Contribution to the ChAImeleon Challenge: Multiple Instance Learning with EfficientNet

Marek Wodziński[1,2]

[1]AGH University of Kraków, Department of Measurement and Electronics, Kraków, Poland

[2]HES-SO Valais-Wallis, Institute of Informatics, Sierre, Switzerland

Motivation:

The method design was motivated by the following factors:

1) The number of training cases was quite limited. The subtasks like e.g. prostate grading, survival prediction, or histology subtype classification directly from radiologic images are extremely difficult. Therefore, the proposed method had to extract good features.
2) It was allowed to use open-source pretrained networks, however, disallowed to use external data.
3) The computational resources were limited. Therefore, it was crucial to propose lightweight models allowing efficient training.

Method:

All the factors above led to solution based on pretrained open-source model and fine-tune it on the provided data using the ChAImeleon platform. Since the number of open source foundation models for 3-D radiology is limited, it was decided to use the open-source EfficientNetV2 (with frozen weights) from TorchVision library and use it in the Multiple Instance Learning (MIL) framework, extracting embedding from each slice and learning the associated weights by learnable attention layers.

Input to the model was normalized to [0-1] range for each subtasks, and additionally resampled to 384x384x128 for the Lung subtask. Then, the inference transforms from EfficientNetV2 were applied to each input. The slices from transversal plane were permuted to the batch dimension and the input was considered as 2-D. The slices were forwarded through the EfficientNetV2 encoder, and then concatenated with the raw features available in the files metadata (e.g. PSA for Prostate, CEA for Rectum). Finally, all the features were forwarded through fully connected layer to calculate the final result. The method is summarized in Figure 1.

Results:

The method achieved final score equal to 0.6353 for the Prostate subtask, 0.6591 for the Lung subtask, and 0.4381 for the Rectum subtask. The final results of Rectum subtask was so bad due to software error that is corrected in the associated repository.

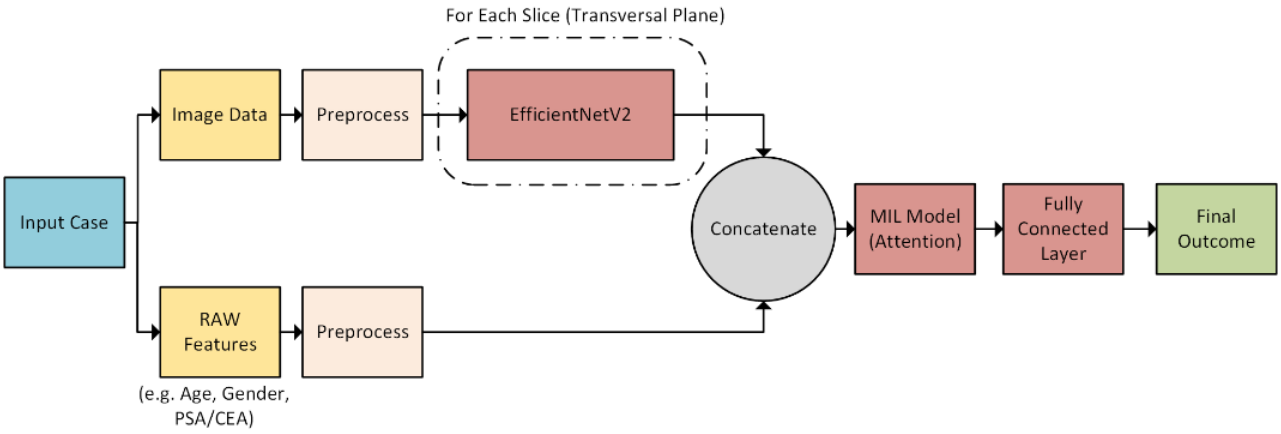

Figure 1. The proposed processing pipeline allowing one to train models with limited amount of training data and computational resources.