# A Chest X-ray Image Retrieval System for COVID-19 Detection using Deep Transfer Learning and Denoising Auto Encoder

**O. F. Layode**
Computer Science Department
Morgan State University
Baltimore, USA
oylay1@morgan.edu

**Md Rahman**
Computer Science Department
Morgan State University
Baltimore, USA
md.rahman@morgan.edu

***Abstract***— The COVID-19 pandemic is the defining global health crisis of our time which is currently challenging families, communities, health care systems, and government all over the world. It is critical to detect and isolate the positive cases as early as possible for timely treatment to prevent the further spread of the virus. It was found in few early studies that patients present abnormalities in chest radiography images that are characteristic of those infected with COVID-19. In the current context, a rapid, accessible and automated screening tool based on image processing of chest X-rays (CXRs) would be much needed as a quick alternative to PCR testing, especially with commonly available X-ray machines and without the dedicated test kits in labs and hospitals. Several classifications based approaches have been proposed recently with encouraging results to detect pneumonia based on CXRs using supervised deep transfer learning techniques based on Convolutional Neural Networks (CNNs). These black box approaches are mainly non-interactive in nature and their prediction represents just a cue to the radiologist. This work focuses on issues related to the development of such an automated system for CXRs by performing discriminative feature learning using deep neural networks with a purely data driven approach and retrieving images based on an unknown query image and performing retrieval evaluation on currently available benchmark datasets towards the goal of realistic comparison and real clinical integration. The system is trained and tested on an image collection of 1700 CXRs obtained from two different resources with encouraging results based on precision and recall measures in individual deep feature spaces. It is hoped that the proposed system as diagnostic aid would reduce the visual observation error of human operators and enhance sensitivity in testing for Covid-19 detection.

***Keywords—X-ray, Covid-19, Computer Aided Diagnosis, Screening Tool, Deep Learning, Image Retrieval***

## I. INTRODUCTION

COVID-19 pandemic is the defining global health crisis of our time and the greatest challenge we have faced since World War II. This global health crisis is challenging families, communities, health care systems, and governments. A critical step in the fight against COVID-19 is effective screening of infected patients, such that those infected can receive immediate treatment and care, as well as be isolated to mitigate the spread of the virus [1]. Testing for COVID-19 involves inserting a 6-inch long swab (like a long Q-tip) into the cavity between the nose and mouth (nasopharyngeal swab) for 15 seconds and rotating it several times. The swabbing is then repeated on the other side of the nose to make sure enough material is collected. The swab is then inserted into a container and sent to a lab for reverse transcriptase-polymerase chain reaction (RT-PCR) testing, which can detect the RNA of the coronavirus called SARS-CoV-2 from respiratory specimens. Currently, tests are prioritized mainly for hospitalized and symptomatic individuals. Health care workers, first responders and other social service employees, plus people exposed to infected individuals in places where COVID-19 risk is high, are also being tested. While the above type of testing is the current gold standard with high specificity and in high demand, it is a very time-consuming and laborious manual process from collection of specimens to obtain the test result. Given the current sensitivity (around 70%) of the nucleic acid tests, many suspected patients have to be tested multiple times several days apart before reaching a confident diagnosis [2].

As a result, there is a crying need for an automated diagnostic tool as an alternative screening method for rapid triaging, availability, accessibility, and portability. In some early studies [2-8], it is found that patients present abnormalities in chest radiography images (e.g., chest X-ray or computed tomography

978-1-7281-7624-6/20/$31.00 ©2020 IEEE
DOI 10.1109/CSCI51800.2020.00301

(CT) imaging) that are characteristic of those infected with novel corona virus including bilateral, multi-focal, ground-glass opacities with a peripheral or posterior distribution, mainly in the lower lobes, in the early stage and pulmonary consolidation in the late stage. Hence, the imaging findings might play a critical role in constraining the viral transmission and fighting against COVID-19. For example, the chest X-ray is one of the most commonly accessible and available radiological examinations for screening and diagnosis of a variety of lung diseases including pneumonia, hence making them a good complement to PCR testing. However, manually interpreting and analyzing radiography images is very subjective in nature and depends on the experience of the radiologists especially where visual indicators most often are very subtle to distinguish COVID-19 from other viral pneumonias. Also, there is a shortage of experts who can interpret X-rays, even when imaging equipment is available. Therefore, a computer aided diagnostic (CAD) system that can aid radiologists to more rapidly and accurately interpret radiography images to detect COVID-19 cases is highly desired. Application of machine learning techniques coupled with radiological imaging can be helpful for the accurate detection of this disease and can also be assistive to overcome the problem of a lack of specialized physicians in remote villages.

The descriptiveness and discriminative power of features to effectively represent the structure and characteristics of images and effectively handle the within-class variation and between-class similarity are critical to achieve good classification and retrieval performances. Recent research in computer vision and pattern recognition has highlighted the capabilities of Convolutional Neural Networks (CNNs) to solve challenging tasks such as classification, segmentation and object detection, achieving state-of-the art performances [9]. This success has been attributed to the ability of CNNs to learn a hierarchical representation of raw input data to capture the intrinsic image features without manual feature design. Deep learning architectures are effectively formed by combination of many linear and nonlinear transformations to obtain more abstract and useful representations of data. Ideally, this data-derived representation discards irrelevant information and preserves only those details that are useful for the intended task. Automatic feature learning or deep learning from image data has thus emerged as a different trend recently, to capture the intrinsic image features and provides an alternative to the hand-crafted features. By varying the objective function of the feature learning method, it might be possible to tailor the features to a specific application, such as classification [10].

Recent studies have shown that the use of deep learning can significantly improve the performance of CAD systems, such as interstitial lung disease (ILD) classification, breast cancer detection, pneumonia detection from chest X-ray images fundus image segmentation, and lung segmentation, and learning the manifold of 3D brain MRI images [11]. Deep learning techniques on chest X-Rays for COVID-19 detection are also getting popularity as it is recently observed that there is a rush in publications from research community. In imaging field, majority of the current studies [2-8] use only a handful of X-ray images from a few available online data sets to classify between COVID-19 and other pneumonia (e.g., bacterial pneumonia and non-COVID-19 viral pneumonia) and healthy subjects.

Deep learning systems could potentially assist clinicians for identifying lung abnormalities related to COVID-19 in triaging and treating high-risk patients, as well as help overcome the scarcity of COVID-19 images available for machine learning development. So far, most of the proposed approaches show promising results based on transfer learning in smaller data sets. For example, three different deep learning models (ResNet50, InceptionV3, and Inception-ResNetV2) were used in [7] to detect COVID-19 infection from X-ray images (50 COVID-19 patients and 50 normal chest X-ray images) and evaluation results show that the ResNet50 model achieves the highest classification performance with 98.0% accuracy, compared to 97.0% accuracy by InceptionV3 and 87% accuracy by Inception-ResNetV2. Linda et al [12] proposes COVID-Net (a deep CNN network) to detect COVID-19 cases in a dataset of 5941 chest X-ray images of four different categories (Fig. 1) and obtains a testing accuracy of 83.5%. Whereas a 3-step technique (COVID-ResNet) is presented in [3] to fine-tune a pre-trained ResNet-50 architecture to improve model performance and reduce training time. Hemdan et al. [8] proposed a DarkCovidNet model by implementing 17 convolutional layers and introduced different filtering on each layer and produced a classification accuracy of 98.08% for binary classes (COVID vs. No-Findings) and 87.02% for multi-class cases (COVID vs. No-Findings vs. Pneumonia) where images were obtained from two different sources. Finally, there are also several recent studies on COVID-19 detection that employed various deep learning models with CT images [13-15].

However, majority of these systems are non-interactive in nature and the prediction represents just a cue without the ability to explain the reasoning of the decision-making. It would be more effective if a radiologist is assisted in the decision making process by means of an interactive approach; where the system retrieves a number of similar images from a database of already diagnosed cases, similar to the one under the analysis in addition to predict the category of that unknown image. Our hypothesis is that by providing with a set of pathologically-confirmed past cases as computer output, it could be utilized to guide them to a precise diagnosis, but not to suggest them a second diagnosis. In the last several years, developing CAD schemes that use content-based image retrieval (CBIR) approach to search for the clinically relevant and visually similar images has been attracting huge research interest [16,17]. CBIR-based CAD schemes have potential to provide radiologists with "visual aid" and increase their confidence in accepting CAD-cued results in the decision making. Furthermore, CBIR would be a useful aid in the training of students, residents, and less experienced radiologists since it would allow them to view images of lesions that appear similar, but may have differing pathology and help them see how the pattern in their current case closely resembles a pattern in cases previously proven to be normal or different disease

categories, thereby improving specificity. Furthermore, an integrated system based on both classification and retrieval would be more effective and robust as compared to using individual classification or retrieval approach alone. To address the limitation of the current CAD in detecting Covid-19 and due to the ongoing success of CBIR to provide clinical decision support for medical images of different modalities, this paper proposes an integrated classification and retrieval based Decision Support System (DSS) for Covid-19 detection with an 'easy to use' user interface by applying fusion and ensemble techniques in deep feature spaces. Deep features are extracted based on using transfer learning in several pre-trained CNNs and deep autoencoder. Furthermore, our CBIR method uses the same deep features by fusing those in different feature combinations using a canonical correlation analysis. Based on image-based visual queries submitted by radiologists, the proposed system would respond by displaying relevant images of chest X-ray images of past cases as well as classifying the image category as different types. Although radiologists play a key role due to their vast experience in this field, the proposed DSS can be assistive to obtain accurate diagnosis and also useful in eliminating disadvantages such as insufficient number of available RT-PCR test kits, test costs, and waiting time of test results.

## II. SYSTEM MODELING

Fig. 1 represents the overall architecture of the proposed retrieval system which is divided into offline and online phases.

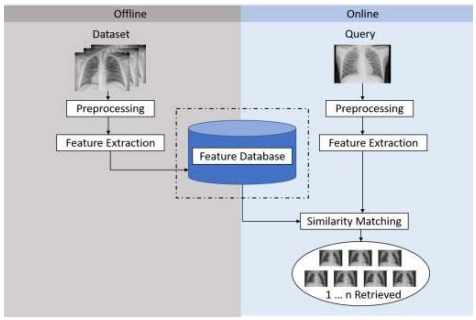


Fig. 1. Workflow diagram of the proposed DSS

The off-line phase (left) involves pre-processing and feature extraction of database images for indexing purpose. The images were preprocessed by resizing and normalizing the pixel values before passing to a feature extraction network. Our feature extraction network consists of a feature extractor module based on training a convolutional denoising Autoencoder for feature generation in a latent space and several other Transfer Learning based deep feature extractors (excludes the last two fully connected layers of CNN), which are pretrained on ImageNet dataset. The system pre-processes all the database images by resizing at first for the respective CNN approaches. Using the transfer learning approach, the deep features of the images in the database are extracted by passing them through the CNNs that are without a classification head. The results obtained after this stage are the features that would have been passed similarity matching layer. During the online phase (the right part of the Fig. 1), an unknown query image is submitted to the system which is also passed to the feature extraction network after which similarity distances between the query features and database features are calculated. The overall network architecture takes query images as input and gives an output of database images corresponding to the top K similarity distances. In the retrieval interface, users also have the option to select different similarity measures and feature fusion approaches in those deep feature spaces of database and query images for both flexibility and effectiveness.

## III. DEEP FEATURE EXTRACTION WITH TRANSFER LEARNING

Feature extraction is a critical component of medical image analysis. In selecting effective features, great research efforts have been focused on capturing the color, texture, and shape properties of images. The descriptiveness and discriminative power of features extracted from X-ray images are critical to achieve good classification and retrieval performances. Instead of using any hand-crafted features, transfer learning techniques can be used to extract features of X-ray image from a relatively small dataset using pretrained CNN models. Transfer learning increases the efficiency of the feature extraction process as it has been consistently proven to boost model accuracy with fewer data and reduce required training time. CNNs trained on large-scale datasets such as ImageNet have demonstrated to be excellent at the task of transfer learning. These networks learn a set of rich, discriminating features to recognize 1,000 separate object classes [18,19]. Using a pretrained CNN as a feature extractor rather than training a CNN from scratch is attractive as it transfers learning (i.e. filters) from other domains where more training data is available and avoids a time-consuming training process [20].

To perform deep feature extraction based on transfer learning, we chose VGG-19 [21], ResNet-50 [22], and Densenet169 [23] as our pre-trained CNN models on ImageNet dataset of 1000 categories. However, these CNNs learns features that are enough to give relatively accurate classification results. VGG - 19 consists of 19 fully connected convolutional layers which has been used in many machine learning applications as a baseline feature extractor [21]. ResNet, short for Residual Networks is another classic neural network which has been inspired by the VGG Net architecture. Typical ResNet models are implemented with double- or triple- layer skips that contain nonlinearities (ReLU) and batch normalization in between [22]. Skipping effectively simplifies the network, using fewer layers in the initial training stages. This speed up learning by reducing the impact of vanishing gradients, as there are fewer layers to propagate through. The network then gradually restores the skipped layers as it learns the feature space. Whereas, traditional CNN with L layers have L connections—one between each layer and its subsequent layer—the Dense

Convolutional Network (DenseNet) has L(L+1) / 2 direct connections. For each layer, the feature-maps of all preceding layers are used as inputs, and its own feature-maps are used as inputs into all subsequent layers [23]. DenseNet layers are very narrow (e.g., 12 filters per layer), adding only a small set of feature-maps to the "collective knowledge" of the network and keep the remaining feature maps unchanged—and the final classifier makes a decision based on all feature-maps in the network.

**Table 1. Pretrained CNN, feature extraction layer and feature dimension**

| Network | Feature Extraction Layer | Feature Dimension |
|---|---|---|
| Resnet50 | Global Average Pooling Layer | 2048 x 1 |
| Densenet169 | Global Average Pooling Layer | 1024 x 1 |
| VGG19 | Fully connected (Dense) Layer | 4096 x 1 |

As shown in Table 1, The outputs from the first layer of the classification head of these pretrained CNNs are saved as features for the retrieval stage. In this work, we loaded all the images from the dataset, extract their features using the above pre-trained CNN models, and store the extracted features keyed on the image id to a new file in HDF5 dataset format those are later loaded and used as inputs for similarity matching.

## IV. Feature Extraction with Convolutional Deep Autoencoder

Autoencoders are a type of unsupervised neural network (i.e., no class labels or labeled data) that consist of an encoder and a decoder model [24,25]. When trained, the encoder takes input data point and learns a latent-space representation of the data. This latent-space representation is a compressed representation of the data, allowing the model to represent it in far fewer parameters than the original data. If we denote our input data as $x$ and the encoder as $E$, then the output latent-space representation, $s$, would be $s = E(x)$. The networks can be trained by minimizing the mean squared error between the original and the reconstructed data. The decoder is responsible for accepting the latent-space representation $s$ and then reconstructing the original input. If we denote the decoder function as $D$ and the output of the detector as $o$, then we can represent the decoder as $o = D(s)$. Using our mathematical notation, the entire training process of the autoencoder can be written as: $o = D(E(x))$.

The required gradient is easily obtained by using the chain rule to back propagate the error derivatives first through the decoder network and then through the encoder network. Since the size of the hidden layer in an autoencoder is smaller than the size of the input data, the dimensionality of input data is reduced to a smaller-dimensional code space at the hidden layer. The outputs from the hidden layer are then reconstructed into the original data at the output layer. Like Principal Component Analysis (PCA), the autoencoders can give mappings in both directions between the data and the code space.

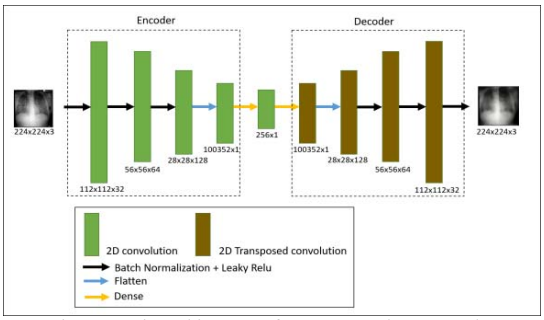


Fig. 2. Basic architecture of our proposed autoencoder

A denoising autoencoder is a feed forward neural network that learns to denoise images. By doing so the neural network learns interesting features on the images used to train it. Once an autoencoder has been trained to encode images, we can use the encoder portion of the network to compute the latent-space representation of each image in our dataset — this representation serves as our feature vector that quantifies the contents of an image. Deep autoencoder built by stacking de-noising autoencoders, i.e. a deep de-noising autoencoder (DDA), has been used for supervised classification and hashing images into binary codes for retrieval purpose [24]. In a similar fashion, our encoder-decoder architecture (Fig. 2) makes use of convolutional layers that take an input image and attempts to reconstruct that image. The encoder region contracts normalized pixel-wise data from input images (224 x 224 x 3 volume) into smaller dimensional feature maps using sequential layers of 2D convolutions, batch normalization and ReLU activation. The output from the convolutional blocks is passed to a fully connected layer that represents a 256 one dimensional feature space. The decoder expands the 256 fully connected output by applying transposed convolutions that up sample the features back to the original input size. Batch normalization and ReLU activation are also added at each step of the transposed convolution sequence and the encoder filter sizes mirror the decoder filter sizes. Once the autoencoder is trained, the feature vectors for each image in our dataset is computed by requiring only a forward-pass of the image through the network where the output of the encoder (i.e., the latent-space representation) serves as our feature vector. The autoencoder will accept our input data, compress it down to the latent-space representation into a vector of much fewer dimensions (256), which is basically a reduction of nearly 99% of the original input data (224 x 224 x 3). After all images are encoded, we can then compare the feature vector from our query image to *all* feature vectors in our dataset (typically you would use either the Euclidean or cosine distance). Images with a smaller distance will be *more similar* than images with a larger distance.

## V. Similarity Matching

Similarity matching is the essential final processing step in the system. For a given query image, a search is made on the

images from the dataset based on the deep features representing each X-ray image. Since, deep learning involves transforming data into distinguishable abstract forms, similarities in extracted features should also infer similarities in the input images. The searching and retrieving result of the CBIR algorithm depends on the effectiveness of the distance metrics to measure the "similarity" level among the selected images. Current CAD schemes using CBIR approaches typically use the *k-nearest* neighbor type searching method which involves searching for the ***k*** most similar reference images (lesions) to the queried image (lesion). The smaller the difference ("distance"), the higher the computed "similarity" level is between the two compared images (lesions). The difference between the feature vector of the query image (lesion) and the feature vectors of lesions of reference images (lesions) is calculated based on using different distance measures, such as Euclidean, Manhattan, and Cosine methods to compute the similarity between the query image and the database. Hence, retrieval accuracy is a function of the efficiency of the distance measures applied to features set. Computational efficiency and speed should also however be considered in choice of a distance measure. In this paper, we applied the Mean Squared Error (MSE), Correlation Distance (COD), and the Euclidean distance (EUD) and evaluated the retrieval accuracies from these measures. The Euclidean measure calculates the vector distance between two vectors ***X*** and ***Y*** as follows:

$$EUD = \sqrt{\Sigma_{i=1}^{n}(X_i - Y_i)^2}$$

The MSE indicates the distance of the regression line between the vector set of points as follows:

$$MSE = \frac{1}{n}\Sigma_{i=1}^{n}(X_i - Y_i)^2$$

A lower MSE value would infer similarity and vice versa.
Whereas the COD measures the linear and nonlinear association between two vectors.

$$v_{xy}^2 = \frac{1}{n^2}\Sigma_{i=1}^{n}(X_i \cdot Y_i)$$

$$\text{COD}, R^2 = \frac{v_{xy}^2}{v_x \cdot v_y}$$

## VI. EXPERIMENTS AND RESULTS

Data collection and experimentation is the final and most important step for validating the effectiveness of any retrieval or classification system. Although there exist large public CXR datasets, due to new infection x-ray images with COVID-19 symptoms are not widely available at present, which greatly hinders the research and development of AI methods. We hypothesized that images of other infections with similar appearances to COVID-19 could be used to train models capable of identifying this new disease [26]. For our experiment in this study, we combined 1550 chest X-ray images from the NIH CXR dataset with 150 publicly available Covid-19 frontal chest X-ray images [27]. The NIH CXR dataset contains 112,120 frontal CXR images of 30,805 unique patients. The dataset contains abnormal pulmonary and cardiothoracic conditions that include nodule, pneumothorax, lung infiltrate, pulmonary atelectasis, pneumonia, hernia, pleural effusion and thickening, edema, cardiomegaly, emphysema, mass, fibrosis, consolidation. With multiple conditions present in one image the dataset was tagged as a multi-label set. The Covid-19 images are labeled as positive COVID-19 detection by professional radiologists and were collected from recently published papers and articles [28].

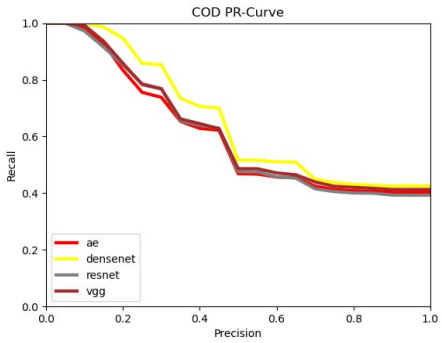


Fig. 3. Precision-Recall (PR) curves for COD distance measure

**Table 2. Average Precision Results**

| CNN | MSE | **EUC** | COD | Average |
|---|---|---|---|---|
| ResNet50 | 0.4204 | 0.4204 | 0.3927 | 0.4112 |
| **DenseNet169** | **0.4328** | 0.4327 | **0.4256** | **0.4304** |
| VGG19 | 0.4099 | 0.4104 | 0.4130 | 0.4111 |
| AutoEncoder | 0.4200 | **0.4368** | 0.4037 | 0.4202 |
| Average | 0.4208 | **0.4251** | 0.4088 | |

In our dataset, 75% images (1275) are used for training autoencoder and 25% (425) are used (for a total of 1700 images) as query images (tests) for retrieval. Table 2 shows the weighted average precision results. Our developed autoencoder gave the best result at an average precision of 0.4368. Feature based on DenseNet169 network however gave the best average precision (0.4304) over the applied similarity measures. Also, the Euclidean distance performed best over the different feature extraction networks. Fig. 3 shows that the deep feature based on DenseNet169 network has the highest precision values for most of the recall levels based on the COD similarity measures. Similar graphs are obtained also for both three the EUD and MSE similarity measures. However, deep feature based on Autoencoder has a little bit better precision at the lower recall levels with Euclidean distance, which also conforms to the result in Table 2.

Fig. 4 shows a snapshot of our retrieval interface where the image at the top left portion is the query images and the system retrieves and displays top 10 most similar images (bottom left) based on using MSE as a distance measure in the deep feature space based on using VGG19 network. Overall, we can see that the interface is very user friendly and flexible for the user where he/she can perform retrieval by selecting different option for similarity measures, feature spaces as well as different datasets.

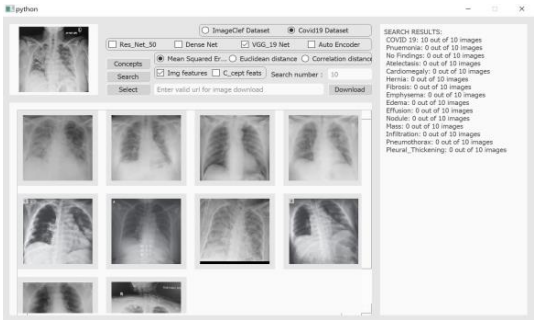

Fig. 4. A snapshot of the retrieval interface

## VII. CONCLUSIONS

In this paper, a deep learning-based retrieval system is proposed for the detection of Covid-19 and other abnormal pulmonary and cardiothoracic conditions in CXRs. It is hypothesized that such a system as a diagnostic aid would greatly improves the decision-making process for both novice and expert radiologists for Covid-19 screening. However, for real clinical integration of such a system it is important to combine imaging data with both clinical manifestations and laboratory examination results to help better screening, detection and diagnosis of COVID-19. It is worth noting that imaging only provides partial information about patients with COVID-19. Hence, in future we will work further in fusing information from these multi-source data, for performing accurate and efficient diagnosis, analysis and follow-up.

## ACKNOWLEDGMENT

This research was supported by an NSF HBCU-UP Research Initiation Award (RIA) (Award Id: 1601044).

## REFERENCES

[1] Muhammad Ilyas, Hina Rehman, Amine Nait-ali, Detection of Covid-19 From Chest X-ray Images Using Artificial Intelligence: An Early Review, Image and Video Processing (eess.IV); Computer Vision and Pattern Recognition (cs.CV); Machine Learning (cs.LG), arXiv:2004.05436 [eess.IV]

[2 ] I. D. Apostolopoulos and T. A. Mpesiana. (2020). "Covid-19: automatic detection from x-ray images utilizing transfer learning with convolutional neural networks," Physical and Engineering Sciences in Medicine, Apr 2020.

[3] M Farooq, A Hafeez Covid-resnet: A deep learning framework for screening of covid19 from radiographs - arXiv preprint arXiv:2003.14395, 2020 - arxiv.org

[4] Ozturk T, Talo M, Yildirim EA, Baloglu UB, Yildirim O, Rajendra Acharya U. Automated detection of COVID-19 cases using deep neural networks with X-ray images. Comput Biol Med. 2020 Apr 28:103792. doi: 10.1016/j.compbiomed.2020.103792. Epub ahead of print. PMCID: PMC7187882.

[5] Alexandru Burlacu, Radu Crisan-Dabija, View ORCID ProfileIolanda Valentina Popa, Bogdan Artene, Vasile Birzu, Mihai Pricop, Cristina Plesoianu, Daniele Generali, Curbing the AI-induced enthusiasm in diagnosing COVID-19 on chest X-Rays: the present and the near-future, doi: https://doi.org/10.1101/2020.04.28.20082776

[6] Shi F, Wang J, Shi J, Wu Z, Wang Q, Tang Z, He K, Shi Y, Shen D. (2020). Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation and Diagnosis for COVID-19. IEEE Rev Biomed Eng. 2020 Apr 16. doi: 10.1109/RBME.2020.2987975. [Epub ahead of print]

[7] Narin, A., Kaya, C., Pamuk, Z. (2020). Automatic Detection of Coronavirus Disease (COVID-19) Using X-ray Images and Deep Convolutional Neural Networks. arXiv preprint arXiv:2003.10849.

[8] E.E.D. Hemdan, M.A. Shouman, M.E. KararCOVIDX-Net: A Framework of Deep Learning Classifiers to Diagnose COVID-19 in X-Ray Images (2020), arXiv preprint arXiv:2003.11055

[9] M. D. Zeiler and R. Fergus, Visualizing and understanding convolutional networks. In: Computer vision–ECCV 2014, pp. 818–833. Springer (2014)

[10] Bengio, Y., Courville, A., Vincent, P. (2012). Representation Learning: A Review and New Perspectives. Technical report, Universit´e de Montr´eal.

[11] Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60-88. doi:10.1016/j.media.2017.07.005

[12] L Wang, A Wong. 2020. COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray ImagesmarXiv preprint arXiv:2003.09871, 2020

[13] Y. Song, S. Zheng, L. Li, X. Zhang, X. Zhang, Z. Huang, ..., Y. Chong. 2020. Deep learning enables accurate diagnosis of novel coronavirus (COVID-19) with CT images ,medRxiv (2020)

[14] S. Wang, B. Kang, J. Ma, X. Zeng, M. Xiao, J. Guo, ..., B. XuA. 2020. Deep learning algorithm using CT images to screen for Corona Virus Disease (COVID-19) medRxiv (2020)

[15] M. Barstugan, U. Ozkaya, S. Ozturk. 2020. Coronavirus (COVID-19) Classification Using CT Images by Machine Learning Methods (2020), arXiv preprint arXiv:2003.09424

[16] H. M¨uller, N. Michoux, D. Bandon, A. Geissbuhler. (2004). "A Review of Content-Based Image Retrieval Systems in Medical Applications Clinical Benefits and Future Directions,"Int J Med Inform., vol.73, pp. 1–23, 2004.

[17] M. Rahman, B. C. Desai, and P. Bhattacharya, Image Retrieval-Based Decision Support System for Dermatoscopic Images, 2006, IEEE Symposium on Computer-Based Medical Systems (CBMS'06). on Modeling, Simulation and Applied Optimization, 2009, 1, 1-6.

[18] O. Russakovsky et al., "ImageNet large scale visual recognition challenge," IJCV, 2015.

[19] J. Donahue et al., "DeCAF: A deep convolutional activation feature for generic visual recognition," ICML, vol. 32, pp. 647–655, 2014.

[20] N. Codella, J. Cai, M. Abedini, R. Garnavi, A. Halpern, and J. R. Smith. 2015. "Deep learning, sparse coding, and SVM for melanoma recognition in dermoscopy images," in MICCAI MLMI, vol. 9352, 2015, pp. 118–126.

[21] Karen Simonyan and Andrew Zisserman. 2014. "Very Deep Convolutional Networks for Large- Scale Image Recognition". In: CoRR abs/1409.1556 (2014). URL: http://arxiv.org/abs/1409.1556 (cited on pages 32, 86, 171).

[22] K. He, X. Zhang, S. Ren, J. Sun. 2016. "Deep residual learning for image recognition", Computer Vision and Pattern Recognition (CVPR), 2016.

[23] Gao Huang, Zhuang Liu, Laurens van der Maaten, Kilian Q. Weinberger, Densely Connected Convolutional Networks CVPR 2017, Proc. Computer Vision and Pattern Recognition.

[24] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, and P.-A. Manzagol, "Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion," The Journal of Machine Learning Research, vol. 11, pp. 3371–3408, 2010

[25] A. Krizhevsky and G. E. Hinton, "Using very deep autoencoders for content-based image retrieval." in ESANN. Citeseer, 2011.

[26] Yi, Paul H. MD,; Kim, Tae Kyung B; Lin, Cheng Ting MD, Generalizability of Deep Learning Tuberculosis Classifier to COVID-19 Chest Radiographs New Tricks for an Old Algorithm?, Journal of Thoracic Imaging: July 2020 - Volume 35 - Issue 4 - p W102-W104, doi: 10.1097/RTI.0000000000000532

[27] Wang X, Peng Y, Lu L, Lu Z, Bagheri M, Summers RM. ChestX-ray8: Hospital-scale Chest X-ray Database and Benchmarks on Weakly-Supervised Classification and Localization of Common Thorax Diseases. IEEE CVPR 2017