Neurologijos seminarai 2022; 26(92): 41–49
DOI: 10.15388/NS.2022.26.6

Machine Learning in Acute Stroke Neuroimaging. A Systematic Literature Review

D. Matuliauskas*
I. Stražnickaitė*
A. Samuilis**
D. Jatužis***

**Faculty of Medicine, Vilnius University, Lithuania*

***Department of Radiology, Nuclear Medicine and Medical Physics, Institute of Biomedical Sciences;*
Faculty of Medicine, Vilnius University, Lithuania

****Center for Neurology, Vilnius University, Lithuania*

Summary. *Background.* Artificial intelligence (AI) in medical imaging is a growing and promising technology that can be applied in stroke diagnosis. The study aims to overview studies that compare diagnostic performance of AI applications in stroke detection and segmentation of stroke lesions with and without human clinicians, appraising the models, study design, and metrics used.

Materials and methods. This systematic review was performed using the PubMed search engine including articles published in the time frame of 2015 January 1 to 2021 July 23. A total of 438 studies were found, out of which 60 were chosen for the review.

Results. Only 2 out of 60 (3.3%) studies were prospective. Minimum unique computer tomography (CT) scans included for validation – 10, maximum – 21586, mean – 599, median – 100, standard deviation – ±2801.1. The training set sizes consisted of minimum 28 CT scans, maximum – 24214, mean – 1279, median – 153, standard deviation – ±5006.7. Most popular software used in the studies were Brainomix (n=12, 20% of studies) and RAPID (n=12, 20%), 6 studies (10%) used convolutional neural networks, and 6 studies did not identify the model or name of software used. The average value of the ROC AUC results reported was 0.884 and the average accuracy was 0.857. The average reported sensitivity and specificity were 0.746 and 0.862, respectively. 27 out of 60 studies used human operators, with the average number of human operators per study being 3.7±2.9.

Conclusions. AI solutions can be widely applied in computation of infarct volumes. Using AI in stroke diagnosis still requires further research with more prospective studies, more expert human operators, and more focus on evaluating secondary outcomes.

Keywords: AI, machine learning, neuroimaging, stroke.

INTRODUCTION

According to the Global Burden of Disease 2019 Study, stroke remains to be the second-leading cause of death and the third leading cause of death and disability in the world combined [1]. The number of people living with stroke is estimated to increase by 27% between 2017 and 2047 in the European Union, mostly due to the aging population and better survival rates [2]. In 2017, the cost associated with stroke was estimated at €45 billion, including direct and indirect costs of care provision and productivity loss [2]. Thus, early stroke detection is an essential part of better diagnostics, prompt intervention, and better overall patient health outcomes.

Artificial Intelligence (AI) is an emerging technology which enables computers to take real-life decisions with no or minimal human input [3]. In recent years, applications of AI in medicine have attracted increasing interest and investments from venture capital funds [4]. This rapidly evolving field provides a promising approach for quicker and more efficient imaging analysis, potentially contributing to quicker patient diagnosis. Machine learning, a subset of AI, can be used at different points of patient care, including ischemic and hemorrhagic stroke detection, triage, segmentation, classification, image quality improvement, Alberta Stroke Program Early CT Score (ASPECTS) grading, and outcome prediction [5].

Address:
Indrė Stražnickaitė
Vilnius University, Faculty of Medicine
M. K. Čiurlionio St. 21, LT- 03101 Vilnius, Lithuania
E-mail: indre.straznickaite@gmail.com

Some advocates of AI even suggest that current medical imaging workflows might be transformed to an extent resulting in staff relay [6]. However, other evidence suggests that AI technology is still in its early phase and most the research related to it tends to be flawed in study design [7]. Some studies even suggest that many computer aided diagnosis (CAD) systems result in additional work-ups for radiologists because of the false-positive cases [8, 9]. Therefore, most AI research must be assessed extremely critically.

STUDY AIM

This systematic review aims to give a contemporary overview of the studies that compare diagnostic AI performance of stroke detection and segmentation of stroke lesions, e.g., hemorrhage or large vessel occlusion, with human clinicians and without, appraising the AI models, study design, and metrics used.

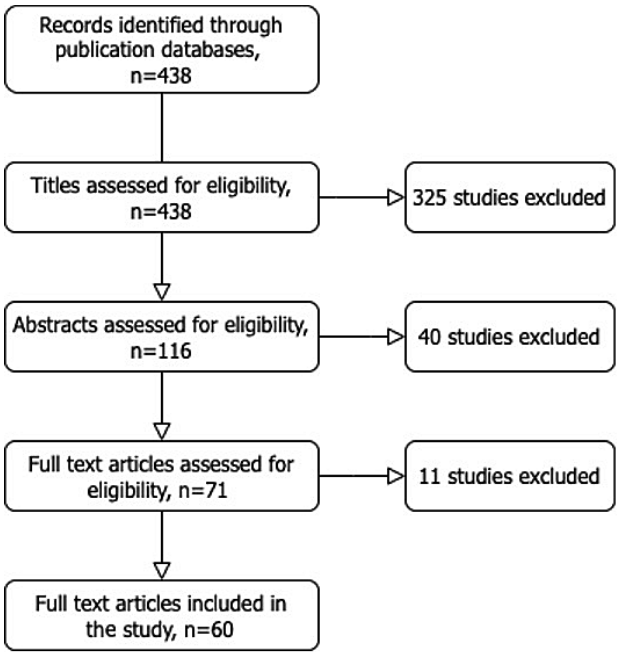

Fig. **Flowchart of literature screening and review protocol**

METHODS

This systematic review was performed using PubMed search engine for peer-reviewed articles using the search terms: “deep learning” or “machine learning” or “artificial intelligence” or “computer aided diagnosis” and “computed tomography” or “CT”, and “stroke” or “ischemic stroke” or “hemorrhagic stroke”. Literature published in the time frame of 2015 January 1 to 2021 July 23 was included in order to review the newest improvements made for AI-aided stroke diagnostics.

Inclusion criteria

Studies were included if they were in English only, had full-articles available, were using deep learning in brain imaging for segmentation and detection only, were using only computed tomography as an imaging modality, and were used for hemorrhagic or ischemic stroke detection.

Exclusion criteria

Studies were excluded from the search only if they did not meet inclusion criteria or compared different software without comparison with human clinicians, were abstracts only, were review articles, were studies involving animals or children, used machine learning to alter pixel values for quality improvement or were studies discussing only technical network architecture matters or predicting patient stroke outcome.

Literature analysis

A total number of 438 studies were found using the Pubmed search engine. Upon removal of studies that did not fit the inclusion criteria, 60 journal articles were chosen for the present review. Identified articles were reviewed independently by two authors. Relevant articles were analyzed to determine the dataset size used for training and validation, if any, the type of stroke discussed in the study (large vessel occlusion, ischemic, intracranial, and intracerebral hemorrhage), the imaging modality used (CTA, CTP, NCCT), the name of the software (if commercially available) or the type of the model used, the metrics used to evaluate the algorithm, and whether the study was conducted prospectively or retrospectively. The data analysis workflow is presented in Fig. The results, covered in this article, are reported only for the best results (e.g., the paper observes several machine learning techniques, but we report the technique that showed the best results).

Data analysis

Data was analysed using MS Excel (2021).

RESULTS

Table 1 shows general characteristics of the 60 studies covered in this article. Only 2 out of 60 (3.3%) included studies were prospective. Out of the 60 articles identified, 17 were on computed tomography angiography, 7 were on computed tomography perfusion, 44 were on non-contrast enhanced computed tomography, 6 were on computed tomography angiography and non-contrast enhanced computed tomography, 1 was on computed tomography angiography and computed tomography perfusion, and 1 was on computed tomography perfusion and non-contrast enhanced computed tomography. 21 studies out of 60 assessed large vessel occlusion, 32 assessed ischemic core volume, 4 assessed intracranial hemorrhage, 7 assessed

Table 1. **Main study characteristics**

Lead author	Publication year	Study type	Software used	Validation dataset size	Training set size
Yahav-Dovrat A. [11]	2021	Retrospective	Viz LVO	1167	–
Sheth SA. [15]	2019	Retrospective	DeepSymNet, Rapid iSchema View	297	–
Chilamkurthy S. [16]	2018	Retrospective	Not stated	21586	290055
Dhar R. [17]	2020	Retrospective	Not stated	308	–
Olive-Gadea M. [18]	2020	Retrospective	MethinksLVO	1453	24214
Nishio M. [19]	2020	Retrospective	You Only Look Once v3 (YOLOv3); Visual Geometry Group 16 (VGG16) CNN	49	189
Xiong Y. [20]	2019	Prospective	RAPID; Olea	120	–
Kasasbeh AS. [21]	2019	Retrospective	Not stated	51	77
Qiu W. [22]	2020	Retrospective	Not stated	100	157
Kniep HC. [23]	2020	Retrospective	Random forest algorithms	69	–
Kuang H. [24]	2020	Retrospective	Not stated	602	–
Li L. [25]	2021	Retrospective	UNet(UNet10), UNet(UNet6), UNet10Gan, UNet6Gan, UNet6Dil, UNet6Flip, UNet6FlipD, UNet6Incep, UNet6Vgg.	–	–
Arab A. [26]	2020	Retrospective	CNNs with deep supervision (CNN-DS)	10	45
Neuhaus A. [27]	2020	Retrospective	Brainomix e-ASPECTS	178	–
Shinohara Y. [28]	2020	Retrospective	Xception	22	–
Kral J. [29]	2020	Prospective	Brainomix	45	–
Guan Y. [30]	2020	Retrospective	Multilayer perceptron, Decision tree, Random Forest, Adaboost, Gradient boosting, Bagging, Bernoulli naive Bayes, Gaussian naive Bayes, Support vector machine, K-nearest neighbor	56	–
Shinohara Y. [31]	2020	Retrospective	Xception	79	–
Cimflova P. [32]	2020	Retrospective	Brainomix, RAPID	81	–
Hoelter P. [33]	2020	Retrospective	Syngo.via Frontier ASPECT Score Prototype V2, Brainomix e-ASPECTS, RAPID ASPECTS	131	–
Öman O. [34]	2019	Retrospective	DeepMedic	30	30
Mah YH. [35]	2020	Retrospective	Support vector machine	1832	–
Ko H. [36]	2020	Retrospective	CNN-LSTM; Xception	727392	4516842
Ironside N. [37]	2019	Retrospective	Analyze 12.0	40	260
Amukotuwa SA. [38]	2019	Retrospective	RAPID 4.9.1	969	–
Kuang H. [39]	2019	Retrospective	Random Forest	100	157
Seker F. [40]	2019	Retrospective	Brainomix e-ASPECTS	43	–
Vargas J. [41]	2018	Retrospective	CNN	40	356
Albers GW. [42]	2019	Retrospective	RAPID ASPECTS	65	–
Austein F. [43]	2019	Retrospective	Brainomix e-ASPECTS, iSchema View RAPID ASPECTS	52	–
Li L. [44]	2020	Retrospective	Frontier ASPECTS	55	–
Sales Barros R. [45]	2020	Retrospective	CNN	396	630
You J. [46]	2020	Retrospective	XGBoost	100	200
Wen X. [47]	2020	Retrospective	Multivariate logistic regression model	39	87
Guberina N. [48]	2018	Retrospective	Brainomix e-ASPECTS	117	–
Amukotuwa SA. [49]	2019	Retrospective	RAPID CTA	477	–
Copelan AZ. [50]	2020	Retrospective	RAPID	38	–
Heit JJ. [51]	2021	Retrospective	RAPID ICH	308	–
Stib MT. [52]	2020	Retrospective	DenseNet-121	116	424
Prevedello LM. [53]	2017	Retrospective	GoogLeNet	50	264
Schultheiss M. [54]	2020	Retrospective	U-net	186	369
Lo CM. [55]	2021	Retrospective	AlexNet	325	1254
Wang C. [56]	2021	Retrospective	3D CNN	194	259
Herweh C. [57]	2016	Retrospective	Brainomix e-ASPECTS	34	–
Nagel S. [58]	2017	Retrospective	Brainomix e-ASPECTS	132	–
Goebel J. [59]	2018	Retrospective	Brainomix, Frontier ASPECTS	150	–
Wu G. [60]	2021	Retrospective	U-net, ResNet, MAP	128	149

Table 1. **Main study characteristics** *(continuation)*

Lead author	Publication year	Study type	Software used	Validation dataset size	Training set size
Naganuma M. [61]	2021	Retrospective	Three-dimensional fully convolutional network-based brain hemisphere comparison algorithm (3D-BHCA)	151	–
Scherer M. [62]	2016	Retrospective	Random Forest	30	28
Rava RA. [63]	2021	Retrospective	Canon AUTOStroke Solution LVO	303	–
Abramova V. [64]	2021	Retrospective	3D U-net	15	61
Pan J. [65]	2021	Retrospective	DL Res Net	58	58
Brinjikji W. [66]	2021	Retrospective	Brainomix e-ASPECTS	60	–
Adhya J. [67]	2021	Retrospective	RAPID-CTA	310	–
Rava RA. [68]	2021	Retrospective	Canon AUTOStroke Solution ICH	302	–
Shi T. [69]	2021	Retrospective	C2MA Network	62	94
You J. [70]	2021	Retrospective	3D dissimilar-siamese-u-net	624	–
Delio PR. [71]	2021	Retrospective	iSchema View software, Rapid	500	–
Seker F. [72]	2021	Retrospective	Brainomix e-CTA	301	–
Tomasetti L. [73]	2021	Retrospective	Random Forest	93	59

intracerebral hemorrhage, 4 studies assessed large vessel occlusion and ischemic core volume, and 1 study assessed intracranial hemorrhage and intracerebral hemorrhage.

When assessing the study dataset size, the study by Ko H et al. 2020 stood out, as it had a significantly greater retrospective database than others (a total of 4516842 images used for training and 727392 for validation). Therefore, we excluded it from the statistics as an outlier. The corrected statistics for training set sizes consisted of minimum 28 CT scans, maximum – 24214, mean – 1279, median – 153, standard deviation – ±5006.7. Test and validation dataset sizes reported in the studies were as follows: minimum unique CT scans included for validation – 10, maximum – 21586, mean – 599, median – 100, standard deviation – ±2801.1. The results of the final statistics are presented in Table 2.

Table 2. **Training and validation dataset sizes**

	Validation	Training
Minimum CT scans used	10	28
Max CT scans used	21586	24214
Mean	599	1279
Standard deviation	2801.1	5006.7
Variance	7846182.5	25067432.5
Median	100	157

Most popular commercially available software used in the studies were Brainomix (n=12, 20% of studies) and RAPID (n=12, 20%), 6 studies (10%) specified using convolutional neural networks, and 6 studies did not identify the type of software, nor the model, nor the neural network architecture used. The other types of machine learning techniques used are shown in Table 1.

A large spectrum of different statistical methods was used to evaluate the performance of the model, making it difficult to compare their results accurately. The most popular methods used in the studies are presented in Table 3. We compared the results of the five most often reported metrics: sensitivity, specificity, accuracy, receiver operating characteristic area under the curve (ROC AUC), and Dice similarity coefficient. Bland-Altman plots were not

Table 3. **Most popular metrics used**

Spearman's rank correlation coefficient	1
Intersection over Union (IoU)	1
The McNemar's test	1
The Somer's delta	1
Krippendorf's alfa	1
^{2} square	1
One-way ANOVA	1
Youden index	1
Matthew's correlation coefficient	1
Change in groin puncture time	1
Concordance coefficient	2
Mann-Whitney U test	2
Fisher exact tests	2
T-test	3
Recall	4
K-fold cross validation	5
Positive predictive value	5
Inter-rater reliability / Interclass correlation coefficient	6
Negative predictive value	6
Precision	6
Cohen's kappa coefficient	8
Pearson correlation coefficient	9
Intra-rater reliability / Intraclass correlation coefficient	11
Dice similarity coefficient	13
Bland-Altman plots	14
ROC AUC	19
Accuracy	21
Specificity	25
Sensitivity	28

Table 4. **Most popular metrics used and their results**

	Sensitivity	Specificity	Accuracy	ROC AUC	Dice similarity coefficient
Publication count with the metric	28	25	21	19	13
Minimum value	0.185	0.570	0.609	0.759	0.317
Maximum value	1.000	1.000	1.000	0.960	0.905
Average value	0.764	0.862	0.857	0.884	0.712
Standard deviation	0.198	0.116	0.105	0.061	0.178
Variance	0.039	0.013	0.011	0.004	0.032

included in the comparison as the results are mostly visual rather than numeric.

The average value of the ROC AUC results is 0.884 and the average accuracy is 0.857, both of which can be considered as excellent. The average sensitivity and specificity of the studies are estimated at 0.746 and 0.862, respectively. The largest standard deviation of 0.198 and the variance of 0.039 between study results were among the results of the sensitivity metric.

Such findings suggest that the reported results show a low variance between different studies, indicating that the different software used perform at a rather similar level. The results are summarized in Table 4.

Human comparator groups were used in 27 out of 60 studies, and these groups were relatively small. The minimum number of people (experts or non-experts) involved in the study was 1, the maximum was 16, with the average number of human operators per study being 3.7±2.9. In most studies, the data were rated independently by most of the human comparators.

DISCUSSION

We have established several findings from our review. First, out of 60 studies reviewed, 58 studies were retrospective and only 2 were prospective. This is an important limitation in studies aimed at testing AI in clinical practice, as prospective studies are more suitable to represent the real clinical environment. AI performance is likely to be less accurate when facing new, real-world data, rather than the data used in algorithm training [10]. The success *in silico* studies and excellent performance metrics do not always translate into clinical functionality, as metrics such as ROC AUC, which is universally used in AI studies, are argued not to be the best metrics to represent clinical success [10].

There is little to no consensus on what metrics to report of AI application in stroke diagnostics should be used. Our findings show a large spectrum of different statistical measures to prove the performance of the model, however, only one study attempted to show the clinical benefit. The study by Yahav-Dovrat et al. (2021) measured how the AI application VIZ LVO reduced the time to the groin puncture [11]. Other studies introduced metrics that do not necessarily reflect any benefit in clinical practice, and their comparison with human operators reflects an *in silico* environment in which clinicians do not typically work.

Second, the studies in our review present a very varied number of dataset sizes with different sample distributions and characteristics. The average number of unique CT scans used for validation of neural networks per study was noted to be 599, however, the validation dataset sizes ranged from 10 to 21586 CT scans. The use of small datasets with less than 100 scans for training and/or validation raises the question of whether the results can be reliable, whether they can be replicated by others, whether they can be replicated in real-life clinical scenarios, or can be generalized in different populations and different regions.

Third, there were few human operators across all studies compared to AI applications, with the average number of human operators per study being 3.7±2.9. Inter-rater variability as well as intra-rater variability among human operators can also be high, therefore future research need to use larger samples of human operators to ensure reliability. In addition, including non-experts and comparing them with AI can make the algorithm better compared to it, making it harder for a human, so comparison with experts would be preferable [12]. Most importantly, studies with human comparators should attempt to move past opposing AI vs. clinicians and move towards observing collaboration between clinicians and machine learning, as the combination of both tends to outperform either alone [13, 14].

However, AI seems to be a promising tool for ridding people of manual and repetitive tasks. For example, AI solutions can be widely applied in computation of infarct volumes since manual delineation tends to be a tedious task. It is widely recognized that medical reporting is a huge responsibility for medical practitioners and takes a significant amount of time for each examination.

Finally, the legal status of any AI application operating in a clinical setting should be established. AI as a standalone decision maker or working on a black box principle does not seem to be a safe and sustainable solution. For example, AI applications in stroke diagnosis still carry a relatively high probability of false negative cases that may result in late or missed diagnosis. Vendors of AI systems must make their applications transparent and focus on the application design that empowers the end-user rather than replaces it. Further research should aim to mention the intended position of the AI system under study in the clinical pathway.

In summary, further studies should focus on including larger patient samples from different geographical regions, including more prospective cases, validating AI models across different care centers, including more expert human operators to account for inter-operator variability among humans, evaluating secondary outcomes, e.g., time savings per procedure and long-term patient outcomes, avoiding *in silico* testing environment to evaluate the performance of human operators compared to the AI systems, and ultimately making datasets and source codes available to other researchers to accelerate AI research globally.

CONLUSIONS

1. Artificial intelligence is a rapidly developing field that promises significant impact in time savings, quality improvement of medical procedures, and even better patient outcomes.
2. However, the use of AI in routine stroke diagnosis still requires further research with more prospective studies, more expert human operators, and more focus on evaluating secondary outcomes.
3. The performance of AI systems in clinical settings should be demonstrated more in tandem with human operators rather than stand-alone systems.

References

1. Feigin VL, Stark BA, Johnson CO, Roth GA, Bisignano C, Abady GG, et al. Global, regional, and national burden of stroke and its risk factors, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet Neurol 2021; 20(10): 795–820. https://doi.org/10.1016/S1474-4422(21)00252-0
2. Wafa HA, Wolfe CDA, Emmett E, Roth GA, Johnson CO, Wang Y. Burden of stroke in Europe: thirty-year projections of incidence, prevalence, deaths, and disability-adjusted life years. Stroke 2020; 51(8): 2418–27. https://doi.org/10.1161/STROKEAHA.120.029606
3. Hamet P, Tremblay J. Artificial intelligence in medicine. Metabolism 2017; 69: S36–40. https://doi.org/10.1016/j.metabol.2017.01.011
4. Zhang D, Saurabh M, Brynjolfsson E, Etchemendy J, Ganguli D, Grosz B, et al. The AI Index 2021 Annual Report [Internet]. [cited 2022 Feb 10]. Available from: https://aiindex.stanford.edu/ai-index-report-2021/
5. Soun JE, Chow DS, Nagamine M, Takhtawala RS, Filippi CG, Yu W, et al. Artificial intelligence and acute stroke imaging. Am J Neuroradiol 2021; 42(1): 2–11. https://doi.org/10.3174/ajnr.A6883
6. AI, radiology and the future of work. Available from: https://www.economist.com/leaders/2018/06/07/ai-radiology-and-the-future-of-work
7. Liu X, Faes L, Kale AU, Wagner SK, Fu DJ, Bruynseels A, et al. A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis. Lancet Digit Health 2019; 1(6): e271–97. https://doi.org/10.1016/S2589-7500(19)30123-2
8. Fenton JJ, Taplin SH, Carney PA, Abraham L, Sickles EA, D'Orsi C, et al. Influence of computer-aided detection on performance of screening mammography. N Engl J Med 2007; 356(14): 1399–409. https://doi.org/10.1056/NEJMoa066099
9. Alcusky M, Philpotts L, Bonafede M, Clarke J, Skoufalos A. The patient burden of screening mammography recall. J Womens Health 2014; 23(S1): S-11–S-19. https://doi.org/10.1089/jwh.2014.1511
10. Kelly CJ, Karthikesalingam A, Suleyman M, Corrado G, King D. Key challenges for delivering clinical impact with artificial intelligence. BMC Med 2019; 17(1): 195. https://doi.org/10.1186/s12916-019-1426-2
11. Yahav-Dovrat A, Saban M, Merhav G, Lankri I, Abergel E, Eran A, et al. Evaluation of artificial intelligence-powered identification of large-vessel occlusions in a comprehensive stroke center. Am J Neuroradiol 2021; 42(2): 247–54. https://doi.org/10.3174/ajnr.A6923
12. Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, et al. Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. BMJ 2020; 368: m689. https://doi.org/10.1136/bmj.m689
13. Briganti G, Le Moine O. Artificial intelligence in medicine: today and tomorrow. Front Med 2020; 7: 27. https://doi.org/10.3389/fmed.2020.00027
14. Tschandl P, Rinner C, Apalla Z, Argenziano G, Codella N, Halpern A, et al. Human-computer collaboration for skin cancer recognition. Nat Med 2020; 26(8): 1229–34. https://doi.org/10.1038/s41591-020-0942-0
15. Sheth SA, Lopez-Rivera V, Barman A, Grotta JC, Yoo AJ, Lee S, et al. Machine learning-enabled automated determination of acute ischemic core from computed tomography angiography. Stroke 2019; 50(11): 3093–100. https://doi.org/10.1161/STROKEAHA.119.026189
16. Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. Lancet 2018; 392(10162): 2388–96. https://doi.org/10.1016/S0140-6736(18)31645-3
17. Dhar R, Falcone GJ, Chen Y, Hamzehloo A, Kirsch EP, Noche RB, et al. Deep learning for automated measurement of hemorrhage and perihematomal edema in supratentorial intracerebral hemorrhage. Stroke 2020; 51(2): 648–51. https://doi.org/10.1161/STROKEAHA.119.027657
18. Olive-Gadea M, Crespo C, Granes C, Hernandez-Perez M, Pérez de la Ossa N, Laredo C, et al. Deep learning based software to identify large vessel occlusion on noncontrast computed tomography. Stroke 2020; 51(10): 3133–7. https://doi.org/10.1161/STROKEAHA.120.030326
19. Nishio M, Koyasu S, Noguchi S, Kiguchi T, Nakatsu K, Akasaka T, et al. Automatic detection of acute ischemic stroke using non-contrast computed tomography and two-stage deep learning model. Comput Methods Programs Biomed 2020; 196: 105711. https://doi.org/10.1016/j.cmpb.2020.105711
20. Xiong Y, Huang CC, Fisher M, Hackney DB, Bhadelia RA, Selim MH. Comparison of automated CT perfusion softwares in evaluation of acute ischemic stroke. J Stroke Cerebrovasc Dis 2019; 28(12): 104392. https://doi.org/10.1016/j.jstrokecerebrovasdis.2019.104392
21. Kasasbeh AS, Christensen S, Parsons MW, Campbell B, Albers GW, Lansberg MG. Artificial neural network computer tomography perfusion prediction of ischemic core.

Stroke 2019; 50(6): 1578–81. https://doi.org/10.1161/STROKEAHA.118.022649

22. Qiu W, Kuang H, Teleg E, Ospel JM, Sohn SI, Almekhlafi M, et al. Machine learning for detecting early infarction in acute stroke with non-contrast-enhanced CT. Radiology 2020; 294(3): 638–44. https://doi.org/10.1148/radiol.2020191193
23. Kniep HC, Sporns PB, Broocks G, Kemmling A, Nawabi J, Rusche T, et al. Posterior circulation stroke: machine learning-based detection of early ischemic changes in acute non-contrast CT scans. J Neurol 2020; 267(9): 2632–41. https://doi.org/10.1007/s00415-020-09859-4
24. Kuang H, Qiu W, Najm M, Dowlatshahi D, Mikulik R, Poppe AY, et al. Validation of an automated ASPECTS method on non-contrast computed tomography scans of acute ischemic stroke patients. Int J Stroke 2020; 15(5): 528–34. https://doi.org/10.1177/1747493019895702
25. Li L, Wei M, Liu B, Atchaneeyasakul K, Zhou F, Pan Z, et al. Deep learning for hemorrhagic lesion detection and segmentation on brain CT images. IEEE J Biomed Health Inform 2021; 25(5): 1646–59. https://doi.org/10.1109/JBHI.2020.3028243
26. Arab A, Chinda B, Medvedev G, Siu W, Guo H, Gu T, et al. A fast and fully-automated deep-learning approach for accurate hemorrhage segmentation and volume quantification in non-contrast whole-head CT. Sci Rep 2020; 10(1): 19389. https://doi.org/10.1038/s41598-020-76459-7
27. Neuhaus A, Seyedsaadat SM, Mihal D, Benson J, Mark I, Kallmes DF, et al. Region-specific agreement in ASPECTS estimation between neuroradiologists and e-ASPECTS software. J NeuroInterventional Surg 2020; 12(7): 720–3. https://doi.org/10.1136/neurintsurg-2019-015442
28. Shinohara Y, Takahashi N, Lee Y, Ohmura T, Kinoshita T. Development of a deep learning model to identify hyperdense MCA sign in patients with acute ischemic stroke. Jpn J Radiol 2020; 38(2): 112–7. https://doi.org/10.1007/s11604-019-00894-4
29. Kral J, Cabal M, Kasickova L, Havelka J, Jonszta T, Volny O, et al. Machine learning volumetry of ischemic brain lesions on CT after thrombectomy – prospective diagnostic accuracy study in ischemic stroke patients. Neuroradiology 2020; 62(10): 1239–45. https://doi.org/10.1007/s00234-020-02419-7
30. Guan Y, Wang P, Wang Q, Li P, Zeng J, Qin P, et al. Separability of acute cerebral infarction lesions in CT based radiomics: toward artificial intelligence-assisted diagnosis. BioMed Res Int 2020; 2020: 8864756. https://doi.org/10.1155/2020/8864756
31. Shinohara Y, Takahashi N, Lee Y, Ohmura T, Umetsu A, Kinoshita F, et al. Usefulness of deep learning-assisted identification of hyperdense MCA sign in acute ischemic stroke: comparison with readers' performance. Jpn J Radiol 2020; 38(9): 870–7. https://doi.org/10.1007/s11604-020-00986-6
32. Cimflova P, Volny O, Mikulik R, Tyshchenko B, Belaskova S, Vinklarek J, et al. Detection of ischemic changes on baseline multimodal computed tomography: expert reading vs. Brainomix and RAPID software. J Stroke Cerebrovasc Dis 2020; 29(9): 104978. https://doi.org/10.1016/j.jstrokecerebrovasdis.2020.104978
33. Hoelter P, Muehlen I, Goelitz P, Beuscher V, Schwab S, Doerfler A. Automated ASPECT scoring in acute ischemic stroke: comparison of three software tools. Neuroradiology 2020; 62(10): 1231–8. https://doi.org/10.1007/s00234-020-02439-3
34. Öman O, Mäkelä T, Salli E, Savolainen S, Kangasniemi M. 3D convolutional neural networks applied to CT angiography in the detection of acute ischemic stroke. Eur Radiol Exp 2019; 3(1): 8. https://doi.org/10.1186/s41747-019-0085-6
35. Mah YH, Nachev P, MacKinnon AD. Quantifying the impact of chronic ischemic injury on clinical outcomes in acute stroke with machine learning. Front Neurol 2020; 11: 15. https://doi.org/10.3389/fneur.2020.00015
36. Ko H, Chung H, Lee H, Lee J. Feasible study on intracranial hemorrhage detection and classification using a CNN-LSTM network. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). Montreal: IEEE, 2020. https://doi.org/10.1109/EMBC44109.2020.9176162
37. Ironside N, Chen CJ, Mutasa S, Sim JL, Marfatia S, Roh D, et al. Fully automated segmentation algorithm for hematoma volumetric analysis in spontaneous intracerebral hemorrhage. Stroke 2019; 50(12): 3416–23. https://doi.org/10.1161/STROKEAHA.119.026561
38. Amukotuwa SA, Straka M, Dehkharghani S, Bammer R. Fast automatic detection of large vessel occlusions on CT angiography. Stroke 2019; 50(12): 3431–8. https://doi.org/10.1161/STROKEAHA.119.027076
39. Kuang H, Najm M, Chakraborty D, Maraj N, Sohn SI, Goyal M, et al. Automated ASPECTS on noncontrast CT scans in patients with acute ischemic stroke using machine learning. Am J Neuroradiol 2019; 40(1): 33–8. https://doi.org/10.3174/ajnr.A5889
40. Seker F, Pfaff J, Nagel S, Vollherbst D, Gerry S, Möhlenbruch MA, et al. CT reconstruction levels affect automated and reader-based ASPECTS ratings in acute ischemic stroke: automated and human ASPECTS with different CT reconstructions. J Neuroimaging 2019; 29(1): 62–4. https://doi.org/10.1111/jon.12562
41. Vargas J, Spiotta A, Chatterjee AR. Initial experiences with artificial neural networks in the detection of computed tomography perfusion deficits. World Neurosurg 2019; 124: e10–6. https://doi.org/10.1016/j.wneu.2018.10.084
42. Albers GW, Wald MJ, Mlynash M, Endres J, Bammer R, Straka M, et al. Automated calculation of Alberta Stroke Program Early CT score: validation in patients with large hemispheric infarct. Stroke 2019; 50(11): 3277–9. https://doi.org/10.1161/STROKEAHA.119.026430
43. Austein F, Wodarg F, Jürgensen N, Huhndorf M, Meyne J, Lindner T, et al. Automated versus manual imaging assessment of early ischemic changes in acute stroke: comparison of two software packages and expert consensus. Eur Radiol 2019; 29(11): 6285–92. https://doi.org/10.1007/s00330-019-06252-2
44. Li L, Chen Y, Bao Y, Jia X, Wang Y, Zuo T, et al. Comparison of the performance between Frontier ASPECTS software and different levels of radiologists on assessing CT examinations of acute ischaemic stroke patients. Clin Radiol 2020; 75(5): 358–65. https://doi.org/10.1016/j.crad.2019.12.010
45. Sales Barros R, Tolhuisen ML, Boers AM, Jansen I, Ponomareva E, Dippel DWJ, et al. Automatic segmentation of cerebral infarcts in follow-up computed tomography images with convolutional neural networks. J NeuroInterventional Surg 2020; 12(9): 848–52. https://doi.org/10.1136/neurintsurg-2019-015471
46. You J, Tsang ACO, Yu PLH, Tsui ELH, Woo PPS, Lui CSM, et al. Automated hierarchy evaluation system of large vessel

occlusion in acute ischemia stroke. Front Neuroinformatics 2020; 14: 13. https://doi.org/10.3389/fninf.2020.00013
47. Wen X, Li Y, He X, Xu Y, Shu Z, Hu X, et al. Prediction of malignant acute middle cerebral artery infarction via computed tomography radiomics. Front Neurosci 2020; 14: 708. https://doi.org/10.3389/fnins.2020.00708
48. Guberina N, Dietrich U, Radbruch A, Goebel J, Deuschl C, Ringelstein A, et al. Detection of early infarction signs with machine learning-based diagnosis by means of the Alberta Stroke Program Early CT score (ASPECTS) in the clinical routine. Neuroradiology 2018; 60(9): 889–901. https://doi.org/10.1007/s00234-018-2066-5
49. Amukotuwa SA, Straka M, Smith H, Chandra RV, Dehkharghani S, Fischbein NJ, et al. Automated detection of intracranial large vessel occlusions on computed tomography angiography: a single center experience. Stroke 2019; 50(10): 2790–8. https://doi.org/10.1161/STROKEAHA.119.026259
50. Copelan AZ, Smith ER, Drocton GT, Narsinh KH, Murph D, Khangura RS, et al. Recent administration of iodinated contrast renders core infarct estimation inaccurate using RAPID software. Am J Neuroradiol 2020; 41(12): 2235–42. https://doi.org/10.3174/ajnr.A6908
51. Heit JJ, Coelho H, Lima FO, Granja M, Aghaebrahim A, Hanel R, et al. Automated cerebral hemorrhage detection using RAPID. Am J Neuroradiol 2021; 42(2): 273–8. https://doi.org/10.3174/ajnr.A6926
52. Stib MT, Vasquez J, Dong MP, Kim YH, Subzwari SS, Triedman HJ, et al. Detecting large vessel occlusion at multiphase CT angiography by using a deep convolutional neural network. Radiology 2020; 297(3): 640–9. https://doi.org/10.1148/radiol.2020200334
53. Prevedello LM, Erdal BS, Ryu JL, Little KJ, Demirer M, Qian S, et al. Automated critical test findings identification and online notification system using artificial intelligence in imaging. Radiology 2017; 285(3): 923–31. https://doi.org/10.1148/radiol.2017162664
54. Schultheiss M, Noël PB, Riederer I, Thiele F, Kopp FK, Renger B, et al. Towards subject-level cerebral infarction classification of CT scans using convolutional networks. PLOS ONE 2020; 15(7): e0235765. https://doi.org/10.1371/journal.pone.0235765
55. Lo CM, Hung PH, Lin DT. Rapid assessment of acute ischemic stroke by computed tomography using deep convolutional neural networks. J Digit Imaging 2021; 34(3): 637–46. https://doi.org/10.1007/s10278-021-00457-y
56. Wang C, Shi Z, Yang M, Huang L, Fang W, Jiang L, et al. Deep learning-based identification of acute ischemic core and deficit from non-contrast CT and CTA. J Cereb Blood Flow Metab 2021; 41(11): 3028–38. https://doi.org/10.1177/0271678X211023660
57. Herweh C, Ringleb PA, Rauch G, Gerry S, Behrens L, Möhlenbruch M, et al. Performance of e-ASPECTS software in comparison to that of stroke physicians on assessing CT scans of acute ischemic stroke patients. Int J Stroke 2016; 11(4): 438–45. https://doi.org/10.1177/1747493016632244
58. Nagel S, Sinha D, Day D, Reith W, Chapot R, Papanagiotou P, et al. e-ASPECTS software is non-inferior to neuroradiologists in applying the ASPECT score to computed tomography scans of acute ischemic stroke patients. Int J Stroke 2017; 12(6): 615–22. https://doi.org/10.1177/1747493016681020
59. Goebel J, Stenzel E, Guberina N, Wanke I, Koehrmann M, Kleinschnitz C, et al. Automated ASPECT rating: comparison between the Frontier ASPECT Score software and the Brainomix software. Neuroradiology 2018; 60(12): 1267–72. https://doi.org/10.1007/s00234-018-2098-x
60. Wu G, Chen X, Lin J, Wang Y, Yu J. Identification of invisible ischemic stroke in noncontrast CT based on novel two-stage convolutional neural network model. Med Phys 2021; 48(3): 1262–75. https://doi.org/10.1002/mp.14691
61. Naganuma M, Tachibana A, Fuchigami T, Akahori S, Okumura S, Yi K, et al. Alberta Stroke Program Early CT score calculation using the deep learning-based brain hemisphere comparison algorithm. J Stroke Cerebrovasc Dis 2021; 30(7): 105791. https://doi.org/10.1016/j.jstrokecerebrovasdis.2021.105791
62. Scherer M, Cordes J, Younsi A, Sahin YA, Götz M, Möhlenbruch M, et al. Development and validation of an automatic segmentation algorithm for quantification of intracerebral hemorrhage. Stroke 2016; 47(11): 2776–82. https://doi.org/10.1161/STROKEAHA.116.013779
63. Rava RA, Peterson BA, Seymour SE, Snyder KV, Mokin M, Waqas M, et al. Validation of an artificial intelligence-driven large vessel occlusion detection algorithm for acute ischemic stroke patients. Neuroradiol J 2021; 34(5): 408–17. https://doi.org/10.1177/1971400921998952
64. Abramova V, Clèrigues A, Quiles A, Figueredo DG, Silva Y, Pedraza S, et al. Hemorrhagic stroke lesion segmentation using a 3D U-Net with squeeze-and-excitation blocks. Comput Med Imaging Graph 2021; 90: 101908. https://doi.org/10.1016/j.compmedimag.2021.101908
65. Pan J, Wu G, Yu J, Geng D, Zhang J, Wang Y. Detecting the early infarct core on non-contrast CT images with a deep learning residual network. J Stroke Cerebrovasc Dis 2021; 30(6): 105752. https://doi.org/10.1016/j.jstrokecerebrovasdis.2021.105752
66. Brinjikji W, Abbasi M, Arnold C, Benson JC, Braksick SA, Campeau N, et al. e-ASPECTS software improves interobserver agreement and accuracy of interpretation of aspects score. Interv Neuroradiol 2021; 27(6): 781–7. https://doi.org/10.1177/15910199211011861
67. Adhya J, Li C, Eisenmenger L, Cerejo R, Tayal A, Goldberg M, et al. Positive predictive value and stroke workflow outcomes using automated vessel density (RAPID-CTA) in stroke patients: one year experience. Neuroradiol J 2021; 34(5): 476–81. https://doi.org/10.1177/19714009211012353
68. Rava RA, Seymour SE, LaQue ME, Peterson BA, Snyder KV, Mokin M, et al. Assessment of an artificial intelligence algorithm for detection of intracranial hemorrhage. World Neurosurg 2021; 150: e209–17. https://doi.org/10.1016/j.wneu.2021.02.134
69. Shi T, Jiang H, Zheng B. C^2MA-Net: cross-modal cross-attention network for acute ischemic stroke lesion segmentation based on CT perfusion scans. IEEE Trans Biomed Eng 2022; 69(1): 108–18. https://doi.org/10.1109/TBME.2021.3087612
70. You J, Yu PLH, Tsang ACO, Tsui ELH, Woo PPS, Lui CSM, et al. 3D dissimilar-siamese-u-net for hyperdense Middle cerebral artery sign segmentation. Comput Med Imaging Graph 2021; 90: 101898. https://doi.org/10.1016/j.compmedimag.2021.101898
71. Delio PR, Wong ML, Tsai JP, Hinson HE, McMenamy J, Le TQ, et al. Assistance from automated ASPECTS software improves reader performance. J Stroke Cerebrovasc Dis

2021; 30(7): 105829. https://doi.org/10.1016/j.jstrokecerebrovasdis.2021.105829
72. Seker F, Pfaff JAR, Mokli Y, Berberich A, Namias R, Gerry S, et al. Diagnostic accuracy of automated occlusion detection in CT angiography using e-CTA. Int J Stroke 2022; 17(1): 77–82. https://doi.org/10.1177/1747493021992592
73. Tomasetti L, Hollesli LJ, Engan K, Kurz KD, Kurz MW, Khanmohammadi M. Machine learning algorithms versus thresholding to segment ischemic regions in patients with acute ischemic stroke. IEEE J Biomed Health Inform 2022; 26(2): 660–72. https://doi.org/10.1109/JBHI.2021.3097591

D. Matuliauskas, I. Stražnickaitė, A. Samuilis, D. Jatužis

DIRBTINIO INTELEKTO PRITAIKYMAS ŪMINIO INSULTO VAIZDINIŲ TYRIMŲ DIAGNOSTIKOJE. SISTEMINĖ LITERATŪROS APŽVALGA

Santrauka

Įvadas. Dirbtinis intelektas (DI) yra sparčiai besivystanti technologija, kuri gali atnešti daug teigiamų pokyčių galvos smegenų insulto diagnostikoje. Tyrimo tikslas – apžvelgti publikacijas, kuriose DI gebėjimas diagnozuoti ūminį insultą iš radiologinių vaizdų ir gebėjimas segmentuoti insulto radiologinius požymius yra lyginamas su žmonėmis vertintojais arba apžvelgiamas be jų. Apžvelgiami naudoti modeliai, studijų dizainas ir taikomi statistikos metodai.

Tiriamieji ir tyrimo metodai. Sisteminė apžvalga buvo atlikta naudojant „Pubmed“ duomenų bazę, į apžvalgą įtraukiant publikacijas nuo 2015 m. sausio 1 d. iki 2021 m. liepos 23 d. Iš viso buvo aptiktos 438 publikacijos, iš kurių apžvalgai atrinkta 60.

Rezultatai. Tik 2 studijos iš 60 (3,3 %) buvo perspektyvinės. Mažiausias unikalių kompiuterinės tomografijos vaizdų skaičius, naudotas DI sistemos validacijai, buvo 10, didžiausias – 21 586, vidurkis – 599, mediana – 100, standartinis nuokrypis – ±2801,1. Mažiausias duomenų kiekis, naudotas neuroninių tinklų mokymui, buvo 28 studijos, didžiausias – 24 214, vidurkis – 1 279, mediana – 153, standartinis nuokrypis – ±5 006,7. Populiariausia naudota programinė įranga buvo „Brainomix“ (n = 12, 20 % visų publikacijų) ir RAPID (n = 12, 20 % visų publikacijų), 6 studijos (10 %) naudojo konvoliucinius neuroninius tinklus ir 6 publikacijos nenurodė nei modelio, nei programinės įrangos pavadinimo. Vidutinis ploto po kreive rodiklis buvo 0,884, o vidutinis tikslumas – 0,857. Jautrumo ir specifiškumo vidurkiai buvo 0,746 ir 0,862. 27 iš 60 atliktų studijų turėjo žmones vertintojus, o žmonių vertintojų skaičiaus vidurkis šiose studijose buvo 3,7 ± 2,9.

Išvados. Dirbtinio intelekto sprendimai gali būti plačiai taikomi automatizuotam galvos smegenų insulto tūrio apskaičiavimui kompiuterinės tomografijos vaizduose. DI taikymas galvos smegenų insulto diagnostikai vis dar reikalauja papildomų tyrimų, perspektyvinių studijų, platesnio palyginimo su žmonėmis vertintojais ir didesnio dėmesio antrinių išeičių vertinimui.

Raktažodžiai: dirbtinis intelektas, insultas, mašininis mokymasis, neuroradiologija.

Gauta: 2022 04 01

Priimta spaudai: 2022 04 12