La radiologia medica (2022) 127:1106–1123
https://doi.org/10.1007/s11547-022-01530-4

COMPUTED TOMOGRAPHY

Accuracy of artificial intelligence for the detection of intracranial hemorrhage and chronic cerebral microbleeds: a systematic review and pooled analysis

Stavros Matsoukas[1] · Jacopo Scaggiante[1] · Braxton R. Schuldt[1] · Colton J. Smith[1] · Susmita Chennareddy[1] · Roshini Kalagara[1] · Shahram Majidi[1] · Joshua B. Bederson[1] · Johanna T. Fifi[1] · J. Mocco[1] · Christopher P. Kellner[1]

Received: 20 October 2021 / Accepted: 12 July 2022 / Published online: 13 August 2022

Abstract

Background Artificial intelligence (AI)-driven software has been developed and become commercially available within the past few years for the detection of intracranial hemorrhage (ICH) and chronic cerebral microbleeds (CMBs). However, there is currently no systematic review that summarizes all of these tools or provides pooled estimates of their performance.

Methods In this PROSPERO-registered, PRISMA compliant systematic review, we sought to compile and review all MEDLINE and EMBASE published studies that have developed and/or tested AI algorithms for ICH detection on non-contrast CT scans (NCCTs) or MRI scans and CMBs detection on MRI scans.

Results In total, 40 studies described AI algorithms for ICH detection in NCCTs/MRIs and 19 for CMBs detection in MRIs. The overall sensitivity, specificity, and accuracy were 92.06%, 93.54%, and 93.46%, respectively, for ICH detection and 91.6%, 93.9%, and 92.7% for CMBs detection. Some of the challenges encountered in the development of these algorithms include the laborious work of creating large, labeled and balanced datasets, the volumetric nature of the imaging examinations, the fine tuning of the algorithms, and the reduction in false positives.

Conclusions Numerous AI-driven software tools have been developed over the last decade. On average, they are characterized by high performance and expert-level accuracy for the diagnosis of ICH and CMBs. As a result, implementing these tools in clinical practice may improve workflow and act as a failsafe for the detection of such lesions.

Registration-URL https://www.crd.york.ac.uk/prospero/ Unique Identifier: CRD42021246848.

Keywords Artificial intelligence · Intracranial hemorrhage · Non-contrast CT scan · AI-assisted diagnosis · Chronic microbleeds · Convolutional neural network

Abbreviations

AI	Artificial intelligence
ICH	Intracranial hemorrhage
NCCT	Non-contrast CT scan
IPH	Intraparenchymal hemorrhage
IVH	Intraventricular hemorrhage
SAH	Subarachnoid hemorrhage
EDH	Epidural hematoma
SDH	Subdural hematoma
CMBs	Chronic microbleeds
SN	Sensitivity
SP	Specificity
PPV	Positive predictive value
NVP	Negative predicted value
AUC	Area under the curve
CNN	Convolutional neural network
RNN	Recurrent neural networks
ANN	Artificial neural networks
2D	2-Dimensional
3D	3-Dimensional

Other Information: Registration and protocol: https://www.crd.york.ac.uk/prospero/ Unique Identifier: CRD42021246848.

✉ Stavros Matsoukas
Stavros.matsoukas@mountsinai.org;
stavrosmatsoukas@hotmail.com

[1] Department of Neurosurgery, Mount Sinai Health System, Annenberg Building, Room 20-86, 1468 Madison Ave, New York, NY 10029, USA

Introduction

Intracranial hemorrhage (ICH) can be associated with high morbidity and mortality, thus requiring diagnosis in a timely fashion [1]. Over the last decade, multiple artificial intelligence (AI) algorithms have been developed for the detection of intracranial hemorrhage (ICH) on non-contrast CT scans (NCCTs), including intraparenchymal hemorrhage (IPH), intraventricular hemorrhage (IVH), subarachnoid hemorrhage (SAH), epidural hematoma (EDH), and subdural hematoma (SDH). Additionally, AI algorithms have been developed for the detection of chronic microbleeds (CMBs) on non-contrast MRI. It is thought that implementing such tools in the healthcare setting will streamline workflow and decrease the time from imaging to treatment [1, 2], ultimately leading to a decrease in morbidity and mortality. However, there is no current work that summarizes the overall diagnostic accuracy of such algorithms.

Machine learning, the driving force behind the AI algorithms for radiographic image analysis, includes shallow learning and deep learning. Training, validation, and testing are required sequential steps for the development of an AI algorithm. Usually, large datasets are used for training purposes. However, training can also be accomplished with small datasets through the use of transfer learning [3], a process in which existing knowledge is transferred from other previously trained AI platforms. Finally, validation refers to the "fine-tuning'' stage of the developing model which takes part in the final stages of training, usually in smaller datasets that focus on the exact patient population of the test [4].

The authors hypothesize that the majority of the available AI algorithms can yield diagnostic accuracy that approaches the caliber of experienced radiologists. The purpose of this systematic review is twofold: First, to identify all currently existing AI algorithms that facilitate ICH detection in NCCTs/MRIs and CMBs in non-contrast MRI scans, along with their corresponding accuracy metrics. Second, to conduct a pooled analysis that provides an estimate of the overall performance of these algorithms. While some narrative reviews describe a small number of the available algorithms, no systematic review of AI ICH detection algorithms in non-contrast CT scans currently exists [5–8]. Additionally, there is currently no review (narrative or systematic) for the detection of CMBs in non-contrast MRI scans. Consequently, this work differs from any previously published reviews, since it includes all the relevant AI-algorithms that have been published to date, along with the first pooled analysis of their overall performance.

Methods

Registration and literature search

This is a systematic review, structured and written in accordance with 2020 PRISMA guidelines [9]. Before performing a literature search, a systematic review protocol was developed according to PRISMA guidelines for review protocols (PRISMA-P 2015) [10] and registered in PROSPERO on May 9th, 2021 [11]. For the purposes of identifying all potentially eligible studies, MEDLINE (PubMed) and EMBASE were accessed through Ovid Search on March 1st, 2021. After identifying mesh terms, search terms included: "intracranial hemorrhage," "intracerebral hemorrhage," "brain hemorrhage," "intraparenchymal hemorrhage," "epidural hemorrhage," "subdural hemorrhage," "intraventricular hemorrhage," "subarachnoid hemorrhage," "cerebral microbleeds," "artificial intelligence," "computational intelligence," "machine intelligence," "AI," "deep learning," "machine learning," and "convolutional neural network." Fig. 1 depicts the search terms used in this study. The same search terms were used to search both EMBASE and MEDLINE.

Inclusion and exclusion criteria

Studies qualified for inclusion if they met the following predetermined inclusion criteria: (1) Study reported any accuracy metric of an AI-algorithm for ICH detection in NCCTs

Fig. 1 MEDLINE and EMBASE preliminary search terms

MEDLINE and EMBASE search:

1. exp Intracranial Hemorrhages/
2. ((Intracerebral or Intraparenchymal or Subarachnoid or Intraventricular or epidural or subdural or intracranial) and (hemorrhage* or haemorrhage* or hematoma*)).mp.
3. cerebral microbleed*.mp.
4. 1 or 2 or 3
5. exp Artificial Intelligence/
6. (Artificial intelligence or Computational Intelligence or Machine Intelligence or AI or Deep learning or Machine learning or convolutional neural network).mp.
7. 5 or 6
8. 4 and 7

or MRIs, (2) study reported any accuracy metric of an AI-algorithm for CMBs detection in MRI scans, (3) Abstracts (for example, conference proceedings) that met these criteria were also included. Studies were excluded in any of the following cases: (1) Study reported semi-automated algorithms and/or computer-assisted techniques, (2) Study reported AI-algorithms only for segmentation of ICH and did not report any accuracy metrics for ICH/CMBs detection. No limits were posed on the year of publication. During the full-text review, additional relevant studies were identified through the included references of each study.

Data endpoints

The purpose of this systematic review was to identify all intracranial ICH and CMBs detection algorithms that utilized AI to either a partial or full degree. Our primary outcome was to gather accuracy metrics for each algorithm, when available, including sensitivity (SN), specificity (SP), positive predictive value (PPV), negative predicted value (NVP), overall accuracy, and area under the curve (AUC).

Screening

Two independent reviewers (S Matsoukas, JS) first screened the studies based on their titles and abstracts. A full-text review was then conducted for the screened studies, and mutual agreement was reached by the reviewers (S Matsoukas, JS) regarding the final decision for inclusion.

Pooled analysis

The accuracy metrics of the included studies were pooled together, and descriptive statistics were provided separately for ICH vs. CMBs algorithms. For each of the data endpoints, the mean, standard deviation, standard error of the mean, range, and 95% confidence interval were calculated. All calculations and graphs were computed using Microsoft Excel 2019.

Results

Included studies

An initial search yielded 588 studies for title and abstract screening, 543 of which were deemed irrelevant and 45 underwent full-text review to be assessed for eligibility. Ultimately, 38 studies met our eligibility criteria and were complemented by 21 additional studies identified through citation search of the reviewed articles, thereby yielding 59 studies eligible for inclusion. Figure 2 depicts the 2020 PRISMA flow diagram for this study.

Almost all of the studies were published between 2015 and 2021. There was one study published in 2001, one in 2013, and one in 2014. In total, 40 of the included studies (25 full-text manuscripts and 15 abstracts) reported AI algorithms for ICH detection in NCCTs/MRIs and included accuracy metrics (Table 1). The earliest published study was in 2001 and used artificial neural networks (ANN) [12]. Most of these studies (18) trained and/or tested a 2D CNN AI algorithm. Three studies used a purely 3D CNN [13–15] (joint CNN-RNN in 2 of these [14, 15]), 2 studies used a hybrid 2D-3D CNN [16, 17], and 1 used an autoencoder [18]. Notably, there was a single study that used an AI algorithm able to identify ICH in MRI scans after an ischemic stroke [19]. The most commonly reported metrics were SN ($n=25/40$ studies, 62.5%), SP ($n=24$, 60%), AUC ($n=24$, 60%), and accuracy ($n=22$, 55%), while PPV and NPV were reported in only 6 studies (15%).

Nineteen studies (14 full-text manuscripts and 5 abstracts) were included that described AI algorithms for CMB detection in non-contrast MRI scans and reported accuracy metrics (Table 2). Within this group, 6 studies used 3D techniques [20–25]. The most commonly reported metrics were SN ($n=17/19$ studies, 89.5%), SP ($n=10$, 52.6%), and accuracy ($n=9$, 47.4%), while PPV and AUC were reported in 1 study (5.3%) and NPV in none. Tables 1 and 2 include the retrospective testing phase or real-life testing metrics. In the event a study reported only training/validation metrics, those were included instead. The full-length versions of tables 1 and 2 are more comprehensive and are available as supplemental material.

Bias assessment

Since all of the included studies were non-randomized cohorts, the ROBINS-I tool was utilized to assess several areas of bias across all studies. A visual representation of the potential biases is provided in Fig. 3a, b (ICH algorithms), and Fig. 4 (CMBs algorithms).

Pooled analysis

The pooled estimates of the mean, standard deviation, and 95% confidence intervals for the AI-driven ICH detection algorithms are as follows: SN: 92.06% (SD = 8.12%; 95% CI: 89.05–95.07%), SP: 93.54% (10.18%; 89.71–97.38%), PPV: 84.5% (10.87%; 75.8–93.2%), NPV: 96.87% (2.79%; 94.63–99.1%), accuracy: 93.46% (18.41%; 86.64–100.28%), AUC: 0.95 (0.05; 0.93–0.97). Likewise, for the CMB-detection algorithms the pooled accuracy metrics were as follows: SN: 91.6% (6.2%; 88.6–94.5%), SP: 93.9% (7.5%; 87.3–100.6%), accuracy: 92.7% (7.8%; 79.8–105.6%). Tables 3 and 4 summarize these findings. Since PPV, NPV,

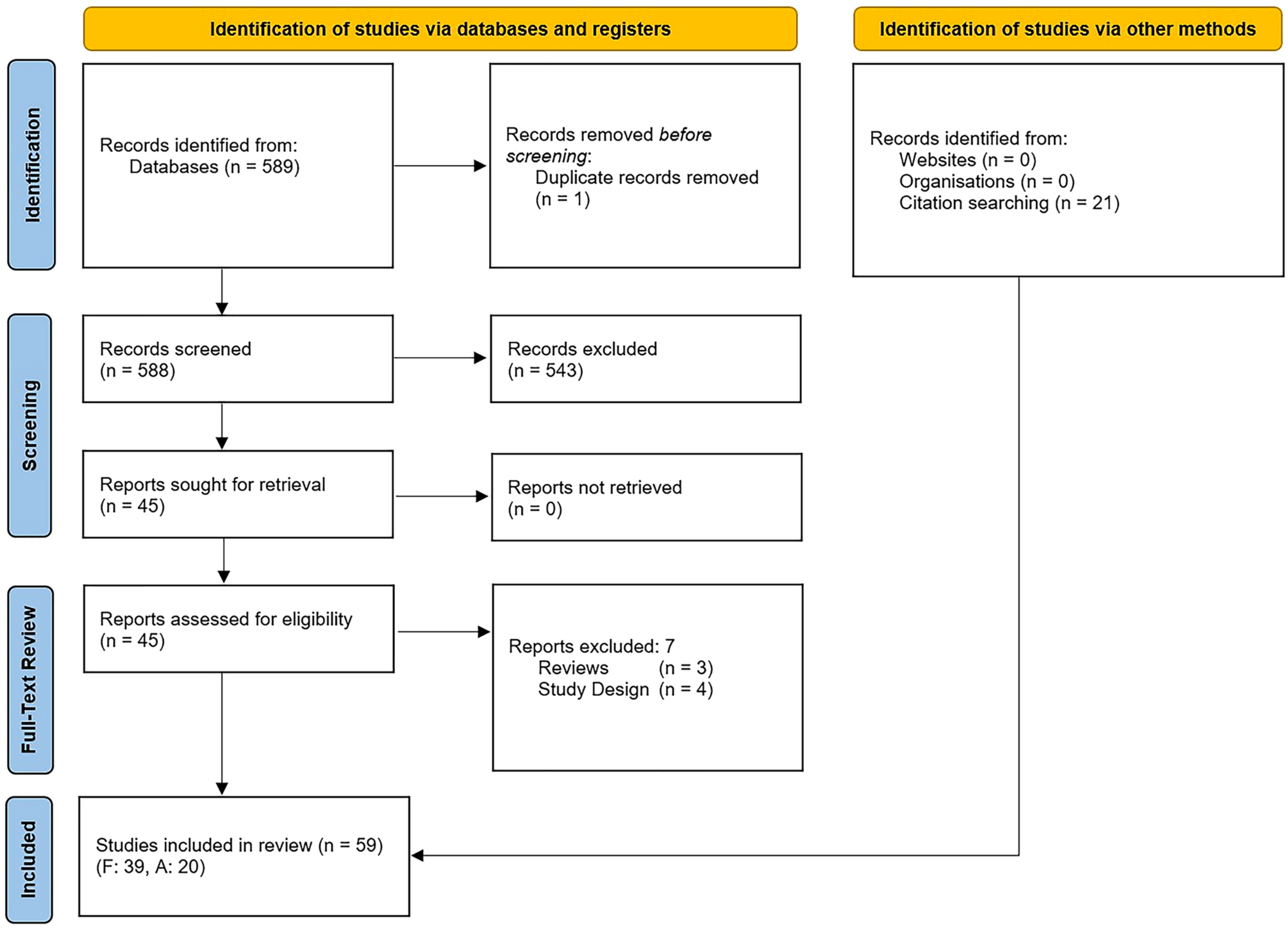


Fig. 2 PRISMA 2020 Flow Diagram, Footnote: A: Available only as an abstract (e.g., conference abstracts), F: Full-text manuscript is available

and AUC were not reported throughout the CMB-studies, it was not possible to calculate their polled estimates.

Figures 5 and 6 provide forest plots depicting the performance of each algorithm from the included studies, along with the corresponding pooled mean and 95% CI of each accuracy metric. Forest plots were designed only for accuracy metrics that were reported in the majority of the included studies.

Discussion

In this systematic review, the authors aimed to gather all the AI algorithms that have been developed for the detection of intracranial hemorrhage in NCCTs/MRIs or cerebral microbleeds in MRIs. The pooled analysis showed that the overall performance of such tools is high, achieving SN, SP, and accuracy > 90% for both ICH and CMBs detection. Furthermore, the NPV for ICH detection reached 97%. According to Kuo et al., SN and SP of experienced radiologists, approach 95–100% for the detection of ICH [26]. Therefore, AI detection algorithms have the potential to improve accuracy and facilitate the diagnosis of ICH detection in an emergent clinical situation [1]. Furthermore, additional AI tools can be used for the purpose of ICH volume quantification [27], ICH segmentation [28], or perihematomal edema quantification [29].

From the authors' perspective, these results highlight the potential of these tools to become a useful adjunct in triaging patients with possible ICH. Specifically, detection algorithms may act as a failsafe to decrease the number of false negatives or re-prioritize scans that need evaluation by radiology. In the narrative review that follows, the authors describe the current status of this technology, categorize the existing AI algorithms, and provide insight into their strengths and limitations.

Limited data availability

Developing and testing AI algorithms entails a complex process that requires very large patient datasets. Preparing these datasets manually is a time-consuming process. However,

Table 1 AI Algorithms for ICH Detection in NCCT/MRI Scans, with reported accuracy metrics

Author	Year	Text	Algorithm classification/method	Sample size	SN	SP	PPV	NPV	Accuracy	AUC
Sinha et al. [12]	2001	F	STATISTICA neural networks; ANNs	348 CTs	82.20%	96%	75.50%	97%	94.30%	0.94
Li et al. [32]	2012	F	Supervised learning	69 CTs	100.00%	92.40%	NR	NR	91.10%	NR
Desai et al. [31]	2017	F	GoogLeNet and AlexNet; 2D CNN	110 CTs	NR	NR	NR	NR	NR	1.00, 0.95
Phong et al. [47]	2017	F	LeNet, GoogLeNet, Inception-ResNet; 2D CNN	340 images	98.4%, 98.6%, 100%	99.4%, 98.1%, 99.3%	NR	NR	99.7%, 98.2% and 99.2%	0.999, 0.998, 1.0
Arbabshirani et al. [1]	2018	F	Stochastic gradient descent	347 "Routine" CTs	70%	87%	NR	NR	NR	0.846
Chang et al. [16]	2018	F	Hybrid 3D/2D mask ROI-based CNN	862 CTs (23,668 images)	95.10%	97.30%	82.90%	99.30%	97%	0.981
Chilamkurthy et al. [2]	2018	F	Qure.ai; 2D CNN	491 CTs	NR	NR	NR	NR	NR	0.94
Jnawali et al. [13]	2018	F	Tensorow 3D CNN (3 different)	40,367 CTs (> 1.5 million 2D images)	NR	NR	NR	NR	NR	0.87
Grewal et al. [15]	2018	A	DenseNet, DenseNet-A and RADnet:A architecture with 3D capabilities (stochastic gradient descent)	77 CTs	NR	NR	NR	NR	72.73%, 80.52%, 81.82%	NR
Majumdar et al. [33]	2018	A	Modified U-Net model	69 CTs	81.00%	98.00%	NR	NR	NR	NR
Ma et al. [19]	2018	A	Kernel spectral regression + neural networks for Hemorrhagic transformation of ischemic stroke in MRI scans	67 CTs	NR	NR	NR	NR	95.10%	NR
Barreira et al. [49]	2018	A	Viz-ICH; Fully Automated CNN	284 CTs	90.20%	99.99%	100%	92.20%	95.5%	0.95
Helwan et al. [18]	2018	A	Autoencoder	2527 images	NR	NR	NR	NR	90.9%	NR
Yune et al. [46]	2018	A	CNN		NR	NR	NR	NR	NR	0.994
Lee et al. [50]	2018	A	Window Setting Optimization (WSO) and CNN	904 CTs	NR	NR	NR	NR	NR	0.976
Lee et al. [30]	2019	F	2D CNN with multiple slices input	200 CTs	98%	95%	NR	NR	NR	0.99
Ginat [38]	2019	F	Aidoc; CNN	2011 CTs	88.70%	94.20%	73.70%	97.70%	93.4%	NR
Kuo et al. [26]	2019	F	Dilated ResNet 38; 2D CNN with multiple slices input, fully convolutional neural network, stochastic gradient descent	200	NR	NR	NR	NR	NR	0.991
Ye et al. [14]	2019	F	3D joint CNN-RNN	299 CTs	NR	NR	NR	NR	NR	0.98

Table 1 (continued)

Author	Year	Text	Algorithm classification/ method	Sample size	SN	SP	PPV	NPV	Accuracy	AUC
Dawud et al. [3]	2019	F	AlexNet, AlexNet with SVM subclassifier and authors' algorithm; CNN	3790 images	NR	NR	NR	NR	92.13%, 93.48%, 90.65%	NR
Ojeda et al. [39]	2019	A	Aidoc; CNN	7112 CTs	95%	99%	NR	NR	98%	NR
Cho et al. [42]	2019	F	Cascade of CNNs and dual FCNs—stochastic gradient descent (SGD) and adaptive moment estimation (ADAM)	5702 CTs (135,974 images)	97.91%	98.76%	NR	NR	98.28%	NR
Ker et al. [43]	2019	F	3D CNN	399 CTs (12,000 images)	93.80%	NR	NR	NR	NR	NR
Barros et al. [51]	2019	A	CNN for SAH detection	70 CTs	NR	NR	NR	NR	96%	NR
Patil et al. [52]	2019	A	Pretrained Inception-ResNet V2 model	100 CTs	90%	100%	NR	NR	91%	NR
Bizzo et al. [53]	2019	A	CNN	2209 CTs	NR	NR	NR	NR	NR	0.95
Hahm et al. [54]	2019	A	Cascade Deep Learning (CDL) automated detection algorithm	5702 CTs	NR	NR	NR	NR	NR	0.989
Li et al. [48]	2020	F	U-Net (Type of CNN)	226 images	NR	NR	NR	NR	NR	0.9859
Danilov et al. [41]	2020	F	ResNexT architecture	752,807 images/401 CTs	NR	NR	NR	NR	NR	0.81
Hssayeni et al. [36]	2020	F	FCN (U-Net)	82 CTs	97.20%	50.40%	NR	NR	NR	NR
Karki et al. [44]	2020	F	Window estimator module attached in deep CNN	2705 CTs	98.22%	98.54%	NR	NR	96.43%	0.994
Yi et al. [34]	2020	A	DAIANA	289 CTs	91.00%	94%	NR	NR	NR	0.97
Herweh et al. [55]	2020	A	Brainomix	160 CTs	91%	89%	NR	NR	NR	0.94
Heit et al. [17]	2020	F	Rapid-ICH; Hybrid 2D-3D CNN	308 CTs	95.60%	95.30%	95.60%	95.30%	NR	NR
Lee et al. [56]	2020	F	ANN	332 CTs	78%	80%	NR	NR	NR	0.859
Ko et al. [57]	2020	F	CNN and long-short term memory (LSTM)	727,392 images	NR	NR	NR	NR	93%	NR
Burduja et al. [58]	2020	F	CNN (ResNeXt-101) and Bidirectional LSTM	100 CTs	96%	98%	NR	NR	97%	0.97
Rao et al. [40]	2021	F	Aidoc; CNN	5585 CTs	NR	NR	NR	NR	NR	NR
Praveen et al. [59]	2021	A	Pretrained deep CNN: modified AlexNet-SVM classifier (framework)	CQ500 dataset (491 CTs)	99.86%	99.86%	NR	NR	99.86%	NR
Gautam et al. [45]	2021	F	Preprocessing method and CNN	43 CTs	NR	NR	NR	NR	98.33%	NR

For studies reporting accuracy metrics for multiple AI algorithms or for different phases of the algorithm development, only the highest number has been included in this table. Please, refer to the online supplemental material, for detailed information. *NR* not reported

A: Available only as an abstract (e.g., conference abstracts)

F: Full-text manuscript is available

a few studies have trained and tested algorithms on small datasets [30–34]. Transfer learning can help decrease the need for large datasets and can lead to the development of algorithms with higher accuracy compared to algorithms that have been developed “from scratch” [3]. The RSNA 2019 Brain CT hemorrhage challenge created a dataset using volumetric data that resulted in the production of the largest publicly available dataset for training AI algorithms. This dataset is comprised of 21,784 exams for training and validation purposes and 3528 for testing purposes (874,035 images in total) [35]. In another work, Hssayeni et al. created a publicly available dataset [36] that consisted of 82 scans (2173 normal slices, 24 IVH slices, 73 with IPH, 18 with SAH, 182 with EDH, and 56 with SDH). To solve this problem of limited data availability, data augmentation is a technique that is frequently utilized. It includes processes such as horizontal flipping, over or undersampling, or purposefully creating balanced datasets (not reflecting the disease’s true prevalence in the population) that function to help solve the “class imbalance’’ problem by providing better training opportunities to the algorithm [7].

AI Performance in the real-life setting

Arbabshirani et al. utilized a 3D deep convolutional neural network (CNN) composed of five convolutional layers and two fully connected layers [1]. The algorithm was initially validated with a dataset of 37,084 studies and then tested on an unseen subset dataset of 9499 studies. The AUC was 0.846 (95% CI: 0.837–0.856), and the authors selected an operating point with a corresponding SN and SP of 0.73 and 0.8, respectively, for prospective real-life testing. In real life, the algorithm yielded 70% SN, 87% SP and overall accuracy of 84%. More specifically, 94 out of 347 cases (26%) were upgraded from “routine’’ (no ICH was detected) to “stat’’ status (ICH was detected, and therefore, urgent review from a radiologist was required). As per the initial radiology dictation, only 60 of these had a hemorrhage. Finally, 5 of them (1.44% of the 347 cases) were deemed as having ICH. The median time to clinical interpretation was 19 min for the scans which were marked as “stat’’ by the algorithm versus 512 (96% reduction in time) for “routine’’ scans ($p < 0.0001$). Overall, AI has been successfully utilized to triage CT scans as urgent or non-urgent by detecting acute neuroradiographic findings [37]. Ginat et al. found that an AI software performed differently in inpatient, outpatient, and emergency cases, having the highest accuracy in emergency cases, which can be attributed to the presence of confounding features in follow-up imaging compared to the acute trauma scans [38]. Among different ICH subtypes, the best accuracy was noted for IPH [38].

Aidoc (Tel Aviv, Israel) is a commercially available, FDA-approved CNN AI algorithm, which was originally trained with a dataset of 50,000 NCCTs from 9 institutions and 17 different scanners during its development [39]. Ojeda et al. retrospectively tested Aidoc in the clinical setting, reporting 95% SN, 99% SP, and 98% accuracy [39]. Reo et al. tested the same algorithm in 5585 real-life retrospective cases, defined as ICH-negative by an automated Natural Language Processing (NLP) algorithm [40]. The AI algorithm was able to detect the 1.6% false negative rate of the radiologists (16 out of 996). The false positive rate of the AI algorithm was reported as 32% [40]. However, due to the design of the study, accuracy, AUC, and other metrics were not calculated. Including the NLP’s performance in the analysis of the total cases, the authors may have ultimately missed “negative by impression’’ cases that were identified as positive by the NLP algorithm. These factors limit this study’s capacity to identify the real sensitivity of the algorithm, but the 1.6% false negative rate of the radiologists is congruent with what other studies have reported, according to the authors [40].

AI in ICH detection and classification

Multiple studies have reported algorithms that can classify ICH according to its subtypes when detected [2, 14, 30, 41–46]. Lee et al. developed and tested a model able to both detect ICH and also classify it into its subtypes (IPH, IVH, SAH, EDH, SDH) [30]. The algorithm yielded 98% SN, 95% SP, and 0.99 AUC in the retrospective dataset, and 92.4% SN, 94.9% SP, and 0.96 AUC in the prospective dataset of the study. Chilamkurthy et al. developed and tested an algorithm in two retrospective datasets coming from different hospitals [2]. In the first one ($n = 21{,}095$), AUC for overall cases was 0.92 and, more specifically, 90% for IPH, 96% for IVH, 92% for SDH, 93% for EDH and 90% for SAH. The numbers for the second dataset ($n = 491$) were 0.94, 95%, 93%, 95%, 97%, and 96%, respectively. This was one of the first studies reporting different accuracies for the 5 different types of hemorrhages. Danilov et al. developed a similar algorithm for classifying ICH with an overall accuracy of 81% [41].

Cho et al. used 2 CNNs and 2 Fully Convolutional Networks (FCNs) to develop an algorithm able to detect and segment (achieved by the use of FCNs) ICH. The algorithm was trained on 135,974 NCCT images (slices), twice: once with the default window and once with the stroke window [42]. Utilizing this technique, false negatives decreased, ultimately resulting in a 1% increase in sensitivity without decreasing specificity [42]. Furthermore, ADAM (adaptive moment estimation) optimization and transfer-learning utilization yielded a higher accuracy (98.28%), SN (97.91%), and SP (98.76%). ICH classification accuracy ranged from 70 to 90%. Taking this one step further, Karki et al. applied a window estimator module (WEM) on a deep CNN algorithm.

Table 2 AI Algorithms for CMBs Detection in MRI Scans, with reported accuracy metrics

Author	Year	Text	Algorithm classification/method	N	SN	SP	PPV	NPV	Accuracy	AUC
Fazlollahi et al. [64]	2015	F	Cascade of binary random forests	66 MRIs	92.65%	99%	NR	NR	NR	NR
Dou et al. [65]	2015	F	Convolutional Independent Subspace Analysis; Unsupervised learning	19 MRIs (161 CMBs)	89.44%	NR	NR	NR	NR	NR
Chen et al. [22]	2015	A	3D CNN trained with stochastic gradient descent	5 MRIs (55 CMBs)	89.13%	NR	NR	NR	NR	NR
Roy et al. [66]	2015	F	MRST + RF	26 MRIs	85.70%	99.50%	NR	NR	NR	NR
Dou et al. [21]	2016	F	3D FCN	50 MRIs (117 CMBs)	93.16%	NR	NR	NR	NR	NR
Hou et al. [67]	2016	A	4-Layer DNN	10 MRIs	93.45%	93.05%	NR	NR	93.23%	NR
Lu et al. [68]	2017	A	CNN	2000 MRIs	97.29%	92.23%	NR	NR	96.05%	NR
Wang et al. [69]	2017	F	5-Layer RAP based CNN	68,847 CMB voxels and 56,582,536 non-CMB voxels	96.94%	97.18%	NR	NR	97.18%	NR
Morrison et al. [70]	2018	F	2D Fast Radial Symmetry Transform	10	86.70%	NR	NR	NR	NR	NR
Tao and Cloutie [71]	2018	F	Genetic Algorithm and BPNN (Back Propagation Neural Network)	69,365 CMB voxels and 69,327 non-CMB voxels	72.90%	72.89%	NR	NR	72.90%	NR
Gunter et al. [72]	2018	A	Conventional image processing followed by a CNN AI Algorithm	4718 pairs	NR	NR	NR	NR	NR	0.974
Zhang et al. [73]	2018	F	SNP + SLFN + LReLU (slicing neighborhood processing, single-hidden layer feed-forward neural-network, leaky rectified linear unit)	10 CADASIL and 10 healthy MRIs (68,847 CMB voxels and 68,854 non-CMB voxels)	93.05%	93.06%	93.06%	NR	NR	NR
Zhang et al. [74]	2018	F	7-Layer SAE DNN (Sparse Autoencoder)	10 CADASIL and 10 healthy MRIs (68,847 CMB voxels and 68,829 non-CMB voxels)	95.13%	93.33%	NR	NR	94.23%	NR
Chen et al. [20]	2019	F	3D Deep Residual Network	12 MRIs	94.70%	NR	NR	NR	90.10%	NR
Wang et al. [62]	2019	F	DenseNet 201 (Densely connected neural network)	10,000 CMB voxels and 10,000 non-CMB voxels	97.78%	97.64%	NR	NR	97.71%	NR
Liu et al. [25]	2019	F	3D Fast Radial Symmetry Transform (FRST) AND A Deep Learning Algorithm	41 MRIs (168 CMBs)	NR	NR	NR	NR	95.80%	NR
Wang et al. [63]	2020	F	CNN with stochastic pooling	20,000 Voxels	97.22%	97.35%	NR	NR	97.28%	NR
Al-Masni et al. [23, 24]	2020	F, A	YOLO and 3D-CNN	107 subjects with 572 CMBs	94.32%	NR	NR	NR	NR	NR

For studies reporting accuracy metrics for multiple AI algorithms or for different phases of the algorithm development, only the highest number has been included in this table. Please, refer to the online supplemental material, for detailed information. *NR* not reported

A: Available only as an abstract (e.g., conference abstracts)

F: Full-text manuscript is available

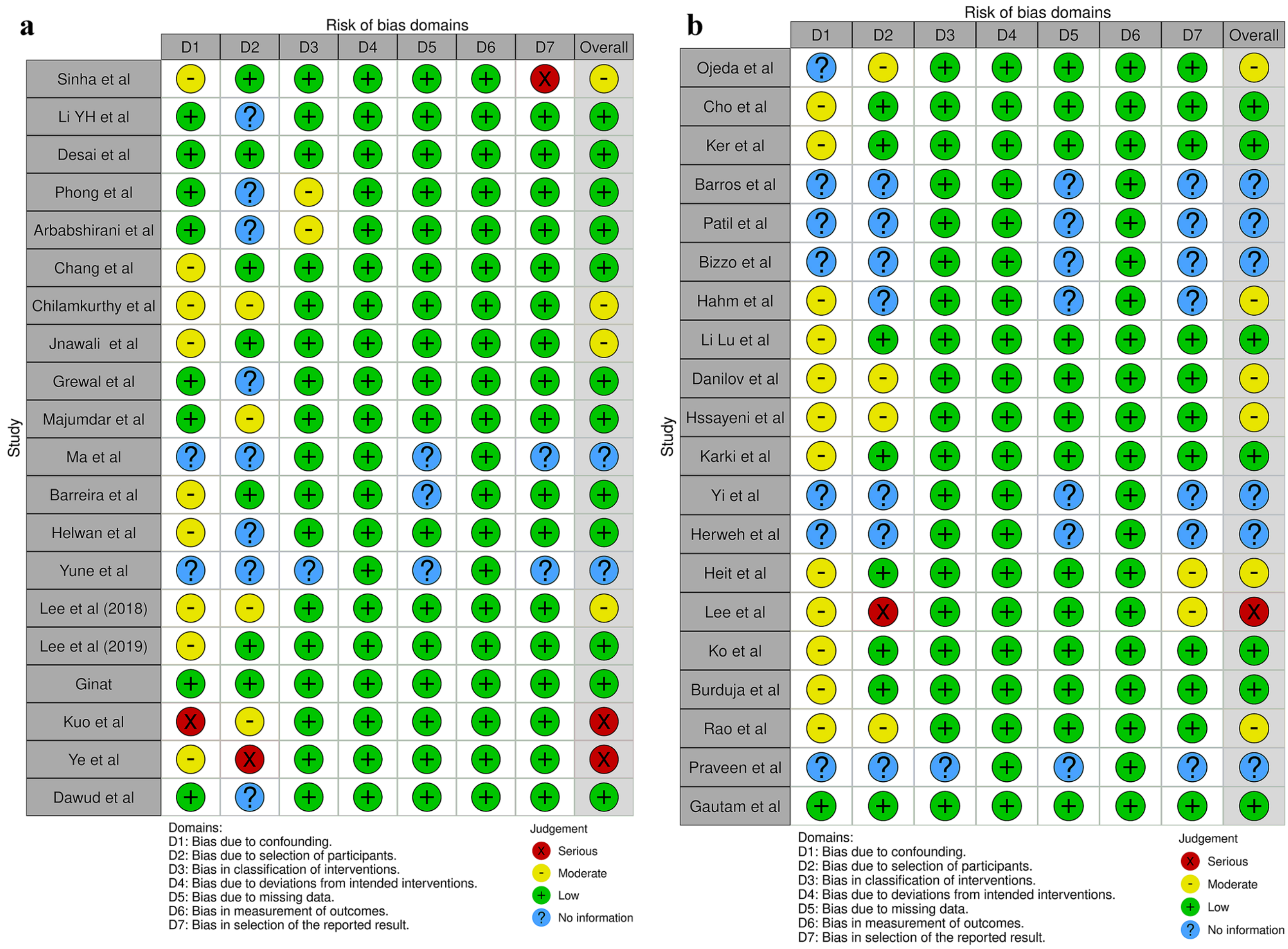


Fig. 3 Bias assessment with the ROBINS-I tool, for studies that reported ICH detection AI-algorithms (**a** Studies 1–20, **b** Studies 21–40)

They compared sensitivity and specificity for ICH detection and IPH, IVH, EDH, SDH, and SAH detection for 3 different window modalities: single window (default or mean), flexible window (initialized or initialized method), and combination window (cascade or aggregate method) [44]. The single window had higher sensitivity for ICH detection and higher specificity for SDH and SAH detection. The flexible window achieved better sensitivity for IPH, IVH, SDH, and SAH detection and better specificity in IVH and EDH detection.

Data augmentation and transfer learning

Utilizing an augmented dataset or pre-training an AI algorithm is common methods for enhancing accuracy. In a study performed by Desai et al., the Augmented GoogLeNet-Pretrained algorithm outperformed the Augmented AlexNet-Untrained algorithm with an AUC of 1.00 versus 0.95, respectively ($p=0.001$), for basal ganglia ICH detection [31]. However, the small sample size, lack of generalizability for all ICHs, and the large volume of each hemorrhage were important limiting factors. Phong et al. compared 3 CNNs after performing data augmentation. Inception-ResNet achieved the highest AUC (1.0) and sensitivity (100%), while LeNet achieved the highest accuracy (99.7%). However, LeNet was considerably slower than the other two (7940 s versus 879 s for GoogLeNet and 867 s for Inception-ResNet) [47]. Li et al. developed a U-net deep learning algorithm with the ability to create and process a mirror image for each slice, and then concatenate it with the original image, based on the principle that hemorrhages do not occur at the same points bilaterally (i.e., that there is asymmetry) [48]. Compared to other contemporary algorithms, the one developed by Li et al. marked a notably higher ability to identify and segment hemorrhages, as evidenced by the Dice score (> 0.4) and the precision–recall curve [48].

Transfer learning is a technique developed to manage the need for large datasets. Dawud et al. tested AlexNet with transfer learning, AlexNet with SVM classifier (support

Fig. 4 Bias assessment with the ROBINS-I tool, for all studies that reported CMBs detection AI-algorithms

vector machine) and their own new algorithm on the same dataset [3]. Ultimately, the pre-trained AlexNet, where transfer learning and new fine-tuning were used, performed significantly better.

Hybrid CNNs and 3D CNNs

Utilization of CNNs (including 2D, 3D, Hybrid 2D/3D, or joint CNN-RNNs) is contended to be ideal for 3D volumetric assessment of NCCT scans [7]. Due to the volumetric nature of NCCT examinations, a combination of CNNs and RNNs has been proposed to solve the challenge of assessing interslice dependencies [14, 15], an ability that 2D CNNs lack [7]. Grewal et al. compared accuracy metrics between a 2D CNN and a joint 2D CNN-RNN [15]. The joint CNN-RNN achieved slightly better accuracy in detecting ICH. Hybrid 2D-3D CNNs can achieve high accuracy [16, 17]. Chang et al. developed, validated, and tested a hybrid 3D/2D region-based CNN algorithm able to both detect and quantify ICH [16]. Accuracy and AUC reached 0.975 and 0.983 in the cross-validation group and 0.97 and 0.981 in the real-life group, respectively, thus outperforming some of the highest accuracies of 2D CNNs. Notably, the algorithm demonstrated greater accuracy for hemorrhages > 5 mL compared to hemorrhages < 5 mL, which was in turn higher than the accuracy for detecting punctuate hemorrhages (< 0.01 mL) [16]. The most difficult to detect bleeds were punctuate subarachnoid ones as well as epidural and/or subdural ones (with accuracy ranging between 0.83 and 0.881).

Because 3D CNNs utilize 3D information as their input, they constitute ideal AI algorithms for ICH detection in NCCTs. However, they require significantly larger datasets for training, which can be especially problematic with supervised learning that requires labeled datasets. Natural Processing Language (NLP) applications serve to help resolve this problem [1, 37]. However, 2D CNNs [2, 26, 30, 47] and CNN-RNNs [13, 15] were able to outperform available 3D CNNs [1, 37] on ICH detection in all but two cases [14, 43]. Interestingly, combining 3 different 3D CNNs with different AUCs (0.85, 0.85, 0.86) into one 3D CNN resulted in a slightly enhanced AUC (0.87) [13]. Ker et al. applied image thresholding in the training and testing phases of a 3D CNN algorithm [43]. As a result, not only was training time significantly decreased, but sensitivities and F1 scores

Table 3 Pooled estimates of performance for ICH detection

Pooled estimates of performance for ICH detection						
Metric	Sensitivity	Specificity	PPV	NPV	Accuracy	AUC
Mean	92.06%	93.54%	84.50%	96.87%	93.46%	0.95
SD	8.12%	10.18%	10.87%	2.79%	18.41%	0.05
SEM	3.01%	3.84%	8.70%	2.23%	6.82%	0.02
95% CI upper bound	95.07%	97.38%	93.20%	99.10%	100.28%	0.97
95% CI lower bound	89.05%	89.71%	75.80%	94.63%	86.64%	0.93
Max	100.00%	100.00%	100.00%	99.70%	99.86%	1.00
Min	70.00%	50.40%	73.70%	92.20%	72.73%	0.81

SD standard deviation, *SEM* standard error of the mean, *CI* confidence interval, *PPV* positive predictive value, *NPV* negative predicted value, *AUC* area under the curve

Table 4 Pooled estimates of performance for CMBs detection

Pooled estimates of performance for CMBs detection			
Metric	Sensitivity	Specificity	Accuracy
Mean	91.6%	93.9%	92.7%
SD	6.2%	7.5%	7.8%
SEM	2.9%	6.7%	12.9%
95% CI Upper Bound	94.5%	100.6%	105.6%
95% CI Lower Bound	88.6%	87.3%	79.8%
Max	97.8%	99.5%	97.7%
Min	72.9%	72.9%	72.9%

SD standard deviation, *SEM* standard error of the mean, *CI* confidence interval

improved radically for SAH and acute SDH detection [43]. Ye et al. trained and tested a 3D joint CNN-RNN, ultimately achieving exceedingly high accuracy (0.98) and outperforming junior radiologists [14].

Chronic cerebral microbleeds

Chronic cerebral microbleeds are small-size parenchymal hemorrhages that are most likely caused by structural abnormalities of small cerebral vessels, especially in the context of underlying cerebrovascular pathology. Therefore, their detection offers risk stratification for future ischemic strokes, ICH, and dementia [60]. AI-driven algorithms seek to ease and improve the laborious process of manually detecting CMBs. Since the ImageNet Large Scale Visual Recognition Challenge in 2012 [61], deep CNNs have demonstrated superior accuracy to classic computer vision methods in computer image classification.

The most challenging task in CMBs detection is reducing the number of false positives (FPs) that can be caused by many sources of susceptibility on SWI. These include calcification, arteries with poor flow compensation, artifacts induced by air–tissue interfaces, and veins with reduced oxygen saturation due to impaired blood perfusion. The most common approach to the latter is to use a two-stage process: First, potential CMB candidates are detected on SWI, and second, false positives are reduced using a CNN.

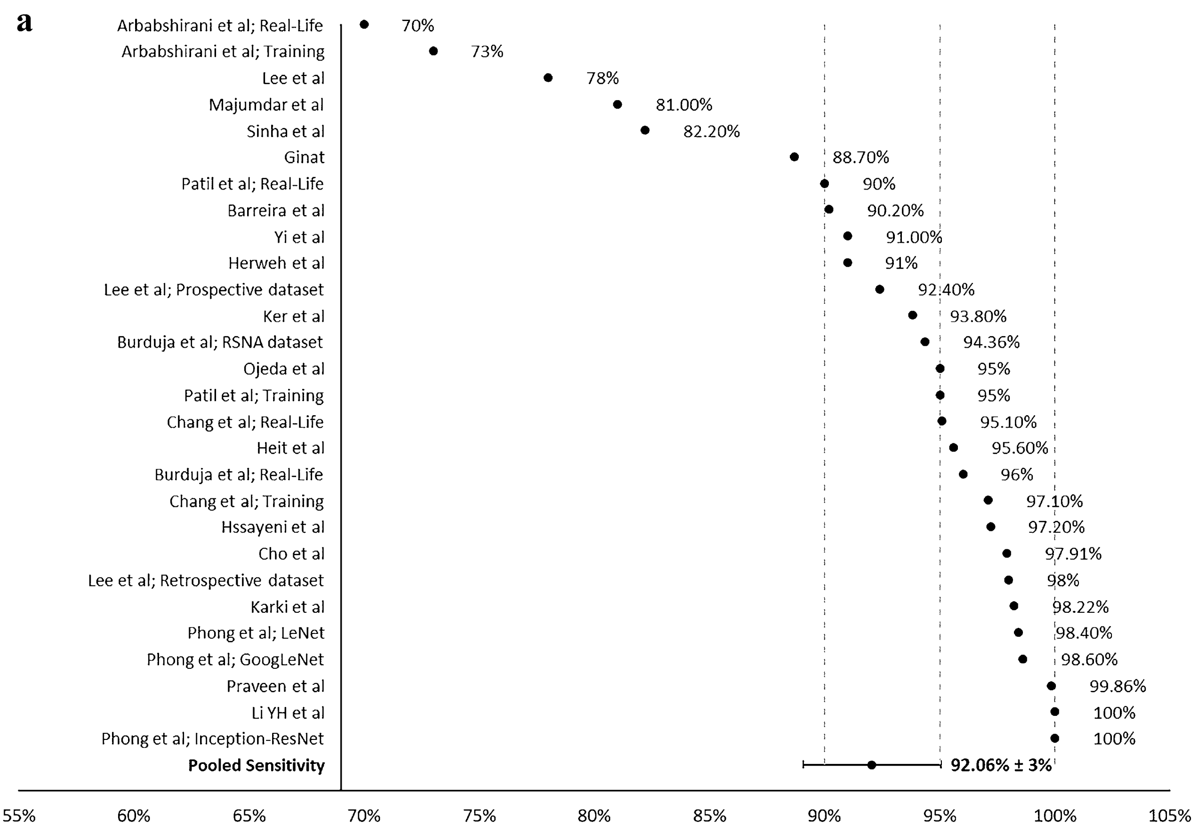
Chen et al. used an auto-encoder for voxel detection of CMBs based on 2D fast radial symmetry transform (FRST) integrated with 3D-ResNet using 7.0 Tesla SWI images. They obtained 94.69% SN and 71.98% precision with a rate of 11.6 false positives per subject (FPavg) at an overall processing time of 2 min [20]. The results of the algorithm were equivalent to the interpretation of radiologists. Similarly, Liu et al. selected CMB candidates based on 3D-FRST of the composite images from SWI and phase images [25]. Then, using a 3D deep learning residual network (3DResNet) for the FPs reduction stage, they achieved 95.80% SN, 70.90% precision, and a 1.6 FPavg rate. Of note, they reported an FP/CMB ratio of 0.39, a normalized metric that is independent of the number of CMB per subject and is helpful in the comparison of different models.

More recently, Al-Masni et al. proposed a fully automated two-stage approach based on the combination of the recent regional-based You Only Look Once (YOLO) technique for potential CMBs candidate detection and a 3D CNN stage for false positives reduction. This approach resulted in 93.62% SN and a 1.42 FPavg rate [23, 24]. Interestingly, using YOLO at the detection stage only required 0.69 s per subject to generate potential candidates, which is 47.9 times faster than the 3D-FRST [23, 24]. On the contrary, Wang et al. proposed to employ DenseNet as the basic algorithm for transfer learning to detect CMBs considering the reported superiority to ResNet on ImageNet classification task [62]. First, three neuroradiologists were employed to mark the CMBs from the subjects manually, then based on DenseNet. Their method achieved 97.78% SN, 97.64% SP, 97.71% accuracy, and 97.65% precision, outperforming previous algorithms' diagnostic accuracy. Lastly, a CNN with stochastic pooling was used by Wang et al., achieving 97.22% SN, 97.35% SP, 97.28% accuracy, and 97.35% precision [63].

Limitations

The inability of the AI algorithms to always provide the physician with a rationale for a given diagnosis/result [16] is a current limitation of applying AI to real-world clinical scenarios. In addition, while it is important to create balanced datasets for training purposes, engineers should consider validating the algorithm on a cohort that is representative of the clinical context in which the algorithm is being implemented [16]. This translates to validating the algorithm with a dataset where ICH prevalence is 10–15%, a rate approaching that of larger hospitals [26].

A possible limitation of this study is the fact that ICH was defined as intracranial hemorrhage and not intracerebral hemorrhage. This was mainly done to reflect the definition that was provided by the majority of the included studies. Most of these studies are published in engineering journals and considered EDH and SDH when developing algorithms. Furthermore, not all accuracy metrics are consistently reported in studies, which is reflected in the included forest plots.



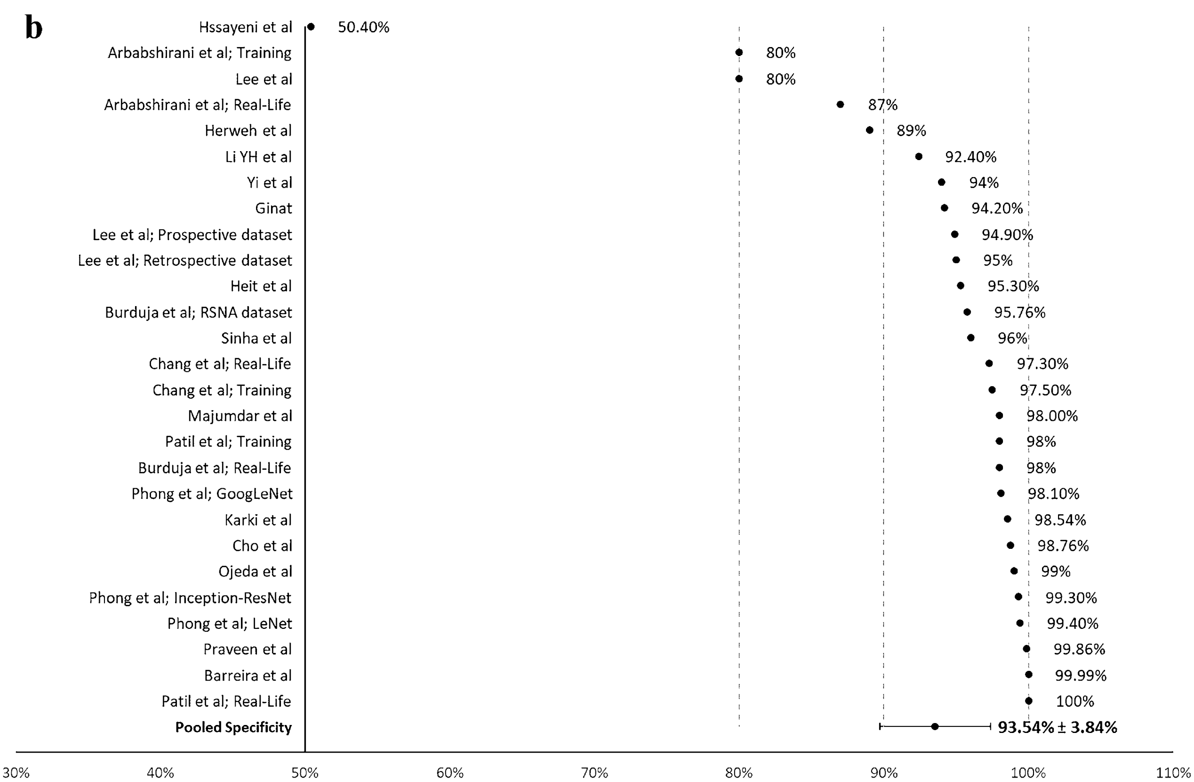


Fig. 5 Forest plots of accuracy metrics (reported and pooled) for studies that reported ICH detection AI-algorithms. **a** Sensitivity, **b** Specificity, **c** Accuracy, **d** Area under the curve (AUC)

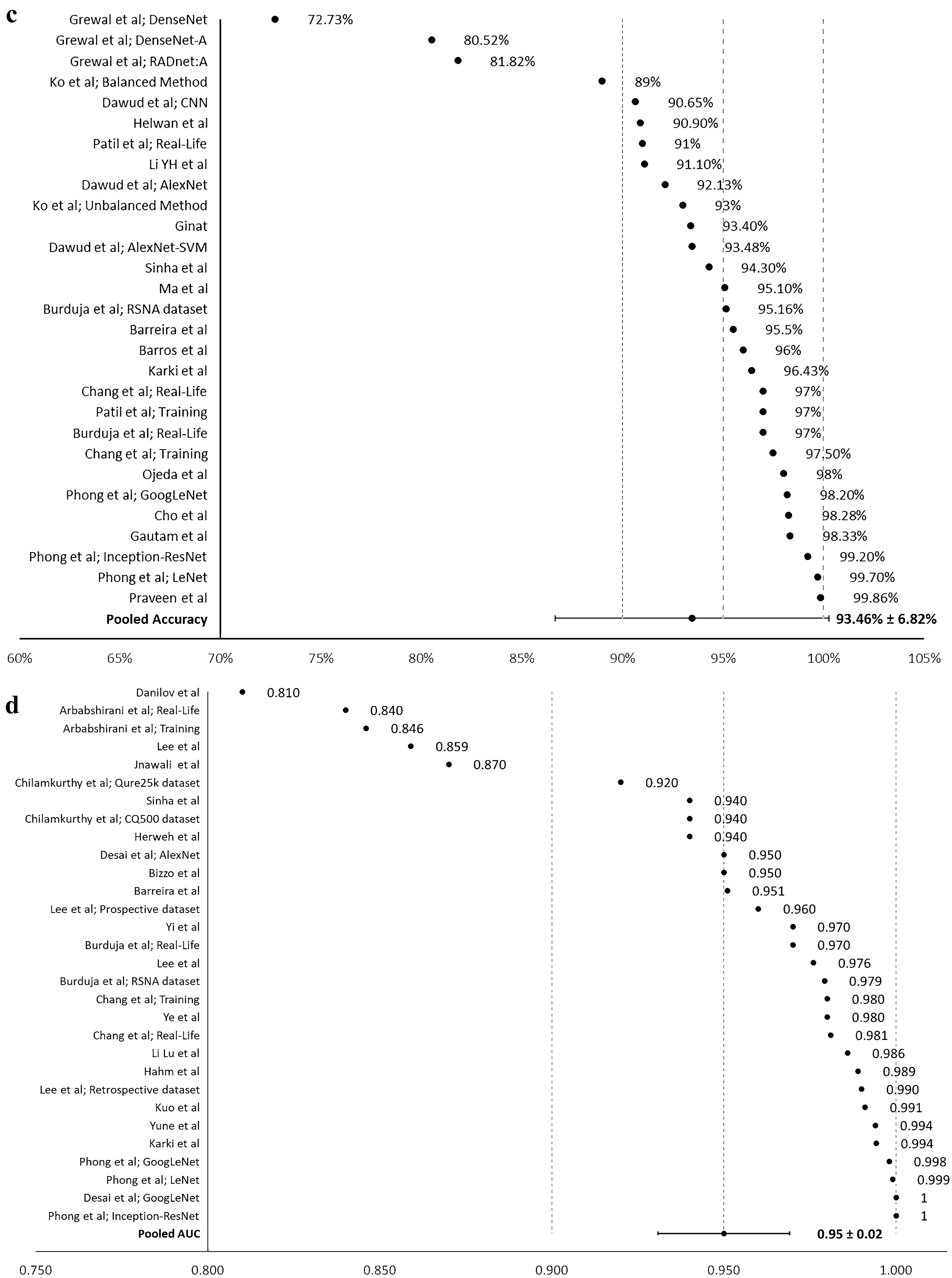


Fig. 5 (continued)

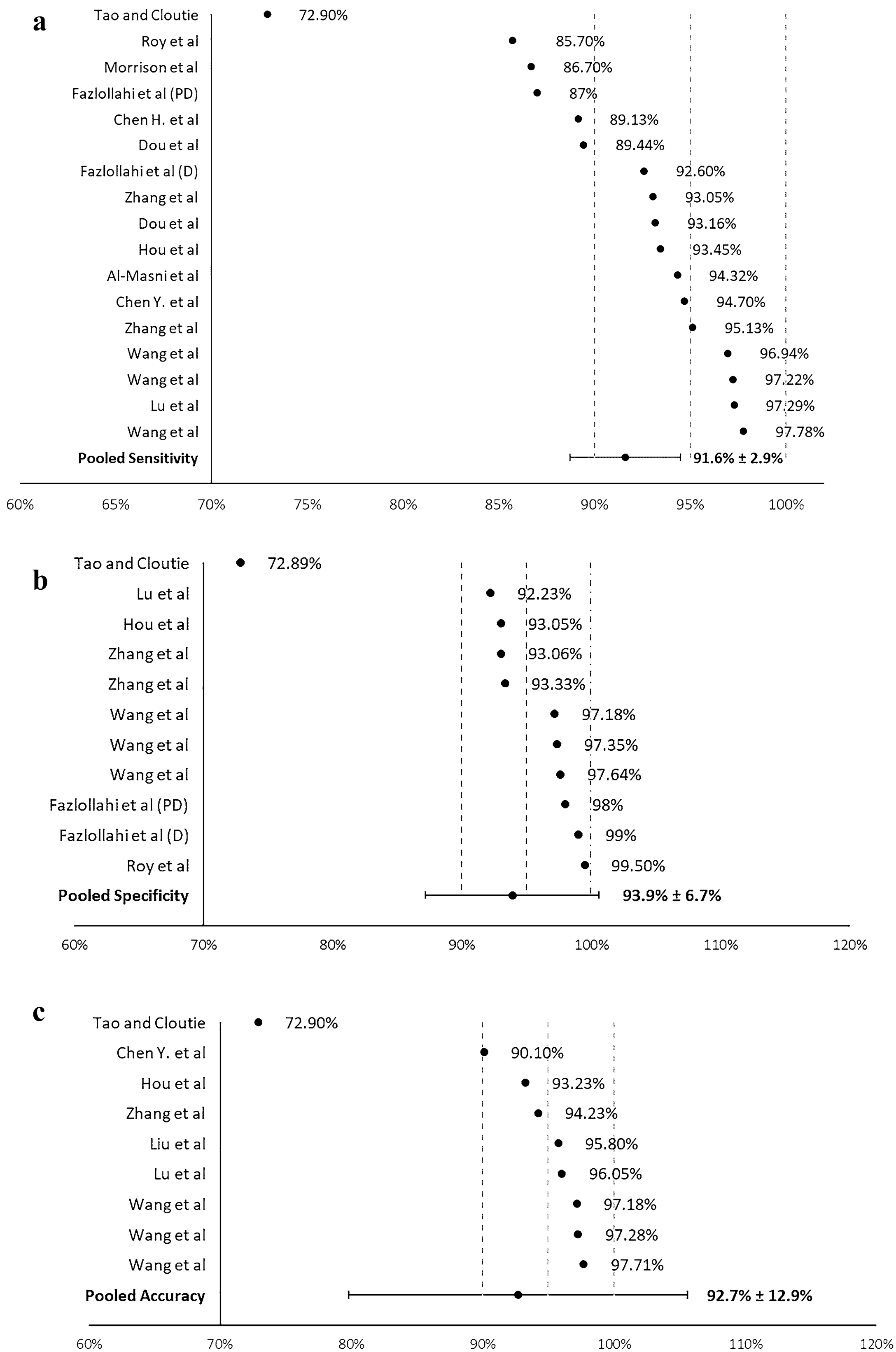


Fig. 6 Forest plots of accuracy metrics (reported and pooled) for studies that reported CMBs detection AI-algorithms. **a** Sensitivity, **b** Specificity, **c** Accuracy

Conclusions

In this systematic review, all the AI algorithms that have been developed to detect intracranial hemorrhage in non-contrast CTs or CMBs in MRIs were gathered. There has been an exponentially increasing availability of such tools with progressively improving performance that may be equivalent to that of experienced radiologists. If utilized properly, such tools may facilitate the timely diagnosis of any ICH subtype in real clinical scenarios by acting as a fail-safe to decrease the number of false negatives or re-prioritize scans that need evaluation by radiology.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s11547-022-01530-4.

Acknowledgements Work was conceptualized by S. Matsoukas and CK. S. Matsoukas and JC performed screening. S. Matsoukas conducted the analysis and generated the tables and graphs. S. Matsoukas drafted and developed Introduction, Methods, Results, and Discussion regarding ICH. CMB part of the discussion was drafted and developed by JS. S. Matsoukas and CK significantly edited the manuscript. Bias assessment was conducted by S Matsoukas and BRS. All authors approved the final version of the article.

Funding There were no funding sources for this study.

Availability of data and material Submitted as supplemental material; Less lengthier tables are included in the manuscript.

Declarations

Conflicts of interest J Mocco is the PI on research trials funded by: Stryker Neurovascular, Microvention, Penumbra, and Genentech and he is an investor in: Cerebrotech, Imperative Care, Endostream, Viseon, BlinkTBI, Myra Medical, Serenity, Vastrax, NTI, RIST, Viz.ai, Synchron, Radical, and Truvic. He serves, or has recently served, as a consultant for: Imperative Care, Cerebrotech, Viseon, Endostream, Vastrax, RIST, Synchron, Viz.ai, Perflow, and CVAid. Cristopher Kellner is the PI on research trials supported by Penumbra, Integra Life Sciences, and Cerenovus; he has received research grants from Viz. AI, Penumbra, Integra LifeSciences, ICE Neurosystems, Minnetronix, Irras, Longeviti Neuro Solutions, Cerebrotech Medical Systems, and Siemens; he has an ownership stake in Borealis, Precision Recovery, and Metis Innovative. Metis Innovative is a venture capital group with investments in Synchron, Fluid Biomed, and Proprio.

Authorship Clarifications BRS has been added as an extra author (3rd), since he contributed significantly during the revision.

References

1. Arbabshirani MR, Fornwalt BK, Mongelluzzo GJ, Suever JD, Geise BD, Patel AA, Moore GJ (2018) Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. NPJ Digit Med 4(1):9. https://doi.org/10.1038/s41746-017-0015-z (**PMID: 31304294; PMCID: PMC6550144**)
2. Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, Mahajan V, Rao P, Warier P (2018) Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. Lancet 392(10162):2388–2396. https://doi.org/10.1016/S0140-6736(18)31645-3 (**Epub 2018 Oct 11 PMID: 30318264**)
3. Dawud AM, Yurtkan K, Oztoprak H (2019) Application of deep learning in neuroradiology: brain haemorrhage classification using transfer learning. Comput Intell Neurosci 3(2019):4629859. https://doi.org/10.1155/2019/4629859 (**Erratum in: Comput Intell Neurosci. 2020 Aug 28; 2020: 4705838. PMID: 31281335; PMCID: PMC6589279**)
4. Park SH, Han K (2018) Methodologic guide for evaluating clinical performance and effect of artificial intelligence technology for medical diagnosis and prediction. Radiology 286(3):800–809. https://doi.org/10.1148/radiol.2017171920
5. Kaka H, Zhang E, Khan N (2020) Artificial intelligence and deep learning in neuroradiology: exploring the new frontier. Can Assoc Radiol J. https://doi.org/10.1177/0846537120954293 (**Epub ahead of print. PMID: 32946272**)
6. Gupta R, Krishnam SP, Schaefer PW, Lev MH, Gilberto GR (2020) An east coast perspective on artificial intelligence and machine learning: part 1: hemorrhagic stroke imaging and triage. Neuroimaging Clin N Am 30(4):459–466. https://doi.org/10.1016/j.nic.2020.07.005 (**Epub 2020 Sep 17. PMID: 33038996**)
7. Yeo M, Tahayori B, Kok HK et al (2021) Review of deep learning algorithms for the automatic detection of intracranial hemorrhages on computed tomography head imaging. J NeuroInterventional Surg 13(4):369–378. https://doi.org/10.1136/neurintsurg-2020-017099
8. Zhu G, Jiang B, Chen H et al (2020) Artificial intelligence and stroke imaging: a west coast perspective. Neuroimaging Clin N Am 30(4):479–492. https://doi.org/10.1016/j.nic.2020.07.001
9. Page MJ, McKenzie JE, Bossuyt PM et al (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. J Clin Epidemiol 134:178–189. https://doi.org/10.1016/j.jclinepi.2021.03.001
10. Moher D, Shamseer L, Clarke M et al (2015) Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev. https://doi.org/10.1186/2046-4053-4-1
11. Matsoukas S, Scaggiante J, Kellner C (2021) Artificial intelligence algorithms for identification of intracranial hemorrhagein non-contrast CT and MRI scans: a systematic review. PROSPERO 2021 CRD42021246848. Available from: https://www.crd.york.ac.uk/prospero/display_record.php?ID=CRD42021246848
12. Sinha M, Kennedy CS, Ramundo ML (2001) Artificial neural network predicts CT scan abnormalities in pediatric patients with closed head injury. J Trauma 50(2):308–312. https://doi.org/10.1097/00005373-200102000-00018 (**PMID: 11242297**)
13. Jnawali K, Arbabshirani M, Rao N, A Patel A (2018) Deep 3D convolution neural network for CT brain hemorrhage classification. In: Proceedings medical imaging 2018: computer-aided diagnosis, vol 10575, 105751C. https://doi.org/10.1117/12.2293725
14. Ye H, Gao F, Yin Y, Guo D, Zhao P, Lu Y, Wang X, Bai J, Cao K, Song Q et al (2019) Precise diagnosis of intracranial hemorrhage and subtypes using a three-dimensional joint convolutional and recurrent neural network. Eur Radiol 29(11):6191–6201. https://doi.org/10.1007/s00330-019-06163-2 (**Epub 2019 Apr 30. PMID: 31041565; PMCID: PMC6795911**)
15. Abstract: Grewal M, Srivastava MM, Kumar P, Varadarajan S (2018) RADnet: radiologist level accuracy using deep learning for hemorrhage detection in CT scans. In: Proceedings of the 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018), Washington, DC, USA, 4–7 April 2018, pp 281–284

16. Chang PD, Kuoy E, Grinband J, Weinberg BD, Thompson M, Homo R, Chen J, Abcede H, Shafie M, Sugrue L et al (2018) Hybrid 3D/2D convolutional neural network for hemorrhage evaluation on head CT. AJNR Am J Neuroradiol 39(9):1609–1616. https://doi.org/10.3174/ajnr.A5742 (**Epub 2018 Jul 26. PMID: 30049723; PMCID: PMC6128745**)
17. Heit JJ, Coelho H, Lima FO et al (2021) Automated cerebral hemorrhage detection using RAPID. AJNR Am J Neuroradiol 42(2):273–278
18. Abstract: Helwan A, El-Fakhri G, Sasani H et al (2018) Deep networks in identifying CT brain hemorrhage. IFS 35:2215–2228
19. Abstract: Ma SJ, Yu S, Liebeskind DS, Yan L, Wang DJ, Scalzo F (2018) Abstract WP60: Kernel spectral regression and neural networks enable regional detection of hemorrhagic transformation on multi-modal MRI for acute ischemic stroke. Stroke 49(Suppl_1). https://doi.org/10.1161/str.49.suppl_1.wp60
20. Chen Y, Villanueva-Meyer JE, Morrison MA, Lupo JM (2019) Toward automatic detection of radiation-induced cerebral microbleeds using a 3D deep residual network. J Digit Imaging 32(5):766–772. https://doi.org/10.1007/s10278-018-0146-z (**Erratum in: J Digit Imaging. 2019 Feb 8; PMID: 30511280; PMCID: PMC6737152**)
21. Dou Qi, Chen H, Lequan Yu et al (2016) Automatic detection of cerebral microbleeds from MR images via 3D convolutional neural networks. IEEE Trans Med Imaging 35(5):1182–1195. https://doi.org/10.1109/TMI.2016.2528129
22. Abstract: Chen H, Yu L, Dou Q, Shi L, Mok VCT, Heng PA (2015) Automatic detection of cerebral microbleeds via deep learning based 3D feature representation. In: 2015 IEEE 12th international symposium on biomedical imaging (ISBI), New York, NY, pp 764-767. https://doi.org/10.1109/ISBI.2015.7163984
23. Al-Masni MA, Kim WR, Kim EY, Noh Y, Kim DH (2020) Automated detection of cerebral microbleeds in MR images: a two-stage deep learning approach. Neuroimage Clin 28:102464. https://doi.org/10.1016/j.nicl.2020.102464
24. Abstract: Al-masni MA, Kim W-R, Kim EY, Noh Y, Kim D-H (2020) A two cascaded network integrating regional-based YOLO and 3D-CNN for cerebral microbleeds detection. In: 2020 42nd Annual international conference of the IEEE engineering in medicine & biology society (EMBC), Montreal, QC, Canada, pp 1055–1058. https://doi.org/10.1109/EMBC44109.2020.9176073
25. Liu S, Utriainen D, Chai C et al (2019) Cerebral microbleed detection using susceptibility weighted imaging and deep learning. Neuroimage 198:271–282. https://doi.org/10.1016/j.neuroimage.2019.05.046
26. Kuo W, Häne C, Mukherjee P, Malik J, Yuh EL (2019) Expert-level detection of acute intracranial hemorrhage on head computed tomography using deep learning. Proc Natl Acad Sci USA 116(45):22737–22745. https://doi.org/10.1073/pnas.1908021116 (**Epub 2019 Oct 21. PMID: 31636195; PMCID: PMC6842581**)
27. Ironside N, Chen CJ, Mutasa S, Sim JL, Ding D, Marfatiah S, Roh D, Mukherjee S, Johnston KC, Southerland AM et al (2020) Fully automated segmentation algorithm for perihematomal edema volumetry after spontaneous intracerebral hemorrhage. Stroke 51(3):815–823. https://doi.org/10.1161/STROKEAHA.119.026764 (**Epub 2020 Feb 12. PMID: 32078476**)
28. Patel A, Schreuder FHBM, Klijn CJM, Prokop M, Ginneken BV, Marquering HA, Roos YBWEM, Baharoglu MI, Meijer FJA, Manniesing R (2019) Intracerebral haemorrhage segmentation in non-contrast CT. Sci Rep 9(1):17858. https://doi.org/10.1038/s41598-019-54491-6 (**PMID: 31780815; PMCID: PMC6882855**)
29. Dhar R, Falcone GJ, Chen Y, Hamzehloo A, Kirsch EP, Noche RB, Roth K, Acosta J, Ruiz A, Phuah CL et al (2020) Deep learning for automated measurement of hemorrhage and perihematomal edema in supratentorial intracerebral hemorrhage. Stroke 51(2):648–651. https://doi.org/10.1161/STROKEAHA.119.027657 (**Epub 2019 Dec 6. PMID: 31805845; PMCID: PMC6993878**)
30. Lee H, Yune S, Mansouri M, Kim M, Tajmir SH, Guerrier CE, Ebert SA, Pomerantz SR, Romero JM, Kamalian S, Gonzalez RG, Lev MH, Do S (2019) An explainable deep-learning algorithm for the detection of acute intracranial haemorrhage from small datasets. Nat Biomed Eng 3(3):173–182. https://doi.org/10.1038/s41551-018-0324-9 (**Epub 2018 Dec 17. PMID: 30948806**)
31. Abstract: Desai V, Flanders A, Lakhani P (2017) Application of deep learning in neuroradiology: automated detection of basal ganglia hemorrhage using 2D-convolutional neural networks. arXiv:1710.03823
32. Li YH, Zhang L, Hu QM, Li HW, Jia FC, Wu JH (2012) Automatic subarachnoid space segmentation and hemorrhage detection in clinical head CT scans. Int J Comput Assist Radiol Surg 7(4):507–516. https://doi.org/10.1007/s11548-011-0664-3 (**Epub 2011 Nov 12. PMID: 22081264**)
33. Abstract: Majumdar A, Brattain L, Telfer B, Farris C, Scalera J (2018) Detecting intracranial hemorrhage with deep learning. Annu Int Conf IEEE Eng Med Biol Soc 2018:583–587. https://doi.org/10.1109/EMBC.2018.8512336 (**PMID: 30440464**)
34. Abstract: Yi T, Pan I, Chen F et al (2020) Identification of intracranial hemorrhage using an original artificial intelligence system. Acad Emerg Med 27(Supplement 1):S194, Netherlands Blackwell Publishing Inc. https://doi.org/10.1111/acem.13961
35. Flanders AE et al (2020) Construction of a machine learning dataset through collaboration: the RSNA 2019 brain CT hemorrhage challenge. Radiol Artif Intell 2(3):e190211
36. Hssayeni MD, Croock MS, Salman AD, Al-khafaji HF, Yahya ZA, Ghoraani B (2020) Intracranial hemorrhage segmentation using a deep convolutional model. Data 5:14
37. Titano JJ, Badgeley M, Schefflein J, Pain M, Su A, Cai M, Swinburne N, Zech J, Kim J, Bederson J et al (2018) Automated deep-neural-network surveillance of cranial images for acute neurologic events. Nat Med 24(9):1337–1341. https://doi.org/10.1038/s41591-018-0147-y (**Epub 2018 Aug 13. PMID: 30104767**)
38. Ginat DT (2020) Analysis of head CT scans flagged by deep learning software for acute intracranial hemorrhage. Neuroradiology 62(3):335–340. https://doi.org/10.1007/s00234-019-02330-w (**Epub 2019 Dec 11. PMID: 31828361**)
39. Abstract: Ojeda P, Zawaideh M, Mossa-Basha M et al (2019) The utility of deep learning: evaluation of a convolutional neural network for detection of intracranial bleeds on non-contrast head computed tomography studies. SPIE 10949. https://doi.org/10.1117/12.2513167
40. Rao B, Zohrabian V, Cedeno P, Saha A, Pahade J, Davis MA (2021) Utility of artificial intelligence tool as a prospective radiology peer reviewer—detection of unreported intracranial hemorrhage. Acad Radiol 28(1):85–93. https://doi.org/10.1016/j.acra.2020.01.035 (**Epub 2020 Feb 24. PMID: 32102747**)
41. Danilov G, Kotik K, Negreeva A, Tsukanova T, Shifrin M, Zakharova N, Batalov A, Pronin I, Potapov A (2020) Classification of intracranial hemorrhage subtypes using deep learning on CT scans. Stud Health Technol Inform 26(272):370–373. https://doi.org/10.3233/SHTI200572 (**PMID: 32604679**)
42. Cho J, Park KS, Karki M, Lee E, Ko S, Kim JK, Lee D, Choe J, Son J, Kim M et al (2019) Improving sensitivity on identification and delineation of intracranial hemorrhage lesion using cascaded deep learning models. J Digit Imaging 32(3):450–461. https://

doi.org/10.1007/s10278-018-00172-1.PMID:30680471;PMCID:PMC6499861
43. Ker J, Singh SP, Bai Y, Rao J, Lim T, Wang L (2019) Image thresholding improves 3-dimensional convolutional neural network diagnosis of different acute brain hemorrhages on computed tomography scans. Sensors (Basel, Switzerland) 19(9):2167. https://doi.org/10.3390/s19092167
44. Karki M, Cho J, Lee E et al (2020) CT window trainable neural network for improving intracranial hemorrhage detection by combining multiple settings. Artif Intell Med 106:101850. https://doi.org/10.1016/j.artmed.2020.101850
45. Gautam A, Raman B (2021) Towards effective classification of brain hemorrhagic and ischemic stroke using CNN. Biomed Signal Process Control 63:102178
46. Abstract: Yune S, Lee H, Do S, Ting D (2018) Case-based learning on artificial intelligence radiology atlas: example of intracranial hemorrhage and urinary stone detection. J Gen Internal Med 33(2 Supplement 1):690–691
47. Phong TD, Duong HN, Nguyen HT, Trong NT, Nguyen VH, Van Hoa T, Snasel V (2017) Brain hemorrhage diagnosis by using deep learning. In: Proceedings of the 2017 international conference on machine learning and soft computing (ICMLSC'17). Association for Computing Machinery, New York, NY, USA, pp 34–39. https://doi.org/10.1145/3036290.3036326
48. Li L, Wei M, Liu B, Atchaneeyasakul K, Zhou F, Pan Z, Kumar S, Zhang J, Pu Y, Liebeskind DS, Scalzo F (2020) Deep learning for hemorrhagic lesion detection and segmentation on brain CT images. IEEE J Biomed Health Inform. https://doi.org/10.1109/JBHI.2020.3028243 (**Epub ahead of print. PMID: 33001810**)
49. Abstract: Barreira C, Bouslama M, Ratcliff J et al (2018) E-078 Advance study: automated detection and volumetric assessment of intracerebral hemorrhage. J Neurointerv Surg 10(Suppl 2):A88–A88
50. Abstract: Lee H, Kim M, Do S (2018) Practical window setting optimization for medical image deep learning. arXiv:1812.00572
51. Abstract: Sales Barros R, van der Steen WE, Ponomareva E et al. Abstract WMP29: detection and segmentation of subarachnoid hemorrhages with deep learning. Stroke 50(Suppl_1):AWMP29–AWMP29
52. Abstract: Patil R, Shreya A, Maulik P, Chaudhury S (2019) Hybrid AI based stroke characterization with explainable model. J Neurol Sci 405:162–163
53. Abstract: Bizzo B, Hashemian B, McNitt T et al. Abstract WP68: interpretable deep learning-based characterization of intracranial hemorrhage on head CT. Stroke 50(Suppl_1):AWP68–AWP68
54. Abstract: Hahm MH, Lee HJ, Lim JK, Lee HS (2021) Clinical usefulness of deep learning-based automated segmentation in intracranial hemorrhage. Technol Health Care. Published online February 26, 2021
55. Abstract: Herweh C, Mokli Y, Bellot P et al (2020) AI-based automated detection of intracranial hemorrhage on non-enhanced CT scans. In: International journal of stroke, vol 15. SAGE Publications Ltd 1 Olivers Yard, 55 City Road, London Ec1y 1sp, England; 2020:295–295
56. Lee JY, Kim JS, Kim TY, Kim YS (2020) Detection and classification of intracranial haemorrhage on CT images using a novel deep-learning algorithm. Sci Rep. https://doi.org/10.1038/s41598-020-77441-z
57. Ko H, Chung H, Lee H, Lee J (2020) Feasible study on intracranial hemorrhage detection and classification using a CNN-LSTM network. Conf Proc IEEE Eng Med Biol Soc 2020:1290–1293
58. Burduja M, Ionescu RT, Verga N (2020) Accurate and efficient intracranial hemorrhage detection and subtype classification in 3D CT scans with convolutional and long short-term memory neural networks. Sensors 20(19):5611. https://doi.org/10.3390/s20195611
59. Abstract: Praveen K, Sasikala M, Janani A, Shajil N, Nishanthi VH (2021) A simplified framework for the detection of intracranial hemorrhage in CT brain images using deep learning. Curr Med Imaging. https://doi.org/10.2174/157340561766621018100641 (Published online ahead of print, 2021 Feb 17)
60. Fan YH, Zhang L, Lam WW, Mok VC, Wong KS (2003) Cerebral microbleeds as a risk factor for subsequent intracerebral hemorrhages among patients with acute ischemic stroke. Stroke 34(10):2459–2462. https://doi.org/10.1161/01.STR.0000090841.90286.81
61. Russakovsky O, Deng J, Su H et al (2015) ImageNet large scale visual recognition challenge. Int J Comput Vis 115:211–252. https://doi.org/10.1007/s11263-015-0816-y
62. Wang S, Tang C, Sun J, Zhang Y (2019) Cerebral micro-bleeding detection based on densely connected neural network. Front Neurosci 13:422. Published 2019 May 17. https://doi.org/10.3389/fnins.2019.00422
63. Wang S, Sun J, Mehmood I, Pan C, Chen Y, Zhang Y (2020) Cerebral micro-bleeding identification based on a nine-layer convolutional neural network with stochastic pooling. Concurr Comput Pract Exp 32
64. Fazlollahi A, Meriaudeau F, Giancardo L, Villemagne VL, Rowe CC, Yates P, Salvado O, Bourgeat P; AIBL Research Group (2015) Computer-aided detection of cerebral microbleeds in susceptibility-weighted imaging. Comput Med Imaging Graph 46 Pt 3:269–276. https://doi.org/10.1016/j.compmedimag.2015.10.001 (**Epub 2015 Oct 24. PMID: 26560677**)
65. Dou Q et al (2015) Automatic cerebral microbleeds detection from MR images via independent subspace analysis based hierarchical features. In: 2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC), Milan, Italy, pp 7933–7936.https://doi.org/10.1109/EMBC.2015.7320232
66. Roy S et al (2015) Cerebral microbleed segmentation from susceptibility weighted images. Proc SPIE 9413:94131E
67. Abstract: Hou X-X, Chen H (2016) Sparse autoencoder based deep neural network for voxelwise detection of cerebral microbleed. In: 22nd International conference on parallel and distributed systems. IEEE, Wuhan, pp 1229–1232
68. Abstract: Lu S, Lu Z, Hou X, Cheng H, Wang S (2017) Detection of cerebral microbleeding based on deep convolutional neural network. In: Proceedings of the 14th international computer conference on wavelet active media technology and information processing (ICCWAMTIP), Chengdu, pp 93–96
69. Wang S, Jiang Y, Hou X, Cheng H, Du S (2017) Cerebral microbleed detection based on the convolution neural network with rank based average pooling. IEEE Access 1–1. https://doi.org/10.1109/ACCESS.2017.2736558
70. Morrison MA, Payabvash S, Chen Y, Avadiappan S, Shah M, Zou X, Hess CP, Lupo JM (2018) A user-guided tool for semi-automated cerebral microbleed detection and volume segmentation: evaluating vascular injury and data labelling for machine learning. Neuroimage Clin 4(20):498–505. https://doi.org/10.1016/j.nicl.2018.08.002 (**PMID: 30140608; PMCID:PMC6104340**)
71. Cloutie RS (2018) Voxelwise detection of cerebral microbleed in CADASIL patients by genetic algorithm and back propagation neural network. Adv Comput Sci Res 65:101–105
72. Abstract: Gunter JL, Spychalla AJ, Ward CP, Graff-Radford J, Huston J, Kantarci K, Knopman DS, Petersen RC, Jack CR (2018) Automating cerebral microbleed detection in support of Alzheimer's disease trials using a convolutional neural network AI. Alzheimer's Dementia 14(7 Supplement):P1530–P1531
73. Zhang YD, Hou XX, Chen Y et al (2018) Voxelwise detection of cerebral microbleed in CADASIL patients by leaky rectified linear unit and early stopping. Multimed Tools Appl 77:21825–21845. https://doi.org/10.1007/s11042-017-4383-9

74. Zhang YD, Zhang Y, Hou XX et al (2018) Seven-layer deep neural network based on sparse autoencoder for voxelwise detection of cerebral microbleed. Multimed Tools Appl 77:10521–10538. https://doi.org/10.1007/s11042-017-4554-8