Sharing is Caring: Concurrent Interactive Segmentation and Model Training using a Joint Model

Ivan MIKHAILOV1,2, Benoît CHAUVEAU2,3, Nicolas BOURDEL1,2,3 and Adrien BARTOLI1,2,3
1Université Clermont Auvergne 2SURGAR SAS 3CHU de Clermont-Ferrand

Classical Data Pipeline

The base supervised machine learning paradigm has three main steps: (1) data annotation, (2) target predictor training, (3) deployment.

1 Annotation

2 Training

3 Deployment

Figure 1. General classical deep learning data annotation and model deployment pipeline.

Archetypical Framework

Archetypical state-of-the-art systems do two preparatory steps to expedite the annotation process: (1) data selection to reduce the quantity of data to annotate; (2) usage of an improvable annotation suggester to pre-annotate the data. They still use classical manual tools for corrections.

Figure 2. Archetypical state of the art framework.

Limitations of the Classical Data Pipeline

Existing systems are suboptimal: (1) annotations are required to produce annotations, - predictor is trained only once a sufficient amount of data has been annotated; (2) data annotation, data selection and target predictor training require different isolated methods and expertise to achieve the same goal - produce annotations; (3) the annotation tools do not improve as more data is annotated.

Figure 3. Two main limitations of the classical data annotation and model training pipeline.

Experimental Results: Metrics

Evaluation: three emulated and one real annotation scenarios.

- Emulated scenarios: (1) in female pelvis MRI segmentation on our dataset; (2) in Liver and (3) in Pancreas CT segmentation on decathlon datasets. Emulated setup: (a) completely annotated datasets; (b) user interaction simulation in order to generate annotations.
- Real scenario: kidney MRI segmentation on the AMOS dataset.

Figure 4. SAIM experimental evaluation results given as mIoU at each iteration, where: green - performance using annotations created in 3D Slicer and MITK, blue - performance using SAIM-created annotations. Dataset color code: Female pelvis MRI, AMOS Kidney MRI, Decathlon Liver CT, Decathlon Pancreas CT.

Figure 5. Two key existing ideas required to realize SAIM.

Figure 6. The intuition behind SAIM: use of deep interactive annotation with the shared model and the use of active learning for efficient data selection using the shared model. The shared model improves at each iteration.

Figure 7. Detailed SAIM schematic: a single iteration is shown, only the input changes between iterations.

Figure 8. SAIM segmentation results in an emulated annotation scenario for female pelvis MRI segmentation dataset, where: uterus - green, bladder - yellow, tumours - red, ovary cavity - pink, and user clicks are in cyan: (a) ground truth; (b) interactive predictor pre-trained on 15% (12 out of 77 series) of the complete training set; (c) the same interactive predictor using the complete training set (77 series), but with the remaining 85% (65 out of 77 series) annotated as a part of SAIM. Performance-wise: rows in green (1-5) - improvement, rows in red (6) - considered a degradation despite overall higher IoU due to the false positive for uterus. Kid is given in the bottom right corner.

Figure 9. SAIM segmentation results in a real annotation scenario for kidney MRI segmentation on the AMOS dataset, where right kidney and left kidney are orange and yellow respectively with user clicks in cyan: (a) ground truth; (b) interactive predictor pre-trained on original AMOS training set (100 series) and (c) the same interactive predictor after doubling the training set as a part of SAIM (90 series). Performance-wise: rows in green (1-4) - improvement, rows in red (5-6) - degradation. Kid is given in the bottom right corner.