Catheter and Line Position Challenge

Fangzhou Ai
Department of Electrical and Computer Engineering
University of California, San Diego
La Jolla, CA 92093
faai@eng.ucsd.edu

Yue Qiao
Department of Electrical and Computer Engineering
University of California, San Diego
La Jolla, CA 92093
yuq021@eng.ucsd.edu

Zunming Zhang
Department of Electrical and Computer Engineering
University of California, San Diego
La Jolla, CA 92093
zuz008@eng.ucsd.edu

Abstract

In this report, we proposed our solution to the Kaggle competition "RANZCR CLiP - Catheter and Line Position Challenge." We used different architectures like Resnet, EfficientNet, DenseNet as the backbone of our solution to classify the line and tube from x-ray photos. We also applied different image augmentation methods to enhance our results. We introduced the learning rate scheduling method for fast convergence during the training, and we also employed mixed-precision training to accelerate the whole procedure while saving the GPU memory. The experiment result shows that...

1 Introduction

Our project will be based on the Code Competition from Kaggle (find more here).

In hospital, malpositioned lines and tubes would cause severe complocations in patients. Nowadays, doctors and nurses frequently use traditional checklists to ensure they follow protocol in managing patients. Yet, these steps can be time consuming and are still prone to human error, especially in stressful situations when hospitals are at capacity.

Nasogastric tube malpositioning into the airways has been reported in up to 3% of cases, with up to 40% of these cases demonstrating complications. Airway tube malposition in adult patients intubated outside the operating room is seen in up to 25% of cases. The likelihood of complication is directly related to both the experience level and specialty of the proceduralist. Early recognition of malpositioned tubes is the key to preventing risky complications (even death), even more so now that millions of COVID-19 patients are in need of these tubes and lines.

In the real world, the gold standard for the confirmation of line and tube positions are chest radiographs. Once getting the chest radiographs (X-Ray images), a physician or radiologist must manully check these chest x-rays to verify the lines and tubes are correctly positioned. Not only does this leave

room for human error, but delays are also common as radiologists can be busy reporting other scans. Leveraging the deep learning, we believe that we could develop the algorithms that may be able to automatically detect malpositioned catheters and lines. Once alerted, clinicians can reposition or remove them to avoid life-threatening complications, which will dramatically reduce the man-power involved.

The Royal Australian and New Zealand College of Radiologists (RANZCR) is a not-for-profit professional organisation for clinical radiologists and radiation oncologists in Australia, New Zealand, and Singapore. The group is one of many medical organisations around the world (including the NHS) that recognizes malpositioned tubes and lines as preventable. RANZCR is helping design safety systems where such errors will be caught.

2 Motivation

Under the COVID-19 padamic, almost all the hospital are at capacity and more potients are in need of correctly positioned catheter lines and tubes. If successful, our efforts may help clinicians save lives. Earlier detection of malpositioned catheters and lines is even more important as COVID-19 cases continue to surge. Quick feedback on catheter and line placement could help clinicians better treat these patients. Even beyond COVID-19, detection of line and tube position will ALWAYS be a requirement in many ill hospital patients.

3 Related Work

Paras Lakhani [4] evaluated the efficacy of deep convolutional neural networks (DCNNs) in differentiating subtle, intermediate, and more obvious image differences in radiography. In the paper, three different datasets were created, which included presence/absence of the endotracheal (ET) tube (n = 300), low/normal position of the ET tube (n = 300), and chest/abdominal radiographs (n = 120). The datasets were split into training, validation, and test. Both untrained and pre-trained deep neural networks were employed, including AlexNet and GoogLeNet classifiers, using the Caffe framework. Data augmentation was performed for the presence/absence and low/normal ET tube datasets.

In[6], Hongyu Wang and Yong Xia proposed a model called ChestNet, which consists of two branches: a classification branch serves as a uniform feature extraction-classification network to free users from troublesome handcrafted feature extraction, and an attention branch exploits the correlation between class labels and the locations of patholog-ical abnormalities and allows the model to concentrate adaptively on the patholog-ically abnormal regions.With this model they achieved SOTA on the Chest X-ray 14 dataset.

In [1], Maayan and etc. suggest a method for training the network, first with synthetic data and then with real X-ray images in a fine-tuning phase, which allows the network to train on thousands of cases without annotating any data. The proposed method was tested on 477 real chest radiography from a public data set and reached AUC of 0.99 in classifying the presence vs. absence of the ET tube, along with outputting high quality ET tube segmentation maps, which could give us more hints on how the training should go on.

4 Method

4.1 Dataset

In this competition, we detect the presence and position of catheters and lines on chest x-rays. Use machine learning to train and test our model on 40,000 images to categorize a poorly placed tube.

The dataset has been labeled with a set of definitions to ensure consistency with labeling. The normal category includes lines that were appropriately positioned and did not require repositioning; The borderline category includes lines that would ideally require some repositioning but would in most cases still function adequately in their current position; The abnormal category included lines that required immediate repositioning. Since there can be multiple tubes in an X-Ray image and different tubes have different usages, the dataset labels are further separated into different tube categories. There are four categories for tubes in the dataset: PAC stands for Pulmonary artery catheter; ETT

stands for endotracheal tube; NGT stands for Nasogastric tube; CVC stands for central venous catheter.

4.2 Architecture

Convolutional neural network (CNN) is one of the most popular architectures to solve the image classification problem. ResNet[2] is one of the most widely used CNN architectures. ResNet features residual learning blocks that help fight against vanishing gradients problem by introducing skip connections. EfficientNet[5] is a recently introduced convolutional neural network. The authors of EfficientNet first design a new baseline network using neural architecture search and later scales it up using a new scaling method that uniformly scales the properties of CNNs using a highly effective compound coefficient. The EfficientNet series archives state-of-the-art accuracy in the ImageNet dataset as of 2020. DenseNet[3] connects each layer to every other layer in a feed-forward fashion. Whereas traditional convolutional networks with L layers have L connections — one between each layer and its subsequent layer — DenseNet network 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. DenseNets have several compelling advantages: they alleviate the vanishing-gradient problem, strengthen feature propagation, encourage feature reuse, and substantially reduce the number of parameters.

In our experiments, we choose ResNet, Densenet and EfficientNet as the backbone CNN for our classification model. Since there are multiple tubes, we proposed a multi-head image classification network for poorly placed tube detection. In our model, we have four different heads. Each of them receives the same flattened CNN features as input and uses a single linear layer for classification output. We use different loss functions for different heads. For ETT and SGC heads, we applied a standard Softmax layer and cross-entropy loss on them. For NGT and CVC heads, we applied a Sigmoid layer and Binary cross-entropy loss on them. We use Binary cross entropy loss on NGT and CVC heads because there could be more than one NGT or CVC tubes in one X-Ray image. Therefore, the result can be both in the Normal category and Abnormal Category. The traditional Softmax layer does not fit this situation, so we use a sigmoid layer and Binary cross-entropy loss.

4.3 Image augmentation

We applied random scale, flip and crop to the image during pre-processing. More details. . .

4.4 Learning rate scheduling

To archive best accuracy, we dynamically reduce the learning rate during training. We use a method called reduce on plateau. More details. . .

4.5 Mixed precision training

As recent CNN models become deeper and deeper, GPU memory requirement increases, and the training speed decreases. We use a new training method called mixed-precision training to accelerate the training speed and reduce GPU memory usage. During training, all convolution and matrix multiplication will be performed under float 16 precision, significantly reducing GPU memory usage by half and increasing training speed. When calculating loss, the precision is scaled back to float 32 to prevent underflow or overflow for loss calculation.

5   Experiments

We split the official dataset into the training set and validation set, and use a fixed random seed to ensure the result is reproducible. 90% percent of data becomes training set, and 10% of data becomes validation set. During the training, we use SGD as the optimizer, set the initial learning rate to 0.001 and momentum to 0.9. We use PyTorch as our training framework and use Tensorboard to monitor and record our training results.

5.1 Results

The validation accuracy for different categories is shown in Figure 1.

(a) ETT Accuracy
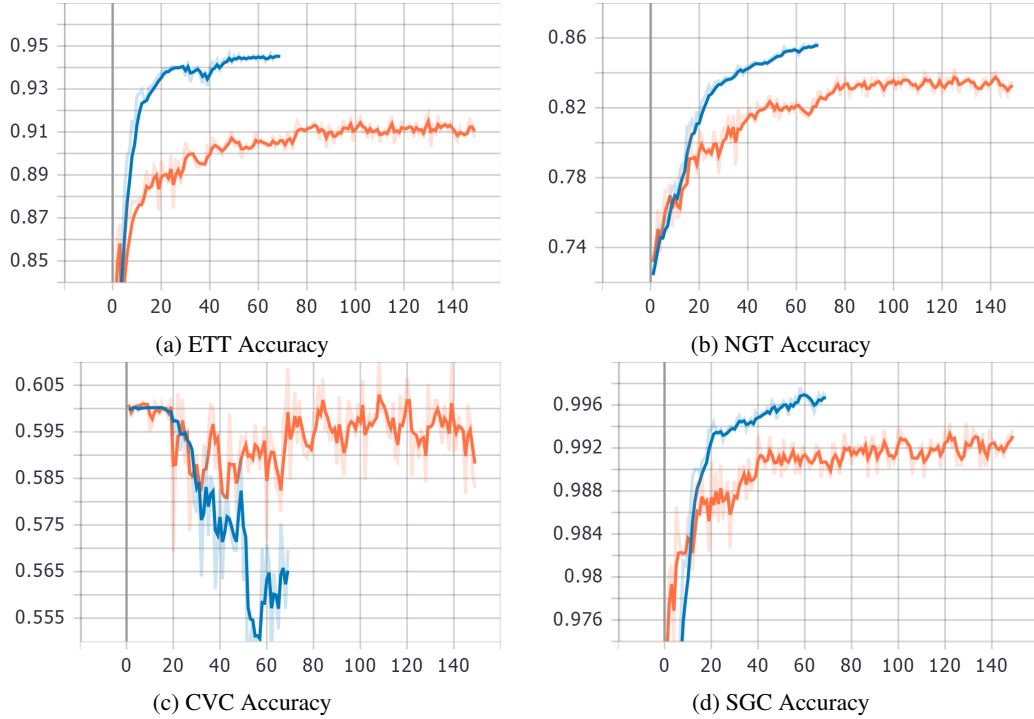
(b) NGT Accuracy

(c) CVC Accuracy

(d) SGC Accuracy

Figure 1: Validation accuracy for different tube categories. Orange line: Model with ResNet50 backbone. Blue line: Model with EfficientNet-b5 backbone.

More experiments in progress...

6 Discussion

We can see that for ETT, NGT and SGC, EfficientNet outperforms ResNet a lot from the Figure 1. That's expected since EfficientNet also outperforms ResNet in ImageNet classification. However, for CVC, EfficientNet has a sharp drop in accuracy during training. We need to do more investigation to find out what happened.

Recent work has shown that convolutional networks can be substantially deeper, more acurate, and efficient to train if they contain shorter connections between layers close to the input and those close to the output. DenseNets take this into consideration and develop the architecture. But from this experiment, we could see that the accuracy is lower than EfficientNet

References

[1] Maayan Frid-Adar, Rula Amer, and Hayit Greenspan. Endotracheal tube detection and segmentation in chest radiographs using synthetic data, 2019.

[2] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015.

[3] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q Weinberger. Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708, 2017.

[4] Paras Lakhani. Deep convolutional neural networks for endotracheal tube position and x-ray image classification: challenges and opportunities. *Journal of digital imaging*, 30(4):460–468, 2017.

[5] Mingxing Tan and Quoc V. Le. Efficientnet: Rethinking model scaling for convolutional neural networks, 2020.

[6] Hongyu Wang and Yong Xia. Chestnet: A deep neural network for classification of thoracic diseases on chest radiography, 2018.