(t) - X_1(t))$$
(3)

where X_1 represents the location of the producer and K represents a parameter for the consumption, it is determined using the levy flight by the following equations:

$$K = \frac{1}{2}\frac{u}{v}, \qquad u \in Norm(0,1), v \in Norm(0,1)$$
(4)

where Norm(0, 1) is a variable generated using the normal distribution with the zero mean and the unit variance.

(b) The update process of the carnivores is performed through the arbitrary customer with several levels of the energy which has an index (l). Such procedure can be modeled as:

$$X_{i}(t+1) = X_{i}(t) + K \cdot (X_{i}(t) - X_{l}(t)),$$
(5a)

$$l = randi([2i - 1]), i = 3, ..., N,$$
(5b)

where *randi*[*a*, *b*] is function used to generate random integer number in [*a*, *b*]].

(c) The position update of omnivores are depends on the producer and as well as the randomly chosen consumer with high level of energy index (*l*) as framed follows:

$$X_{i}(t+1) = X_{i}(t) + K \cdot (rand_{3} \cdot (X_{i}(t) - X_{1}(t))) + (1 - rand_{3}) \cdot (X_{i}(t) - X_{l}(t)), \quad (6a)$$

$$= randi([2i-1]), i = 3, ..., N,$$
 (6b)

3. Decomposition process: This represents the last phase in the biological system in which each agent passes on and the remaining parts are separated. This step refers to the exploitation of AEO and it is formulated as in [41]:

$$X_{i}(t+1) = X_{i}(t) + D \cdot (e \cdot X_{n}(t) - h \cdot X_{i}(t)), i = 1, ..., N,$$
(7a)

$$D = 3u, \quad u \in N(0, 1) \tag{7b}$$

$$e = rand_4 \cdot randi(\begin{bmatrix} 1 & 2 \end{bmatrix}) - 1, \tag{7c}$$

$$h = 2 \cdot rand_4 - 1, \tag{7d}$$

In Equation (7), the parameter *D* refers to the decomposition factor, *h* and *e* represent the weight parameters. $rand_4$ is random number generated from [0,1].

The steps of AEO are given in Algorithm 1. Moreover, The AEO has some advantages over the other MH techniques such as no parameters need to be determined during the optimization process. In addition, it has a high ability to balance between the exploration and exploitation which leads to improve convergence and avoids stuck at local optima. Consequently, improve the quality of the output. These characteristics made it more suitable to combine with DNN.

Algorithm 1 The AEO algorithm steps [41].	
Inputs : <i>N</i> the number of solution and T_{max} : total number of iterations.	
Generate initial ecosystem X (solutions).	
Compute the fitness value Fit_i , and X_1 is the best solution.	
t = 1.	
repeat	
Update X_1 using Equation (2).	▷ Production
for $i = 2,, n$ do	▷ Consumption
if <i>rand</i> < 1/3 then	
Update X_i using Equation (3),	⊳ Herbivore
else if 1/3 < <i>rand</i> < 2/3 then	
Update X_i using Equation (6),	⊳ Omnivore
else	
Update X_i using Equation (5),	▷ Carnivore
Compute the fitness of each X_i .	
Find the best solution X_1 .	▷ Decomposition
Update X_i using Equation (7).	
Compute the fitness of each X_i .	
Update the best solution X_1 .	
t = t + 1.	
until ($t < T_{max}$)	
Return X ₁ .	

3. Proposed MobileNet-AEO for Chest X-ray Classification Approach

The proposed MobileNet-AEO approach starts by extracting deep features from raw Chest X-ray images using MobileNet as discussed in Section 2.1. About 50 K features are produced, which represent the output of the last layer $7 \times 7 \times 1024$.

Then a selection of the optimal subset of relevant features is performed using binary AEO as a feature selection technique. The AEO set the initial value for *N* solutions each of them has dimension equal to the total number of extracted features *Dim*. This process is formulated as:

$$U_i = LB_i + \alpha_i \times (UB_i - LB_i), j = 1, 2, \dots, Dim$$
(8)

where $\alpha_i \in [0, 1]$ is a randomly generated number [46]. The $LB_i = 1$ and $UB_i = 0$ indicates the bottom and top boundary of the domain of searching.

Thereafter, each solution U_i is converted into binary vector (BU_i) using Equation (9).

$$BU_{ij} = \begin{cases} 1 & if \ U_{ij} > 0.5\\ 0 & otherwise \end{cases}$$
(9)

The binary vector considered as the main step to determine the relevant features. Those extracted features that corresponding to zeros will be removed and the rest features represent the relevant features. For example, assume $U_i = [0.58, 0.94, 0.72, 0.21, 0.12, 0.78]$. By using Equation (9), $BU_i = [1, 1, 1, 0, 0, 1]$. Then, the 3rd and 4th features will be removed while the remaining features will be kept as the relevant ones.

To validate the quality of the selected features using the current solution U_i , the following formula is used which represents the fitness function.

$$Fit_i = \lambda \times \gamma_i + (1 - \lambda) \times \left(\frac{|BU_i|}{Dim}\right)$$
(10)

In Equation (10), $|BU_i|$ is the total number of the selected features. λ is applied to balance between the ratio of the selected features (the second part of Equation (10)) and the error of classification (γ_i) (first part). In this stage, KNN classifier is used to assess the performance of the selected features from the training set (here, it is 80% from the whole dataset).

The next process is to find the best solution U_b which has the smallest Fit_b . Followed by updating the solutions using operators of traditional AEO as discussed in Section 2.2. The process of updating the solutions is repeated until reaching the total number of iterations. The best solution U_b is the output from this stage, to evaluate its quality, the testing set (represents 20% from the dataset) is reduced according to it. Again, KNN is used to predict the target of the testing set and the classification metric is computed. The outline of the proposed MobileNet-AEO approach is presented in Figure 2.



Figure 2. Flow chart of MobileNet-AEO approach.

4. Datasets and Evaluation

4.1. Dataset Description

The dataset we used is called Shenzhen dataset [47]. It was collected by Shenzhen Hospital, in China. The chest X-rays were collected from outpatient clinics, they were captured daily- bases routine within 1-month, mostly in 2012, using a specialized medical Diagnostic system. It contains 662 frontal chest X-ray images, of which 326 images represent the benign cases, while 336 images represent TB (malignant) cases. All image file names have the same naming template: CHNCXR_####_X.png, where X can be "0" for Non-TB (benign) or "1" for Tuberculosis (malignant) X-ray. A clinical report is available for each X-ray in a file with the same format, which contains the patient's age, gender, and abnormality seen in the lung if any.

In order to validate the proposed approach, we apply it on another medical application that works also with chest X-ray to detect Pneumonia.We also used a second, unbalanced and more recently collected dataset [48], which we call *Dataset 2*. It was collected and labeled by physicians. It consists of 5232 children's chest X-ray images, 3883 of them labeled as showing the signs of pneumonia (2538 bacterial and 1345 viral), where the rest 2538 labeled as normal. Both datasets represents a binary classification problem, where the proposed approach has to classify each input image into malignant or benign. Table 1 shows some examples for the two datasets.

Table 1. Samples from Shenzhen dataset (top) and Dataset 2 (bottom).



Table 1 illustrates the variation in each dataset, in terms of morphology, structure, shape, and zoom level. The variation in size (height and width) is from 200 to 4000 pixels, also each of them has a different image file format.

4.2. Evaluation

We used accuracy (Acc), sensitivity (Sens), specificity (Spec) [49] and time consumption as fitness measures. These are defined as follows:

$$Accuracy = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$
(11)

$$Sensitivity = \frac{TP}{TP + FN}$$
(12)

$$Specificity = \frac{1N}{TN + FP}$$
(13)

where TP (true positives) is the number of the TB samples that were labeled correctly, TN (true negatives) is the number of the not-TB samples that were labeled correctly. FP (false positives) is the number of the TB samples that were labeled incorrectly as being not TB samples, and "FN" (false negatives) is the number of the not-TB samples that were miss classified as the TB samples.

We also calculate the standard deviation (STD) of the fitness measures as follows:

$$STD = \sqrt{\frac{1}{r-1} \sum_{i=1}^{r} (Fit_i - \mu)^2}, \ \mu = \frac{1}{r} \sum_{i=1}^{r} Fit_i$$
(14)

where *r* is the run numbers. Fit_i denotes a fitness function value. μ represents the average of the fitness value overall *r*.

4.3. Implementation Environment

The proposed approach was implemented in Python 3 on Windows 10 64 bit using a Core i7 CPU and 8 GB RAM, besides Google Colaboratory "Colab" [50]. The model was developed using Keras library [51] with Tensorflow backend [52]. For feature selection, the experiment was performed using Matlab 2018b on a computer Core i5 and 8 GB of RAM running with Windows 10.

5. Results and Discussion

5.1. Parameters

The proposed approach has been trained on 80% of samples (537 for Shenzhen dataset and 4160 for Dataset 2), while the rest 20% (134 for Shenzhen dataset and 1040 for Dataset 2) were used for testing (external validation) the model's performance. There is no overlap between any of the two sets. Also, all results are given on the testing set. In this work, we adopt the following parameters for building MobileNet. A learning rate of 0.0001 and a mini-batch of size 20 and binary cross-entropy as a loss function were used. Also, we adopt Rmsprop as the optimization algorithm [53]. In total, the model has 4,253,864 trainable parameters.

5.2. Performance

We compare the performance of AEO against other metaheuristic techniques including Harris Hawks optimization (HHO) [54], Henry gas solubility optimization (HGSO) [55], Whale optimization algorithm (WOA) [56], grey wolf optimization (GWO) [57], sine-cosine algorithm (SCA) [58], teaching learning based optimization (TLBO) [59] and traditional AEO.

The parameters of each algorithm are set to default. The input to all feature selection algorithms is the features extracted by Mobilenet.

Tables 2 and 3 show the results of the feature selection process over the two datasets, Shenzhen and Dataset 2. From the presented results, it can be noticed that our method based on AEO excels other methods. For example, in terms of accuracy, it achieves the first rank, followed by the SCA and TLBO that allocate second and third rank, respectively. Similarly, in terms of sensitivity, AEO performs well as has the highest best, worst, and means sensitivity values. In terms of specificity, AEO shows better best value than all other methods. However, the SCA and TLBO provide better mean and worst specificity when compared to other methods. These results indicates the high ability of the proposed AEO algorithm to find the relevant features and this reflected from its Best value of accuracy, sensitivity, and specificity. In addition, by analysis the value of the worst case in terms of performance measures, one can be noticed that the AEO outperforms other FS methods in sensitivity value. However, in accuracy and specificity allocates the third rank after SCA and TLBO, respectively. For the stability of the FS algorithms as measured using STD, it can be noticed that the WOA, AEO, and TLBO are more stable in terms of accuracy, sensitivity and specificity. Moreover, by analysis the results in terms of Best, Worst, and STD value for Datatset 2, it cab be noticed that AEO has high superiority over other models except in terms of STD of accuracy it has be observed that the GWO is more stable.

		AEO	HHO	HGSO	WOA	SCA	GWO	TLBO
	Best	0.9023	0.8195	0.8120	0.8120	0.8872	0.8421	0.8872
1	Mean	0.8617	0.7880	0.7835	0.7744	0.8436	0.8150	0.8526
Acc	Worst	0.8045	0.7368	0.7669	0.7293	0.8271	0.7744	0.8195
	STD	0.0400	0.0374	0.0171	0.0340	0.0252	0.0294	0.0253
	Best	0.9194	0.8923	0.8448	0.8636	0.8676	0.8500	0.9077
Carra	Mean	0.8839	0.8092	0.8069	0.7848	0.8294	0.8033	0.8369
Sens	Worst	0.8548	0.7231	0.7759	0.7273	0.8088	0.7667	0.7846
	STD	0.0239	0.0612	0.0256	0.0540	0.0246	0.0361	0.0456
	Best	0.9014	0.8382	0.8133	0.8358	0.9385	0.8630	0.8824
Creation	Mean	0.8423	0.7676	0.7653	0.7642	0.8585	0.8247	0.8676
Spec	Worst	0.7465	0.6765	0.7333	0.6567	0.8000	0.7123	0.8529
	STD	0.0609	0.0627	0.0335	0.0711	0.0503	0.0653	0.0104

Table 2. Comparison of performance for Shenzhen dataset. The best values are shown in bold.

Table 3. Comparison of performance for Dataset 2. The best values are shown in bold.

		AEO	ННО	HGSO	WOA	SCA	GWO	TLBO
	Best	0.9418	0.9152	0.9041	0.9187	0.9307	0.9349	0.9050
A	Mean	0.9360	0.8964	0.8955	0.8991	0.9199	0.9322	0.8997
Acc	Worst	0.9307	0.8767	0.8759	0.8690	0.9110	0.9289	0.8955
	STD	0.0048	0.0153	0.0116	0.0189	0.0081	0.0022	0.0038
	Best	0.8722	0.8291	0.8306	0.8384	0.8630	0.8642	0.8148
Carra	Mean	0.8518	0.7905	0.7792	0.7848	0.8327	0.8463	0.7872
Sens	Worst	0.8307	0.7025	0.7264	0.6768	0.8017	0.8210	0.7609
	STD	0.0174	0.0538	0.0393	0.0632	0.0247	0.0183	0.0191
	Best	0.9708	0.9495	0.9501	0.9500	0.9612	0.9704	0.9495
Spec	Mean	0.9668	0.9357	0.9370	0.9438	0.9561	0.9652	0.9380
	Worst	0.9637	0.9061	0.9268	0.9310	0.9491	0.9609	0.9288
	STD	0.0027	0.0178	0.0094	0.0075	0.0054	0.0043	0.0091

Figure 3 shows a comparison between the proposed approach and other feature selection algorithms. From this figure, it can be seen that the proposed approach shows advantages compared to the other algorithms in terms of accuracy, sensitivity and specificity on all experiments (best, mean and worst).









⁽c) Sensitivity



Table 4 depicts the performance of AEO in terms of computation time and number of selected features for Shenzhen dataset and Dataset 2. For Shenzhen dataset, it can be seen that AEO allocates the fourth rank by selecting 24.6 features on average, compared to WOA, which ranked first in selecting the smallest subset features, with only 10.8 features selected on average. However, the extracted features by WOA are less efficient as they show lower performance compared to AEO, as seen in Tables 2 and 3. In contrast, the proposed AEO provides the smallest and most significant feature set among other methods, on dataset 2.

		AEO	нно	HGSO	WOA	SCA	GWO	TLBO
	Features	24.6	9.6	11.6	10.8	42.2	58.8	19.6
Datacet1	Best time (s)	8.8713	9.1562	8.1690	4.1322	5.1606	9.0793	9.1075
Dataset1	Mean time (s)	9.7251	9.6065	8.5953	4.3891	5.2591	9.5839	9.5879
	STD time (s)	0.9068	0.3497	0.2658	0.1653	0.0912	0.3606	0.4349
	Features	19	33.8	31.6	52.6	31.8	72	30.8
Datasat	Best time (s)	137.7945	250.7114	671.3892	125.0128	251.9918	422.1539	247.6316
Datasetz	Mean time (s)	162.4733	281.7812	743.5204	128.3362	257.9852	467.4882	297.0471
	STD time (s)	7.9267	30.0768	50.8346	2.2845	9.4944	26.7742	33.5236

Table 4. Mean number of features selected and time performance. Best values are shown in bold.

5.3. Comparison with Other CNN Models

Here we compare the proposed approach (MobileNet-AEO) to MobileNet and other CNN models based on the classification evaluation criteria, (i.e., accuracy, specificity and sensitivity). Table 5 shows a comparison between our approach and MobileNet.

Table 5. Number of features and performance for Shenzhen dataset and Dataset 2. Best values are shown in bold.

Shenzhen	Features	Percentage	Accuracy	Specificity	Sensitivity
MobileNet	50176	100%	0.89	0.89	0.90
Proposed approach	~ 25	0.05 %	0.902	0.901	0.914
Dataset 2	Features	Percentage	Accuracy	Specificity	Sensitivity
MobileNet	50176	100%	0.842	0.846	0.846
Proposed approach	19	0.038 %	0.941	0.97	0.872

In Table 5, we compare the features extracted from MobileNet and those extracted by our method. Only 0.05% and 0.038% of MobileNet features were extracted from Shenzhen and Dataset 2, respectively. The MobileNet-AEO method uses only 24 and 19 features (for Shenzhen dataset and Dataset 2, respectively) shows better performance in all classification measures than the basic MobileNet feature set that has some 50 K features. Figure 4 presents a comparison between the proposed approach and different efficient CNN architectures such as VGGNet (VGG 16 and VGG 19) [34], ResNet [35], NasNet [36], MobileNet [37], Inception [38] and Xception [39].



Figure 4. Performance (accuracy, sensitivity and specificity) of models on Shenzhen (**left**) and Dataset 2 (**right**).

As shown in Figure 4 (top), the proposed MobileNet-AEO approach outperforms all other CNN architectures in both accuracy and sensitivity with a slight advantage than the basic version of MobileNet, while VGG 19 comes first in specificity with 91% compared to ours, 90%. In Figure 4 (right), which represents Dataset 2, the proposed approach shows an advantage among all CNNs in all classification criteria. Also, ResNet comes last with both datasets for accuracy, sensitivity and specificity. It is noted that NasNet was excluded from the second dataset experiments due to resource limitations as it contains 88 K parameters, which produces about 487 K of features, which put it first as the deepest CNN.

5.4. Comparison with Related Works

Here we compare our results with relevant works. Table 6 shows the most recent published works on both Shenzhen and Dataset 2.

Shenzhen Dataset Jaeger et al. [17] Hwang et al. [21] Lopes et al. [18] Sivaramakrishnan et al.

Proposed approach

Kermany et al. [48]

Rajaraman et al. [61]

Proposed approach

Dataset 2

	Feature Extraction	Classifier	Accuracy (%)
	Manually	SVM	84.10
	Deep features by CNN	KNN	83.70
	ResNet, VGG and GoogLeNet	SVM	84.60
[60]	VGG16 with optimal features	CNN	85.5
-	Deep features by MobileNet,		

CNN

Classifier

N/A

Residual CNN

Inception

CNN

90.2

Accuracy (%)

92.8

91

88.6

94.1

Table 6. Comparison of results with related works on Shenzhen and Dataset 2. Best values are shown in bold.

feature selection by AEO

Feature extraction

Deep features

Deep features

by CNN architectures

Deep features by MobileNet,

feature selection by AEO

As seen in Table 6 (top), the proposed approach has an advantage in performance (accuracy) over other recent works. Also, the first three models with the highest performance used CNN as a feature extractor, which means that CNN can extract the most informative features that improve the model's performance. Although Jaeger et al. [17] extracted features from chest x-ray images then adopted various classification methods, the results reported were achieved using low-level image features and linear logistic regression classifiers. Also, Lopes et al. [18] adopted the bags of features method on features extracted from GoogLeNet, ResNet, and VGG networks and then applied an ensemble of individual SVM classifiers.

In Dataset 2, as shown in Table 6 (button), the proposed method come also first among other previous works on the same dataset. Although in [61], the authors claim that they achieved 94.3% of classification accuracy, but they provided no more details about the model they proposed, they named their model as Sequential CNN.

6. Conclusions

In this paper, a new hybrid method for Tuberculosis X-ray image classification was introduced. The method is based on extracting features from the chest X-ray images using a MobileNet deep neural network, and then filtering the produced huge number of features using the recently proposed Artificial Ecosystem-based Optimization (AEO) algorithm to include only the relevant features and exclude the irrelevant features.

The classification approach (MobileNet-AEO) was validated using two publicly available datasets, Shenzhen Dataset and Dataset 2, which both contain chest X-ray images. MobileNet-AEO performed well by achieving high classification accuracy. Also, the complexity was reduced, which positively affects the computation time. MobileNet-AEO was successful in reducing the number of features from 50 K to only 25 and 19 and achieving an accuracy of 90.2% and 94.1% for Shenzhen dataset and Dataset 2, respectively, while increasing the performance at the same time. The proposed MobileNet-AEO approach outperforms all published works on the two datasets, as well as showed an advantage when compared to other convolutional network models. Our future work will include building a hybrid approach that combines a transfer learning model and a swarm optimisation algorithm to build a classification model for COVID-19 diagnostics from the chest radiographs.

Author Contributions: Conceptualization, A.T.S., M.A.E., A.T.J. and R.D.; formal analysis, A.T.S., M.A.E., R.D. and O.F.H.; data curation, A.T.S., M.A.E. and A.T.J.; writing–original draft preparation, A.T.S. and M.A.E.; writing–review and editing, A.T.S., M.A.E., R.D. and O.F.H.; funding acquisition, R.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

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

References

- 1. Pasa, F.; Golkov, V.; Pfeiffer, F.; Cremers, D.; Pfeiffer, D. Efficient Deep Network Architectures for Fast Chest X-Ray Tuberculosis Screening and Visualization. *Sci. Rep.* **2019**, *9*, 6268. [CrossRef] [PubMed]
- 2. Anderson, L.; Dean, A.; Falzon, D.; Floyd, K.; Baena, I.; Gilpin, C.; Glaziou, P.; Hamada, Y.; Hiatt, T.; Char, A.; et al. Global tuberculosis report 2015. *WHO Libr. Cat. Data* **2015**, *1*, 1689–1699.
- Sui, Y.; Wei, Y.; Zhao, D. Computer-Aided Lung Nodule Recognition by SVM Classifier Based on Combination of Random Undersampling and SMOTE. *Comput. Math. Methods Med.* 2015, 2015, 1–13. [CrossRef] [PubMed]
- 4. Santosh, K.C.; Antani, S. Automated Chest X-Ray Screening: Can Lung Region Symmetry Help Detect Pulmonary Abnormalities? *IEEE Trans. Med. Imaging* **2018**, *37*, 1168–1177. [CrossRef] [PubMed]
- Khan, M.A.; Rubab, S.; Kashif, A.; Sharif, M.I.; Muhammad, N.; Shah, J.H.; Zhang, Y.D.; Satapathy, S.C. Lungs cancer classification from CT images: An integrated design of contrast based classical features fusion and selection. *Pattern Recognit. Lett.* 2020, 129, 77–85. [CrossRef]
- Filho, P.; Barros, A.; Ramalho, G.; Pereira, C.; Papa, J.; de Albuquerque, V.; Tavares, J. Automated recognition of lung diseases in CT images based on the optimum-path forest classifier. *Neural Comput. Appl.* 2019, 31, 901–914. [CrossRef]
- Woźniak, M.; Połap, D. Bio-inspired methods modeled for respiratory disease detection from medical images. *Swarm Evol. Comput.* 2018, 41, 69–96. [CrossRef]
- 8. Gupta, N.; Gupta, D.; Khanna, A.; Rebouças Filho, P.; de Albuquerque, V. Evolutionary algorithms for automatic lung disease detection. *Meas. J. Int. Meas. Confed.* **2019**, *140*, 590–608. [CrossRef]
- Połap, D.; Woźniak, M.; Damaševičius, R.; Wei, W. Chest radiographs segmentation by the use of nature-inspired algorithm for lung disease detection. In Proceedings of the 2018 IEEE Symposium Series on Computational Intelligence (SSCI), Bangalore, India, 18–21 November 2018; pp. 2298–2303. [CrossRef]
- Abiyev, R.; Ma'aitah, M. Deep Convolutional Neural Networks for Chest Diseases Detection. J. Healthc. Eng. 2018, 2018. [CrossRef]
- 11. Hosny, A.; Parmar, C.; Quackenbush, J.; Schwartz, L.; Aerts, H. Artificial intelligence in radiology. *Nat. Rev. Cancer* **2018**, *18*, 500–510. [CrossRef]
- Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; Van Der Laak, J.A.; Van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* 2017, 42, 60–88. [CrossRef] [PubMed]
- 13. Rouhi, R.; Jafari, M.; Kasaei, S.; Keshavarzian, P. Benign and malignant breast tumors classification based on region growing and CNN segmentation. *Expert Syst. Appl.* **2015**, *42*, 990–1002. [CrossRef]
- Woźniak, M.; Połap, D.; Capizzi, G.; Sciuto, G.; Kośmider, L.; Frankiewicz, K. Small lung nodules detection based on local variance analysis and probabilistic neural network. *Comput. Methods Programs Biomed.* 2018, 161, 173–180. [CrossRef] [PubMed]
- 15. Chouhan, V.; Singh, S.; Khamparia, A.; Gupta, D.; Tiwari, P.; Moreira, C.; Damaševičius, R.; de Albuquerque, V. A novel transfer learning based approach for pneumonia detection in chest X-ray images. *Appl. Sci.* **2020**, *10*, 559. [CrossRef]
- 16. Melendez, J.; Sánchez, C.I.; Philipsen, R.H.; Maduskar, P.; Dawson, R.; Theron, G.; Dheda, K.; Van Ginneken, B. An automated tuberculosis screening strategy combining X-ray-based computer-aided detection and clinical information. *Sci. Rep.* **2016**, *6*, 25265. [CrossRef]
- 17. Jaeger, S.; Karargyris, A.; Candemir, S.; Folio, L.; Siegelman, J.; Callaghan, F.; Xue, Z.; Palaniappan, K.; Singh, R.K.; Antani, S.; et al. Automatic tuberculosis screening using chest radiographs. *IEEE Trans. Med. Imaging* **2013**, *33*, 233–245. [CrossRef]
- 18. Lopes, U.; Valiati, J.F. Pre-trained convolutional neural networks as feature extractors for tuberculosis detection. *Comput. Biol. Med.* 2017, *89*, 135–143. [CrossRef]
- 19. Vajda, S.; Karargyris, A.; Jaeger, S.; Santosh, K.; Candemir, S.; Xue, Z.; Antani, S.; Thoma, G. Feature selection for automatic tuberculosis screening in frontal chest radiographs. *J. Med. Syst.* **2018**, *42*, 146. [CrossRef]
- 20. Lakhani, P.; Sundaram, B. Deep learning at chest radiography: Automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology* **2017**, *284*, 574–582. [CrossRef]

- Hwang, S.; Kim, H.E.; Jeong, J.; Kim, H.J. A novel approach for tuberculosis screening based on deep convolutional neural networks. In SPIE Medical Imaging, Proceedings of the Medical Imaging 2016: Computer-Aided Diagnosis, San Diego, CA, USA, 27 February–3 March 2016; International Society for Optics and Photonics: Bellingham, WA, USA, 2016; Volume 9785, p. 97852W.
- 22. Islam, M.T.; Aowal, M.A.; Minhaz, A.T.; Ashraf, K. Abnormality detection and localization in chest X-rays using deep convolutional neural networks. *arXiv* **2017**, arXiv:1705.09850.
- Shin, H.C.; Roberts, K.; Lu, L.; Demner-Fushman, D.; Yao, J.; Summers, R.M. Learning to read chest x-rays: Recurrent neural cascade model for automated image annotation. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2497–2506.
- Deng, J.; Dong, W.; Socher, R.; Li, L.J.; Li, K.; Fei-Fei, L. Imagenet: A large-scale hierarchical image database. In Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition, Miami, FL, USA, 20–25 June 2009; pp. 248–255.
- 25. Shin, H.C.; Roth, H.R.; Gao, M.; Lu, L.; Xu, Z.; Nogues, I.; Yao, J.; Mollura, D.; Summers, R.M. Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. *IEEE Trans. Med. Imaging* **2016**, *35*, 1285–1298. [CrossRef] [PubMed]
- 26. Ho, T.K.K.; Gwak, J. Multiple feature integration for classification of thoracic disease in chest radiography. *Appl. Sci.* **2019**, *9*, 4130. [CrossRef]
- Abdeldaim, A.M.; Sahlol, A.T.; Elhoseny, M.; Hassanien, A.E. Computer-aided acute lymphoblastic leukemia diagnosis system based on image analysis. In *Advances in Soft Computing and Machine Learning in Image Processing*; Springer: Berlin, Germany, 2018; pp. 131–147.
- Sahlol, A.T.; Abdeldaim, A.M.; Hassanien, A.E. Automatic acute lymphoblastic leukemia classification model using social spider optimization algorithm. In *Soft Computing*; Springer: Berlin, Germany, 2018; pp. 1–16.
- 29. Ke, Q.; Zhang, J.; Wei, W.; Połap, D.; Woźniak, M.; Kośmider, L.; Damaševičius, R. A neuro-heuristic approach for recognition of lung diseases from X-ray images. *Expert Syst. Appl.* **2019**, *126*, 218–232. [CrossRef]
- 30. Qi, G.; Luo, J. Small Data Challenges in Big Data Era: A Survey of Recent Progress on Unsupervised and Semi-Supervised Methods. *arXiv* **2019**, arXiv:1903.11260.
- Sharif Razavian, A.; Azizpour, H.; Sullivan, J.; Carlsson, S. CNN features off-the-shelf: An astounding baseline for recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, Columbus, OH, USA, 23–28 June 2014; pp. 806–813.
- 32. Donahue, J.; Jia, Y.; Vinyals, O.; Hoffman, J.; Zhang, N.; Tzeng, E.; Darrell, T. Decaf: A deep convolutional activation feature for generic visual recognition. In Proceedings of the International Conference on Machine Learning, Beijing, China, 21–26 June 2014; pp. 647–655.
- Nguyen, L.D.; Lin, D.; Lin, Z.; Cao, J. Deep CNNs for microscopic image classification by exploiting transfer learning and feature concatenation. In Proceedings of the 2018 IEEE International Symposium on Circuits and Systems (ISCAS), Florence, Italy, 27–30 May 2018; pp. 1–5.
- 34. Simonyan, K.; Zisserman, A. Very deep convolutional networks for large-scale image recognition. *arXiv* **2014**, arXiv:1409.1556.
- 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.
- Blog, G. AutoML for Large Scale Image Classification and Object Detection. 2017. Available online: https://research.googleblog.com/2017/11/automl-for-large-scaleimage.html (accessed on 31 May 2020).
- Howard, A.G.; Zhu, M.; Chen, B.; Kalenichenko, D.; Wang, W.; Weyand, T.; Andreetto, M.; Adam, H. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv* 2017, arXiv:1704.04861.
- Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
- 39. 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.
- 40. Li, Z.; Wang, C.; Han, M.; Xue, Y.; Wei, W.; Li, L.J.; Fei-Fei, L. Thoracic disease identification and localization with limited supervision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Palace Convention Cetner, Salt Lake City, UT, USA, 28–23 June 2018; pp. 8290–8299.

- 41. Zhao, W.; Wang, L.; Zhang, Z. Artificial ecosystem-based optimization: A novel nature-inspired meta-heuristic algorithm. *Neural Comput. Appl.* **2019**. [CrossRef]
- 42. Ioffe, S.; Szegedy, C. Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv* 2015, arXiv:1502.03167.
- 43. Bengio, Y.; Courville, A.; Vincent, P. Representation learning: A review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* **2013**, *35*, 1798–1828. [CrossRef]
- 44. Bolboacă, S.D.; Jäntschi, L.; Sestraş, A.F.; Sestraş, R.E.; Pamfil, D.C. Pearson-Fisher chi-square statistic revisited. *Information* **2011**, *2*, 528–545. [CrossRef]
- 45. Sahlol, A.T.; Kollmannsberger, P.; Ewees, A.A. Efficient Classification of White Blood Cell Leukemia with Improved Swarm Optimization of Deep Features. *Sci. Rep.* **2020**, *10*, 1–11. [CrossRef]
- 46. Bálint, D.; Jäntschi, L. Missing data calculation using the antioxidant activity in selected herbs. *Symmetry* **2019**, *11*, 779. [CrossRef]
- 47. Jaeger, S.; Candemir, S.; Antani, S.; Wáng, Y.X.J.; Lu, P.X.; Thoma, G. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. *Quant. Imaging Med. Surg.* **2014**, *4*, 475–477. [PubMed]
- Kermany, D.S.; Goldbaum, M.; Cai, W.; Valentim, C.C.; Liang, H.; Baxter, S.L.; McKeown, A.; Yang, G.; Wu, X.; Yan, F.; et al. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 2018, 172, 1122–1131. [CrossRef] [PubMed]
- 49. Zhu, W.; Zeng, N.; Wang, N. Sensitivity, specificity, accuracy, associated confidence interval and ROC analysis with practical SAS implementations. *NESUG Proc.* **2010**, *19*, *67*.
- 50. Bisong, E. Building Machine Learning and Deep Learning Models on Google Cloud Platform; Springer: Berlin, Germany, 2019.
- 51. Chollet, F. Keras. 2015. Available online: https://github.com/fchollet/keras (accessed on 31 May 2020).
- 52. Abadi, M.; Agarwal, A.; Barham, P.; Brevdo, E.; Chen, Z.; Citro, C.; Corrado, G.S.; Davis, A.; Dean, J.; Devin, M.; et al. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems. 2015. Available online: tensorflow.org (accessed on 31 May 2020).
- 53. Wichrowska, O.; Maheswaranathan, N.; Hoffman, M.W.; Colmenarejo, S.G.; Denil, M.; de Freitas, N.; Sohl-Dickstein, J. Learned optimizers that scale and generalize. In Proceedings of the 34th International Conference on Machine Learning, Sydney, Australia, 6–11 August 2017; Volume 70, pp. 3751–3760.
- 54. Heidari, A.A.; Mirjalili, S.; Faris, H.; Aljarah, I.; Mafarja, M.; Chen, H. Harris hawks optimization: Algorithm and applications. *Future Gener. Comput. Syst.* **2019**, *97*, 849–872. [CrossRef]
- 55. Hashim, F.A.; Houssein, E.H.; Mabrouk, M.S.; Al-Atabany, W.; Mirjalili, S. Henry gas solubility optimization: A novel physics-based algorithm. *Future Gener. Comput. Syst.* **2019**, *101*, 646–667. [CrossRef]
- 56. Mirjalili, S.; Lewis, A. The whale optimization algorithm. Adv. Eng. Softw. 2016, 95, 51–67. [CrossRef]
- Ibrahim, R.A.; Elaziz, M.A.; Lu, S. Chaotic opposition-based grey-wolf optimization algorithm based on differential evolution and disruption operator for global optimization. *Expert Syst. Appl.* 2018, 108, 1–27. [CrossRef]
- 58. Elaziz, M.A.; Oliva, D.; Xiong, S. An improved opposition-based sine cosine algorithm for global optimization. *Expert Syst. Appl.* **2017**, *90*, 484–500. [CrossRef]
- 59. Allam, M.; Nandhini, M. Optimal feature selection using binary teaching learning based optimization algorithm. *J. King Saud Univ. Comput. Inf. Sci.* 2018. [CrossRef]
- 60. Sivaramakrishnan, R.; Antani, S.; Candemir, S.; Xue, Z.; Abuya, J.; Kohli, M.; Alderson, P.; Thoma, G. Comparing deep learning models for population screening using chest radiography. In SPIE Medical Imaging, Proceedings of the Medical Imaging 2018: Computer-Aided Diagnosis, Houston, TX, USA, 10–15 February 2018; International Society for Optics and Photonics: Bellingham, WA, USA, 2018; Volume 10575, p.105751E.
- 61. Rajaraman, S.; Candemir, S.; Kim, I.; Thoma, G.; Antani, S. Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. *Appl. Sci.* **2018**, *8*, 1715. [CrossRef] [PubMed]



© 2020 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 (http://creativecommons.org/licenses/by/4.0/).