Transfer learning for domain adaptation

Adapting Semantic Segmentation Models to Medical Imaging for Retinal Layer Analysis in Diabetic Macular Edema

Project Proposal Deliverable

Team Members:
Xiaoquan Liu (xl395)
Yuanjing Zhu (yz792)

Advisor:
Hai Li, Ph.D.

Project Goals

What are you trying to do?

In this project, we aim to adapt existing image segmentation models, initially designed for other domains such as autonomous driving cars, to analyze retinal layers in optical coherence tomography (OCT) scans from patients with diabetic macular edema (DME). We aim to work with a real-world medical image dataset[1], making sure the data is thoroughly cleaned and preprocessed for analysis. Then we plan to use the mm-segmentation library[2] to explore various pre-trained segmentation models and and select the best method to fine-tune them for our specific task. By experimenting with different model structures and fine-tuning techniques including loss function optimization and data augmentation, we aim to effectively adapt models initially designed for tasks like autonomous driving to the healthcare domain, enhancing their performance in identifying pathologies from medical images.

Background

How is it done today?

Semantic segmentation is the process of classifying pixels in order to isolate certain items of interest from pictures that present a comprehensive picture of the scene. A technique called transfer learning, where a pre-trained neural network model, often trained on a large dataset for a related task, is fine-tuned or adapted to a specific dataset, is usually used in semantic segmentation due to several key advantages that make it a popular approach in the field. These advantages include addressing the limited availability of labeled data, facilitating knowledge transfer from pre-trained models, speeding up convergence during training, improving model accuracy, reducing the risk of overfitting, and enabling the reusability of well-performing models for various related tasks.

While many existing segmentation models are initially trained on datasets from domains like autonomous driving, where their task involves segmenting vehicles, pedestrians, and obstacles within camera images from vehicles, the growing demand for precise image segmentation in healthcare has driven the adoption of semantic segmentation techniques within the healthcare field. Accurate image segmentation is fundamental for diagnosing and monitoring a wide range of medical conditions, including the detection of tumors, lesions, or abnormalities within organs and tissues. It enables precise quantification of anatomical structures and their volumetric analysis, facilitating treatment planning, surgical interventions, and radiation therapy targeting. Moreover, segmentation is essential for tracking disease progression over time and assessing treatment outcomes, offering valuable insights into the effectiveness of medical interventions. In essence, medical image segmentation empowers healthcare professionals with the tools needed for early diagnosis, personalized treatment, and improved patient care, ultimately leading to better clinical decision-making and patient outcomes.

Methodology

Your approach and why do you think it will be successful? How long will it take? What are the final "exams" to check for success?

To segment the retinal layer from optical coherence tomography (OCT) scans obtained from patients with diabetic macular edema (DME), we would process and analyze the DME dataset to visualize a subset of the images and record the essential statistics (e.g., total number of images, image resolution) as the preliminary study. The next step would be using semantic segmentation models on the ophthalmic dataset and evaluating their performance without fine-tuning. Based on the results, we would then fine-tune the model to achieve better segmentation outcomes, evaluated by metrics such as the intersection-of-union (IoU) score, the dice coefficient (F1 score), etc. Data augmentation techniques, such as random shift, random flip, and random brightness change, etc., as well as the use of other loss functions, such as dice loss and intersection over union (IoU)-balanced loss, would be taken into consideration to further improve the model's performance. The project is scheduled for **SIX WEEKS** with the joint efforts of the team members.

- **Semantic Segmentation Models**

Here are the two segmentation models from mm-segmentation library that we plan to experiment with the medical dataset.

Fully Convolutional Networks (FCNs)[3] have revolutionized the way we approach pixel-wise classification tasks in computer vision, marking a major milestone in the field of semantic segmentation. FCNs are specifically developed for the task of dense semantic segmentation, in contrast to typical convolutional neural networks, which are intended for classification. They introduced a significant change in architecture: instead of using fully connected layer-based architecture, they used one composed entirely of convolutional layers. Because of this architectural change, FCNs are now able to produce pixel-by-pixel segmentation maps, which makes them ideal for applications like medical image analysis, object recognition, and scene comprehension.

MobileNet[4] is based on a streamlined architecture to construct lightweight deep neural networks.The core idea behind MobileNet is the use of depthwise separable convolutions, which break down a traditional convolution into a depthwise convolution followed by a pointwise convolution. This simplification significantly reduces the number computations while still preserving a substantial amount of representational capability. MobileNet's efficiency and adaptability make it a popular choice for healthcare applications, particularly in the context of semantic segmentation of medical images. For instance, MobileNet can be employed for the segmentation of different anatomical structures or abnormalities in medical imaging modalities such as X-rays, MRI, CT scans, and ultrasound images. In this project, we will explore and evaluate both MobileNet v2 and MobileNet v3.

- **Evaluation Metrics**

During the fine-tuning process, we intend to test two evaluation metrics to assess the model's performance. First is **m**ean **I**ntersection **o**ver **U**nion (**mIoU**). The Intersection over Union (IoU) score, also known as the Jaccard Index, is a commonly used metric to evaluate the performance of image segmentation algorithms. It quantifies how well the predicted segmentation aligns with the ground truth by calculating the ratio between the area of overlap and the area of union of the two regions. In multi-class segmentation, the mean IoU is calculated by taking the average of the IoU scores for each class. It ranges from 0 to 1, where higher mIoU scores suggest better segmentation performance.

$$IoU = \frac{Area\ of\ Intersection}{Area\ of\ Union}$$

$$mIoU = \frac{1}{N}\sum_{i=1}^{N} IoU_i$$

Another evaluation metric we plan to experiment is the dice score, also known as **D**ice **S**imilarity **C**oefficient (**DSC**). It is a statistical metric used to evalucate the similarity between two sets or samples, and is also commonly used in image segmentation to quantify the performance of a model. Like IoU, the dice score also ranges between 0 and 1, with a score of 1 signifying a perfect match between the predicted segmentation and the ground truth. However, the distinction lies in their sensitivity: IoU tends to penalize both under- and over-segmentation more harshly compared to the DSC.

$$Dice\ Score = \frac{2 * Area\ of\ Overlap}{Total\ Area}$$

Potential Risks

What are the risks?

1. Semantic segmentation model can be computationally intensive, which requires extensive GPU resources, memory, as well as prolonged fine-tuning time.
2. We are also cnocerned about low mIoU (<0.5) and dice score, since significant domain difference between medical images and images from autonomous driving can lead to poor transferability of learned features.

References

[1] Dataset: *https://www.kaggle.com/datasets/paultimothymooney/chiu-2015/data*

[2] mm-segmentation library: *https://mmsegmentation.readthedocs.io/en/latest/*

[3] Long, J., Shelhamer, E., & Darrell, T. (2015, March 8). *Fully convolutional networks for semantic segmentation*. arXiv.org. https://arxiv.org/abs/1411.4038

[4] Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.