OBJECT DETECTION NETWORKS FOR LOCALIZATION AND CLASSIFICATION OF INTRACRANIAL HEMORRHAGES

Jakub Nemcek
Doctoral Degree Programme (1), FEEC BUT
E-mail: nemcek@vutbr.cz

Supervised by: Roman Jakubicek
E-mail: jakubicek@feec.vutbr.cz

Abstract: Intracranial hemorrhages represent life-threatening brain injuries. This paper presents two state-of-the-art object detection systems (Faster R-CNN and YOLO v2) which are trained to localize and classify hemorrhages in axial head CT slices by providing labelled rectangular bounding boxes. Publicly available datasets of head CT data and ground truth bounding boxes are used to evaluate and compare the performance of both detectors. The Faster R-CNN shows better results by achieving an average Jaccard coefficient of 58.7 %.

Keywords: intracranial hemorrhage, convolutional neural network, localization, classification, computed tomography, Faster R-CNN, YOLO v2

1 INTRODUCTION

Intracranial hemorrhages (ICHs) represent critical medical event as they belong to life-threatening brain injuries with high incidence (25 cases per 100,000 persons per year), which can originate in physical trauma or other non-traumatic causes (hemorrhagic stroke) [1]. According to its location, there are five types of ICH: intraparenchymal (IPH), intraventricular (IVH), subdural (SDH), epidural (EDH) and subarachnoid hemorrhage (SAH).

Significance of this pathology is given by the risk of lasting severe consequences and high mortality (40 %), hence early and accurate diagnosis is crucial [2]. The x-ray computed tomography (CT) is the most frequent ICH diagnostic method, however, the traditional examination of 3D CT data by radiologists is either time-consuming and requires high level of concentration. Nowadays, computer-aided diagnosis (CAD) systems are being developed to reduce the examination time and to prevent possible omissions in case of small or indistinct bleedings [3].

Novel deep learning methods have become state-of-the-art in many applications including image analysis. Several approaches for analysis of ICH in brain CT data using convolutional neural networks (CNN) have been recently published. In some papers [4, 5] simple 2D classification CNN is used for ICH detection and classification. Some authors use combination of 2D CNN architecture with other methods to extend the detection and classification results to whole 3D scan. In these cases, random forests [6] and recurrent neural networks (long short-term memory - LSTM) [3, 7] were used for the extension. More precise localization of ICH is done by segmentation. The authors in [8] proposed method using U-Net-like architecture, but the ICH type is not determined. A cascade of classification CNN ahead of a segmentation fully convolutional network was used in [9]. In [2] authors proposed mask R-CNN model for ICH segmentation. Although precise voxel-level segmentation might be favourable for a radiologist, training of such algorithms demands voxel-level annotations. Gaining of such annotations for sufficiently large dataset is, however, very time-consuming. In [10] a set of 2D CNN-based classifiers with cascade-parallel architecture was proposed, that enables ICH classification and localization by labelled 3D bounding box (BB).

In this paper, two automatic methods for the localization and classification of ICHs are proposed and compared. In comparison to other recently published methods, two detection networks are trained to predict labelled 2D BBs for a delineation of the bleedings and for the ICH types determination in axial slices, using publicly available datasets of images and annotations. Such an automatic method might lead to a significant decrease in the duration of the diagnostic process (manual detection can take more than 5 minutes) and false negative diagnostic conclusions.

2 METHODS

The combination of a CNN and a region of interest prediction network has recently proved to be an efficient solution to object detection [13, 14]. Thus two state-of-the-art detection networks are chosen to localize and classify ICHs in axial slices of a CT image.

2.1 EXPERIMENTAL DATA

The head CT data from the publicly available dataset CQ500 [6] was used with respect to several conditions: only non-contrast soft tissue kernel series with thick slices (3 mm or more) containing at least one type of hemorrhage were chosen (191 CT scans). The dataset is, however, imbalanced, considering the amount of data with different ICH types. In case of EDH, the amount is insufficient (only 13 scans). Manually extracted axial slice-level bounding box annotations marked by three neuroradiologists are obtained from BHX dataset [11] publicly available on the PhysioNet [12].

A simple pre-processing is applied on individual axial slices of the 3D image in the form of contrast enhancement with respect to three special radiological windows (i.e., brain, subdural and bone) to form three-channel input into the detection networks. The slices are resized to the size of 224×224 pixels.

2.2 DESIGN OF DETECTORS

Two modern detection network architectures are chosen for the purpose of ICH localization and classification: Faster R-CNN [13] and YOLO v2 [14]. Both networks consist of a feature extraction CNN and a detection part. Faster R-CNN uses a region proposal network (RPN) to generate region proposals from the feature space, that are then refined and classified [13]. YOLO v2 is a single neural network applied to the full image which divides an image into regions and then predicts BBs and classes for them [14]. Both architectures provide results in the form of labelled rectangular bounding boxes denoting the precise positions of the ICHs in the image.

Table 1: Training options for both detection networks: number of epochs, mini-batch size, L2 regulation factor and learning rate drop period with the factor of 0.1

	Epochs	**Mini-batch**	**L2 reg.**	**lr. drop**
Faster R-CNN	24	12	0.0005	15
YOLO v2	35	42	0.0005	30

2.3 IMPLEMENTATION DETAILS

The proposed methods were implemented in Matlab. For both detectors, ResNet-50 [15] pre-trained on the ImageNet database was used as a feature extraction CNN. The training dataset consisted of axial slices from 80 % of the available CT scans. An augmentation of the training images is applied in the form of slight rotation, translation, crop, scaling, reflection, and random noise adding. The Adam optimizer [16] with an initial learning rate of 0.001 ($\beta_1 = 0.9$ and $\beta_2 = 0.999$) was used for training of both networks. Other training parameters are shown in Table 1.

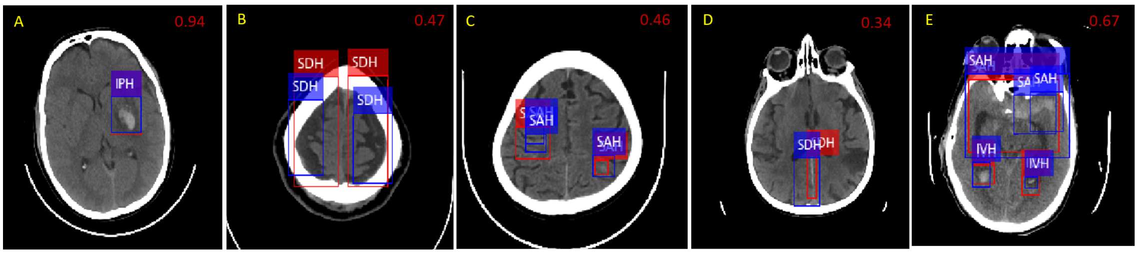

Figure 1: The visual display of predicted labelled BBs (blue) compared to ground truth (red) together with the IoU value (in the right upper corner). A – BB of IPH matching well with the ground truth, B – results of SDH localization on both sides of the head with the total IoU for both bleedings, C – two localized SAH bleedings, D – poorly localized SDH with insufficient overlap of the BB and the ground truth, E – localization of multiple hemorrhage types in one slice with the total IoU for both types jointly.

3 RESULTS AND DISCUSSION

Axial slices from twenty percent of the available CT scans were randomly selected for the evaluation of the detection models regarding the distribution of individual ICH types in the training dataset. The Jaccard coefficient (known as the intersection over union – IoU), which expresses the relative overlap of the ground truth and predicted BB, was chosen as the evaluation metric together with the sensitivity (Se) and the positive predictive value (PPV). To determine the number of true positive (TP), false positive (FP) and false negative (FN) results, threshold value of the IoU was chosen as 0.4 IoU taking into account the properties of the metric – a slight difference in the position of BBs leads to a distinct decrease in the IoU value, although subjectively it seems to be a correct localization (Fig. 1 B C D).

Table 2: Achieved results of the proposed detection networks on the testing dataset and their comparison to recently published 3D BB predicting method. The second row represents relative amount of the ICH type among the testing images – the bold value (multiple types might be present in an image). The sensitivity (Se) and positive predictive value (PPV) are calculated with a threshold value of 0.4 IoU. The bold values in the last column represent the average value through all the testing images.

		IPH	**IVH**	**SDH**	**SAH**	**Total ICHs**
	cases	32 %	5 %	61 %	29 %	**766**
Faster R-CNN	**IoU [%]**	67.4	60.9	62.4	47.4	**58.7**
	Se$_{0.4}$ [%]	78.0	69.5	74.9	49.6	**69.1**
	PPV$_{0.4}$ [%]	82.2	80.8	79.0	62.5	**76.5**
YOLO v2	**IoU [%]**	59.7	53.8	54.3	41.6	**49.1**
	Se$_{0.4}$ [%]	71.3	60.0	61.6	28.6	**57.3**
	PPV$_{0.4}$ [%]	82.3	60.0	80.7	31.9	**69.9**
Nemcek et al. [10]	**IoU [%]**	59.0	53.5	43.2	52.5	**53.7**
	Se$_{0.4}$ [%]	65.0	62.5	44.4	50.0	**57.3**
	PPV$_{0.4}$ [%]	81.7	83.3	78.0	61.5	**77.1**

Regarding the achieved results (Table 2), the Faster R-CNN shows better ability to correctly locate and classify ICH, than the YOLO v2. However, considering the achieved sensitivities of both networks, it is evident that the FN cases may occur, thus the ICHs are not detected, or more frequently, the localization is inaccurate (Fig. 1 D). Despite the insufficient overlap, such a result would be highly beneficial for a radiologist as it would call the attention to the bleeding. In comparison to 3D BB-predicting method published in [10], Faster R-CNN generally achieved better results (Table 2).

The detectors can localize and classify various ICH types even in one slice (Fig. 1 E) despite their different sizes, shapes, and positions. The high variability of shapes and possible locations of SAH (Fig. 1 C E) probably affected the detection results (Table 2) of both detectors in comparison to other ICH types.

The trained detectors show an ability to roughly delineate and classify the bleeding, and thus they might minimize the chance of missing an ICH by oversight. The process of manual ICH diagnostics consists in axial slices evaluation, hence the 2D BB-based localization in individual axial slices appears to be convenient for radiologists. In contrast to 3D BB output (as in [10]), 2D BBs may delineate the superior and inferior margins of one bleeding more precisely. Moreover, both real-time object detection systems provide very fast performance. [13, 14] An implementation of such an algorithm in a CAD system might lead to a distinct decrease in diagnostic time, hence, to saving lives or preventing from lasting consequences.

4 CONCLUSION

This paper presents an implementation of two automatic ICH localization and classification systems, that provide labelled 2D rectangular bounding boxes as an output. Two state-of-the-art object detection networks (Faster R-CNN and YOLO v2) were trained, evaluated, and compared using publicly available datasets of head CT image and annotations. An average Jaccard coefficient of 58.7 % was achieved by the Faster R-CNN, that globally showed better results on the available dataset.

ACKNOWLEDGEMENT

The Titan Xp GPU used for this research was donated by the NVIDIA Corporation.

REFERENCES

[1] CACERES, J. Alfredo a Joshua N. GOLDSTEIN. Intracranial Hemorrhage. *Emergency Medicine Clinics of North America* [online]. 2012, 30(3), 771-794 [cit. 2021-03-08]. ISSN 07338627. doi:10.1016/j.emc.2012.06.003

[2] CHANG, P.D., E. KUOY, J. GRINBAND, et al. Hybrid 3D/2D Convolutional Neural Network for Hemorrhage Evaluation on Head CT. *American Journal of Neuroradiology* [online]. 2018, 39(9), 1609-1616 [cit. 2021-03-08]. ISSN 0195-6108. doi:10.3174/ajnr.A5742

[3] GREWAL, Monika, Muktabh Mayank SRIVASTAVA, Pulkit KUMAR a Srikrishna VARADARAJAN. RADnet: Radiologist level accuracy using deep learning for hemorrhage detection in CT scans. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)* [online]. IEEE, 2018, 2018, s. 281-284 [cit. 2021-03-08]. ISBN 978-1-5386-3636-7. doi:10.1109/ISBI.2018.8363574

[4] LEE, Hyunkwang, Sehyo YUNE, Mohammad MANSOURI, et al. An explainable deep-learning algorithm for the detection of acute intracranial haemorrhage from small datasets. *Nature Biomedical Engineering* [online]. 2019, 3(3), 173-182 [cit. 2021-03-08]. ISSN 2157-846X. doi:10.1038/s41551-018-0324-9

[5] KARKI, Manohar, Junghwan CHO, Eunmi LEE, et al. CT window trainable neural network for improving intracranial hemorrhage detection by combining multiple settings. *Artificial Intelligence in Medicine* [online]. 2020, 106 [cit. 2021-03-08]. ISSN 09333657. doi:10.1016/j.artmed.2020.101850

[6] CHILAMKURTHY, Sasank, Rohit GHOSH, Swetha TANAMALA, et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. *The Lancet* [online]. 2018, 392(10162), 2388-2396 [cit. 2021-03-08]. ISSN 01406736. doi:10.1016/S0140-6736(18)31645-3

[7] PATEL, Ajay, Sil. C. VAN DE LEEMPUT, Mathias PROKOP, Bram VAN GINNEKEN a Rashindra MANNIESING. Image Level Training and Prediction: Intracranial Hemorrhage Identification in 3D Non-Contrast CT. *IEEE Access* [online]. 2019, 7, 92355-92364 [cit. 2021-03-08]. ISSN 2169-3536. doi:10.1109/ACCESS.2019.2927792

[8] PATEL, Ajay, Floris H. B. M. SCHREUDER, Catharina J. M. KLIJN, et al. Intracerebral Haemorrhage Segmentation in Non-Contrast CT. *Scientific Reports* [online]. 2019, 9(1) [cit. 2021-03-08]. ISSN 2045-2322. doi:10.1038/s41598-019-54491-6

[9] CHO, Junghwan, Ki-Su PARK, Manohar KARKI, et al. Improving Sensitivity on Identification and Delineation of Intracranial Hemorrhage Lesion Using Cascaded Deep Learning Models. *Journal of Digital Imaging* [online]. 2019, 32(3), 450-461 [cit. 2021-03-08]. ISSN 0897-1889. doi:10.1007/s10278-018-00172-1

[10] NEMCEK, Jakub, Roman JAKUBICEK, Jiri CHMELIK. Localization and Classification of Intracranial Hemorrhages in CT Data. *8th European Medical and Biological Engineering Conference* [online]. Springer International Publishing, 2021, 2021-11-30, s. 767-773 [cit. 2021-03-08]. IFMBE Proceedings. ISBN 978-3-030-64609-7. doi:10.1007/978-3-030-64610-3_86

[11] REIS, E. P., F. Nascimento, M. Aranha, et al. Brain Hemorrhage Extended (BHX): Bounding box extrapolation from thick to thin slice CT images (version 1.1). *PhysioNet* [online]. 2020, [cit. 2021-03-08]. https://doi.org/10.13026/9cft-hg92

[12] GOLDBERGER, A., L. Amaral, L. Glass, et al. PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation* [Online]. 2000, 101 (23), pp. e215-e220 [cit. 2021-03-08].

[13] REN, Shaoqing, Kaiming HE, Ross GIRSHICK a Jian SUN. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* [online]. 2017, 39(6), 1137-1149 [cit. 2021-03-12]. ISSN 0162-8828. doi:10.1109/TPAMI.2016.2577031

[14] REDMON, Joseph a Ali FARHADI. YOLO9000: Better, Faster, Stronger. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [online]. IEEE, 2017, 2017, s. 6517-6525 [cit. 2021-03-12]. ISBN 978-1-5386-0457-1. doi:10.1109/CVPR.2017.690

[15] HE, Kaiming, Xiangyu ZHANG, Shaoqing REN a Jian SUN. Deep Residual Learning for Image Recognition. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* [online]. IEEE, 2016, 2016, s. 770-778 [cit. 2021-03-12]. ISBN 978-1-4673-8851-1. doi:10.1109/CVPR.2016.90

[16] Kingma, D.P., Ba, J.: Adam: A method for stochastic optimization. CoRR abs/1412.6980 (2015) [cit. 2021-03-12]