The Importance of Health Informatics in Public Health during a Pandemic
J. Mantas et al. (Eds.)

doi:10.3233/SHTI200572

Classification of Intracranial Hemorrhage Subtypes Using Deep Learning on CT Scans

Gleb DANILOV [a,1], Konstantin KOTIK [a], Anna NEGREEVA [b], Tatiana TSUKANOVA [a], Michael SHIFRIN [a], Natalya ZAKHAROVA [b], Artem BATALOV [a], Igor PRONIN [a] and Alexander POTAPOV [a]

[a] *Laboratory of Biomedical Informatics and Artificial Intelligence, National Medical Research Center for Neurosurgery named after N.N. Burdenko, Moscow, Russian Federation*

[b] *Department of neuroradiology, National Medical Research Center for Neurosurgery named after N.N. Burdenko, Moscow, Russian Federation*

Abstract. Intracranial hemorrhage is a pathological condition that requires fast diagnosis and decision making. Recently, a neural network model for classification of different intracranial hemorrhage types was proposed by a member of our research group Konstantin Kotik as part of the machine learning competition at Kaggle. Our current pilot study aimed to test this model on real-world CT scans from patients with intracranial hemorrhage treated at N.N. Burdenko Neurosurgery Center. The deep learning model for intracranial hemorrhage classification based on ResNexT architecture showed an accuracy of detection greater than 0.81 for every subtype of hemorrhage without any tuning. We expect further improvement in the model performance.

Keywords. Intracranial hemorrhage, deep learning, neurosurgery, computed tomography

1. Introduction

Intracranial hemorrhage is a pathological condition that requires fast diagnosis and decision making [1]. The common subtypes of intracranial hemorrhage are subarachnoid hemorrhage (SAH), intraparenchymal hemorrhage (IPH), subdural hemorrhage (SDH), epidural hemorrhage (EDH) and intraventricular hemorrhage (IVH). Traumatic brain injury, cerebrovascular pathology, arterial hypertension, and even surgical intervention may be complicated by any type of bleeding or their combination. Computed tomography (CT) is a fast and reliable diagnostic tool for acute hemorrhage detection. The automation of hemorrhage detection in CT scans with computer-assisted technologies might be helpful in objectifying findings and speeding up the decision on treatment. However, the amount of neuroimaging data available for the development of these solutions is commonly limited [1].

[1] Corresponding Author, Gleb Danilov, N.N. Burdenko Neurosurgery Center, Moscow, 4th Tverskaya-Yamskaya str. 16, 125047, Russian Federation; E-mail: glebda@yandex.ru.

Recently, a neural network model for classification of different intracranial hemorrhage types was proposed by a member of our research group Konstantin Kotik as part of the machine learning competition at Kaggle (31st place out of 1345 possible, https://www.kaggle.com/c/rsna-intracranial-hemorrhage-detection/overview). The aim of our pilot study was to test this model on real-world consecutive CT scans from patients with intracranial hemorrhage treated at N.N. Burdenko Neurosurgery Center.

2. Methods

2.1. Model description

The solution introduced for the Kaggle competition was based on ResNexT model - a simple, highly modularized network architecture for image classification [2]. A special instance of ResNexT (se_resnext50_32x4d) was used with Adam optimizer and multi-label logarithmic loss function (negative log-likelihood function for several classes). "SE" in the model name stands for the "Squeeze-and-Excitation" block, which adaptively recalibrates channel-wise feature responses by explicitly modeling interdependencies between channels. Stacking these blocks together enables the construction of architectures that have only a few hyper-parameters to set and generalize well across challenging datasets. This strategy exposed a new dimension called "cardinality" (the size of the set of transformations), as an essential factor in addition to the dimensions of level and width. The final model was designed as the average of two models stacked and fitted during 3 epochs.

In the Kaggle competition, this solution was developed on a training dataset of 674.258 labeled DICOM images (CT scans). The data were preprocessed (corrupted Rescale Intercept Attribute of DICOM file set to -1000, images of 512x512 size rescaled), and appropriate window level (40 for brain, 80 for blood, 40 for soft tissues) and window width (80 for brain, 200 for blood and 380 for soft tissues) were applied. The image augmentations (horizontal flip, random crop, brightness, contrast, random rotate within 30 degrees) and test-time augmentation were used. The model was tested in a private dataset of 752.807 DICOM files resulted in a weighted multi-label logarithmic loss equal to 0.05098.

2.2. Real-world clinical dataset

A dataset of 300 CT series (32-64 images in each) from patients with hemorrhage and 101 CT series with patients with no intracranial bleeding and minor intracranial pathology were consecutively selected from the PASC archive of N.N. Burdenko Neurosurgery Center. The initial consecutive selection of data for screening was performed by a data engineer using context search (with keywords related to hemorrhage defined by a radiologist) to exclude any selection bias related to the pathology preferences coming from doctors. One junior radiologist did the screening of CT scans under the supervision of a senior doctor. The output of expert screening was the classification of 401 CT series by five subtypes of intracranial hemorrhage: epidural, subdural, subarachnoid, intraparenchymal, intraventricular. The data scientists were blinded to the complete diagnosis, full image annotations, and other medical data until the end of the test. The numbers of positive (POS, series with hemorrhage) and negative (NEG, series with no hemorrhage) classes, accuracy (ACC), precision (PREC, also referred to as positive predictive value), recall (REC, also referred to as sensitivity), area under ROC-curve (AUC) and logarithmic loss (LOG_LOSS) were computed to evaluate model performance.

3. Results

The statistics on five hemorrhage subtypes classes in our dataset and quality metrics of model predictions are shown in Table 1.

Table 1. The quality metrics of classification for different types of intracranial hemorrhage.

Hemorrhage subtype	POS	NEG	ACC	PREC	REC	AUC	LOG_LOSS
Epidural	99 (25%)	302 (75%)	0.828	0.660	0.626	0.762	0.849
Subdural	83 (21%)	318 (79%)	0.818	0.566	0.518	0.711	0.896
Subarachnoid	118 (29%)	283 (71%)	0.820	0.829	0.492	0.748	0.881
Intraventricular	137 (34%)	264 (66%)	0.893	0.952	0.723	0.804	0.890
Intraparenchymal	188 (47%)	213 (53%)	0.835	0.868	0.766	0.803	0.941

The overall quality of classification of each hemorrhage subtype is shown in Figure 1.

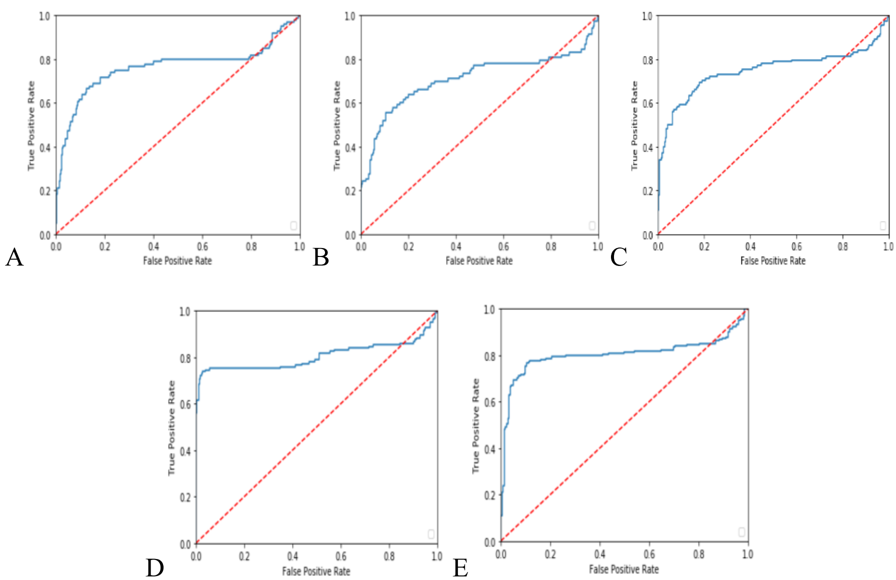

Figure 1. The relationship between true positive and false positive rate for different types of intracerebral hemorrhage: A – Epidural, B – Subdural, C – Subarachnoid, D – Intraventricular, E – Intraparenchymal

Hemorrhage was verified in 300 studies performed once or repeatedly in 260 patients (avg. age 41.8 ± 18.8 years, 148 (56.9%) men, 243 (81.0%) operated during the hospital stay). Seventeen studies were done in non-operated patients, 24 - before the day of the first surgery, 63 - on the surgery day and 195 - in the postoperative period. The primary diagnosis in most of the cases was related to brain tumors (n = 96), cerebrovascular pathology (aneurysms, arteriovenous malformations, hypertensive stroke, n=78), and traumatic brain injury (n = 69). Since the selection of images was made with broad inclusion criteria, the other lesions and artifacts were found on 141 (35.2%) CT scans in addition to hemorrhagic signals including ventricular shunts and drainages – in 60 (15.0%) cases, intracranial pressure sensors - in 15 (3.7%), artifacts from other metal devices – in 47 (11.7%), other artifacts – in 6 (1.25%), poor series quality – in 6 (1.5%) cases.

4. Discussion

We presented the direct pilot implementation of the ResNexT-based model for intracranial hemorrhage classification. We designed our experiment on the real-world dataset without special optimization to understand whether the model derived from Kaggle competition could be generalizable and demonstrate the efficiency in clinical routine, as well as explore the straightaway quality of such an "out-of-box" solution. In contrast to Kaggle data, our raw CT scans contained specific signals the model had probably never seen before. However, our first results are consistent with those from other similar studies (using Convolutional Neural Networks, Long-Short Term Memory Networks) in terms of precision (more than 0.66) and recall (more than 0.75) [1], [3]–[6]. There is still room for tuning the solution (windowing, test-time augmentation, dataset exploration), so an increase in its performance is expected.

5. Conclusions

The model for intracranial hemorrhage classification tested on real-world CT scans without any tuning showed an accuracy of detection greater than 0.81 for every subtype of hemorrhage. The further improvement of the model is expected.

References

[1] Ker J, Singh SP, Bai Y, Rao J, Lim T, Wang L. Image Thresholding Improves 3-Dimensional Convolutional Neural Network Diagnosis of Different Acute Brain Hemorrhages on Computed Tomography Scans. Sensors (Basel) 2019;19: 2167.

[2] Xie S, Girshick R, Dollár P, Tu Z, He K. Aggregated residual transformations for deep neural networks. Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017 2017; 5987–5995.

[3] Jnawali K, Arbabshirani MR, Rao N, Alpen AP. Deep 3D convolution neural network for CT brain hemorrhage classification. Proc.SPIE 2018: 10575.

[4] Padosch SA et al. Medicolegal aspects of witnessed suicide due to gunshot to the head. II. Legal medicine aspects and examination of the firing hand 2004;214: 149–162.

[5] Dou Q et al. Automatic Detection of Cerebral Microbleeds from MR Images via 3D Convolutional Neural Networks. IEEE Trans. Med. Imaging 35: 1182–1195.

[6] Grewal M, Srivastava MM, Kumar P, Varadarajan S. RADNET: Radiologist Level Accuracy using Deep Learning for HEMORRHAGE detection in CT Scans; 2017.