

Team Project
CV-Group-12
16, February, 2022

1 Exercises

1.1 Loss function for segmentation model

The main task of a deterministic segmentation model is to classify every single pixel of images. Thus, the loss function can be the averaged binary cross entropy loss of every pixel.

Furthermore, some loss weights can be added to optimizer the loss function if some specific pixels should be attended more. In this way, the averaged performance is not cared but the model will perform extra well for the specific pixels.

Standard image classification is really similar to deterministic segmentation model because the task of these two models is same. The only difference is that standard images classification aims to classify the whole images to a certain category, whilst segmentation model tends to classify every single pixel of input images.

1.2 Difference between aleatoric and epistemic uncertainty

Aleatoric uncertainty, also known as data uncertainty, roots in the noise of input data, whereas epistemic uncertainty, also called model uncertainty, comes from missing training data.

Aleatoric and epistemic uncertainty describe confidence in the input data and the confidence of the prediction respectively. Epistemic uncertainty can be reduced by adding more training data, but aleatoric uncertainty cannot.

Unlike aleatoric uncertainty can be directly learned using neural network, epistemic uncertainty is much more challenging to be estimated and one possible solution is Bayesian neural network.

1.3 Dropout

For Bayesian neural network, every weight is stochastic rather than deterministic and obeys certain probability distribution. Instead of running the network once, run the neural network T times during testing. Even though the input is the same at each time we run the network, different results will be predicted given that weights are random variables and T different sets of weights $\{W\}_{t=1}^T$ are generated. The averaged results is regarded as the final output and the variance is a measure of epistemic uncertainty. The corresponding mathematical formula is shown below:

$$E(\hat{Y}|X) = \frac{1}{T} \sum_{t=1}^T f(X|W_t)$$
$$Var(\hat{Y}|X) = \frac{1}{T} \sum_{t=1}^T f(x)^2 - E(\hat{Y}|X)^2$$

The higher variance is, the larger epistemic uncertainty will be.