Unsupervised Clustering of COVID-19 Chest X-Ray Images with a Self-Organizing Feature Map

Bayley King
Dept. of EECS
University of Cincinnati
Cincinnati, OH, USA
king2b3@mail.uc.edu

Siddharth Barve
Dept. of EECS
University of Cincinnati
Cincinnati, OH, USA
barvesh@mail.uc.edu

Andrew Ford
Dept. of EECS
University of Cincinnati
Cincinnati, OH, USA
fordaj@mail.uc.edu

Rashmi Jha
Dept. of EECS
University of Cincinnati
Cincinnati, OH, USA
rashmi.jha@uc.edu

***Abstract*—Machine learning approaches are gaining popularity in the medical field for diagnostics, predictive analytics and general research. With data often being unlabeled or sparse to collect, there is a need for unsupervised learning networks in the medical field. Self-Organizing Feature Maps (SOFM) are a common application of unsupervised networks and allow for the use of unlabeled data in their training. We applied chest x-ray images of COVID-19 patients to an SOFM network and found a distinct classification between sick and healthy patients with an average euclidean distance of 1.1 between 1st and 2nd winning neurons in our testing set. We were also able to show which features of the input space had the highest weight on the classification, to study saliency of features on this unsupervised network. This work shows that unsupervised learning is able to extract features of medical data, specifically chest x-rays of COVID-19 patients, while also accurately classifying the image. This SOFM network can be found at https://github.com/king2b3/SOFM.**

***Index Terms*—self-organization, machine learning (ML), artificial intelligence (AI), unsupervised learning, covid-19**

I. Introduction

In late December 2019, a cluster of pneumonia cases in Wuhan, China were brought forward to the World Health Organization (W.H.O.), which later identified the coronavirus as causative virus behind the outbreak. The virus was officially declared a pandemic on March 11, 2020 and the W.H.O in return called for research to focus on accelerating diagnostics, vaccines and therapeutics for the virus [1]. There is an urgent need for more research but there is a limit on the amount of available data, especially in the field of diagnostic [2]. Some of the difficulties with applying machine learning to biological data is that the datasets are often very small and can often be unlabeled.

Recently, many authors have shown that conventional image processing techniques, like Convolutional Neural Networks and Deep Learning, can help detect features in COVID-19 chest x-ray images [3]–[12]. The major downside of these networks is that they rely on availability of labeled data. COVID-19 databases often have incomplete and inaccurate labels and new data is constantly being added to these existing datasets [13]. Unsupervised learning's biggest strength is that labeled data is not needed in the training process, because the network will cluster the given input space based on similar features. Researchers in the past have applied unsupervised learning techniques to various other biological images, like liver CT scans and other x-ray images, to extract features and have shown promising results [14]–[16]. Shi et al. [13] in their general literature survey over published diagnostic COVID-19 AI techniques showed that only deep learning and other supervised techniques have been published.

Our literature survey also found this lack of publications applying unsupervised techniques to classify COVID-19 chest x-ray images. Unsupervised learning has the benefit of not needing labeled data to still train itself, it makes a great candidate for any kind of medical application within machine learning. We chose to implement the Self-Organizing Feature Map (SOFM) algorithm for its ability to be efficiently deployed for continuous and real-time learning. The simplicity of this algorithm also allows it to be implemented as a hardware algorithm on an application specific integrated circuit (ASIC) [17], which would allow for low powered learning and inferencing of patient data. Our goal with this research is to show that a SOFM is an effective technique to cluster an input set of COVID-19 chest x-ray images while also extracting which features caused the clustering of each classification.

II. The SOFM Algorithm

The SOFM was originally proposed by Kohonen [18] and offers an approach to unsupervised clustering of unlabeled data. The output of an SOFM are usually one or two dimensional lattices of neurons, but higher dimensional lattices are sometimes used in vector quantization problems. The network then uses competitive learning to selectively tune the output neurons to the classes of the input patterns. These output neurons then cluster their weights in locations respective to each other based off of feature similarities. The output map is then a representation of the topographic mapping of the input space [19].

There are three parts to the training algorithm of an SOFM: network initialization and presentation of input vectors, finding the best matching neuron or unit (BMU), and updating the corresponding weights.

A. Network Initialization and Presentation of Input Vectors

There are a few parameters that need to be defined before the training can begin. First the network size needs to be

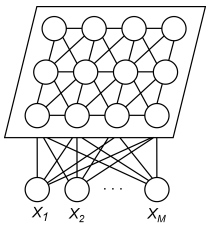

Fig. 1: Visualization of an SOFM 2-D map

defined for the input and output space. The input space is defined as

$$\mathbf{x} = [x_1, x_2, ..., x_M]^T \tag{1}$$

For example, if the dataset consisted of images that are *n*x*n* pixels, then $M = n^2$. Then, the dimension of the output space needs to be defined; in most cases the output is a 2-dimensional grid, which allows for simple visualization. Fig. 1 shows a visual representation of a 2-dimensional map with *M* inputs.

After the network is defined, the initial weights need to be randomized so that training can start. One benefit of using an SOFM is that after training is complete the output space will be a topographic product of the input space, so no translation is needed to prepare the inputs. The input space can consist of any data, for example, ranging from images converted to matrices of their gray-scaled pixel values, biological data like ECG sensor data, or simply binary classifications. Depending on the need of the network the dataset is normally divided into training and testing portions. The training set does not need to be labeled, but the testing set is often labeled for easier visualization after training.

B. Finding the Best Matching Neuron

After applying the first input vector to the network, the BMU must be found. The BMU is whatever output neuron's weights are closest to that input. The BMU, $i*$, is found by

$$i^* = argmin(\sum (w_i - x^q)^2) \tag{2}$$

where w_i is the weight vector of neuron i and x^q is the input pattern.

C. Updating the Weights

After the BMU is found, the corresponding weights w_{ij} need to be updated as follows

$$\Delta w_{ij} = \eta(t)\Lambda(i, i^*, t)(x_j^q - w_{ij}) \tag{3}$$

where η is the learning rate which varies by

$$\eta(t) = \eta_o exp(-\frac{t}{\tau_L}) \tag{4}$$

where τ_L is a user defined constant. The learning rate is usually floored near 0.01 so that learning continues throughout the whole training process. Λ is the neighborhood rate, where a typical choice is

$$\Lambda(i, i^*, t) = exp[\frac{-\mid r_i - r_{i^*} \mid^2}{2\sigma^2(t)}] \tag{5}$$

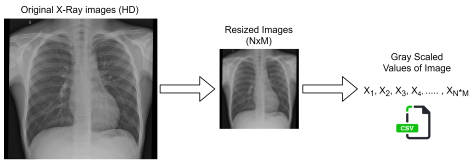

Fig. 2: Dataset generation process

The decay in the neighborhood rate is controlled by τ_N just like the decay of the learning rate.

$$\sigma(t) = \sigma_o exp(-\frac{t}{\tau_N}) \tag{6}$$

The neighborhood rate is normally floored at 1, so that each BMU updates at least it's neighbors in the output map.

Often the training process is broken down into two separate sections: the Self-Organization Phase and the Convergence Phase. The Self-Organization Phase is what trains the weights from their initial randomization to the general structure of the input space. The Convergence Phase is then when the neighborhood rate and learning rate floor to their minimum values, and then the network refines the learned weights to better match the input space. This phase generally runs for 500 times the number of neurons in the output map, but this might be unfeasible for certain map sizes [19]. For a larger datasets the number of epochs can be decreased to allow for a more reasonable training time. If so, then τ_N and τ_L need to be modified to allow for the Convergence Phase to still be roughly two-thirds of the training.

III. Methodology

A. Data Set Generation

We used Cohen, Morrison, and Dao's database [20] of COVID-19 and other chest related illness x-ray images to generate our testing and training datasets. The database features chest x-ray images of patients progressing with Severe Acute Respiratory Syndrome-related Coronavirus (SARS) or SARSr-CoV-1, SARSr-CoV-2 or COVID-19, *Streptococcus spp.*, *Pneumocystis spp.*, and Acute Respiratory Distress Syndrome (ARDS).

The images are divided into two classes, either infected or not infected. First, each image must be resized to the same dimensions due to the varying resolutions of the images. We used OpenCV [21] in Python 3.7 to generate these new resized images, and then return their grayscaled values as a Python list as shown in Fig. 2 We then randomly selected 80% of the images to be used as a training set and the remaining 20% as a testing set. The labels are then removed from the training set, but left for the testing set.

B. Feature Extraction

In a fully trained self organizing feature map, the weights of each neuron take the form of various inputs in the data set belonging to the cluster the neuron represents. Therefore, a mean of the weights of neurons in a cluster can provide an estimate for the mean of the corresponding cluster in the data

TABLE I: SOFM Parameters

Max Epochs	η_0	σ_0	Map Size	τ_N	τ_L
7000	0.1	100	10x10	1000	1000

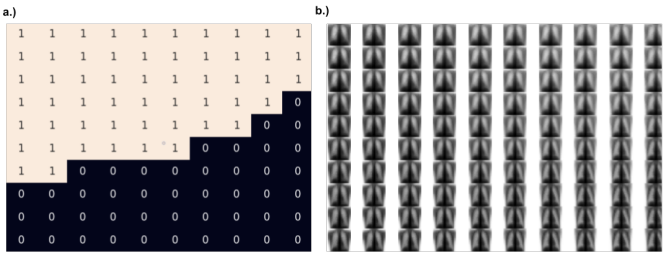

Fig. 3: (a) Output map showing 0 and 1 as healthy and sick patients respectively, (b) Corresponding weight map.

set. Additionally, the variance of the corresponding cluster in the data set can also be estimated by computing the variance of the weights of neurons in the cluster. The means and variance of each cluster allows us to fit a normal distribution model for each attribute in each cluster. In this case, the pixels for the x-ray images were treated as independent and an independent normal distribution was modeled for every individual pixel. This was to allow for regions to be analyzed independently. An overlapping coefficient was used to compare the distributions of pixels in the infected cluster with the corresponding pixels in the non-infected cluster. The overlapping coefficient (OVL) is a measure of the area of overlap between two normal distributions; in this case it is representative of the probability that an input pixel falls within both the infected and non-infected distributions. A lower OVL is indicative of pixels that are able to differentiate between the two clusters since there is little overlap. An acceptance threshold can be applied to the OVL measures of each pixel to identify its significance in the differentiation between the two clusters: in this case we used 0.3.

C. Explainable Results

When introducing an input to the self-organizing feature map, the SOFM outputs the neuron with the closest matching weight map. The input can then be compared to the BMU to see if the output of the neuron is justifiable. A low difference between the BMU and the input image indicates the BMU has captured the features of the input sufficiently. An acceptance threshold can be defined by the user to validate the BMU's representation of the input. Once the cluster that the input the is associated with is known via the cluster of the BMU, the input can be compared to the distribution of the cluster described in the the previous subsection. A Z-score can be computed using the cluster mean and variance for each pixel in the input image. An acceptance region for the normal distribution can then be defined to identify any anomalies in the input image. The input image can also be compared to the distributions of the other clusters. The Z-score for the significant features distinguishing the clusters should be lower

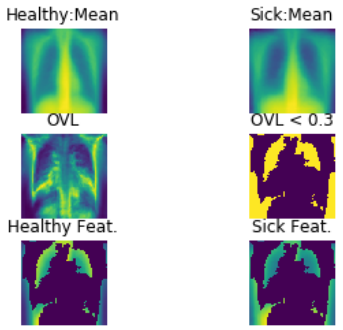

Fig. 4: Significant features distinguishing the clusters

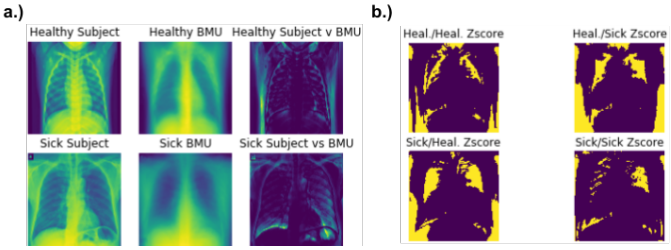

Fig. 5: (a) Healthy and sick input images compared to corresponding BMU's, (b) Z-score of each in pixel in input images in healthy and sick distribution

for the cluster associated with the input and higher for the remaining clusters. There would then be a higher confidence in the classification as well as explainable results.

IV. Results and Discussion

Our training set has 148 images which means by conventional SOFM standards for two phase learning that we should run for approximately 51 thousands epochs, which we found to be unnecessary due to the inherent similarities in the dataset. Although there are two classes of data, the chest x-ray of a sick patient vs. a healthy patient are not very different in terms of the general layout of the input space. We still set parameters for the Self-Organization Phase to run for 1000 epochs, and then have our Convergence Phase run for 6000 epochs. The parameters that we used for our training and testing can be seen in table I. Fig. 3 shows the weight map of the network and output map for the fore-mentioned trained network on the testing set. The healthy cluster made up 45% of the map while the sick cluster made up the remaining 55%. The final euclidean distance between the 1st and 2nd BMUs after training was 1.1 neurons, indicating that the neighbors of each BMU reflect proper clustering.

The significant features distinguishing the two clusters were extracted using the process described in Section III.B. using an OVL acceptance threshold of 0.3. Visually inspecting the cluster means for the healthy and sick cluster, illustrated in Fig. 4, we can identify the top portion of the lungs for the sick cluster is darker and the bottom portion of the sick cluster is wider. Performing the OVL-based feature extraction, we can see that the features extracted distinguishing the two

clusters agree with the features that were identified by visual inspection.

The input of a healthy subject and sick subject were randomly selected for testing the explainability of the SOFM. Visually inspecting the inputs and their corresponding BMU in Fig. 5, shows that the input and the BMU match fairly well aside from the details of the ribs smoothed out in the BMU. The image illustrating the difference between the input and BMU agrees with the visual inspection showing very little error or difference. Comparing the subject's x-ray images to the cluster distributions in Fig. 5 with 2-tailed Z-test acceptance region with α = 0.005, shows that there is a lesser quantity of anomalies in the significant regions specified in Fig. 4 for the cluster associated with the subject. These results increase the confidence of the output classification of the SOFM for that input, because the BMU properly captures the input and the input falls better in the acceptance region of the associated cluster, especially in the significant regions.

V. Conclusion and Future Works

We have shown that unsupervised learning, specifically with an SOFM network, can cluster COVID-19 chest x-ray images and extract their features successfully. The major strength of the network is that labeled data is not a requirement for training, which allows for the network to be trained on a larger quantity of data since labeled data is often scarce. Convolutional Neural Networks have a high computational complexity as well as a large memory requirement to implement. SOFMs utilize a much simpler algorithm and the memory requirement is bound by the size of the network which can be reduced at the expense of some resolution penalty. The SOFM also allows for lifelong learning and enables for it to continuously learn as more COVID-19 patient x-rays are taken. This would also enable it to adapt as new symptoms may arise or if the virus begins to mutate. The distributions produced for each cluster also allow for anomaly detection using Z-scores and acceptance regions as demonstrated in Fig. 5. This could potentially allow for experts to be notified of any other infection that a patient may have or that may be spreading, which may include any further mutations.

We would like to investigate incorporating an additional set of different inputs or attributes for each patient: symptoms, age, gender, etc. We would like to investigate finding correlation between these sets of attributes and various conditions and infection in addition to the current COVID-19 pandemic. This would allow for the network to classify using more easily available measurements since x-rays are not easily accessible especially during the current crisis. We would also like to incorporate more clusters for various conditions instead of just healthy and sick. This may result in a network which is able to track various diseases and classify which cluster subjects fall within in addition to identifying anomalies.

Acknowledgment

We would like to thank National Science Foundation for Award No. # CCF 1718428 (REU supplement) and Air Force Research Lab (AFRL) Wright Patterson, Ohio for Award No. FA8650-18-C-1191 P00005 for funding our research. We would also like to thank Alex Jones, Joshua Mayersky, Wayne Stegner, and Tyler Westland for their help and feedback throughout our research.

References

[1] W. H. Organization, "Coronavirus disease (covid-2019) r and d," Apr 2020. [Online]. Available: https://www.who.int/blueprint/priority-diseases/key-action/novel-coronavirus/en/

[2] Y.-R. Guo *et al.*, "The origin, transmission and clinical therapies on coronavirus disease 2019 (covid-19) outbreak–an update on the status," *Military Medical Research*, vol. 7, no. 1, pp. 1–10, 2020.

[3] L. Li, L. Qin, Z. Xu, Y. Yin, X. Wang, B. Kong, J. Bai, Y. Lu, Z. Fang, Q. Song *et al.*, "Artificial intelligence distinguishes covid-19 from community acquired pneumonia on chest ct," *Radiology*, p. 200905, 2020.

[4] F. M. Salman, S. S. Abu-Naser, E. Alajrami, B. S. Abu-Nasser, and B. A. Alashqar, "Covid-19 detection using artificial intelligence," *International Journal of Academic Engineering Research*, vol. 4, pp. 18–25, 2020. [Online]. Available: http://dspace.alazhar.edu.ps/xmlui/handle/123456789/587

[5] A. E. Hassanien, L. N. Mahdy, K. A. Ezzat, H. H. Elmousalami, and H. A. Ella, "Automatic x-ray covid-19 lung image classification system based on multi-level thresholding and support vector machine," *medRxiv*, 2020.

[6] C. Zheng *et al.*, "Deep learning-based detection for covid-19 from chest ct using weak label," *medRxiv*, 2020.

[7] K. Hammoudi *et al.*, "Deep learning on chest x-ray images to detect and evaluate pneumonia cases at the era of covid-19," *arXiv preprint arXiv:2004.03399*, April 2020.

[8] X. Li, C. Li, and D. Zhu, "Covid-mobilexpert: On-device covid-19 screening using snapshots of chest x-ray," *arXiv:2004.03042*, April 2020.

[9] O. Gozes, M. Frid-Adar, N. Sagie, H. Zhang, W. Ji, and H. Greenspan, "Coronavirus detection and analysis on chest ct with deep learning," *arXiv preprint arXiv:2004.02640*, April 2020.

[10] L. O. Hall, R. Paul, D. B. Goldgof, and G. M. Goldgof, "Finding covid-19 from chest x-rays using deep learning on a small dataset," *arXiv preprint arXiv:2004.02060*, April 2020.

[11] E. E.-D. Hemdan, M. A. Shouman, and M. E. Karar, "Covidx-net: A framework of deep learning classifiers to diagnose covid-19 in x-ray images," *arXiv preprint arXiv:2003.11055*, April 2020.

[12] A. Narin, C. Kaya, and Z. Pamuk, "Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks," *arXiv preprint arXiv:2003.10849*, April 2020.

[13] F. Shi *et al.*, "Review of artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for covid-19," *arXiv preprint arXiv:2004.02731*, 2020.

[14] A. Afifi and T. Nakaguchi, "Unsupervised detection of liver lesions in ct images," in *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. IEEE, 2015, pp. 2411–2414.

[15] H. Greenspan and A. T. Pinhas, "Medical image categorization and retrieval for pacs using the gmm-kl framework," *IEEE Transactions on Information Technology in Biomedicine*, vol. 11, no. 2, pp. 190–202, 2007.

[16] P. Fischer, T. Pohl, A. Faranesh, A. Maier, and J. Hornegger, "Unsupervised learning for robust respiratory signal estimation from x-ray fluoroscopy," *IEEE transactions on medical imaging*, vol. 36, no. 4, pp. 865–877, 2016.

[17] J. R. Mann and S. Gilbert, "An analog self-organizing neural network chip," in *Advances in neural information processing systems*, 1989, pp. 739–747.

[18] T. Kohonen, "Self-organized formation of topologically correct feature maps," *Biological cybernetics*, vol. 43, no. 1, pp. 59–69, 1982.

[19] S. Haykin, *Neural networks - A comprehensive foundation*, 2nd ed. Prentice-Hall, 1999.

[20] J. P. Cohen, P. Morrison, and L. Dao, "Covid-19 image data collection," *arXiv 2003.11597*, 2020. [Online]. Available: https://github.com/ieee8023/covid-chestxray-dataset

[21] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.