Interpretable Saab Subspace Network for COVID-19 Lung Ultrasound Screening

*

Dennis Hou
Research and Development
Caloudi Corporation
San Francisco, CA 94131
dennis@caloudi.com

Raymond Hou, M.D.
Department of Emergency Medicine & Critical Care
Northshore University Hospital
Manhasset, NY 11030
Rhou@northwell.edu

Janpu Hou
Applied Data Research
Institute of Data Learning
Las Vegas, NV 89180
janpu@ieee.org

***Abstract*—In addition to bedside Point-of-care diagnosis, lung ultrasound imaging classifier has been used to triage of COVID-19 symptomatic patients after emergency room admission. There is a more urgent need for a simple and low cost portable ultrasound device for each elderly resident at risk in retirement communities and independent living facilities to monitor their lung conditions, to find out if their lung condition is getting worse with COVID-19 during self-quarantine phase or if it's getting better during their recovery phase. Various complicated convolutional neural networks have been developed with very high accuracy but health professionals find it hard to understand and trust such complex models due to the lack of intuition and explanation of their predictions. In this work, we proposed to use an interpretable Subspace Approximation with Adjusted Bias (Saab) multilayer network to screen the lung ultrasound images. Such subspace representations learned from a successive subspace network will provide more invariance to intra-class variability and thus give better discrimination for a task such as classification. We demonstrated the advantage of using Saab Subspace Network to design a low-complexity, low-cost, low-power-consumption solution for interpreting and visualizing features of the lung ultrasound images to confirm the classifier recommendation. Since both training and inference can be done on-device, make it a potential solution to deliver personalized healthcare to underserved senior communities without any internet connection.**

***Index Terms*—Digital Health, Saab Transform, Edge Computing, Subspace Learning**

I. Introduction

The 2019 novel coronavirus, SARS-CoV-2, (COVID-19) is an emerging pathogen of critical significance to international public health.[1] Many community efforts were put in to understand the impact, risks and mitigation of such pandemics.[2] Lung ultrasound can have several advantages, such as reduced health worker exposition to infected patients, repeatability during follow-up, low-costs and easier application in low-resource settings. There are many reports on the possible use of lung ultrasound for the evaluation of patients with suspected COVID-19 infection. Development of fast, accurate and objective machine learning methods for COVID-19 lung ultrasound image evaluation is still at an early stage. Ultrasound is very helpful for assessing the content of

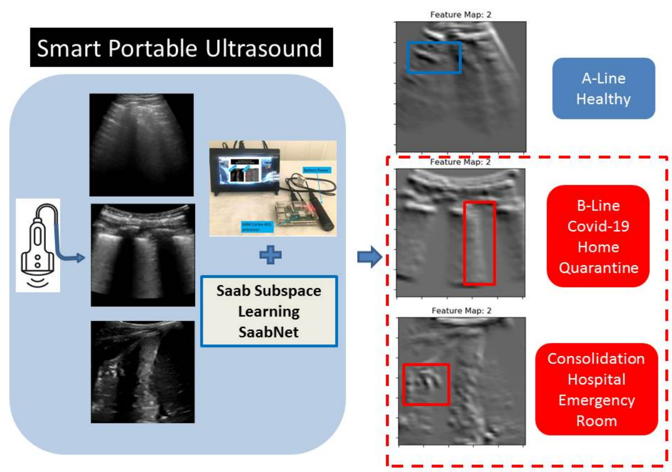

Fig. 1. Smart Portable Ultrasound with On-Device AI

the pleural space and the assessment of pleural fluid versus consolidated lung. The A-line is a horizontal artifact indicating a normal lung surface. The B-line is a kind of comet-tail artifact indicating sub-pleural interstitial edema. With the new General Data Protection Regulation (GDPR) in healthcare, the importance for easy interpretability of the machine learning model becomes essential. A convolution neural network (CNN) architecture has multiple stages stacked on top of each other, followed by a supervised classifier. Each stage generally comprises "three key components" – a convolutional filter banks, a nonlinear activation with rectified linear unit (ReLU) and a max pooling to reduce dimension. [3] To obtain weights of the filter bank, we need to use more intense back-propagation computation. Most of times we need additional AI hardware acceleration chips for both training and inference. Increasing the speed of medical CNN and maintaining the accuracy is still a challenging task. One possible solution is texture feature coding method (TFCM) [4], which is a texture analysis scheme which transforms an original image into a text feature image whose pixels represent texture feature numbers. Additional sets of tool and frameworks such as Explainable AI (XAI) is needed to interpret and visualize machine learning models before deploy them with confidence for critical mission such as healthcare. [5,6]

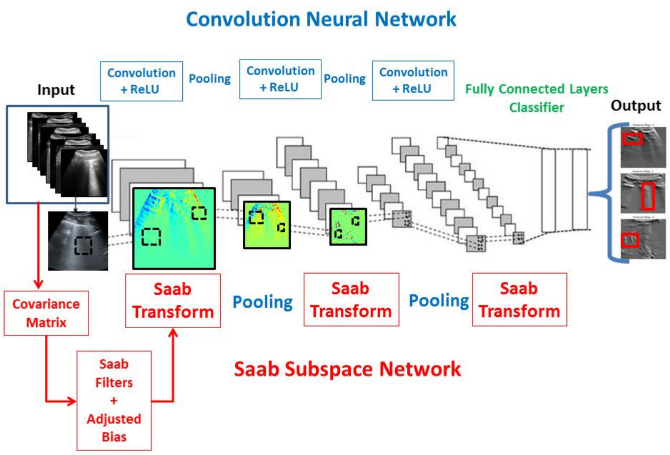

Fig. 2. Subspace Approximation Adjusted Bias (Saab) Successive Subspace Learning Network with CNN Architecture

To develop a fast and interpretable deep learning model with pre-defined filters with wavelet functions, Mallat et al. introduced the invariant Scattering Convolution Networks to both supervised and unsupervised classification of signals and images [7]. Scattering transforms extract invariant and stable subspace representations through a non-linear successive transform, which approximated image representation with scattering decomposition. Each stage computed with "three key components" - a convolutional network where filter coefficients are given by a pre-defined wavelet operator, a nonlinear activation with wavelet modulus operators, and a pre-defined scale function for pooling on averaging. Since the filter banks are pre-defined, there is no need for extra hardware to do back-propagation computation.

To further develop data-defined filter banks with fast and interpretable deep learning model, Kuo et al. [8,9] proposed a similarity decomposition and composition from covariance matrix called "Saab (subspace approximation with adjusted bias) transform" to both supervised and unsupervised image and point cloud classification. The multi-stage Saab transforms[10], which approximated the image representation with successive subspace learning[11]. Each stage has "three key components" - a convolutional layer with data-driven covariance operator to find kernels as filter bank, a nonlinear activation with adjust bias, and a max pooling to reduce dimension in between each layers. Similar to CNN, Saab subspace network learn the filters weights from the training data. But instead of using back-propagation to update the filter weights, the filter weights are updated in a much less computation intensive feed-forward manner. Saab subspace network has a very similar architecture to Scattering Convolution Networks and produce subspace coefficients as feature maps by each layer.

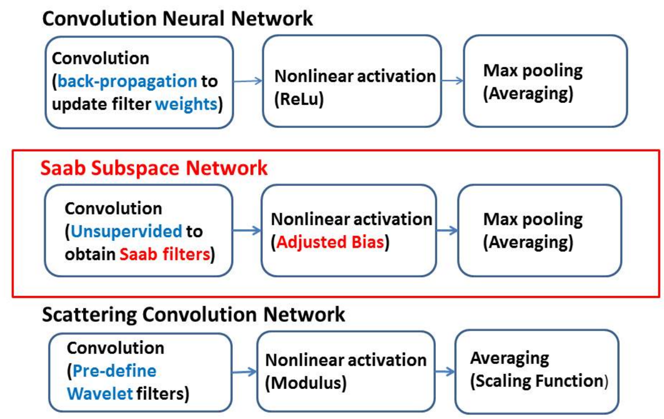

Fig. 3. Similarity and Difference among CNN, Saab Subspace Network and Scattering Convolutional Network

The goal is to integrate a low-power ARM based SoCs with on-device training capability into the ultrasound scanner. When tested positive, the COVID-19 patients can use such device to monitor their lung status during quarantine even without internet connection. The CNN can provide very accurate classification, but due the complexity and size of the model, it will require the patients sending their data to the cloud for training and inferencing. During the pandemics, even the tele-medicine facilities have limited bandwidth to handle massive data analysis. If we can train the model at the edge where the data were collected, we can reduce the workload in the health cloud. In this work we apply the Saab transform based successive subspace learning model for COVID-19 lung ultrasound image screening. The general framework for the classifier is displayed in Fig.1. In this work we demonstrate the concept by training the classifier on Raspberry Pi. The lung ultrasound image acquired by portable ultrasound scanner will feed into a battery powered raspberry pi platform with ARM Cortex-52 processor then classified by Saab subspace network without any additional AI accelerator hardware.

Normally, patients presented with fever, cough, fatigue, dyspnea symptom and with known COVID-19 contact need to use Real-Time Reverse Transcriptase (RT)–PCR diagnostic panel for testing. While waiting for the results, pulse oximetry and POCUS lung ultrasound should be taken. [12]

In reality, researchers found that COVID-19 patients were most likely already infected before or as they developed symptoms. Study also found that 5% of infected people took at least 2 weeks to develop symptoms, suggesting they may have been infectious for a long time.[13] Therefore there is a need for low-cost easy to use portable ultrasound for high risk individual, e.g. elderly in retirement home, to daily monitor their progress without share the ultrasound scanner

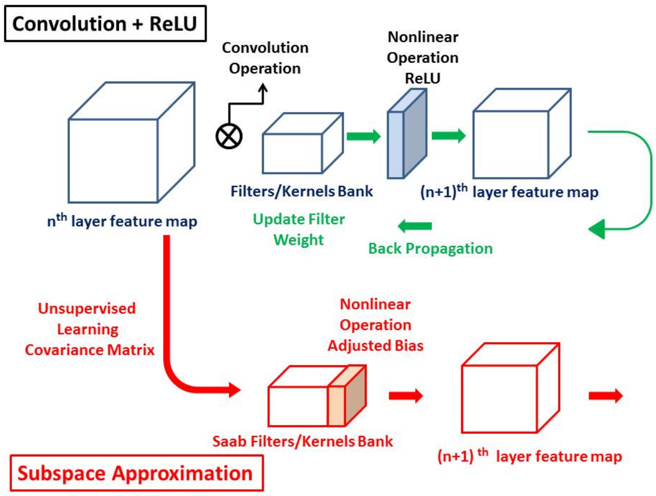

Fig. 4. Data-driven Saab Filters and Adjusted Bias Application

with others. If A-line is presented, most likely the lung is still healthy. When B-line starts to present, self-quarantine and continues monitoring is a must. With B-line patient may not need supplemental oxygen but viral testing is a must. When B-line increased and supplemental oxygen is needed admit to hospital quickly. If consolidation already occurred, not only supplemental oxygen is needed but emergency room visit is necessary. [14-16]

Assisted living homes do not provide medical services yet have the greatest risk for severe illness from COVID-19 among those aged 65 or older. There are also other factors that can increase their risk for severe illness, such as having underlying medical conditions. If all the elderly residents in retirement communities and independent living facilities have their own low cost AI-powered portable ultrasound device to continuously monitor their lung condition, we can reduce the developing of more serious complications from COVID-19 illness. The nursing staff can take evidence-based action to eliminate potential factors that put they at an increased risk.

Born et al. demonstrated the role of ultra-sound images for COVID-19 detection and aggregated data in an open source repository. [17]

II. Methodology

The basic concepts underlying pyramid and Saab decompositions techniques are applicable to other multiscale decomposition methods. Similar to traditional image processing, the ultrasound Image representations can be obtained by Saab transforms. On the other hand, ultrasound image features can be extracted by kernels or filters. By applying Saab transform, we can not only obtain the ultrasound image representation but also extract the features with Saab filters as shown in Fig.2.

In this feed-forward design, a target pattern is typically represented as a linear combination of responses of a set of orthogonal Saab filters. Once the Saab filters were learnt from given image set, any additional image from the similar distribution set will be represented by the similar Saab coefficients. We can consider our ultrasound image dataset as one major set with three subsets: A-line, B-line and Consolidation. Each subset of has slight different characteristics.

Input: n^{th} layer of feature $\{ x_1, x_2, x_3 \ldots x_m \}$

Step 1: Compute covariance matrix Σ

$$\Sigma = \frac{1}{m}\sum_{i=1}^{m}(x_i - xm_{ean})(x - xmea_n)^T \qquad x_{mean} = \frac{1}{m}\sum_{i=1}^{m}(x_i)$$

Step 2: Find Saab Kernels = Eigenvectors of covariance matrix Σ

DC kernels a_o = eigenvector with zero eigenvalue
AC kernels a_k = eigenvectors with (K-1) largest eigenvalues

Step 3: Find Adjusted Bias = $b_k \geq \max \|\| x \|\|$

Step 4: Perform Saab Transform

$$y = x^T a_o + \sum_{k=1}^{K-1} x^T a_k + b_k$$

Output: $(n+1)^{th}$ layer of feature $\{ y_1, y_2, y_3 \ldots y_{mk} \}$

Fig. 5. Functional block from one layer of features map into next layer of features

All the images in the major set can be represented by the same set of Saab filters with different coefficients. Different image from different subset can be discriminated from Saab coefficients.

As indicated by Hinton et.al on developing deep learning network, high-dimensional data can be converted to low-dimensional codes by training a multilayer neural network with a small central layer to reconstruct high-dimensional input vectors. [18] Similarly, SaabNet use a multilayer structure to reduce the high dimensional data and discover deeper latent features. The idea of successive subspace learning is to discover multiple stages of representation, that higher-stage features represent more abstract semantics of the data. Since here is no need to update the filter weights with backpropagation, thus save computation complexity so it can fit into simple battery powered portable device.

The similarity and difference among CNN, SaabNet and ScatNet are illustrated in Fig.3.

A. Compute covariance matrix and find Saab filters

In the traditional convolutional neural network, the hidden layers work as trainable feature extractor. Training with back-propagation allows CNN to optimize the weights of its convolutional filters, hence learning the filters weights by minimizing the error, as shown in Fig. 4 upper section.

In Saab-transform based approach, we need to compute the covariance matrix among the images under training first. Because the eigenvectors of the covariance matrix are orthogonal to each other, they become the natural choice of anchor vectors for Saab transform. The eigenvector with the largest eigenvalue is the direction along which the data set has the maximum variance. If we select the largest n eigenvectors it will form a basis of the n-dimensional subspace which retains the most variance of the original data.

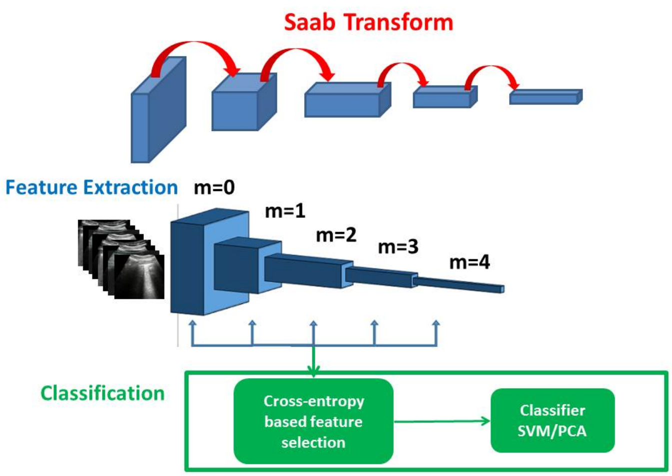

Fig. 6. Illustrated processing flow of SaabNet

In Saab notation, the first (K-1) largest eigenvectors are selected as AC kernels, and the eigenvector of zero eigenvalue is selected as DC kernels. Together it will transform the original dataset into a K-dimensional subspace, as shown in Fig. 4 lower section.

The CNN filter, similar to a matched filter encountered in signal processing, provides a measure for similarity of a patch of input and a feature. On the other hand, the Saab filters obtained from the covariance matrix based on similarity measure are the match filter banks for extract features.

Conventionally, convolutional neural networks process different images with the same set of filters. However, the variations in images pose a challenge to this fashion and force the usage of large number of filters with large number of layers to approximate the images. With Saab transform, we generate sample-specific filters in the forward pass. Since the filters are generated on-the-fly, the model becomes more flexible and can better fit the training data with much less training time compared to traditional CNNs.

B. Non-linear activation with Adjusted-Bias

Neural network activation functions are a crucial component of deep learning. Activation functions also have a major effect on the convergence speed of the neural network. The function is attached to each neuron in the network, and based on the relevance to the subject of interest to determine whether it should be activated or not. Mathematically, the activation function is a gate in between the input feeding the current neuron and its output going to the next layer.

Anchoring and adjustment is a cognitive heuristics where human starts off with an initial idea and adjusts their beliefs based off of this starting point. Same idea is used in Saab Transform based subspace network.

Anchoring is down with Saab Filters as anchor vector, than the adjustment is made by adjusted bias as shown in Fig. 4 lower section. Deep neural network models use non-linear activation functions to create complex mappings between the network's inputs and outputs. So is in Saab-Transform based successive subspace learning network. Between layers the adjusted-bias is applied to make sure all the necessary features are mapping into next layer.

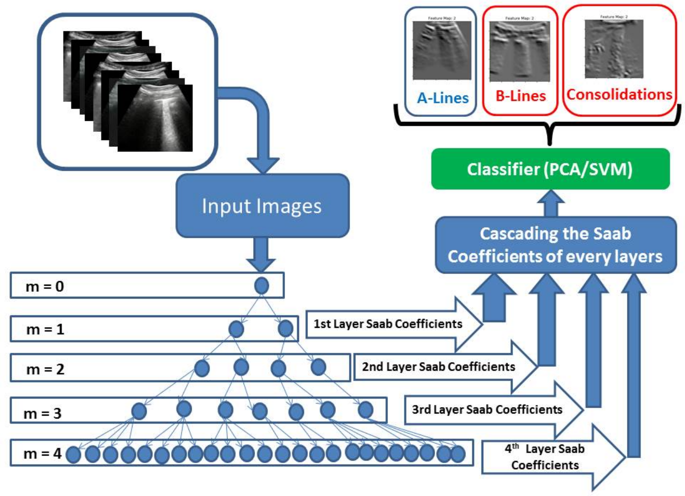

Fig. 7. Application of SaabNet for COVID-19 Lung Ultrasound Screening

C. Max pooling and complete Saab Transform

Although Saab subspace network has similar architecture as scattering convolution network, the way they solve the translation invariance problem is slightly different. The scattering convolution network use the scale function as low pass filter to do the averaging and use additional complex modulus to obtain the stable pooling. The Saab subspace network just uses the same max pooling as traditional convolution neural network.

The max pooling not only creates translation invariance, but also improves deformation invariance. This has the effect of making the resulting down sampled feature maps more robust to changes in the position of the feature in the image. All the computation steps from one layer to the next layer have been illustrated in Fig.5. Once the multi-layer feature extractor was build, the classifier can be assembled as display in Fig.6.

III. Results and discussion

A dataset of 2800 images were used for this study, majority of the dataset were from images and video clips provided by Born[18] with sources of the data from grepmed.com, thepocusatlas.com, butterflynetwork.com, and radiopaedia.org. The data set consisted of 740 A-line images, 1150 B-line images and 910 consolidation images) extracted from video clips and published research works. Data augmentation techniques were also used to diversify the data. 560 images were used for testing and 2240 images were used for training. We need to build the image embedding from the dataset. If a simple convolutional neural network such as AlexNet with 5 layers of convolution layer and 3 layers of fully

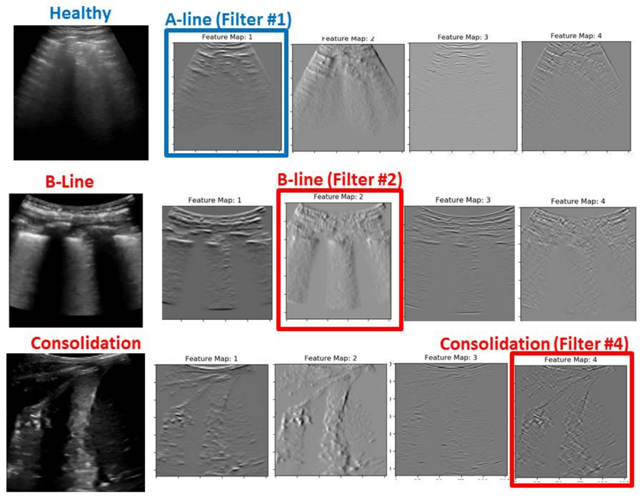

Fig. 8. Discrimination of different patterns with same set of Saab filters

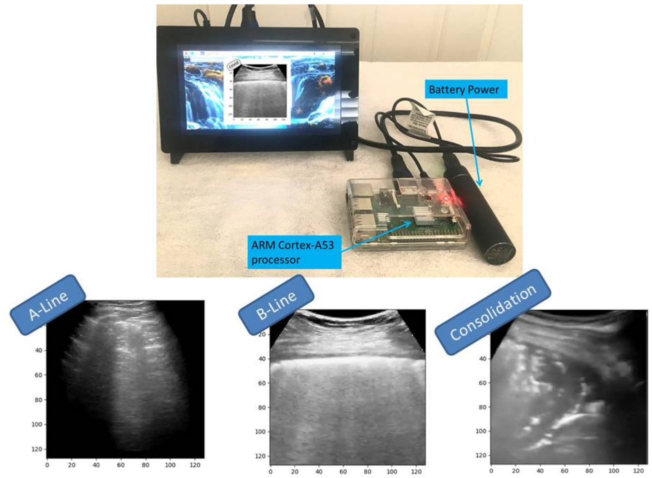

Fig. 11. Classifier under testing

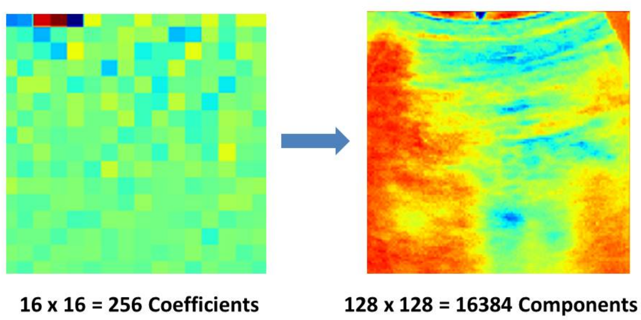

Fig. 9. Reconstruction of Image from Inverse Saab Transform

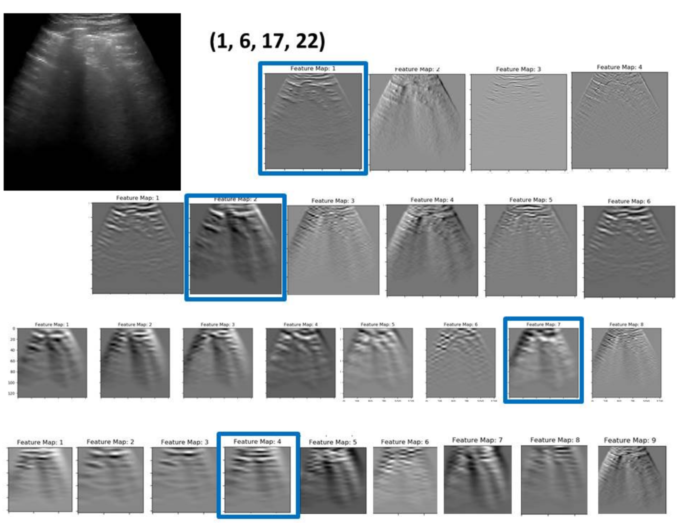

Fig. 10. Reconstruction of Image from different layer coefficients

connected layers were used to yield an embedding vector of 1000 elements, we need to solve the 62,378,344 unknown weights and biases. After computing the embedding vector, we can use a softmax layer for classification [19]. The softmax is a generalization of logistic regression that can be used for multi-class classification but require time consuming end-to-end backpropagation computation. These high unknown parameters of CNN model required special hardware for both training and inference make it inadequate to fit into a small embedded platform like Raspberry Pi. On the other hand the SaabNet only need to solve 2,800 eigenvalues to yield an embedding vector of 1183 elements and can be done on any low cost simple board computers. When datasets arise from a multivariate normal distribution, we can perform accurate inference on its mean vector and covariance matrix. The Saab filter bank recovered from the covariance matrix is used for all the images to extract coefficients. The images from different distribution may respond to certain filters than others. We found that when we reach m=4, we already have over 95% classification accuracy as shown in Fig.7.

We can consider Saab filters as matched filter, some match A-line better than others, the other match B-line better, for example. We show the four top-ranked images that have the strongest responses with respect to a certain Saab filter in the each layer of the SaabNet and their corresponding feature responses in Fig. 8.

One advantage of SaabNet over traditional CNN is easy interpretation. By using the Saab Inverse transform, one layer feature map can inverse transform back to the early layer, and eventually reconstruct the original image. Similar to the auto-encoder, with Saab transform as encoder and inverse Saab transform as decoder. A reconstruction from the hidden layer with 256 coefficients can re-construct the image of 16348 pixels in Fig.9. The reconstructed image from different layers of one lung ultrasound image has been shown in Fig. 10.

Such Saab transform based successive subspace learning model can easily been applied to build a powered portable COVID-19 lung ultrasound image screening classifier.

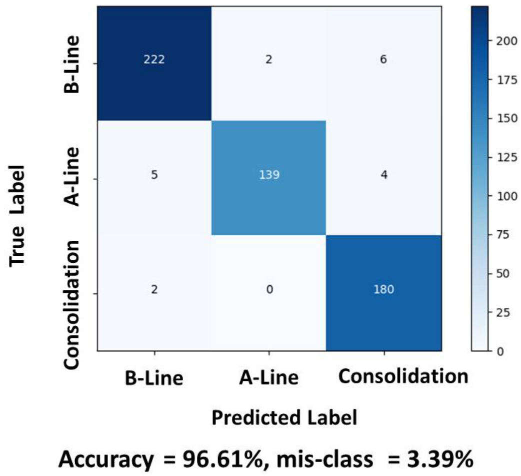

Fig. 12. Classifier confusion matrix

The lung ultrasound image acquired by portable ultrasound scanner will feed into a battery powered raspberry pi platform with ARM Cortex-52 process then classified by Saab subspace network without any additional AI accelerator hardware as shown in Fig. 11. Over 96% accuracy of testing data can be obtained as shown in Fig. 12.

IV. Conclusion

The lack of on-device intelligence based energy efficient portable lung ultrasound slow the adoption of digital health customization for high risk elderly patients in underserved areas, and assistance with monitoring and managing COVID-19 lung conditions. The backpropagation during the training of deep learning model encapsulate the most computationally intensive part of the training calculation. Improved energy efficient performance can be achieved by replacing the default standard framework such as Pytorch or Tensorflow with a more optimized alternative such as Saab subspace learning network.

Each feature extraction layer is computed by Saab transform, which utilize the correlation between the images of different symptoms. The difference between each type of images, the shape and size of different Saab filters and the irregularity of lung all can be translated into Saab coefficients. We demonstrated that Saab subspace learning Network is suitable for portable device because it does not need long iteration of backpropagation to update filter weights. Covid-19 ultrasound lung images like A-line, B-line and consolidation can achieve 96% detection accuracy. Such on-device AI can mitigate network limitations, reduce energy consumption, due to the training of learning model is done right at device where the data is. With the low cost aspect, this covid-19 screening classifier can be applied to more nursing homes with limited budgets.

References

[1] Hou R, Koti S, DePeralta DK "Covid-19 Complicated by Hemorrhagic Necrotizing Pancreatitis: A Case Report." J Gastrointest Hepat Surg 3(1):32-36, 2020

[2] Getz, Michael, Yafei Wang, G. An, Andrew Becker, C. Cockrell, N. Collier, M. Craig, Courtney L. Davis, J. Faeder, A. F. Versypt, Juliano F. Gianlupi, J. Glazier, Sara Hamis, R. Heiland, T. Hillen, Dennis Hou, M. A. Islam, Adrianne Jenner, Furkan Kurtoglu, B. Liu, F. Macfarlane, Pablo Maygrundter, P. Morel, A. Narayanan, J. Ozik, E. Pienaar, P. Rangamani, J. Shoemaker, A. M. Smith and P. Macklin. "Rapid community-driven development of a SARS-CoV-2 tissue simulator." Published Biology, Medicine (bioRxiv) 2020

[3] R. Marconi Ramos, C. Ghedini Ralha, T. M. Kurc, J. H. Saltz and G. Teodoro, "Increasing Accuracy of Medical CNN Applying Optimization Algorithms: An Image Classification Case," 2019 8th Brazilian Conference on Intelligent Systems (BRACIS), Salvador, Brazil, 2019, pp. 233-238

[4] J. Liang, X. Zhao, R. Xu, C. Kwan and C. -. Chang, "Target detection with texture feature coding method and support vector machines," 2004 IEEE International Conference on Acoustics, Speech, and Signal Processing, Montreal, Que., 2004, pp. ii-713

[5] M. A. Ahmad, A. Teredesai and C. Eckert, "Interpretable Machine Learning in Healthcare," 2018 IEEE International Conference on Healthcare Informatics (ICHI), New York, NY, 2018, pp. 447-447

[6] T. Karatekin et al., "Interpretable Machine Learning in Healthcare through Generalized Additive Model with Pairwise Interactions (GA2M): Predicting Severe Retinopathy of Prematurity," 2019 International Conference on Deep Learning and Machine Learning in Emerging Applications (Deep-ML), Istanbul, Turkey, 2019, pp. 61-66

[7] J. Bruna and S. Mallat, "Invariant Scattering Convolution Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 35, no. 8, pp. 1872-1886, Aug. 2013

[8] Y. Ding et al., "Saak Transform-Based Machine Learning for Light-Sheet Imaging of Cardiac Trabeculation," in IEEE Transactions on Biomedical Engineering

[9] C.-C. J. Kuo. "Understanding convolutional neural networks with a mathematical model." J. Visual Communication and Image Representation, 41:406–413, 2016

[10] T. Ramanathan, A. Manimaran, S. You and C. J. Kuo, "Robustness of Saak Transform Against Adversarial Attacks," 2019 IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, 2019, pp. 2531-2535

[11] C.-C. J. Kuo, M. Zhang, S. Li, J. Duan and Y. Chen, "Interpretable Convolutional Neural Networks via Feedforward Design." arXiv preprint arXiv:1810.02786 (2018).

[12] D. Dong et al., "The role of imaging in the detection and management of COVID-19: a review," in IEEE Reviews in Biomedical Engineering

[13] Kucharski AJ Russell TW Diamond C et al. "Early dynamics of transmission and control of COVID-19: a mathematical modelling study." Lancet Infect Dis. March 11 2020

[14] Peng, Q., Wang, X. & Zhang, L. Findings of lung ultrasonography of novel corona virus pneumonia during the 2019–2020 epidemic. Intensive Care Med (2020).

[15] Huang, Yi and Wang, Sihan and Liu, Yue and Zhang, Yaohui and Zheng, Chuyun and Zheng, Yu and Zhang, Chaoyang and Min, Weili and Zhou, Huihui and Yu, Ming and Hu, Mingjun, A "Preliminary Study on the Ultrasonic Manifestations of Peripulmonary Lesions of Non-Critical Novel Coronavirus Pneumonia (COVID-19)" (February 26, 2020).

[16] S. Roy et al., "Deep Learning for Classification and Localization of COVID-19 Markers in Point-of-Care Lung Ultrasound," in IEEE Transactions on Medical Imaging, vol. 39, no. 8, pp. 2676-2687, Aug. 2020.

[17] Jannis Born, Gabriel Brändle, Manuel Cossio, Marion Disdier, Julie Goulet, Jérémie Roulin, and Nina Wiedemann. Pocovid-net: Automatic detection of covid-19 from a new lung ultrasound imaging dataset (pocus). arXiv preprint arXiv:2004.12084, 2020.

[18] G. E. Hinton, R. R. Salakhutdinov. "Reducing the Dimensionality of Data with Neural Networks", Science 28 Jul 2006: Vol. 313, Issue 5786, pp. 504-507

[19] Ayhan, B.; Kwan, C.; Budavari, B.; Kwan, L.; Lu, Y.; Perez, D.; Li, J.; Skarlatos, D.; Vlachos, M. "Vegetation Detection Using Deep Learning and Conventional Methods." Remote Sens. 2020, 12, 2502.