ADAGSS: Automatic Dataset Generation for Semantic Segmentation

L. Palladino, B. Maris, P. Fiorini
University of Verona, Computer Science, Verona, Italy

Keywords Dataset Creation, Semantic Segmentation, Image Classification

Purpose

A common issue in medical deep learning research is the creation of dataset for training the neural networks. Medical data collection is also tied-up by privacy laws and even if a lot of medical data are available, often their elaboration can be time demanding. This problem can be avoided using neural networks architectures that can achieve a good predicting precision with few images (e.g. U-Net). In the case of semantic segmentation, the dataset generation is even more cumbersome since it requires the creation of segmentation masks manually. Some automatic ground-truth creation techniques may be employed like filtering, thresholding and Self Organized Maps[1] (SOM).

These automatic methods can be very powerful and useful, but they always have a bottle-neck phase: data validation. Due to algorithm reliability (that sometimes can fail), data needs to be validated manually before they can be included in a dataset for training. In this work, we propose a method to automatize this phase by moving manual intervention to an easier task: instead of creating masks and then validate them manually, we train a convolutional neural network to classify segmentation quality. Therefore, the validation is performed automatically. An initial manual phase is still required, but the classification task requires a smaller number of elements in the dataset that will feed a network employed for classification. After this phase, similar dataset creations will require less effort. This procedure is based on the fact that to obtain a high classification precision, fewer data are required than the data that are needed to obtain high precision in semantic segmentation. High classification score, can automatize validation procedure in dataset creation, being able to discard failure case in dataset creation. Being able to produce bigger dataset in less time can led to higher precision in semantic segmentation.

Methods

Automatic dataset generation for semantic segmentation (ADAGSS) method is divided in steps. Although this method can be applied generally in every segmentation situation, our test case is the segmentation of the prostate in ultrasound images taken from a multimodal phantom (CIRS 070, Computerized Imaging Reference Systems, Inc. (CIRS), Norfolk, Virginia, US). We tested the method on coronal images acquired with a convex probe and on sagittal images taken with a linear probe. Phase one consists in the segmentation of the dataset. Segmentation is achieved by applying the following methods: spatial filtering (median, top-hat and closing), thresholding (threshold depends on data and can be tuned manually). Threshold operation produces a preliminary binarization version of the input image. At this point, binarization is refined by feeding binarized images to a Self-Organizing Map (SOM) network that perform pixel-clustering, refining binarization boundaries. After this, morphological transformations such as opening and closing are applied. Final step is de-blobbing in which eventual outliers (blobs of pixel residual of segmentation) are removed. De-blobbing is performed applying the heuristic that assumes that prostate blob is the most centered one. Blob centroid distance from the center of the image is computed and only the most central one remains. After this step, segmentation is done. If fine-tuned, this method grants a satisfying result in most of the samples. Sometimes, noise is too much and prevent segmentation to produce precise masks and/or sometimes de-blobbing lead to blank images.

After segmentation has been made, a preliminary manual classification is required. Using a user interface with 3 buttons (Figure 2) it’s possible to manually classify segmentation masks in three semantic classes:

valid segmentations, segmentation to be discarded (e.g. blank ones most of the times) and fixable segmentation. Fixable segmentations are, in proportion, very less than the previous two: they are the segmentation masks that are not correct, but with a very quick and easy manual intervention can be made "valid" and added to the dataset. An example is when de-blobbing phase fails, and a blob still has to be deleted.

This manual classification is the base to a second dataset of labeled images which will be used to train a second neural network. The second network is a classifier and will be trained to classify segmentations mask, going to automatize the manual process previously described. The classifier has a typical convolutional neural network structure for classification task (Figure 1). The last layer of the network has a softmax function for the classification of the input in three classes: valid, fixable or discarded. After having achieved a classification accuracy that is high enough (> 90%), dataset generation can be fully automatized by applying the classifier to validate generated data. After an initial effort, automatic dataset creation can lead to far more big dataset generations than the one achieved by manual intervention.

Results

We've applied this method in 2 case scenarios: one with a high number of training data (around 75% of the total) and one with only a few training images (around 2.5% of the total). The first one achieved 92% precision in classification task of ground-truth masks classification of 908 coronal plane images. Network has been trained on a dataset of 3630 labeled images. By using data augmentation, the result reach an accuracy of 96%. The second one achieved 79% precision (82% with some data-augmentation) in classifying 3630 coronal plane images. The network has been trained on 100 manually classified images.

No. of training data	No. of validation data	Precision	With data augmentation
3630	908	92%	96%
3630	100	79%	82%

Conclusion

The use of neural networks for medical image segmentation is a powerful tool but its use is limited by the necessity to have big dataset as ground-truth. The creation of the ground-truth is most of the time manual. Automatic tools are also available, and they are based on classical image processing algorithms. Their result requires always manual supervision, which is again time consuming for large sets of data. We propose here a convolutional network to supervise the automatic generation of ground-truth data. The results of this method (automatic segmentation plus automatic validation) can be itself a segmentation method, but, used in conjunction with a semantic segmentation neural network (e.g. U-net) can be even a more powerful tool, considering that the neural network allows for a more accurate segmentation obtained in less time.

We proved here that with a low number of input images for a classification neural network (2.5% of the total) we can achieve a high volume of data validated as ground-truth dataset for a semantic segmentation neural network. Procedure can be re-iterated until a higher precision is reached.

References

[1] Zaim, A.. (2005). Automatic Segmentation of the Prostate from Ultrasound Data Using Feature-Based Self Organizing Map. Lecture Notes in Computer Science. 3540. 1259-1265. 10.1007/11499145_127.

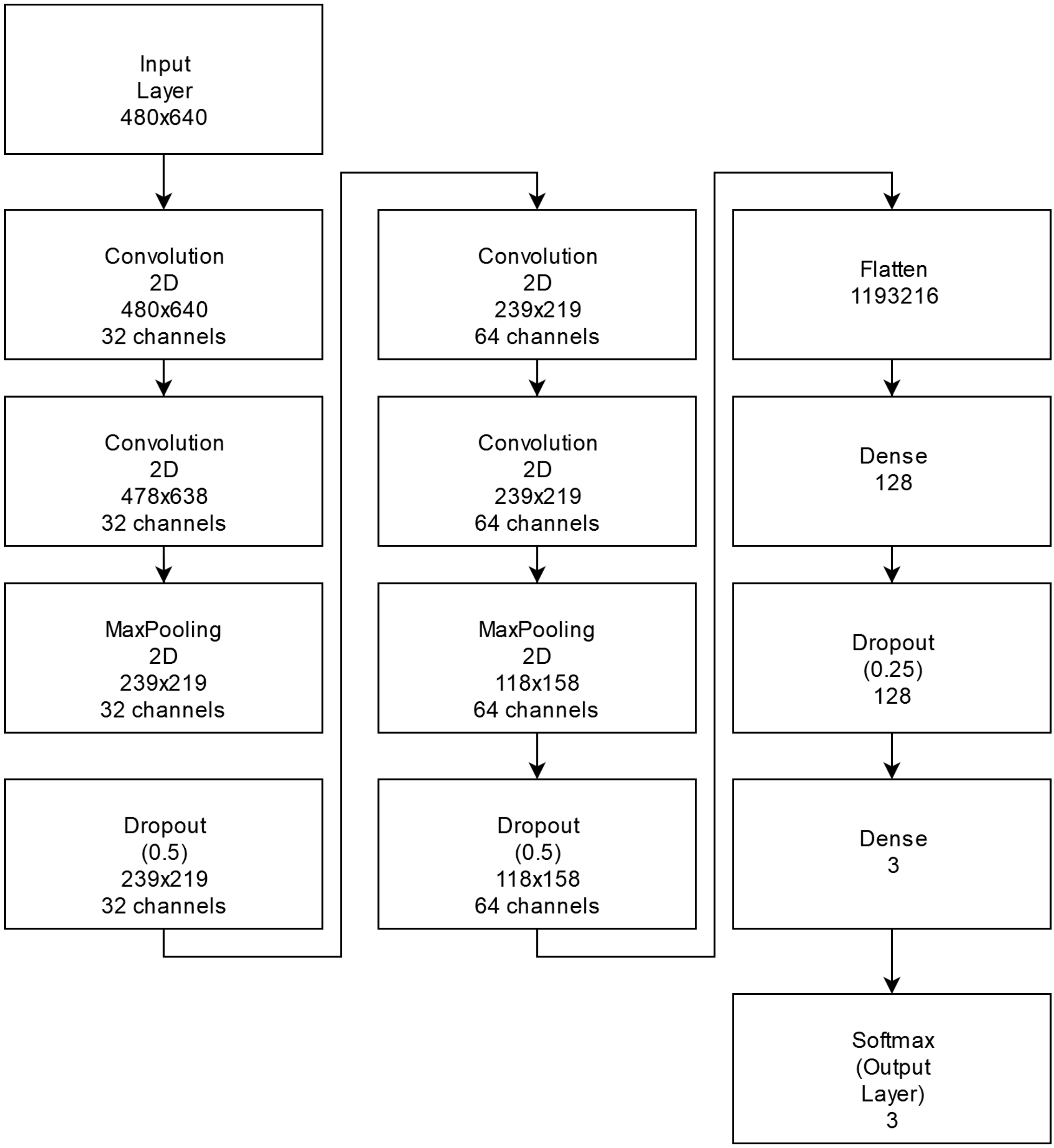

Figure 1: Classifier Architecture

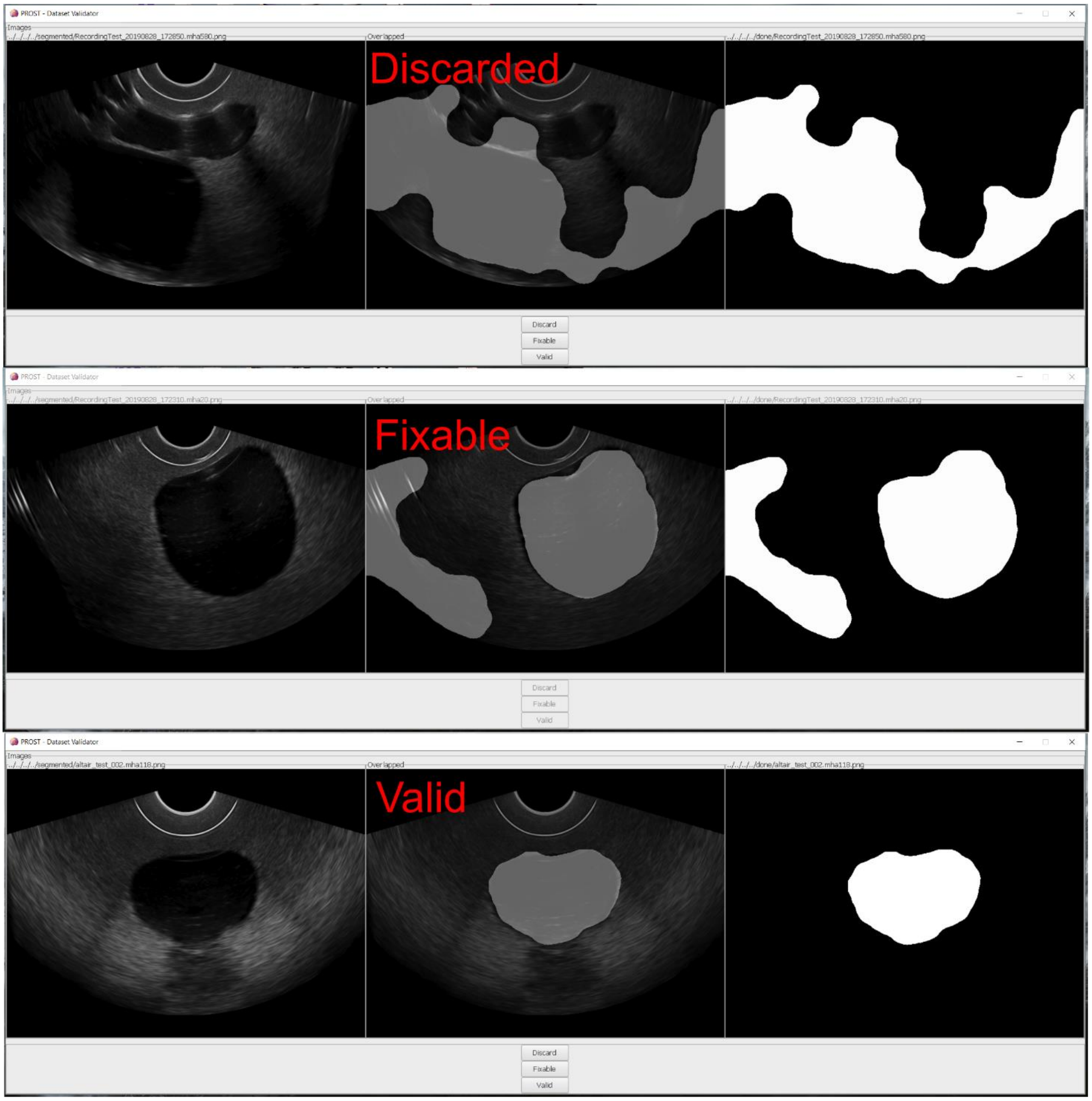

Figure 2: Manual validation interface. In order: Valid, Fixable and Discarded case.