Project Design Phase-II

Third-Party API's

Date	06 May 2023
Team ID	NM2023TMID18418
Project Name	CancerVision: Advanced Breast Cancer Prediction with Deep Learning

The description of breast cancer requires interpretation of the complex and rich clinical information provided by breast imaging from the macroscopic level to the microscopic level. With the fast increase in medical data scale and the development of imaging technology, analyzing large-scale highdimensional breast images with artificial intelligence (AI) holds great promise in improving the accuracy and efficiency of clinical procedures. Current AI is typically represented by deep learning (DL), which has made remarkable achievements over the past decade and has been widely adopted in various fields such as image or speech recognition (LeCun et al., 2015). Compared with conventional computer-aided diagnosis techniques that rely on hand-engineered features.

Deep Learning Methods for Breast Cancer Analysis :

This section will introduce the major deep learning techniques used in breast cancer imaging. For a more detailed review of deep learning, we refer the readers to Goodfellow et al. (2016). We will first introduce the formulations and some majorly used deep learning models by categorizing breast cancer image analysis into three basic tasks, i.e., classification, detection, and segmentation, according to the output types. We will then introduce the widely applied deep learning paradigms, in cluding supervised learning, semi-supervised learning, weaklysupervised learning, unsupervised learning, transfer learning, and multimodal learning

Classification :

Classification aims to give discrete predictions to categorize the whole inputs, e.g., 1 to indicate that a breast image contains cancer and 0 to indicate that the image does not contain cancer. A classification model can be regarded as a mapping function $f : X \longrightarrow Y$, where X is the domain of images or features and $Y \in R$ is usually a one-hot representation of the disease existence. Formally, given x an input, y the target output, and ˆy the model output, the classification models are typically optimized by minimizing the cross entropy between ˆy and y:

$$L = -y\log\hat{}$$

To model f , earlier studies would utilize artificial neural networks (ANNs) (McCulloch and Pitts, 1943) that are constructed by several fully-connected layers and take as input hand-crafted features. Convolutional neural network (CNN) (LeCun et al., 1989) gets rid of feature engineering and makes the classification problem on images fully end-to-end. In 2012, the success of AlexNet (Krizhevsky et al., 2012), a 5-layer CNN powered by graphic processing unit (GPU), kicked off the era of deep learning with its outstripping performance on the ImageNet challenge (Deng et al., 2009). VGG (Simonyan and Zisserman, 2014) extended the depth of CNNs with smaller kernels and auxiliary losses. Residual networks (ResNet) (He et al., 2016) further deepened CNNs to hundreds of layers and conquered the gradient vanishing problem with skip connections. Apart from AlexNet, VGG, and ResNet, many other networks like Densely Connected Network (DenseNet) (Huang et al., 2017) and the Inception series (Szegedy et al., 2015, 2016, 2017) have all been widely used in breast cancer imaging.

Recently, vision transformer (Dosovitskiy et al., 2020), a type of deep neural networks that are mostly based on attention mechanism (Vaswani et al., 2017), has also shown great potential in image processing.

Deep Learning Paradigms :

There are diverse options of deep learning paradigms to apply the models to different scenarios, given the availability of the data and labels. Supervised learning requires all training samples to be labeled exactly in the form of targeted outputs, e.g., masks for the segmentation task or bounding boxes for the detection task. Supervised learning is the most common form of deep learning, and a large proportion of studies reviewed in this paper fall into this category. However, deep learning is notoriously data-hungry and labeling medical images is time-consuming, and expertise-depending. Hence, supervised learning may not be the optimal solution for many practical medical image analysis scenarios. Weakly-supervised learning (WSL) is applied when the given label is not in the format of the targeted output. For example, using image-level annotations for detection or segmentation. In breast cancer imaging, the mostly used weaklysupervised learning methods are class activation map (CAM) (Zhou et al., 2016; Selvaraju et al., 2017) and multiple instance learning (MIL) (Carbonneau et al., 2018). CAM is often used for rough detection of targeted lesions, which is computed as the feature maps weighted by corresponding gradients. Higher values on a CAM indicate the regions that contributes more to the final prediction. MIL treats an input image as a bag of instances (i.e., image patches) which is negative only when all instances are negative. The goal of MIL is often to develop a bag-level classifier, which is quite a common strategy in processing whole slide pathology images which are of giga-pixel scale. Like CAM, MIL can also be used to roughly localize the lesions by highlighting the mostly contributed instances. Semi-supervised learning (SSL) can be regarded as another type of WSL, which enables utilizing a large amount of unlabeled data together with limited labeled data. Typical SSL methods are based on graph, entropy minimization, pseudo labeling, generative modeling, or consistency learning. Recently, consistency-based approaches have shown great success in SSL, which inject a regularization on the model that the predictions on different perturbated versions of a model should remain consistent. Unsupervised learning leverages unlabeled data for model training, often aiming at clustering or dimension reduction. In the literature of deep learning-based breast cancer image analysis, two major directions in unsupervised learning have gained research attention: generative modeling and self-supervised learning. The former uses generative methods, such as the generative adversarial network (GAN) (Goodfellow et al., 2020) or the variational auto-encoder (VAE) (Kingma et al., 2019) to model the data distribution and generate new samples, which is also quite often used in SSL. Self-supervised learning trains a neural network on the unlabeled images to learn representations for the supervised downstream tasks (Jing and Tian, 2020).