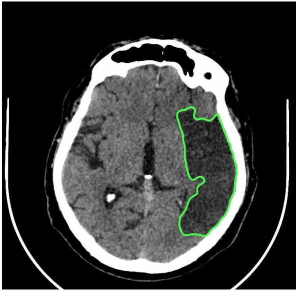
Training of deep neural networks with incomplete training information on the example of recognition of tomographic images

Kalmutskiy Kirill   Vladimir Berikov

May 2021

Formulation of the problem



A large dataset of tomographic images of the brain is given. Each image is a section of the brain. Manual marking of such a dataset is very expensive, since only a highly qualified specialist can do this.

In this regard, it is necessary to create a model that will predict areas of stroke and select them on tomographic images of the brain. The model will be used as a "decision support system."

Complexity of the problem and scientific novelty

Complexity

- Very little labeled data available: only 34 images, which is clearly not enough for training classical deep convolutional neural networks;
- It is necessary not only to properly segment the image, but also to have high recognition accuracy;
- The model should be interpretable, otherwise doctors will not understand its decisions and will not trust it.

Scientific novelty

- All current solutions require a huge amount of training data;
- The interpretability of neural networks is poorly understood.

Initial data

- Images in DICOM format, labels in JPG format;
- Only 34 labeled images of brain slices;
- More than 30 GB of unlabeled images.

Segmentation task

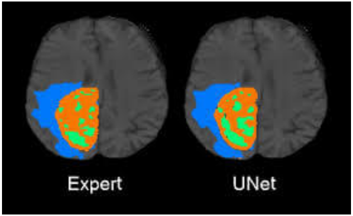
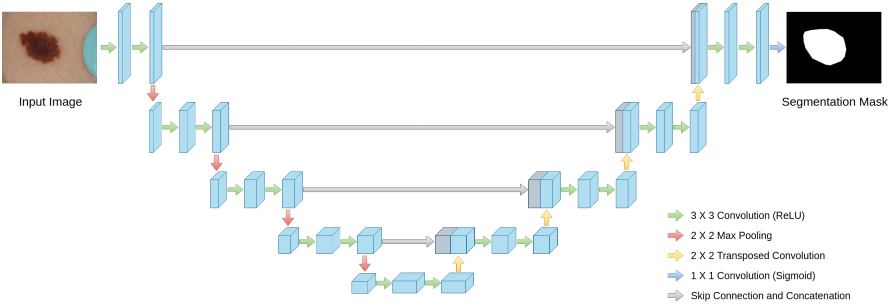


Image segmentation is the process of partitioning a digital image into multiple segments (sets of pixels, also known as image objects).

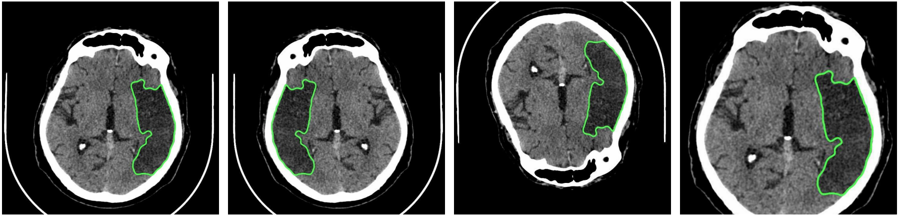
- **Automatic segmentation**: bad accuracy, doesn't work in medical image recognition tasks;
- **Semi-automatic segmentation**: requires the participation of a specialist during segmentation process, poorly suited for creating a decision-making system;
- **Trainable segmentation**: the most promising approach.

Trainable segmentation



- The model allows you to achieve high metrics;
- A lot of data is needed for training;
- Transfer learning doesn't work well for medical image recognition tasks;
- Inability to interpret model decisions.

Augmentations

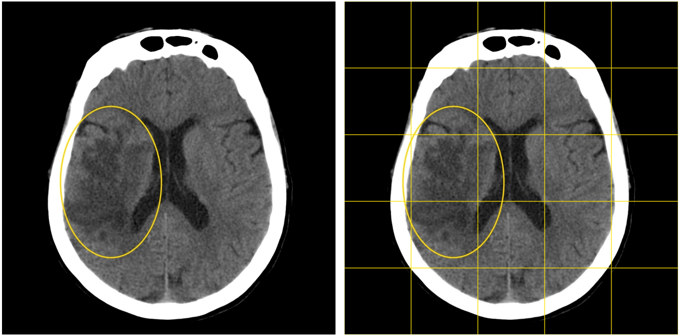


The only augmentations that have improved the quality:

- horizontal flips;
- vertical flips;
- small zoom;

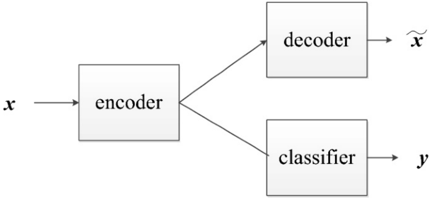
Instead of a random crop, which is a very popular augmentation in such problems, a slightly different approach was used.

Cutting the image into sub images



- Learning on crops smaller than the original image;
- To predict, we cut the picture into crop again, predict for them and collect the image back;
- We increase the dataset 16 times using a 4x4 grid.

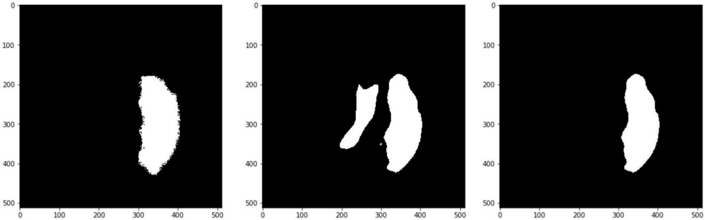
Network Architecture



To improve the accuracy of stroke diagnosis, the model includes two parts: a segmentator and a classifier. Thus, segmentation is performed only if the classifier reveals a stroke in the image.

- **Segmentator**: an architecture including an encoder and a decoder with skip-connections and higher dilation parameter;
- **Classifier**: a linear layer with sigmoid at the output of the encoder;
- **Loss function**: sum of the weighted cross-entropy of the classifier and the weighted cross-entropy of the segmentator.

Model Prediction Processing



Left image - ground truth, middle image - all components, right image - only the maximum connected component

- **Using TTA**: augmentations (horizontal flips) are applied not only during training, but also during prediction, then the predictions for the images are averaged;
- To increase the quality of segmentation, the model **leaves only the maximum connected component** in the resulting segmented image;
- In addition to this, **the remaining component is removed if it is less than a certain size** (currently less than 1200 pixels).

Results of Experiments

Table: Validation and Test Results

Metric	Validation	Test
Accuracy	0.922	0.890
IoU	0.727	0.675

- **Validation**: 8 (4+4) images (128 sub images);
- **Test**: 6 images (92 sub images).

Results of Experiments

Table: Comparison of the proposed algorithm with the baseline

Approach	IoU	FPR
U-Net (baseline)	0.541	1
Proposed algorithm	0.727	0.2

Conclusion

We managed to build a model for the diagnosis of stroke with good quality, having only 34 labeled images. To do this, we used various techniques:

- cutting the image into sub images;
- a convolutional neural network with a high dilation parameter, consisting of a segmenter and a classifier;
- processing model predictions by cutting off small components of connectivity.

To increase the interpretability of the model, it is assumed in the future to use well-interpreted algorithms, such as: decision tree, k-nearest neighbors, linear regression etc instead of the classifier layer.

Bibliography

Yadav, S.S., Jadhav, S.M. "Deep convolutional neural network based medical image classification for disease diagnosis," J Big Data 6, 113, 2019.

J. Seetha, and S.S. Raja "Brain tumor classification using convolutional neural networks," Biomedical & Pharmacology Journal, 11(3), pp.1457– 1461. 2018.

M. Polsinelli, L. Cinque, G. Placidi, "A Light CNN for detecting COVID-19 from CT scans of the chest," arXiv, 2004.12837, 2020

S.M. Abulnaga, and J. Rubin "Ischemic stroke lesion segmentation in CT perfusion scans using pyramid pooling and focal loss," in International MICCAI Brainlesion Workshop, pp. 352–363. Springer, Cham. September 2018.

H. Kuang, B.K. Menon, and W. Qiu "Segmenting hemorrhagic and ischemic infarct simultaneously from follow-up non-contrast CT images in patients with acute ischemic stroke," IEEE Access, 7, pp.39842– 39851. March 2019.

M. Raghu and M. Zhang and J. Kleinberg and S. Bengio "Transfusion: Understanding Transfer Learning for Medical Imaging," arXiv, 1902.07208, 2019

I. Rizwan I Haque, J. Neubert "Deep learning approaches to biomedical image segmentation, Informatics in Medicine Unlocked," ISSN 2352-9148, 2020.

H. Du, R. LaLonde, R. van Mechelen, S. Zhang "Performing Semantic Segmentation on an Extremely Small Dataset," MICS, 2016.