Camelyon17 challenge

Nicolas Pinchaud

nicolas.pinchaud@gmail.com

Model

In the CAMELYON17 challenge the objective is to make a patient level prediction based on information from several whole slide images. Our proposed method has two stages. First the tumor segmentation for each slide is computed using a deep convolutional neural network. Secondly, geometrical properties from these segmentation maps are extracted and used as features for a classification model predicting the slide level metastasis grading. Patient level gradings are finally inferred according to the rules.

We propose an ensemble approach: we combine several segmentation models learned on different pixel resolutions in a directed acyclic graph (DAG) structure as described in (1) (figure 1). The segmentation model being a pixel wise classification model (as to compare with a patch wise classification model), its output and input are in the same domain, this allows to combine different models together by concatenating the output of one (or several) model to the input of another.

Since we can learn a model on any slide level resolution, this architecture allows integrate the information available from these different levels. It also has the benefits of ensemble learning, we can learn the individual model with different strategies or hyper-parameters inducing different modeling expressiveness that can be integrated in their combination. Another advantage of such approach is that the resources spent to train the models are cumulative. The classical non ensemble approach would spend resources to learn different models in order to choose the best hypothesis among different architectures or hyper parameter sets, and then choose the best candidate and disregard the others. This is resource wise expensive because discarded models only contribute in the choice of the best candidate. In our ensemble approach the models are combined and they contribute by providing statistical information about the joint distribution between pixels and annotation data to another model.

Our approach uses a composition of three Deeplab(2) models trained on three different resolutions: 0.5, 1 and 2 micrometer per pixels (mpp). The model is trained to segment cancerous areas at pixel resolution. The segmentation is then used to predict the slide's largest tumor type (Normal, ITC, Micro or Macro) using a random forest on hand-crafted features. Finally, the pN-stage for each patient is computed according to the rules.

Training the segmentation model

Dataset. We used all the annotated slides from camelyon16 and camelyon17 for training.

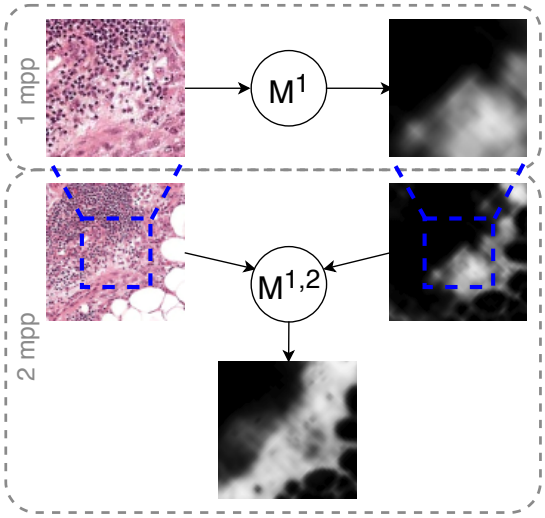

Fig. 1. An illustration of our model compounding principle. A first model M^1 is trained on 1 mpp. A second model $M^{1,2}$ is trained on 2 mpp by taking the output of M^1 as an extra channel. This principle can be extended to compose any number of model in a directed acyclic graph structure.

Augmentation. We augmented the data using rotation, color jittering and elastic deformation as shown in figure **??**. This was made on the fly during training, on a separate GPU, and allowed to train the model on a virtually infinite source of data variation.

Quasi Online Hard Example Mining. The training pipeline is illustrated on Figure 2. It has two processes running synchronously and in parallel. This allows to perform *quasi online hard example mining* on two levels. First, on the WSI level, the most difficult slides are sampled more frequently. Secondly, within a slide, the model is trained from patches extracted from the most difficult regions.

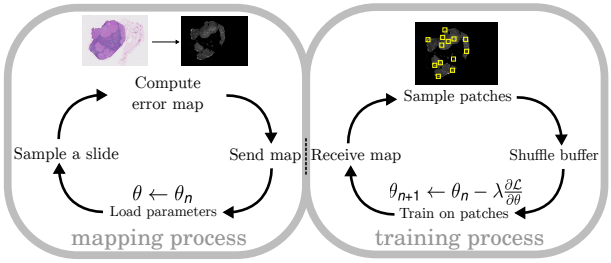

Fig. 2. Dynamic sampling cycle. A training process samples patches using an error map that indicates the regions where the model makes the most mistakes. Error maps are computed and provided by a mapping process. Both processes work synchronously and in parallel.

Slide sampling. The *mapping process* (Figure 2) chooses the next slide to sample patches from. This is done using a slide level sampling distribution that gives more probability on