

Evaluating Convolutional Neural Networks for COVID-19 classification in chest X-ray images

Leonardo Gabriel Ferreira Rodrigues, Larissa Ferreira Rodrigues,
Danilo Ferreira da Silva, João Fernando Mari
Instituto de Ciências Exatas e Tecnológicas
Universidade Federal de Viçosa - UFV
Caixa Postal 22 - 38.810-000 - Rio Paranaíba - MG - Brasil
Email: {leonardo.g.rodrigues, larissa.f.rodrigues, danilo.f.silva, joaof.mari}@ufv.br

Abstract—Coronavirus Disease 2019 (COVID-19) pandemic rapidly spread globally, impacting the lives of billions of people. The effective screening of infected patients is a critical step to struggle with COVID-19, and treating the patients avoiding this quickly disease spread. The need for automated and scalable methods has increased due to the unavailability of accurate automated toolkits. Recent researches using chest X-ray images suggest they include relevant information about the COVID-19 virus. Hence, applying machine learning techniques combined with radiological imaging promises to identify this disease accurately. It is straightforward to collect these images once it is spreadly shared and analyzed in the world. This paper presents a method for automatic COVID-19 detection using chest X-ray images through four convolutional neural networks, namely: AlexNet, VGG-11, SqueezeNet, and DenseNet-121. This method had been providing accurate diagnostics for positive or negative COVID-19 classification. We validate our experiments using a ten-fold cross-validation procedure over the training and test sets. Our findings include the shallow fine-tuning and data augmentation strategies that can assist in dealing with the low number of positive COVID-19 images publicly available. The accuracy for all CNNs is higher than 97.00%, and the SqueezeNet model achieved the best result with 99.20%.

Keywords—COVID-19; coronavirus; chest X-ray; convolutional neural networks; data augmentation; fine-tuning.

I. INTRODUCTION

Coronavirus Disease 2019 (COVID-19) caused by a novel coronavirus, officially named Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) [1], is a pandemic that first emerged in the Chinese city of Wuhan, and rapidly spread to other countries, affecting Italy, Iran, Spain, Brazil, Russia, India, and United States severely. All countries affected by COVID-19 imposed a nationwide lockdown in an attempt to slow the spread of the virus, causing a profound overall impact on the lives of billions of people from a health, safety, and economic perspective [2] [3]. The COVID-19 can cause illness to the respiratory system leading to inflammation of the lungs and pneumonia [4]. There is no known specific therapeutic drugs or vaccine for COVID-19, and the impact in the healthcare system is also high due to the number of people that needs intensive care unit (ICU) admission and breathing machine for long periods [5].

The most common test technique currently used for COVID-19 diagnosis is the reverse transcription-polymerase chain reaction (RT-PCR). However, considering the difficulties

of distributing the kits and collecting the samples and the waiting time for results, auxiliary diagnostics methods are welcome to assist the medical team decision making. In this context, the development of computer-aided diagnosis systems based on machine learning is essential and widely applied in several fields of medicine [6] [7] [8].

Early studies demonstrated that many patients infected with COVID-19 present abnormalities in chest X-ray images [9] [10] [11]. These images can be easily collected, shared, and analyzed around the world. Moreover, the task of COVID-19 identification is not easy, and the specialist reviewing the chest X-ray needs to look for white patches in the lungs, i.e., air sacs filled with pus or water. However, these white patches can also be confused with diseases such as tuberculosis, bronchitis, and other types of pneumonia caused by different viruses or bacteria.

In this study, we aim to explore the identification of COVID-19 using chest X-ray images due to its reduced cost, fast result, and general availability. Our principal goal is to reach the best possible identification rate among COVID-19 and other types of pneumonia. We applied a pure deep learning approach comparing four Convolutional Neural Networks (CNNs): AlexNet, VGG-11, SqueezeNet and DenseNet-121, and we evaluated the performance using a k-fold cross-validation procedure over the training and test sets. Moreover, we carried a confusion matrix analysis to measure accuracy, precision, recall, and F1-score indices.

Furthermore, the chest X-ray image dataset used in this work contains a few positive images on the COVID-19. To deal with this, we applied Shallow Fine-Tuning (SFT) training and data augmentation based on random rotation and shifting to balance the class distribution and the performance of the classification. In addition, our approach is fast and simple producing high performing system. As far as we know, our result is the best obtained for COVID-19 identification in chest X-ray images. We believe that our proposed method can contribute to future researches intended to help healthcare workers to identify COVID-19 and to manage patient's conditions.

The remaining of this paper is organized as follows: Section II surveys related work; Section III describes the material and methods; In Section IV we present and discuss the results obtained. Finally, conclusions and future work are presented in Section V.

II. RELATED WORK

The COVID-19 has been attracting much attention from the image analysis research community due to its severity. In this sense, Narin et al. [12] compared three different CNN architectures (ResNet50, Inception-V3, and InceptionResNetV2) to identify COVID-19 in chest X-ray images. They used a dataset composed of fifty COVID-19 images taken from the open-source GitHub repository shared by Dr. Joseph Cohen [13] and fifty healthy lung images from Kaggle repository ‘‘Chest X-Ray Images (Pneumonia)’’ [14]. The ResNet-50 obtained the best result achieving an accuracy of 98%.

Hemdan et al. [15] proposed a COVIDX-Net composed of seven popular CNN models and used the same dataset considered by [12] and achieved 90% in terms of accuracy. However, only 25 samples of COVID-19 positive and 25 samples of negative images were considered.

Sethy and Behera [16] also considered the same dataset of [15]. Their study states that the ResNet-50 as a feature extractor and Support Vector Machine (SVM) classifier provided the best performance obtained an accuracy of 95.38%.

Wang and Wong [17] proposed a COVID-Net architecture, an open-source CNN created to detect COVID-19 on chest X-ray images. The authors used a dataset created exclusively to support COVID-Net experimentation, which obtained 93.3% accuracy in classifying normal, non-COVID pneumonia, and COVID-19 classes.

Apostolopoulos and Mpesiana [18] adopted different pre-trained network architectures to address the task of classification of COVID-19 in chest X-ray images and achieved 96.78% of accuracy with MobileNet v2 model.

Khan et al. [19] proposed the CoroNet deep CNN with 71 layers, inspired by Xception (Extreme Inception) and trained on the ImageNet dataset [20]. According to the authors, the CoroNet was evaluated using a dataset not publicly available for download and achieved an average accuracy of 89.60% for the COVID-19 identification.

Ozturk et al. [21] proposed an approach for early detection of COVID-19 cases using the DarkNet model as a classifier YOLO object detection system and obtained an accuracy of 98.08% for binary classes and 87.02% for multi-class cases. Ucar and Korkmaz [22] proposed a method based on SqueezeNet [23] architecture with Bayes optimization and achieved 98.30% of accuracy.

Pereira et al. [24] utilized texture descriptors, fusion techniques and a pre-trained Inception-V3 model to identify COVID-19 obtained 89.00% in terms of F1-score. However, the validation methodology used in [22] and [24] is a simple hold-out technique that has a certain probability of building biased sets, which may achieve abnormal accuracy results, mainly in small datasets.

The automated classification of COVID-19 in X-ray images is a hot topic nowadays due to the growing pandemic, and new works are emerging every day. In contrast to the previous works, we explore training based on SFT, data augmentation strategy, and our approach is a promising alternative by delivering a simple and efficient that allows achieved better results with a low computational cost.

III. MATERIAL AND METHODS

The main goal of this paper is to evaluate the performance of different architectures of CNNs to classify COVID-19 in chest X-ray images. More precisely, we find the best possible identification rate among COVID-19 and other types of pneumonia. Fig. 1 illustrates the steps of the methodology adopted here.

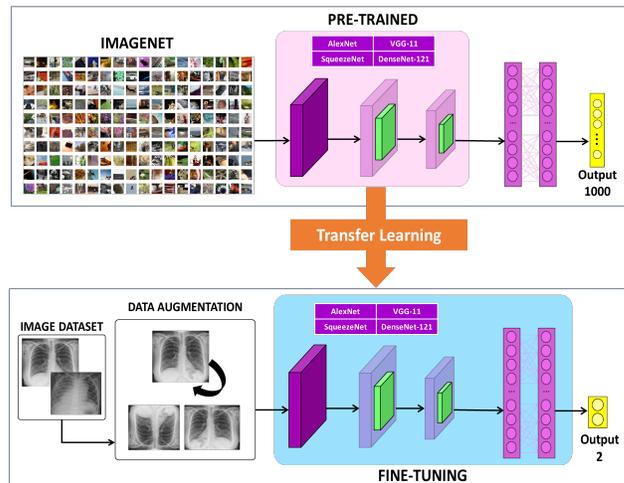


Fig. 1. Steps of proposed method.

A. Image dataset

The images used in this work were obtained from two datasets of chest X-ray images. The first dataset contains 108 images of COVID-19 positive and was taken from the GitHub repository shared by Dr. Joseph Cohen, at the University of Montreal [13] (last accessed April 10, 2020). We selected 299 images of COVID-19 negative, corresponding to 20% of viral pneumonia images selected randomly from the Chest X-Ray Images (Pneumonia) dataset available in Kaggle repository [14]. Note that only about 20% of the viral pneumonia was selected in order to avoid imbalance between the classes or bias the classification performance.

The information about images is summarized in Table I. To illustrate the dataset resulting from the combination of the two datasets previously mentioned, sample images from each class are presented in Fig. 2.

TABLE I. DISTRIBUTION OF THE CHEST X-RAY IMAGES.

COVID-19	Samples	Source
Positive	108	GitHub (Dr. Joseph Cohen) [13]
Negative	299	Kaggle (X-ray images of Pneumonia) [14]
Total	407	

B. Pre-processing

All images were resized to 224×224 pixels based on bilinear interpolation. The resize allows adapting each image for the input of the CNN architectures used in this work.

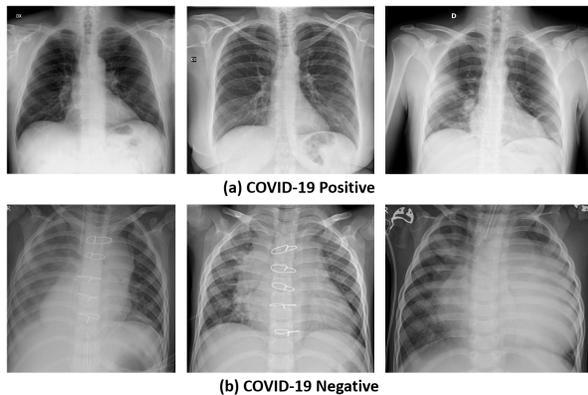


Fig. 2. Examples of image instance for each class.

As one of the main obstacles in this study is the lack of images, we applied data augmentation strategy to increase the training data artificially without introducing labeling costs [25]. All training images were augmented by using vertical and horizontal flips, and rotating each of the original images around its center through randomly chosen angles of between -10° and 10° .

C. Convolutional Neural Networks (CNNs)

Convolutional Neural Networks (CNNs) are a multi-stage image classification technique that incorporates spatial context and weight sharing between pixels in order to extract high-level hierarchical representations of the data [26] [27]. Thus, CNN is able to extract features during training. In this work, four CNNs architectures are tested: AlexNet [25], VGG-11 [28], SqueezeNet [23], and DenseNet-121 [29]. All CNNs were selected based on their success in previous image classification tasks.

The AlexNet was the champion of ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012 [20]. This CNN consists of five convolutional layers, three max-pooling layers, two fully connected layers with a final softmax layer. In order to reduce overfitting, AlexNet uses dropout connections and Rectified Linear Unit (ReLU) activation function [25].

The VGG won the identification and classification tasks in the ILSVRC 2014. In order to reduce the computational cost, this CNN adopted sequences of convolutional filters of size 3×3 [28]. In this paper, we use the VGG-11 architecture with batch normalization, due to its simplicity and robustness. It is important to note that batch normalization is very effective to overcome the challenges of deep training [30].

SqueezeNet is a compact CNN with approximately 50 times fewer parameters than AlexNet model. It is composed of a stand-alone convolution layer followed by eight fire modules and a final convolution layer [23]. Each fire module contains only a filter of size 1×1 inputting into an expanded layer composed by convolutional filters of size 1×1 and 3×3 . In this way, the modules are able to perform the same functions of fully connected and dense layers.

DenseNet-121 architecture uses dense blocks to concatenate a number of convolutional layers reducing the number of features through average pooling. In the present study, we use

the DenseNet with 121 layers: one initial convolutional layer followed by max-pooling, 116 convolutional layers followed by batch normalization, and ReLU functions, interpolated with three transition blocks, and a last average pooling before the start of fully connected layers [29].

D. Shallow Fine-Tuning

The Shallow Fine-Tuning (SFT) [31] [32] is a training strategy based on the concept of transfer learning and is suitable for small data sets. This approach is used to train deep learning models in which the network is pre-trained for a classification task using a huge dataset such as ImageNet [20].

The weights in all convolutional layers are initialized with the corresponding values from the pre-trained model. These layers are considered more general and retain information about texture, color, and shape. SFT performs fine-tuning only in the last fully connected layer, which is more specialized. Usually, this strategy is the most common allowing weights of the last layers to adapt to the classification problem.

E. Training strategy

The training of the CNNs models is defined as an optimization problem in order to optimize the quality of the prediction. In this work, we considered the objective function as the cross-entropy defined by $\mathcal{L}(W)$. Equation 1 show that $\mathcal{L}(W)$ is computed over a set of training samples X_j considering the tuned weights W , parameters $f(x_j)$, and the known classes y_j , where j represents the classes COVID-19 positive and negative.

$$\mathcal{L}(W) = \frac{1}{n} \sum_{j=1}^N \ell(y_j, f(x_j; W)) \quad (1)$$

In this way, to minimize $\mathcal{L}(W)$, we applied the Stochastic Gradient Descent (SGD) [33] optimization algorithm with momentum of 0.9, learning rate of 0.001, and batch size of eight. All CNNs were trained for 30 epochs.

F. Evaluation methodology

Due to the very small number of positive images of COVID-19 available, we have decided not to perform hyperparameter optimization nor early stop strategy. These procedures require a validation set, which further reduces the number of images available for testing the models. For this reason, all CNN models were trained and tested using the stratified k -fold cross-validation method [34]. All chest X-ray images were randomly partitioned into ten folds. Then, the model is trained with $k - 1$ folds and tested on the remaining fold. The training and testing procedures are repeated k times, alternating the testing folds. Thus, we guarantee that each image will participate in the training process ($k - 1$ times) and will also be part of the test group (1 time). Finally, the results from the k testing sets are averaged to produce a single and trustworthy estimation.

The metrics used to assess the classification performance include accuracy, precision, recall, and F1-score indices. All

indices are based on the number of true positives (TP), true negatives (TN), false positives (FP), and false-negative (FN) classifications obtained from the confusion matrix [35]. Also, we measure the standard deviation in order to assess the confidence of results, where smaller values represent high-reliability.

- Accuracy: is the ratio between the correct classifications and total samples (Eq.2).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

- Precision: is the ratio between TP and the total of positives classification (Eq. 3)

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

- Recall: is the harmonic average of recall and precision (Eq. 4).

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

- F1-Score: is the weighted average of the precision and recall (Eq. 5)

$$F1-Score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

Also, we used the Receiver Operating Characteristic (ROC), and the Area Under ROC (AUC) as a reliable classification performance measure of all possible classification thresholds.

IV. RESULTS AND DISCUSSION

All experiments were programmed using Python (version 3.6) and PyTorch (version 1.4) deep learning framework [36]. This study investigated the performance of four CNNs architectures to classify chest X-ray images on COVID-19 positive and COVID-19 negative (pneumonia) classes.

A. Comparison of architectures

One of the most challenges in training CNNs architectures for classification tasks is the lack of large enough datasets for adjust a large number of model parameters. We adopted SFT fine-tuning training strategy to overcome this problem, instead of training the models from scratch (as described in Section III-D). Thus, the weights in the initial layers of CNN (simpler features) were kept, but the weights of the deeper layers (more specialized features) were adapted to the problem of classifying X-ray images generating greater specialization in the deep layers. It is important to mention that number of positive COVID-19 images in the public repository is very small. To overcome this issue, we applied data augmentation strategies aim to increase the size of the training set.

Table II presents the average classification performance considering the results of accuracy, precision, recall and F1-score indices concerning each CNN evaluated. The chart in Fig. 3 illustrate the variation of each performance values based on the results presented in Table II. Interestingly, SqueezeNet achieved the best performance, followed by

AlexNet, DenseNet-121, and VGG-11. Also, SqueezeNet is recognized as having a low computational cost and designed for embedded systems. Therefore, the result suggested that our proposed could be used to evaluate chest X-ray images using mobile devices. Moreover, the standard deviation value obtained for each CNN was small and indicates that our results are reliable.

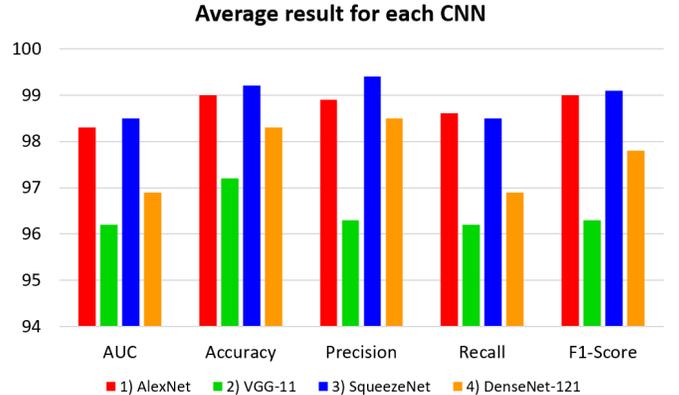


Fig. 3. 10-fold average values of the performance measures for each CNN model. 1) AlexNet; 2) VGG-11; 3) SqueezeNet; and 4) DenseNet-121.

As in each fold the images in the testing sets do not repeat, we consolidate the confusion matrices from each fold by adding the values from each confusion matrix. Therefore, the confusion matrix presented in Table III summarizes results for all ten folds, and presents the prediction for all images in the dataset.

The confusion matrices of each CNN model allow observing several aspects of the classification problem investigated in this work. Note that for all CNNs, the COVID-19 positive and COVID-19 negative is well identified. In particular, AlexNet and SqueezeNet models were able to classify the most positive cases of COVID-19 correctly. The results suggest that models have been able to preserve in the feature maps important information about visual patterns of diagnostic positive. It is important to mention that number of positive COVID-19 images in the public repository is very small. In this study, to compare the performance of different CNN architectures, we focused to obtain reliable and trustfully results.

With the lack of data, we still cannot recommend these methods as a diagnostic aid system, but our results support that the use of CNN models is a promising technique to assist the early diagnosis of COVID-19 in conjugation with other standard tests. However, the number of available images tends to grow as studies advance, and with more accurate researches, it will be possible to understand the capability of CNNs in helping detecting COVID-19.

B. Comparison with literature

The best result achieved in this study in terms of accuracy and F1-score is compared with other state-of-art work in the literature. The best result in our work was obtained with SqueezeNet, trained with SFT using augmented data, which scored 99.20% of accuracy and 99.10% of F1-score (as shown in Table II). The best results reported in the literature are

TABLE II. 10-FOLD AVERAGE VALUES AND STANDARD DEVIATION OF THE PERFORMANCE MEASURES FOR EACH CNN MODEL.

CNN	AUC (%)	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
AlexNet	98.30 ± 0.02	99.00 ± 0.01	98.90 ± 0.02	98.60 ± 0.02	99.00 ± 0.01
VGG-11	96.20 ± 0.04	97.20 ± 0.03	96.30 ± 0.04	98.60 ± 0.04	99.00 ± 0.04
SqueezeNet	98.50 ± 0.02	99.20 ± 0.01	99.40 ± 0.01	98.50 ± 0.02	99.10 ± 0.01
DenseNet-121	96.90 ± 0.03	98.30 ± 0.02	98.50 ± 0.01	96.90 ± 0.03	97.80 ± 0.02

TABLE III. 10-FOLD VALUES OF CONFUSION MATRIX FOR EACH CNN MODEL.

AlexNet			VGG-11		
	Positive	Negative		Positive	Negative
Positive	105	3	Positive	102	6
Negative	1	298	Negative	6	293

SqueezeNet			DenseNet-121		
	Positive	Negative		Positive	Negative
Positive	105	3	Positive	102	6
Negative	0	299	Negative	1	298

presented in Table IV for the same COVID-19 dataset. It can be seen that our best score is upper to the best state-of-the-art technique reported in the literature.

TABLE IV. HIGHEST ACCURACY OF OTHER CLASSIFICATION METHODS USING THE COVID-19 DATASET FROM GITHUB [13].

Method	Accuracy (%)
Narin et al. [12]	98.00
Hemdan et al. [15]	90.00
Sethy and Behera [16]	95.38
Wang and Wong [17]	93.30
Apostopoulos and Mpsiana [18]	96.78
Khan et al. [19]	89.60
Ozturk et al. [21]	98.08
Our work (SqueezeNet + SFT + data aug.)	99.20

	F1-Score (%)
Pereira et al. [24]	89.00
Our work (SqueezeNet + SFT + data aug.)	99.10

V. CONCLUSION

The results presented in this paper point to a promising using of CNN models to classify COVID-19 cases based on chest X-ray images. We compared the performance of four CNN architectures to classify X-ray images in COVID-19 positive and negative (pneumonia) classes, and our training strategy consists of applying transfer learning with SFT and different data augmentation approaches. Our best result of 99.20% was upper to the highest accuracy score presented in the literature; this result was obtained with SqueezeNet model.

Although there are few COVID-19 positive chest X-ray images, we designed our experiments to minimize the effects of CNN training with small data sets. The training using fine-tuning and data augmentation aim to increase the classification rate, while the stratified k-fold cross validation allows more reliable results than simple hold-out. Now, it is necessary to wait for more images of positive COVID-19 to be available in order to train more reliable models that may confirm the positive perspectives demonstrated by this study.

The presented results open new opportunities towards better machine learning based on deep CNNs for automated detection of COVID-19 and developing of new computer-aided diagnosis applications. Moreover, exciting opportunities and future works raise such as testing other CNN models, evaluating more data augmentation strategies, and applying some hyperparameter optimization and combining classifications techniques.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of NVIDIA Corporation with the donation of the TITAN Xp GPU used for this research. We would like to thanks CAPES and FAPEMIG (Grant number CEX - APQ-02964-17) for the financial support. This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001.

REFERENCES

- [1] World Health Organization. Coronavirus disease 2019 (covid-19): situation report, 73. Technical documents, 2020-04-02.
- [2] Peng Zhou, Xing-Lou Yang, Xian-Guang Wang, Ben Hu, Lei Zhang, Wei Zhang, Hao-Rui Si, Yan Zhu, Bei Li, Chao-Lin Huang, Hui-Dong Chen, Jing Chen, Yun Luo, Hua Guo, Ren-Di Jiang, Mei-Qin Liu, Ying Chen, Xu-Rui Shen, Xi Wang, Xiao-Shuang Zheng, Kai Zhao, Quan-Jiao Chen, Fei Deng, Lin-Lin Liu, Bing Yan, Fa-Xian Zhan, Yan-Yi Wang, Geng-Fu Xiao, and Zheng-Li Shi. A pneumonia outbreak

- associated with a new coronavirus of probable bat origin. *Nature*, 579(7798):270–273, 2020.
- [3] Graziano Onder, Giovanni Rezza, and Silvio Brusaferro. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *JAMA*, 03 2020.
 - [4] Wei-jie Guan, Zheng-yi Ni, Yu Hu, Wen-hua Liang, Chun-quan Ou, Jian-xing He, Lei Liu, Hong Shan, Chun-liang Lei, David S.C. Hui, Bin Du, Lan-juan Li, Guang Zeng, Kwok-Yung Yuen, Ru-chong Chen, Chun-li Tang, Tao Wang, Ping-yan Chen, Jie Xiang, Shi-yue Li, Jin-lin Wang, Zi-jing Liang, Yi-xiang Peng, Li Wei, Yong Liu, Ya-hua Hu, Peng Peng, Jian-ming Wang, Ji-yang Liu, Zhong Chen, Gang Li, Zhi-jian Zheng, Shao-qin Qiu, Jie Luo, Chang-jiang Ye, Shao-yong Zhu, and Nan-shan Zhong. Clinical characteristics of coronavirus disease 2019 in china. *New England Journal of Medicine*, 382(18):1708–1720, 2020.
 - [5] Giacomo Grasselli, Antonio Pesenti, and Maurizio Cecconi. Critical Care Utilization for the COVID-19 Outbreak in Lombardy, Italy: Early Experience and Forecast During an Emergency Response. *JAMA*, 323(16):1545–1546, 04 2020.
 - [6] Loris Nanni, Stefano Ghidoni, and Sheryl Brahnam. Ensemble of convolutional neural networks for bioimage classification. *Applied Computing and Informatics*, 2018.
 - [7] Daniel S. Kermany, Michael Goldbaum, Wenjia Cai, Carolina C.S. Valentim, Huiying Liang, Sally L. Baxter, Alex McKeown, Ge Yang, Xiaokang Wu, Fangbing Yan, Justin Dong, Made K. Prasadha, Jacqueline Pei, Magdalene Y.L. Ting, Jie Zhu, Christina Li, Sierra Hewett, Jason Dong, Ian Ziyar, Alexander Shi, Runze Zhang, Lianghong Zheng, Rui Hou, William Shi, Xin Fu, Yaou Duan, Viet A.N. Huu, Cindy Wen, Edward D. Zhang, Charlotte L. Zhang, Oulan Li, Xiaobo Wang, Michael A. Singer, Xiaodong Sun, Jie Xu, Ali Tafreshi, M. Anthony Lewis, Huimin Xia, and Kang Zhang. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell*, 172(5):1122 – 1131.e9, 2018.
 - [8] Larissa Ferreira Rodrigues, Murilo Coelho Naldi, and João Fernando Mari. Comparing convolutional neural networks and preprocessing techniques for hep-2 cell classification in immunofluorescence images. *Computers in Biology and Medicine*, 116:103542, 2020.
 - [9] Yicheng Fang, Huangqi Zhang, Jicheng Xie, Minjie Lin, Lingjun Ying, Peipei Pang, and Wenbin Ji. Sensitivity of chest ct for covid-19: Comparison to rt-pcr. *Radiology*, 0(0):200432, 0. PMID: 32073353.
 - [10] Tao Ai, Zhenlu Yang, Hongyan Hou, Chena Zhao, Chong Chen, Wenzhi Lv, Qian Tao, Ziyong Sun, and Liming Xia. Correlation of chest ct and rt-pcr testing in coronavirus disease 2019 (covid-19) in china: A report of 1014 cases. *Radiology*, 0(0):200642, 0. PMID: 32101510.
 - [11] Geoffrey D. Rubin, Christopher J. Ryerson, Linda B. Haramati, Nicola Sverzellati, Jeffrey P. Kanne, Suhail Raouf, Neil W. Schluger, Annalisa Volpi, Jae-Joon Yim, Ian B.K. Martin, Deverick J. Anderson, Christina Kong, Talissa Altes, Andrew Bush, Sujal R. Desai, Jonathan Goldin, Jin Mo Goo, Marc Humbert, Yoshikazu Inoue, Hans-Ulrich Kauczor, Fengming Luo, Peter J. Mazzone, Mathias Prokop, Martine Remy-Jardin, Luca Richeldi, Cornelia M. Schaefer-Prokop, Noriyuki Tomiyama, Athol U. Wells, and Ann N. Leung. The role of chest imaging in patient management during the covid-19 pandemic: A multinational consensus statement from the fleischner society. *Chest*, 2020.
 - [12] Ali Narin, Ceren Kaya, and Ziyinet Pamuk. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks, 2020.
 - [13] Joseph Paul Cohen, Paul Morrison, and Lan Dao. Covid-19 image data collection. *arXiv*, 2020.
 - [14] P Mooney. Chest x-ray images (pneumonia). *kaggle*, Marzo, 2018.
 - [15] Ezz El-Din Hemdan, Marwa A. Shouman, and Mohamed Esmail Karar. Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images, 2020.
 - [16] Prabira Kumar Sethy and Santi Kumari Behera. Detection of coronavirus disease (covid-19) based on deep features. *Preprints*, 2020030300:2020, 2020.
 - [17] Linda Wang and Alexander Wong. Covid-net: A tailored deep convolutional neural network design for detection of covid-19 cases from chest x-ray images, 2020.
 - [18] Ioannis D. Apostolopoulos and Tzani A. Mpesiana. Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. *Physical and Engineering Sciences in Medicine*, pages 1–6, Apr 2020. PMC7118364[pmcid].
 - [19] Asif Iqbal Khan, Junaid Latief Shah, and Mohammad Mudasar Bhat. Coronet: A deep neural network for detection and diagnosis of covid-19 from chest x-ray images. *Computer Methods and Programs in Biomedicine*, 196:105581, 2020.
 - [20] J. Deng, W. Dong, R. Socher, L. J. Li, Kai Li, and Li Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, June 2009.
 - [21] Tulin Ozturk, Muhammed Talo, Eylul Azra Yildirim, Ulas Baran Baloglu, Ozal Yildirim, and U. [Rajendra Acharya]. Automated detection of covid-19 cases using deep neural networks with x-ray images. *Computers in Biology and Medicine*, 121:103792, 2020.
 - [22] Ferhat Ucar and Deniz Korkmaz. Covidiagnosis-net: Deep bayes-squeezenet based diagnosis of the coronavirus disease 2019 (covid-19) from x-ray images. *Medical Hypotheses*, 140:109761, 2020.
 - [23] Forrest N. Iandola, Song Han, Matthew W. Moskewicz, Khalid Ashraf, William J. Dally, and Kurt Keutzer. Squeezenet: Alexnet-level accuracy with 50x fewer parameters and 0.5mb model size, 2016.
 - [24] Rodolfo M. Pereira, Diego Bertolini, Lucas O. Teixeira, Carlos N. Silla, and Yandre M.G. Costa. Covid-19 identification in chest x-ray images on flat and hierarchical classification scenarios. *Computer Methods and Programs in Biomedicine*, 194:105532, 2020.
 - [25] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
 - [26] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
 - [27] M. A. Ponti, L. S. F. Ribeiro, T. S. Nazare, T. Bui, and J. Collomosse. Everything you wanted to know about deep learning for computer vision but were afraid to ask. In *2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI-T)*, pages 17–41, Oct 2017.
 - [28] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.
 - [29] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.
 - [30] Sergey Ioffe and Christian Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift, 2015.
 - [31] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang. Convolutional neural networks for medical image analysis: Full training or fine tuning? *IEEE Transactions on Medical Imaging*, 35(5):1299–1312, 2016.
 - [32] Mohammadhassan Izadyyazdanabadi, Evgenii Belykh, Michael Mooney, Nikolay Martirosyan, Jennifer Eschbacher, Peter Nakaji, Mark C. Preul, and Yezhou Yang. Convolutional neural networks: Ensemble modeling, fine-tuning and unsupervised semantic localization for neurosurgical cle images. *Journal of Visual Communication and Image Representation*, 54:10 – 20, 2018.
 - [33] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11):2278–2324, Nov 1998.
 - [34] Pierre A. Devijver and Josef Kittler. *Pattern Recognition: A Statistical Approach*. Prentice-Hall, 1982.
 - [35] Richard O. Duda, Peter E. Hart, and David G. Stork. *Pattern Classification (2Nd Edition)*. Wiley-Interscience, New York, NY, USA, 2000.
 - [36] Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, Alban Desmaison, Andreas Kopf, Edward Yang, Zachary DeVito, Martin Raison, Alykhan Tejani, Sasank Chilamkurthy, Benoit Steiner, Lu Fang, Junjie Bai, and Soumith Chintala. Pytorch: An imperative style, high-performance deep learning library. In *Advances in Neural Information Processing Systems 32*, pages 8026–8037. Curran Associates, Inc., 2019.