Learnable Quantization and Lesion Area Segmentation for COVID-19 Lungs' CT Images

[1]Hanyou Cao, [1]Libo Zhang, [2]Sihan Lyu, [2]Sizhe Kuang

[1]Department of Electrical and Computer Engineering
[2]Department of Biomedical Engineering
Duke University
Durham, NC 27705
hanyou.cao, libo.zhang, sihan.lyu, sizhe.kuang@duke.edu

Abstract

Recently the most infectious disease is the novel Coronavirus disease (COVID 19), which creates a devastating effect on public health in more than 200 countries in the world. [1] Right now, RT-PCR is the most popular diagnosis for COVID-19 disease, but it is still too time-consuming and error prone. These defects may be further amplified when billions of tests are required. The alternative detection method is to find suspected lesion region in lung CT images, which is more reliable since viral infection in the lung is one of the most common early indicators of this disease. However, there are still several challenges restricting the application of this method, including high variation in lesion characteristics and low contrast between lesion areas and healthy tissues. Here, we provide a U-Net model with a trainable quantization layer to automatically segment the suspected lesion region from the lungs' CT images, which is dedicated to relieve the heavy-loaded work for CT scan detection. We propose a trainable physical layer and a U-Net to segment lesion regions from lungs' image scan. With this model, we can help agents to diagnose coronavirus disease accurately and efficiently. We have 2184 lungs' CT images as the training data and 545 images as testing data. With the trainable quantization layer, the performance is dramatically improved. This results can help the design of CT detectors' quantizers.

1 Introduction

Starting from 2019, the emerging cases of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) population represents a tremendous medical and economic crisis. [2] Such coronavirus will cause serious symptoms including fever, dry cough, and shortness of breath. These acute respiratory syndromes have caused a worldwide sudden and substantial increase in hospitalizations for pneumonia with multiorgan disease. [3]5% of patients with COVID-19, and 20% of those hospitalized, experience severe symptoms necessitating intensive care. More than 75% of patients hospitalized with COVID-19 require supplemental oxygen. [3] Within those syndromes, bilateral distribution of ground glass opacities (GGO) with or without consolidation in posterior and peripheral lungs was the cardinal hallmark of COVID-19. [4]

Currently, RT-PCR is the most common strategy to diagnose patients with COVID-19. However, it is proved that compared to traditional RT-PCR strategy, chest CT scan is far more sensitive which can provide more reliable diagnostic results. Several studies reported that chest CT scans show typical imaging features in all patients with COVID-19. This high sensitivity and initial presentation in CT chest can be helpful in rectifying false negative results obtained from RT-PCR. [5] Due to the strong

transmission ability of COVID-19 virus, Billions of diagnostic tests may be required every day. If we try to apply CT chest scan to help diagnose COVID-19, it is an extremely heavy loaded work for doctors to view such large numbers of CT photos. To help diagnosis, we plan to design a specific python U-Net to classify, localize, and segment the strongly affected part of the lung.

We decided to implement a U-Net with a quantization physical layer that performs the functionality of quantizers to modify the original dataset and segment the lesion region from patients' lungs. In this project, we would like to compare predicted results from the following experiment setups:

(a) non-quantized training data, and quantized testing data without quantization layer.

(b) quantized training data and quantized testing data without quantization layer

(c) quantized training data and quantized training data with trainable quantization layer.

We will also compare the dice coefficients calculated from different models. Without a doubt, all the clues prove that our U-Net and trainable physical layer can segment lesion regions from injured lungs' CT photo accurately and efficiently, meanwhile saving substantial computational cost.

2 Related Work

The paper related to our project is 'CARes-U-Net: Content-aware residual U-Net for lesion segmentation of COVID-19 from chest CT images'. In this paper, Xinhua Xu and his group decided to create CARes-U-Net to segment the lesion area of COVID-19 from the chest CT slices. In their CARes-U-Net, the residual connection was used in the convolutional block, which alleviated the degradation problem during the training. Then, the content-aware up sampling modules were introduced to improve the performance of the model while reducing the computation cost. Moreover, to achieve faster convergence, an advanced optimizer named Ranger was utilized to update the model's parameters during training. Finally, they employed a semi-supervised segmentation framework to deal with the problem of lacking pixel-level labeled data. [6]

We decided to apply a traditional U-Net which was included in Lab8 as the basic format. On the other hand, to improve accuracy, we used Xinhua's loss function as a reference. We also designed a trainable quantization physical layer serving as an old camera to further reduce the storage memory requirement for large datasets and improve the model's segmentation performance.

3 Method

3.1 U-Net Model

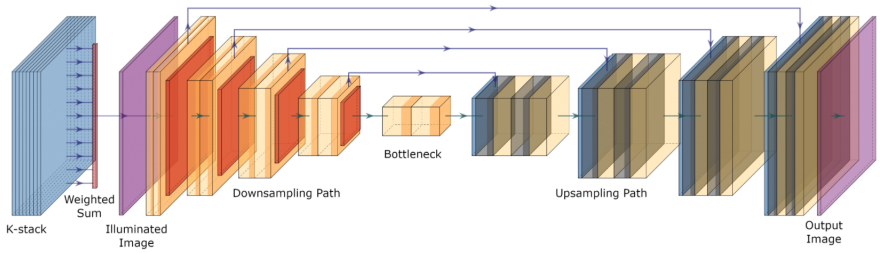


Figure 1: U-Net

We decided to build a U-Net model to complete the basic segmentation task. Based on Figure 1, Our U-Net architecture is referenced from BME548L Week 8 Lab Session. The down-sampling path of the U-Net model functions as an encoder to compress feature information. After the bottleneck which consists of two convolutional layers, the up-sampling path works as a decoder to restore the original image size with predicted lesion area segmentation.

However, we have to make some modifications to make it suitable for our project. Firstly, instead of a weighted sum of illuminated images, we have quantized images before the convolutional layers.

Secondly, instead of the Sparse Categorical Cross-Entropy loss appropriate for multi classification problems, for this project, we define our own Segmentation loss function taking both Dice loss and Binary Cross-Entropy loss into consideration, because Dice coefficient is a very important metric to evaluate segmentation performance. Therefore, instead of Softmax activation, we utilize Sigmoid activation function for the output layer.

$$DICE_{discrete} = \frac{2TP}{2TP + FP + FN} \tag{1}$$

$$DICE_{continuous} = \frac{2 \times \sum(\hat{y}y)}{\sum(\hat{y} + y)} \tag{2}$$

$$L_{DICE} = 1 - DICE_{continuous} \tag{3}$$

$$L_{Segment} = L_{DICE} + L_{BCE} \tag{4}$$

The discrete Dice coefficient equation (Eq. 1) is important when we try to evaluate the segmentation performance. However, the transition from discrete Dice (Eq. 1) to continuous Dice (Eq. 2) is necessary for training our U-Net model because the Dice loss has to be differentiable for backpropagation optimization. We define our Dice loss function as one minus the continuous Dice (Eq. 3) because we always want to minimize the loss. Finally, we formulate our Segmentation loss as a linear combination of Dice loss and Binary Cross-Entropy loss (Eq. 4).

3.2 Physical Layer

We propose a trainable quantization layer. In the layer, we have a selection array of the same size as the image. In the array, one indicates that the corresponding pixel is selected to quantize and zero indicates it's not selected. They are discrete parameters, so we must use continuous variables to train them. To this end, we sample the selection from a relaxed Bernoulli distribution. The trainable parameters become the probability matrix of the Bernoulli distribution. With a higher probability, it's more likely to get one and quantize the pixel.

Quantization is performed according to the following formula (Eq. 5). In our experiment, k is set to 7, which means that after quantization, there are only two gray value levels left for each pixel.

$$X_{qntz} = floor(X/2^k) \times 2^k, \tag{5}$$

where the X_{qntz} is the gray value after quantization and x is the original value.

In the experiment, the number of pixels to quantize must be controlled for comparison. Since the selection of pixels is sampled from Bernoulli distribution, the constraint should be the sum of the probabilities of the Bernoulli variables (Eq. 6). It makes sense because the expectation of Bernoulli distribution is the probability of the positive event.

$$\mathbb{E}\, Q = \sum_{i=1}^{n} p_i, \tag{6}$$

where Q is the number of pixels quantized, n is the number of all the pixels and p is the corresponding probability to quantize.

4 Results

For dataset, we utilize 2729 grayscale images (512 by 512) with pixel values ranging from 0 to 255 (8 bits) and we set 80%-20% train-test split ratio, which means 2184 observations for training, and 545 images for testing. Based on Figure 2, starting with 0.0005, we set a decreasing learning rate for

the last 30 epochs, and we train our baseline U-Net model for 60 epochs with Adam optimizer.

The baseline U-Net model is trained by the non-quantized images. To demonstrate the influence of quantization, we implement the hard-coded quantization on all pixels, transferring the original data from 8-bits gray level into other 7 different gray levels, as shown in Figure 3a. We directly feed the quantized images into the baseline U-Net model to check the model's segmentation performance when predicting different gray-level images. One segmentation result is shown in Figure 3b. From the result, the segmentation performance declines when input images' gray level decreases and it hardly provides any output when the input images are 2-bits gray level and 1-bit gray level. However, when we retrain the U-Net by quantized images, the model has a more accurate segmentation area when it deals with the input images which have the same gray level with the training data. It should be noted that there is no output for the 1-bit data training model because the training data lose too much contrast between healthy tissue and lesion area. One example is shown in the Figure 3c. We use dice coefficients to measure and compare the performance of different models, as shown in Figure 3d.

Then, we apply the trainable quantization layer into the U-Net structure. There are 262144 weights in this physical layer, which is the same number as the number of pixels. Each weight represents the pixel probability of quantizing. We still use the whole dataset to train the U-Net with trainable quantization. We train the new structure U-Net for 200 epochs with Adam optimizer. We set a bigger learning rate, starting from 0.001 and it will start to decrease from 150 epochs.

After the training, we pull out the quantization layer weights to find the location with higher quantized probabilities. As it shows in Figure 4a , we can roughly recognize a lung's shape from the weight. To visualize more clearly, we set two thresholds to figure out the regions that the model prefers to quantize and reserve.

In Figure 4b, the probabilities smaller than 0.3 are set as 0 and others are set as 1. Most of the black part focuses on the inner part of the lung because the inner lungs provide more details for segmentation. From Figure 4c, the probabilities smaller than 0.7 are set as 0 and others are set as 1. The edge of the lung mainly represents white. Based on this, we assume that the model takes the quantization implementation as a method to improve the contrast to achieve better segmentation performance.

To prove the enhancement of performance is achieved through quantization at specific regions, not just through uniform quantization, we design a comparison group which randomly quantizes pixels in the image. We set the constraint as 0.5 for the trainable quantization group which means in this case, we expect to quantize 50% pixels. On the other hand, we hard code the comparison group (random quantization group) to quantize 50% pixels. In Figure 5, the trainable quantization layer generates

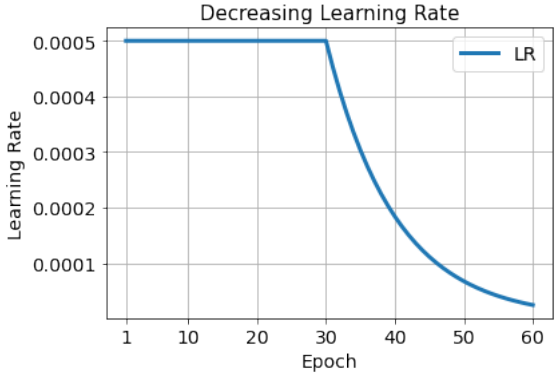
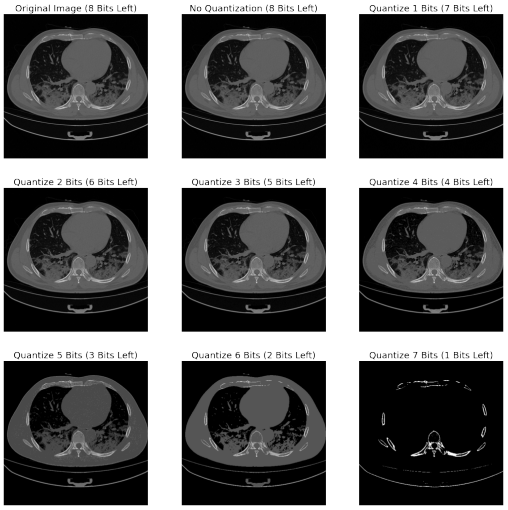
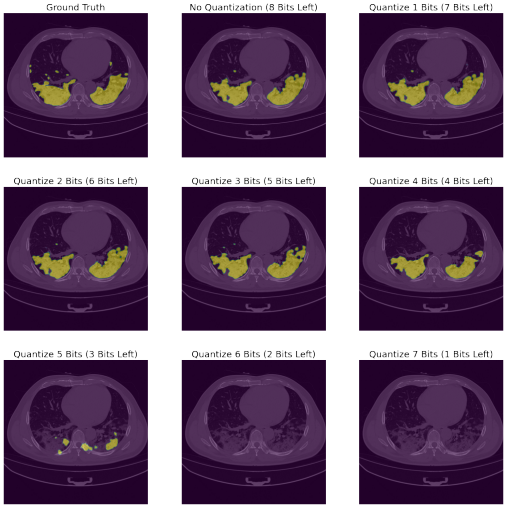


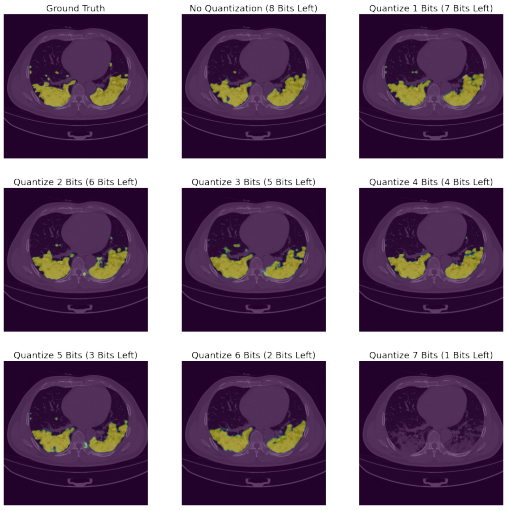
Figure 2: Learning Rate



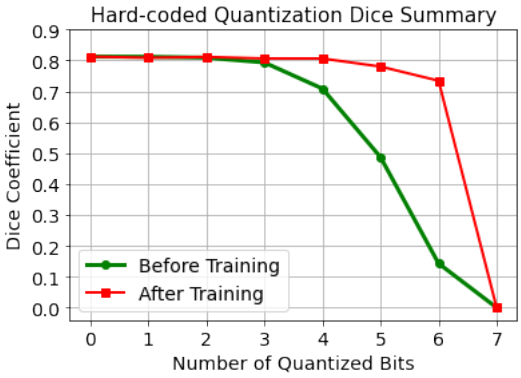
(a) Quantized Images



(b) Before Hard-coded Quantization Training

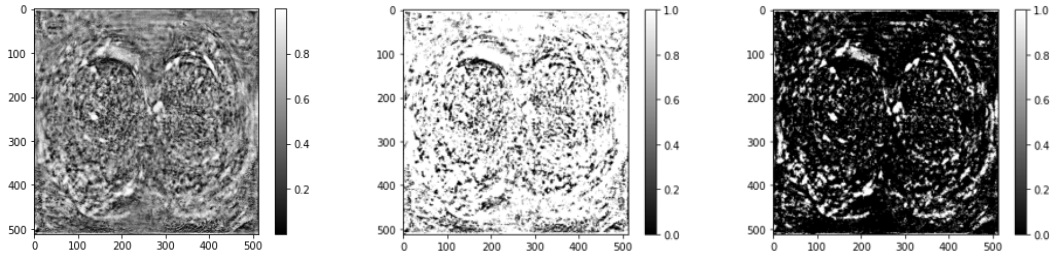


(c) After Hard-coded Quantization Training



(d) Hard-coded Quantization Dice Summary

Figure 3: Quantization Input and Predicted Results from Different Hard-Coded model



(a) Original Weights (b) Weights with 0.3 as Threshold (c) Weights with 0.7 as Threshold

Figure 4: Quantization Layer Weights

a much better segmentation result than the random selection one. From Figure 6, the U-Net with trainable quantization layer achieve the DICE coefficient as 0.88, even better than the baseline U-Net.

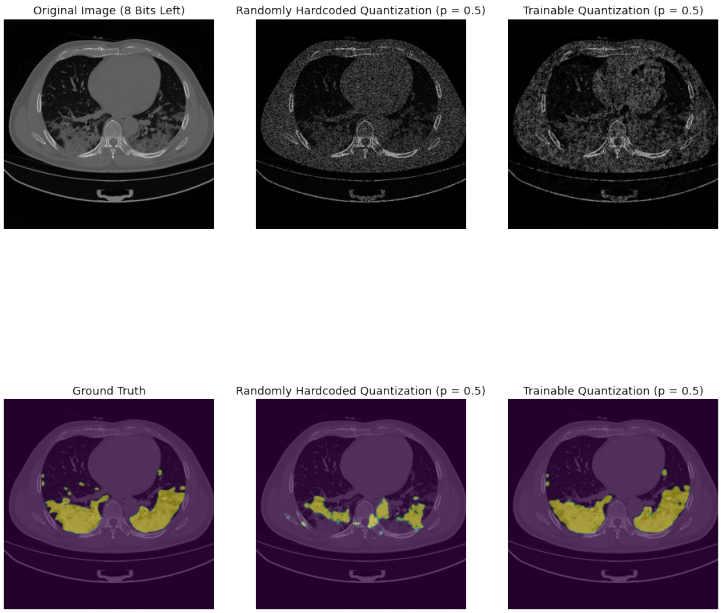


Figure 5: Results from Untrainable and Trainable Physical Layer

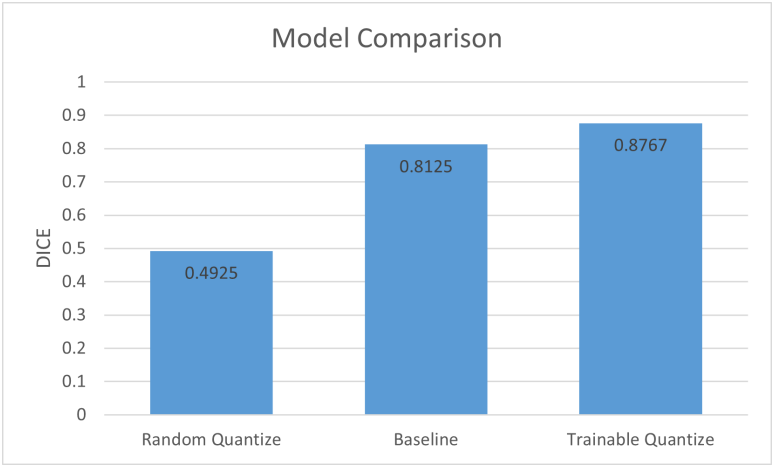


Figure 6: Dice coefficients at different situation

5 Discussion

In the project, we've comprehensively studied the influence of quantization on the segmentation task of Covid-19 lungs. Our work can be summarized as follows:

(a) We quantize the images with different number of bits, train the U-Net only and evaluate the performance degradation.

(b) We propose a trainable quantization layer and come up with an approach to adding constraints on expected number of pixels to quantize.

(c) Lastly, we insert the trainable quantization layer before the U-Net and train them on the COVID-19 lungs' CT data set.

The results show that the neural network doesn't pay attention to the background. However, it captures the edge of the lungs. Along the edge, it tends to quantize the pixels on one side and not to quantize on the other side. In this way, the contrast of the lesion areas is increased so that

performance can be improved. The result can help the CT community to design adapted quantizers in the CT detectors. The automatic diagnosis of U-Net segmentation can be more accurate while the storage for CT images can be reduced.

There is still some meaningful future work.

(a) We can extend from binary choices of quantization to multiple choices, which means that different pixels can be quantized with different numbers of bits instead of having only one bit when quantized.

(b) We can extend the neural network input to 3D images.

(c) We can examine the performance when different percentages of pixels are quantized.

References

[1] Islam, M. R. (2022). Complex features extraction with deep learning model for the detection of COVID19 from CT scan images using ensemble based machine learning approach. Expert Systems with Applications.

[2] L.Schultze, J. (2021). COVID-19 and the human innate immune system. Cell, 1.

[3] Wiersinga, W. J. (2020). Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19). JAMA Network, 1.

[4] Ye, Z. (2020). Chest CT manifestations of new coronavirus disease 2019 (COVID-19): a pictorial review. SpringerLink, 1. Retrieved from springer Link.

[5] Alsharif, W. (2021). Effectiveness of COVID-19 diagnosis and management tools: A review. Radiography, 682. .

[6] Xu, X. (2021). CARes-U-Net: Content-aware residual U-Net for lesion. Medical Physics.