DOI: *http://dx.doi.org/10.18502/cjn.v24i1.???*

Current Journal of Neurology

Review Paper

Curr J Neurol 2025; 24(1): 41-69

Podcast

Applications of deep learning in intracranial aneurysm imaging: A scoping review of detection, risk prediction, and emerging prognostic models

Received: 05 Sep.2024
Accepted: 13 Nov. 2024

Ali Jafarpour[#1], Matin Akhbari[#2], Kamyar Khorsand[#3], Kiana Naghavi[4], Mohammad Moein Ghaemmaghami[5], Elaheh Ghaderi[6], Elham Yadegarifard[7], Shamimeh Arabgol[8], Rozhan Esmaeili-Benvidi[9], Behnaz Mahmoudvand[10], Seed Amirabbas Shahidi-Marnani[11], Saina Hasany[12], Amirhossein Rigi[13], Nazanin Sarvi[14], Pooria Sobhanian[15], Mahsa Asadi-Anar[16], Melika Arab-Bafrani[17], Farbod Khosravi[18]

[1] Department of Neurosurgery, Sina Hospital, Tehran University of Medical Sciences, Tehran, Iran
[2] Faculty of Medicine, Istanbul Yeni Yuzyil University, Istanbul, Turkey
[3] School of Medicine, Ahvaz Jondishapur University of Medical Sciences, Ahvaz, Iran
[4] Student Research Committee, School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[5] School of Medicine, Najafabad Branch, Islamic Azad University, Najafabad, Iran
[6] Kermanshah University of Medical Sciences, Kermanshah, Iran
[7] Student Research Committee, School of Medicine, Hormozgan University of Medical Sciences, Bandar Abbas, Iran
[8] School of Medicine, Iran University of Medical Sciences, Tehran, Iran
[9] School of Medicine, Kashan University of Medical Sciences, Kashan, Iran
[10] Student Research Committee, School of Medicine, Iran University of Medical Sciences, Tehran, Iran
[11] School of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran
[12] Department of Biotechnology, Islamic Azad University, Tehran Medical Branch, Tehran, Iran
[13] Department of Radiology, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[14] Department of Medicine, Islamic Azad University, Yazd Branch, Yazd, Iran
[15] Student Research Committee, School of Medicine, Mazandaran University of Medical Sciences, Sari, Iran
[16] Student Research Committee, Shahid Beheshti University of Medical Sciences, Tehran, Iran
[17] Students' Scientific Research Center (SSRC), Tehran University of Medical Sciences, Tehran, Iran
[18] School of Medicine, Shahid Beheshti University of Medical Sciences, Tehran, Iran

Keywords
Deep Learning; Intracranial Aneurysm; Neurosurgery

How to cite this article: Jafarpour A, Akhbari M, Khorsand K, Naghavi K, Ghaemmaghami MM, Ghaderi E, et al. Applications of Deep Learning in Intracranial Aneurysm Imaging: A Scoping Review of Detection, Risk Prediction, and Emerging Prognostic Models. Curr J Neurol 2025; 24(1): 41-69.

Corresponding Author 1: Ali Jafarpour
Email: ali.jafarpour@hotmail.com
Corresponding Author 2: Mahsa Asadi-Anar
Email: mahsa.boz@gmail.com

[#]Contributed equally to this work.

Abstract

Background: Intracranial aneurysms (IAs) pose a significant risk of rupture and subarachnoid hemorrhage, necessitating early, accurate detection and risk stratification. With advances in artificial intelligence, deep learning (DL) has emerged as a transformative tool in neurovascular imaging. However, the clinical translation of DL applications remains constrained by variability in model design, data sources, and validation strategies. The aim of the present study was to systematically map and evaluate the landscape of DL applications in the detection, segmentation, risk prediction, and outcome assessment of IAs, with attention to methodological rigor, clinical utility, and translational limitations.

Methods: We conducted a scoping review of studies indexed in PubMed, Scopus, and Web of Science up to August 2023, following PRISMA-ScR guidelines. Eligible studies employed DL algorithms for IA-related diagnostic or prognostic tasks using radiological imaging. Data extraction included model architecture, imaging modality, validation strategy, performance metrics, and thematic focus. Study quality was assessed using the Joanna Briggs Institute (JBI) critical appraisal tools.

Results: Forty-two studies met the inclusion criteria, encompassing over 10,000 patients across diverse imaging platforms and DL architectures. Convolutional neural networks (CNNs) were the most commonly used models, with reported sensitivities ranging from 73% to 99% and AUCs frequently exceeding 0.85. Despite promising results in IA detection and rupture risk prediction, only a minority of studies conducted external validation or addressed post-treatment outcomes. Major gaps include a lack of benchmarking across models, limited explainability, and regulatory or ethical frameworks.

Conclusion: DL algorithms demonstrate strong diagnostic and predictive performance in IA imaging but face critical barriers to clinical integration, including interpretability challenges, dataset heterogeneity, and limited generalizability. Future research should prioritize multicenter validation, explainable AI techniques, and outcome-focused modeling to advance safe and effective deployment in neurosurgical care.

Introduction

Deep learning (DL) is a subset of machine learning (ML) that aims to extract high-level representations, analyze them, and learn relevant information from raw data using hierarchical architectures.[1,2] It consists of various algorithms used to develop complex generalized systems capable of solving problems and providing accurate predictions. ML and DL algorithms have become popular tools for addressing various challenges in medical imaging fields.[3]

These algorithms use supervised or unsupervised methods and rely on detailed datasets to predict early signs of disease.[4] There are various potential applications of DL technology in medical imaging that can improve the healthcare system and patient outcomes.[3]

The use of DL to predict neurosurgical outcomes is still in its infancy. While profound learning studies have shown promise, promoting the validity and reproducibility of DL models requires more data and model interpretability.[5]

Surgeries related to the brain are high-risk procedures that carry a considerable risk of morbidity and mortality.[2] To improve clinical treatment outcomes and minimize postoperative disability, the recovery process can incorporate DL and microscopic imaging to reduce risks and potential patient loss. Studies demonstrated that this approach can significantly enhance the neurosurgery nursing process.[6]

To revolutionize neurosurgery, AI, ML, and DL are combined to provide insights into the patient's condition and assist neurosurgeons in making more effective decisions during surgical interventions, improving diagnostic and prognostic outcomes.[2] AI can be practical in diagnosing complicated neurological disorders like intracerebral hemorrhage (ICH) and cerebral aneurysms.

AI uses ML and DL algorithms that are more efficient than radiologists in detecting intracranial aneurysms (IAs) and anticipating their extent using computed tomography (CT) angiograms and non-contrast CT.[7,8]

IAs are relatively common, occurring in approximately 4% of the population.[9] IAs can be classified into four types: saccular, fusiform, dissecting, and mycotic. Saccular aneurysms account for 90% of all IAs.[10] Most aneurysms do not cause any symptoms and may not rupture, but they can expand unpredictably and always carry a risk of rupture.[11] In some cases, an aneurysm may cause symptoms due to its mass effect.[9]

If an aneurysm ruptures, it can result in a subarachnoid hemorrhage, which has a high mortality rate and significant disease burden.[9,12] Therefore, accurate and early detection of IAs in clinical practice is crucial.[13] The size and location of an aneurysm affect its rupture risk.[14] A 1998 study called the International Study of Unruptured

Intracranial Aneurysms (ISUIA) found that for patients with aneurysms smaller than 10 mm and no prior SAH, the annual risk of rupture was 0.05%. For aneurysms larger than 10 mm, the risk was 1% per year. Study data from 2003 showed a 0% and 2.5% five-year rupture risk for small aneurysms in the anterior and posterior circulation, respectively.[15]

To diagnose and monitor IAs, various imaging techniques are used, including intra-arterial digital subtraction angiography (IADSA), which is the gold standard for diagnosis of cerebral aneurysms, computed tomography angiography (CTA), magnetic resonance angiography (MRA), and transcranial Doppler ultrasonography.[9] The detection rate of asymptomatic unruptured intracranial aneurysms (UIAs) has remarkably increased over the past 15 years. This coincides with increased CTA/MRA imaging.[16] Developing more precise imaging modalities is crucial for assessing the risk of UIA rupture and improving conservative treatment options, such as medication.[16,17]

Artificial intelligence (AI)-based algorithms can enhance the detection rate early and minimize intra- and inter-rater variability.[18] A CNN is a type of DL architecture which can aid clinicians in diagnosing IAs with high sensitivity. It has been shown to improve clinicians' performance by providing dependable and accurate predictions, thereby optimizing patient care.[12] Our objective in this scoping review was to assess the effectiveness of DL algorithms in detecting IAs and their subsequent neurosurgical outcomes.

Research Question: "How have DL algorithms been applied in the detection of IAs and the prediction of neurosurgical outcomes across diverse clinical populations and imaging contexts?"

PCC Framework

Population: Patients with diagnosed or suspected IAs

Concept: Application of DL algorithms

Context: Diagnostic imaging and neurosurgical outcome prediction in clinical and experimental settings

Materials and Methods

This review was conducted in accordance with the PRISMA-ScR (Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews) guidelines.[19]

Given the heterogeneity in study designs, imaging modalities, and model architectures applied in the domain of DL for IA detection, a scoping review was deemed more appropriate than a systematic review. Scoping reviews allow for mapping the breadth of evidence, identifying key concepts, and clarifying working definitions and gaps in research.[20] This approach aligns with the updated methodological guidance by the Joanna Briggs Institute (JBI) for conducting evidence synthesis in complex and emerging fields.

Search strategy: A comprehensive literature search was conducted across PubMed, Web of Science, and Scopus up to August 2023. The strategy incorporated a combination of controlled vocabulary (e.g., MeSH terms) and free-text terms related to "deep learning," "intracranial aneurysm," and "cerebral aneurysm." The full search strategies used for each database are detailed in table 1. The reference lists of included articles were also manually screened to identify additional eligible studies. Duplicate records were removed using EndNote, and screening was performed using the Rayyan platform.

Eligibility criteria: This scoping review included original research articles that reported on the application of DL algorithms for the detection, segmentation, risk prediction, or neurosurgical outcome assessment of IAs. Eligible studies involved human subjects with diagnosed or suspected IAs and utilized radiological imaging modalities such as CT, CTA, magnetic resonance imaging (MRI), MRA, or digital subtraction angiography (DSA). We included studies employing DL models, such as CNNs, artificial neural networks (ANNs), and other advanced architectures, provided they reported at least one quantitative performance metric—such as sensitivity, specificity, accuracy, AUC, or Dice coefficient—and conducted any form of model validation.

We included retrospective, prospective, diagnostic, and experimental studies published in English with full-text availability. Studies were excluded if they were preclinical or non-human studies, review articles, editorials, case reports, conference abstracts, or methodological commentaries without original data. Additionally, studies were excluded if they did not utilize DL methods or failed to report any measurable outcomes related to model performance or clinical applicability. Articles lacking accessible full texts or published in languages other than English were also excluded.

Study selection and data extraction: All identified articles were imported into EndNote for deduplication and, then, uploaded to Rayyan, a web-based tool designed to facilitate systematic review screening.

Table 1. Curated search strategies for each chosen database

Database	Search strategy
PubMed	("Deep Learning"[tiab] OR "hierarchical learning"[tiab] OR "deep machine learning"[tiab] OR "deep structured learning"[tiab] OR "machine learning"[tiab] OR "reinforcement learning"[tiab] OR "supervised learning"[tiab] OR "unsupervised learning"[tiab] OR "action-based learning"[tiab] OR "actor-critic methods"[tiab] OR "actual learning"[tiab] OR "adversarial training"[tiab] OR "algorithmic learning"[tiab] OR "apprenticeship learning"[tiab] OR "artificial neural network"[tiab] OR "artificial neural network*"[tiab] OR "autoencoders"[tiab] OR "automated learning"[tiab] OR "backpropagation"[tiab] OR "bayesian network*"[tiab] OR "bayesian optimization"[tiab] OR "computational intelligence paradigm"[tiab] OR "computer vision"[tiab] OR "computer-based learning"[tiab] OR "convolutional neural network*"[tiab] OR "data-driven learning"[tiab] OR "Deep Learning"[Mesh] OR "Machine Learning"[Mesh] OR "Supervised Machine Learning"[Mesh] OR "Unsupervised Machine Learning"[Mesh] OR "Artificial Intelligence"[Mesh] OR "Neural Networks, Computer"[Mesh]) AND ("Intracranial Aneurysm"[tiab] OR "Intracranial Aneurysm*"[tiab] OR "cerebral aneurysm"[tiab] OR "cerebral aneurysm*"[tiab]
WOS	((((((((((((((((((((((((((((((((((((((TS=("Deep Learning")) OR TS=("hierarchical learning")) OR TS=("deep machine learning")) OR TS=("deep structured learning")) OR TS=("machine learning")) OR TS=("reinforcement learning")) OR TS=("supervised learning")) OR TS=("unsupervised learning")) OR TS=("action-based learning")) OR TS=("actor-critic methods")) OR TS=("actual learning")) OR TS=("adversarial training")) OR TS=("algorithmic learning")) OR TS=("apprenticeship learning")) OR TS=("artificial neural network")) OR TS=("artificial neural network*")) OR TS=("auto encoders")) OR TS=("automated learning")) OR TS=("back propagation")) OR TS=("bayesian network*")) OR TS=("bayesian optimization")) OR TS=("computational intelligence paradigm")) OR TS=("computer vision")) OR TS=("computer-based learning")) OR TS=("convolutional neural network*")) OR TS=("data-driven learning")) AND TS=("Intracranial Aneurysm")) OR TS=("Intracranial Aneurysm*")) OR TS=("cerebral aneurysm")) OR TS=("cerebral aneurysm*")) OR TS=("brain aneurysm")) OR TS=("brain aneurysm*"))
Scopus	(TITLE-ABS-KEY ("Deep Learning") OR TITLE-ABS-KEY ("hierarchical learning") OR TITLE-ABS-KEY ("deep machine learning") OR TITLE-ABS-KEY ("deep structured learning") OR TITLE-ABS-KEY ("machine learning") OR TITLE-ABS-KEY ("reinforcement learning") OR TITLE-ABS-KEY ("supervised learning") OR TITLE-ABS-KEY ("unsupervised learning") OR TITLE-ABS-KEY ("action-based learning") OR TITLE-ABS-KEY ("actor-critic methods") OR TITLE-ABS-KEY ("actual learning") OR TITLE-ABS-KEY ("adversarial training") OR TITLE-ABS-KEY ("algorithmic learning") OR TITLE-ABS-KEY ("apprenticeship learning") OR TITLE-ABS-KEY ("artificial neural network") OR TITLE-ABS-KEY ("artificial neural network*") OR TITLE-ABS-KEY ("auto encoders") OR TITLE-ABS-KEY ("automated learning") OR TITLE-ABS-KEY ("back propagations") OR TITLE-ABS-KEY ("bayesian network*") OR TITLE-ABS-KEY ("bayesian optimization") OR TITLE-ABS-KEY ("computational intelligence paradigm") OR TITLE-ABS-KEY ("computer vision") OR TITLE-ABS-KEY ("computer-based learning") OR TITLE-ABS-KEY ("convolutional neural network*") OR TITLE-ABS-KEY ("data-driven learning") AND (TITLE-ABS-KEY ("Intracranial Aneurysm") OR TITLE-ABS-KEY ("Intracranial Aneurysm*") OR TITLE-ABS-KEY ("cerebral aneurysm") OR TITLE-ABS-KEY ("cerebral aneurysm*") OR TITLE-ABS-KEY ("brain aneurysm") OR TITLE-ABS-KEY ("brain aneurysm*")

Two independent reviewers screened titles and abstracts against the eligibility criteria. Full texts of potentially relevant studies were subsequently retrieved and reviewed in detail for inclusion. Discrepancies in study selection were resolved through discussion between the two reviewers, and a third reviewer was consulted when consensus could not be reached. Inter-rater agreement during the initial screening phase was measured using Cohen’s kappa coefficient to ensure consistency in study selection.

For data extraction, a standardized charting form was developed a priori and piloted on a subset of included studies to ensure clarity and comprehensiveness. This form included the following variables: first author, year of publication, country or region, study design, sample size, imaging modality, aneurysm characteristics (location, size, type), type of DL architecture used, data preprocessing methods, model validation approach (internal or external), performance metrics (e.g., sensitivity, specificity, AUC), and primary outcomes related to detection accuracy or clinical utility. Data extraction was conducted independently by two reviewers. Any discrepancies in data abstraction were resolved through discussion, and if necessary, adjudicated by a third reviewer to ensure data accuracy and reliability.

Quality assessment: Quality assessment was performed for each included study using the JBI critical appraisal tools, with the version selected according to the study design. For diagnostic accuracy studies, the JBI Critical Appraisal Checklist for Diagnostic Test Accuracy Studies was applied. For cohort and cross-sectional studies, the JBI Checklist for Analytical Cross Sectional Studies and the JBI Checklist for Cohort Studies were used as appropriate. Two reviewers assessed each study independently, with discrepancies resolved by consensus.

Thematic analysis approach: To synthesize the heterogeneity of study objectives, methodologies, and outcomes, we conducted a thematic analysis of the included studies. Themes were derived inductively, based on patterns observed during data extraction and synthesis, rather than being predefined. Two reviewers independently examined the extracted study variables—such as model architecture, clinical application, validation strategies, and imaging modality—and grouped them into emergent thematic categories. Discrepancies in theme assignment were resolved through discussion, and consensus was reached in all cases. No formal coding software was used; however, the process was guided by principles of qualitative content analysis and aimed to achieve thematic saturation. The final set of themes—encompassing clinical applications, model architectures, validation rigor, imaging inputs, geographic distribution, performance reporting, and focus on rupture status—was reviewed by a third senior reviewer to ensure conceptual clarity and coherence with the review's aims.

Data synthesis and visualization: To provide a comprehensive overview of the evidence landscape, we conducted several quantitative and thematic syntheses of the extracted data. Following data extraction, all study-level variables—including model architecture, clinical application, country of origin, and performance metrics—were coded in a structured spreadsheet and independently verified for accuracy.

Thematic mapping was performed to categorize each study according to its primary DL application domain (e.g., detection/segmentation, rupture risk prediction, treatment outcome prediction, or EMR/NLP-based identification). This enabled the creation of a thematic evidence map, visualized as a matrix, which illustrates the breadth and concentration of research activity across key clinical domains.

To examine global research distribution, we recorded the country of origin for each study based on first or corresponding author affiliations. Multinational collaborations were attributed to all participating countries. The frequency of studies per country was then visualized on a world map (bubble map), with bubble size reflecting the total number of included studies from each country, thereby highlighting geographic disparities and research clusters.

For methodological benchmarking, we further cross-tabulated the model architecture (e.g., CNN-based, hybrid ML/DL, ANN/DNN, or NLP) against the principal clinical application of each study. This was visualized as a stacked bar chart, illustrating how different DL architectures are distributed across the major clinical tasks addressed in the literature.

All visualizations were generated using Python (matplotlib, seaborn, Basemap, and pandas libraries) based on the manually curated extraction table. These figures are presented in the results section to facilitate transparent comparison and thematic and geographic research gaps identification.

Results

Study characteristics and geographic distribution: This scoping review included 42 studies published between 2011 and 2023, encompassing a broad range of study designs, including retrospective analyses, cross-sectional studies, diagnostic investigations, prospective cohorts, and translational projects. The majority of the included studies were retrospective, and most were conducted in China, followed by the United States, South Korea, Germany, and Japan (Table 2). Collectively, the studies represented data from over 10000 patients with confirmed or suspected IAs, reflecting a substantial and growing global interest in leveraging artificial intelligence for neurovascular diagnostics (Figures 1 and 2).

Quality assessment results: The overall methodological quality of the included studies was variable. While most studies met the majority of relevant JBI criteria, frequent limitations included lack of external validation, absence of control groups, and lack of reporting of patient selection or blinding.

Table 2. Summary findings of studies included in this scoping review (Part I)

Author	Country	Type of study	Population	DL model	Outcome	Conclusion
Chen et al.[21]	China	Translational	75 Intracranial aneurysm patients 37 control samples	The study employed ML to identify the diagnostic significance of key genes related to IA	Both the high mitochondrial dysfunction group and the high necroptosis group had increased levels of mitochondrial pathways, necroptosis pathways and immune pathways. The upregulation of mitochondria-induced necroptosis emerges as a potential and novel target for predictive diagnosis.	Involvement of mitochondria-induced necroptosis in the formation of IAs was shown.
Feng et al.[22]	China	Cross-Sectional	Training set: 898 patients Test set: 253 patients	A three-dimensional CNN was used to automatically perform aneurysm detection, segmentation, and morphological feature extraction. Following the process of dimensionality reduction, three classification models were developed and assessed using the area under the receiver operating characteristic curve: SVM, RF, and MLP	The method found 13 features associated with aneurysm rupture.	The proposed method had high diagnostic efficiency in identifying between ruptured and unruptured aneurysms.
Ham et al.[23]	South Korea	Retrospective	Internal validation: 154 patients External validation: 113 patients	CNN with 3D TOF-MRA	High adequate performance of the proposed method in aneurysm segmentation was shown.	Utilizing 3D patches in brain 3D TOF-MRA, enables rapid and accurate aneurysm detection, supporting quick diagnosis.
Jiang et al.[24]	USA, China	Retrospective	Training set: 102 patients Test set: 10 patients	Computational fluid dynamics simulations and geometrical analyses were conducted, and 3D velocity vector fields within the IA dome were processed for velocity-informatics. Four ML methods were employed (SVM, RF, generalized linear model, and GLM with Lasso or elastic net regularization) to evaluate the effectiveness of the proposed velocity-informatics.	The prediction was improved with velocity-informatics metrics.	Including velocity-informatics from aneurismal velocity data can enhance the overall rupture status characterization of an IA.
Liu et al.[25]	China	Retrospective	Training set: 80 patients Test set: 10 patients	The study utilizes the DeepMedic platform, employing a 3D CNN architecture for automatic segmentation and detection of IAs from CTA images.	The DeepMedic platform's DL architecture, which uses a 3D CNN model, can segment and detect IAs from CTA images with high sensitivity and reliability.	The 3D CNN system shows accurate intracranial aneurysm detection and segmentation from CTA images.

Table 2. Summary findings of studies included in this scoping review (Part I) (continue)

Author	Country	Type of study	Population	DL model	Outcome	Conclusion
Patel et al.[26]	USA	Retrospective	Training set: 27 Validation: 30 Test set: 20	A DL model (3D DeepMedic) utilized for cerebrovasculature segmentation from CTA	The DeepMedic model successfully delineated all IAs and showed lower error rates for IA morphometrics compared to human raters. The results demonstrate the ability of CTA scans to produce precise visualizations of cerebral vasculature and abnormalities, including IA.	Architecture performed exceptionally well in the segmentation of cerebral vessels and IA
Shao et al.[27]	Australia	-	-	The method comprises two stages: unsupervised learning and downstream tasks. In the first stage, augmentation is applied to each point cloud, creating pairs of augmented samples with differing poses and in the second stage, the trained model's unsupervised representations are concatenated and used as input for downstream tasks to assess the effectiveness of unsupervised learning.	The unsupervised method demonstrates comparable or superior performance compared to state-of-the-art supervised techniques, particularly excelling in the detection of aneurysmal vessels.	Unsupervised representation learning method was effective in the classification and segmentation of 3D IAs.
Wang et al.[13]	China	Retrospective	Training set: 1110 patients Internal validation: 139 patients Test set:134 patients	The DAResUNet network, employed for training, utilizes a 3D-CNN with an encoder-decoder architecture similar to 3D-U-Net.	The multiphase analysis demonstrated higher sensitivity compared to the single-phase analysis in internal validation, test, and independent validation data.	Automated detection of IAs with high sensitivity was made possible using a multiphase fusion DL model with automatic phase selection.
Allgaier et al.[28]	Germany	-	No patients were included. There were 4 individuals who performed a specific type of VR.	The work simulation focuses on craniotomy and head placement. This study chose to create a virtual OR and use a VR HMD in order to create a more immersive experience than existing simulations that use haptic devices and stationary stereoscopic displays.	Craniotomy was generally accepted but could benefit from improvements in hand and arm positioning and the ability to mill the sphenoid bone.	The provision of a VR system for craniotomy, utilizing an HMD to create an immersive training experience, is shown.
Lei and Yang[29]	China	Diagnostic	40 Subjects	This study introduces 2 models for diagnosing IAs, the first model is a 3D U-Net algorithm designed to quickly diagnose and label potential intracranial aneurysm locations in 3D TOF MRA image sequences and the second model is a 3D CNN for intracranial aneurysm classification with a simple structure to avoid overfitting.	Both proposed methods were able to locate and detect aneurysm successfully with U-Net's better performance in diagnosing and 3D CNN's better performance in positioning.	The study employed DL, specifically the U-Net and 3D CNN network models, to automatically label intracranial aneurysm MRA images. With the U-Net algorithm showing agreement with manual labeling.

Table 2. Summary findings of studies included in this scoping review (Part I) (continue)

Author	Country	Type of study	Population	DL model	Outcome	Conclusion
Li et al.[30]	China	Retrospective	Training set: 120 patients Test set: 107 patients	-	Model A+B demonstrated a slightly higher AUC than individual models, while Model A+C did not show a notable improvement. Combining radiomics and traditional morphological features proved effective in identifying intracranial aneurysm instability. Relying solely on Radiomics-derived morphological features is not recommended. Notably, the Radiomics-based models did not outperform the model using traditional morphological features	Combining radiomics and traditional morphological features proved effective in identifying intracranial aneurysm instability.
Tian et al.[31]	China	Retrospective	Control group: 393 Complication Group: 48	Three machine learning algorithms (ANN, RF, and LR) were trained on the expanded training set using ten-fold cross-validation and grid search for hyperparameter optimization.	With ANN showing the best performance among other algorithms, this study found significant features for the prediction of periprocedural complications.	Machine learning algorithms may accurately predict periprocedural problems.
Wu et al.[32]	China		Training set: 1205 CTA images Test set: 303 CTA images	The study utilized a cascade model for aneurysm detection, initially employing a fine-tuned feature pyramid network (FPN) for candidate detection. Machine learning and deep learning-based rupture classification methods were employed to distinguish between ruptured and unruptured aneurysms.	The findings suggest the feasibility of the pipeline for potential clinical use, aiding radiologists in aneurysm detection and the classification of ruptured and unruptured aneurysms.	Multichannel information can improve the performance of aneurysm detection.
Kim et al.[33]	Korea	Retrospective	343 Patients	Explainable artificial intelligence (XAI) was used to analyze the contribution of risk factors on the development of CAV.	In the proposed model, the relationship between aneurysm size, age, and CAV in individuals with aSAH was quantitatively examined.	Aneurysm size and age were identified as the most significant influencers.
Ou et al.[34]	China	Prospective	182 Patients	Some traditional ML algorithms like SVM, K-Nearest Neighbors, Decision Tree, Artificial Neural Network, RF, and Naïve Bayes were used. Also, the automated machine learning named TPOT was used.	Aneurysm size, use of SAC, and posterior circulation were significant factors in predicting recanalization. The AutoML-derived model outperformed other models. The performance of autoML might outperform that of conventional statistical and manually constructed machine learning models.	By demonstrating that the AutoML-derived model accurately predicts treatment outcomes, the study established the viability of employing AutoML for aneurysm treatment outcome prediction.

Table 2. Summary findings of studies included in this scoping review (Part I) (continue)

Author	Country	Type of study	Population	DL model	Outcome	Conclusion
Pennig et al.[35]	Germany	Retrospective	Training set: 68 patients Test set: 104 patients	Three 3D CNNs based on DeepMedic were utilized in the study, namely DLM-Orig, DLM-Vess, and DLM-LDim, each trained on CTA datasets with distinct inputs.	The statistical analysis showed a significant improvement with DLM assistance. Results imply that incorporating deep learning assistance might be a beneficial addition that improves the precision of aneurysm identification in patients with aSAH.	The DLM considerably increased radiologists' aneurysm detection in patients with aSAH, particularly for secondary aneurysms.
Afzal et al.[36]	USA	-	-	The proposed Biomed-Summarizer introduces a novel framework that combines a prognosis quality recognition model with a clinical context-aware model for intelligent and context-enabled summarization of biomedical text. It employs a DNN for quality recognition, a bidirectional long-short term memory recurrent neural network for clinical context awareness, and calculates similarity between query and PICO text sequences.	Multiclass classifier had better performance than traditional machine-learning in classifying categories.	Evaluation results indicate superior performance compared to existing approaches.
Chen et al.[37]	China	Retrospective	Training set: 807 patients Internal validation: 200 patients External validation: 108 patients	With the aim to predict the individual rupture status of UIAs, models based on traditional LR and ML algorithms combining clinical, morphological, and hemodynamic information are built and, then, tested in internal and external validation datasets.	The study indicates varying performance for different models across different datasets in predicting the risk of rupture related to aneurysms. In prediction models integrating clinical, aneurysm morphological, and hemodynamic characteristics, ML techniques did not outperform traditional LR in determining the rupture state of UIAs.	The models' ability to make predictions is significantly influenced by hemodynamic factors.
Chen et al.[38]	China	Diagnostic	Training dataset : 76 patients Internal test dataset: 20 patients External test dataset: 35 patients	The paper introduced a CAD system designed for cerebral aneurysms in TOF-MRA. The system offers clinicians a fully automated process, generating (1) a three-dimensional mesh model of the intracranial artery for hemodynamic analysis and (2) identification of suspected aneurysm areas using an FCN-based network.	According to the proposed method's results in the internal and external test, the method has the potential to detect aneurysm. There is potential for routine physical examinations to screen for aneurysms using this technique.	Using contrast-unenhanced time-of-flight MRA images, the proposed computer-assisted detection system may locate possible aneurysm sites on its own.

Table 2. Summary findings of studies included in this scoping review (Part I) (continue)

Author	Country	Type of study	Population	DL model	Outcome	Conclusion
Detmer et al.[39]	USA, Japan, Colombia		Training set: 1061 patients External validation: 203 patients	This study compared alternative ML classifiers for aneurysm rupture status discrimination to a previously constructed aneurysm rupture logistic regression probability model (LRM).	Despite variations in rankings, statistical tests did not find significant differences in variable importance among the classifiers. Additional data, such as those relevant to the aneurysm wall, may be required in order to further enhance the predictions.	The LRM demonstrated a comparable level of performance to other ML classifiers, indicating its potential for assessing aneurysm rupture.
Duan et al.[40]	Japan	Diagnostic	4 ICA Patients and 11 Healthy people	Deep Learning This study designed a mapping function by combining a U-net model with a single convolution.	DL-MRA was comparable to time-of-flight MRA (TOF-MRA), and both were superior to linear-MRA.	MRA generated through DL from 3D synthetic MRI data effectively visualized major intracranial arteries, comparable to time-of-flight MRA (TOF-MRA).
Jin et al.[41]	China	Retrospective	Model development set: 347 patients Test set: 146 patients	The network structure is based on a general U-shaped design for medical image segmentation and detection. The network includes a fully convolutional technique to detect aneurysms in high-resolution DSA frames.	The system was shown to be highly sensitive in identifying cerebral aneurysms.	DNN techniques have been successfully applied to the automatic segmentation and detection of aneurysms in 2D DSA pictures.
Lv et al.[42]	China	Cross-Sectional	65 Patients	The models include RFs, Neural Networks, Generalized Linear Model, Partial Least Squares, Gradient Boosting Machines, SVM, Linear Discriminant Analysis, Mixture Discriminant Analysis, and K Nearest Neighbors.	Gradient boosting had the best performance among machine learning models in predicting wall enhancement.	Size ratio, PHASES score, and mean wall shear stress at the aneurysm wall were identified as crucial predictors for wall enhancement in cerebral aneurysms using a machine learning approach.
Ou et al.[43]	China	Retrospective	374 Patients	Machine learning methods, including SVM, artificial neural network, and XGBoost, along with conventional logistic regression, were used to create prediction models.	XGBoost had the best performance and key predictors for rupture included location, size ratio, and triglyceride level. The results point to the possibility of improving the treatment of unruptured aneurysms	Utilizing a machine learning model to assess the risk of aneurysm rupture is feasible.
Podgorsak et al.[44]	USA	Retrospective	Training set: 250 DSA Images Test set: 100 DSA Images	A CNN architecture was implemented using Keras for semantic segmentation.	In parametric imaging procedures, CNN segmentation of aneurysms and the surrounding vasculature from DSA images is a non-inferior method compared to manual contouring of aneurysms.	CNN can effectively and accurately segment saccular aneurysms and surrounding vasculature from DSA images

Table 2. Summary findings of studies included in this scoping review (Part I) (continue)

Author	Country	Type of study	Population	DL model	Outcome	Conclusion
Poppenberg et al.[45]	USA	Translational	Training set: 94 subjects Test set: 40 subjects	The study utilized LASSO, K-Nearest Neighbors, RF, and SVMs.	In the study, feature selection using LASSO identified 37 IA-associated transcripts in the training cohort. The RF model outperformed others in both training and testing cohorts. Importantly, comorbidities and demographics did not significantly impact IA prediction.	Predictive mode enhancements were done by employing LASSO for feature selection and robust machine learning techniques.
Rajabzadeh-Oghaz et al.[46]	USA	Retrospective	47 Patients	The study developed a rupture discriminator model for IAs, identifying 3 significant features: aneurysm size ratio, time-averaged normalized WSS, and OSI.	The RRS is not a predictor of rupture but serves as a data-driven model assessing the similarity of UIAs to ruptured ones in morphology and hemodynamics.	The study underscores RRS's clinical utility as an adjunctive tool for managing UIAs in real-world scenarios.
Shi et al.[47]	China	Retrospective	Internal Validation sets: 2355 subjects (There are 5 separate cohorts in this group) External Validation sets: 674 subjects (There are 3 separate cohorts in this group)	The study introduced DAResUNet, a 3D CNN, designed for the segmentation of IAs from digital subtraction CTA images.	When compared to human experts, the suggested DL-based model for automated intracranial aneurysm diagnosis and segmentation showed higher patient-level and lesion-level sensitivity.	The proposed DL-based model for automated detection and segmentation of IAs demonstrated higher patient-level sensitivity and lesion-level sensitivity compared to human experts, suggesting its potential to reduce their workload.
Bhurwani et al.[48]	USA	Retrospective	163 Patients	A DNN was trained to predict the binary outcome of IA occlusion (occluded/unoccluded).	The results indicate the possibility to forecast the outcome of an intervention in real time during surgery by comparing API characteristics with blood flow.	API data was analyzed with DNNs, suggesting the potential to correlate API parameters with blood flow and predict intervention success in real-time during surgery.
Wu et al.[49]	USA	Retrospective	Training set: 436 patients Validation set: 50 patients Internal testing: 60 patients Testing: 670 patients	A two-step model was developed for IA detection: a 3D RPN to locate IAs and 3D DenseNets for classification. DPN was used for the detection step, and DenseNet for probability prediction at suspicious locations.	Compared to an available model, the new model showed statistically higher patient-level accuracy, sensitivity, and specificity.	CADIA displayed commendable diagnostic performance for detecting and localizing IAs.

Table 2. Summary findings of studies included in this scoping review (Part I) (continue)

Author	Country	Type of study	Population	DL model	Outcome	Conclusion
Xia et al.[50]	China	Retrospective	Training set: 485 patients Internal testing: 122 patients External Validation: 202 patients	The study utilized a RF machine learning algorithm for predicting clinical outcomes after the rupture of ACoA aneurysms.	Patient factors such as age, ventilated breathing status, WFNS grade, and Fisher grade are identified as significant predictors of poor outcomes after the rupture of ACoA aneurysms.	Poor outcomes are found to be significantly associated with patient age, breathing status, WFNS grade, and Fisher grade, while morphological parameters of the aneurysm are not independent predictors.
Yang et al.[51]			103 Patients	The study introduced a surface-based DL framework that combines human intervention with automated processes. The system samples 3D vessel surface fragments, classifies them using the PointNet++ DL network to distinguish those with and without aneurysms, and applies surface segmentation (SO-Net) to fragments containing aneurysms.	While there are still issues with small aneurysms, the surface-based segmentation method performs better than the volume-based approach in most situations. Overall, the method improves segmentation accuracy by efficiently filtering out non-aneurysmal components.	Employing a two-step approach, involving classification and segmentation using advanced point-based DL networks, the proposed framework outperforms existing volume-based methods.
Zeng et al.[52]	China		300 original sequences with 263 aneurysms	This paper utilizes a DL approach for intracranial aneurysm detection in 3D-RA, employing a SIF method.	This approach allows training on a 2D CNN directly, avoiding the computationally expensive 3D-CNN, by leveraging time series with evident frame-to-frame correlation. The results demonstrate the practicality and effectiveness of the SIF feature.	Evaluation revealed effective improvement in aneurysm detection accuracy with SIF features, but careful consideration of the upper limit of scale is necessary to avoid introducing redundant information.
Zhu et al.[53]	China	Retrospective	1897 ICA	Three machine learning models-RF, SVM, and feedforward ANN-were developed for assessing IA stability.	Machine learning models outperformed statistical LR and the PHASES score and resulted in the potential of machine learning to enhance clinical decision-making for IA stability assessment.	Machine learning models surpassed traditional statistical methods (LR) and the PHASES score in assessing intracranial aneurysm stability.
Duan et al.[54]	China	Diagnostic	Training set: 241 subjects Test set: 40 subjects	In this study, a two-stage CNN-based detection network was developed to implement the automatic detection of intracranial aneurysm on DSA images.	The proposed method had better accuracy than classical DIP.	CAD architecture is able to help physicians quickly and effectively diagnose IAs.
Hanaoka et al.[55]	Japan		-	Linear SVM was utilized for classifying.	Lung nodules and cerebral aneurysms were successfully identified using the suggested technique.	Two medical lesion identification applications demonstrated the high general versatility of the HoTPiG image feature set.

Table 2. Summary findings of studies included in this scoping review (Part I) (continue)

Author	Country	Type of study	Population	DL model	Outcome	Conclusion
Liu et al.[56]	China	Retrospective	368 ICA	The whole Morphology Prediction Models were built using general linear and ridge regression, and were dubbed the GLM model and the ridge model. The partial morphology model was created using Lasso regression and was given the term Lasso model.	Lasso regression identified flatness as the most crucial morphological feature for predicting aneurysm stability. For unstable aneurysms, spherical disproportion was higher in patients with hypertension.	Flatness was identified as a key determinant for predicting aneurysm stability. Machine learning models, especially with data from multiple centers, could enhance the predictive accuracy of aneurysm stability.
Liu et al.[57]	China	Retrospective	594 ICA	A two-layer feed-forward ANN was constructed to predict the rupture risk of ACOM aneurysms.	In training, validation, testing, and overall datasets, the ANN performance was assessed using ROC analysis, which showed strong classification abilities for both ruptured and unruptured samples.	The management of unruptured ACOM aneurysms may be made easier by this ANN's good performance and useful tool for predicting rupture risk in ACOM aneurysms.
Castro et al.[58]	USA	Retrospective	5,589 patients were classified as having aneurysms, and 54,952 controls were matched to those patients. There were 300 patients for validation.	Utilizing NLP in conjunction with the EMR, patients with cerebral aneurysms and their matched controls were accurately identified.	Compared to models that exclusively used coded or NLP variables, the suggested model performed better. The study demonstrates how a combination method utilizing NLP and ICD codes can correctly identify and categorize patients with cerebral aneurysms.	Using NLP and EMR to collect a substantial group of patients with IAs and corresponding controls and proposed algorithms has the potential to be adapted for various diseases.
Meuschke et al.[59]	Germany	-	-	-	The tool was user-friendly for all experts, and they expressed a willingness to use it for analyzing cerebral stress tensors.	This study introduced a framework for assessing the potential rupture risk of cerebral aneurysms and aims to facilitate the introduction of wall stress into clinical discussions by offering novel glyph visualizations of tensor information.
Haraguchi et al.[60]	Japan	-	-	A new mechanical coil insertion system was developed for the single-operator control.	The developed coil insertion system operated smoothly without issues.	The developed mechanical coil insertion system demonstrated successful endovascular embolization of IAs in an in vitro experiment without any issues.

Table 2. Summary findings of studies included in this scoping review (Part I) (continue)

Author	Country	Type of study	Population	DL model	Outcome	Conclusion
Johnson et al.[61]	USA	-	Three patient-specific cerebral aneurysm models were created.	The paper introduces an innovative method to model weakened cerebral aneurysm walls by creating an equivalent wall thickness.	The results show that the use of the equivalent wall thickness provides a more accurate rupture site prediction than utilizing a uniform wall thickness.	A novel approach for estimating equivalent wall thickness in cerebral aneurysm models involves parameterizing surfaces and deforming a healthy model mesh to match an aneurysm's shape.

IAs: Intracranial aneurysms; CNN: Convolutional neural network; SVM: Support vector machine; RF: Random forest; MLP: Multi-layer perceptron; TOF-MRA: Time-of-flight magnetic resonance angiography; OR: Operating room; HMD: Head-mounted display; VR: Virtual reality; CAV: Cerebral angiographic vasospasm; HU: Hounsfield units; TPOT: Tree-based Pipeline Optimization; UIAs: Unruptured intracranial aneurysms; ML: Machine learning; CAD: Computer-aided diagnosis; FCN: Fully convolutional network; DSA: Digital subtraction angiography; RRS: Rupture risk score; WSS: Wall shear stress; OSI: Oscillatory shear index; DNN: Deep neural network; RPN: Region proposal network; DPN: Dual-pass network; ACoA: Anterior communicating artery; WFNS: World Federation of Neurosurgical Societies; 3D-RA: 3D-Rotational angiography; SIF: Spatial information fusion; DIP: Digital image processing; CAD: Computer-aided diagnosis; ICA: Internal carotid artery; PCoA: Posterior communicating artery; NLP: Natural language processing; EMR: Electronic medical record; DL: Deep learning; MRA: Magnetic resonance angiography; ML: Machine learning

Table 2. Summary findings of studies included in this scoping review (Part II)

Author	Quality of evidence	Specificity	Sensitivity	AUC	Preprocessing algorithm	Aneurysm details
Wang et al.[13]	Moderate risk	-	Internal validation: 0.942 Test set: 0.970	-	Convert CT angiograms from DICOM to numpy matrices; Normalize grayscale values using DICOM window width and level; Preprocess with Dr. Wise-CTA for arterial tree extraction and vessel naming; Extract 3D image patches along the arterial vasculature; Crop images to 80x80x80 pixels; Design patches to cover most aneurysms	Aneurysm type: Saccular, Fusiform Size: < 3 mm: 34n 3-7 mm: 946n 7-10 mm: 454n > 10 mm: 408n Location: ACA, ACoA, MCA, PCA, PCoA, ICA, Basilar A, Vertebral A, and Others
Chen et al.[21]	Moderate risk	-	-		-	-
Feng et al.[22]	Low risk	-	-	Training set SVM: 0.86 RF: 0.85 MLP: 0.90 Test set SVM: 0.85 RF: 0.88, MLP: 0.86	-	Size: - Location: ACA, ICA, MCA, PCA

Table 2. Summary findings of studies included in this scoping review (Part II) (continue)

Author	Quality of evidence	Specificity	Sensitivity	AUC	Preprocessing algorithm	Aneurysm details
Ham et al.[23]	Moderate risk	Internal dataset: 0.893 external datasets: 0.856 (with a 2:1 ratio of normal to aneurysmal patches)	Internal dataset: 0.926 external datasets: 0.879 (with a 2:1 ratio of normal to aneurysmal patches)	-	Skull-stripping, signal intensity normalization, and N4 bias correction	Size: Mean size of 2.6 mm in Internal Validation group Location: -
Jiang et al.[24]	Moderate risk	-	-	SVM:0.86 GLM: 0.82 GLMNet: 0.83 RF: 0.78	-	All aneurysms were saccular aneurysms. Size: 4-25 mm Location: ICA, MCA, ACA
Liu et al.[25]	Moderate risk	-	92.3%	-	Conversion from DICOM to NIfTI format; Manual segmentation of aneurysms; Determination of intracranial artery boundaries; Image cropping based on determined boundaries; Normalization of cropped images using MATLAB	Size: Average diameter 7.1 mm Location: ICA (Anterior circulation aneurysm), MCA, Posterior circulation aneurysm
Patel et al.[26]	Moderate risk	-	-	-	Dataset was first preprocessed to generate co-registered, re-sampled, ROIs of the major arteries of the circle of Willis, which was followed by ground truth generation and data normalization.	Size: 6.01 mm mean Location: ICA, MCA, PComA, OpthA, AComA
Shao et al.[27]	-	-	-	-	-	-
Allgaier et al.[28]	-	-	-	-	-	-
Lei and Yang[29]		MRA: 100% DSA: 86.01%	MRA: 95.87% DSA: 91.46%	-	-	-
Li et al.[30]	Low risk	Model A: 87.3% Model B: 73.5% Model C: 65.1% (Only test set was reported)	Model A: 77.8% Model B: 61.1% Model C: 41.2% (Only test set was reported)	Model A: 0.909 Model B: 0.739 Model C: 0.552 (Only test set was reported)	-	Unruptured saccular aneurysm Size: median of the maximal diameter of the aneurysm 3.9 mm Location: ICA/PCOM, AC, PC, MCA
Tian et al.[31]	Moderate risk	-	-	ANN: 0.761 RF: 0.735 LR: 0.668 (Only test set was reported)	-	Size: 6.92 mm Location: Anterior circulation, Posterior circulation, Distal aneurysm

Table 2. Summary findings of studies included in this scoping review (Part II) (continue)

Author	Quality of evidence	Specificity	Sensitivity	AUC	Preprocessing algorithm	Aneurysm details
Wu et al.[32]	Moderate risk	-	90% for 1 false positive per image	0.906	Truncate intensities of all CTA images between HU Resample each CTA image into isotropic resolution using B-spline interpolation.	Size Training set (6.2 mm) Test set (6.9 mm) Location: -
Kim et al.[33]	Moderate risk	0.77	0.78	0.88	-	Aneurysmal subarachnoid hemorrhage (aSAH) Size: - Location: Ophthalmic A, Distal ICA, PCoA, Anterior choroidal Artery, ICA bifurcation, M1 (first segment of the middle cerebral artery), MCA bifurcation, A1 (first segment of the anterior cerebral artery), ACoA, Distal ACA, Vertebral A, Posterior inferior cerebellar A, Basilar tip
Ou et al.[34]	Low risk	-	1.000	-	Normalization	Size: 5.3 mm Location: ICA, MCA, ACA and AComA, PComA, Posterior circulation
Pennig et al.[35]	High risk	-	85.7%	-	Brain extraction with SPM8; Image standardization and intensity normalization; Multi-scale vessel enhancement filter application; Normalization of CTA image and vessel enhanced images;	Size: mean volume 129.2 mm^3 Location: AC, ICA, ACA, MCA, PC
Afzal et al.[36]	High risk	-	-	-	-	-
Chen et al.[37]	Low risk	LR: 74.6% RF: 81.8% MLP: 76.4% SVM: 83.6% (Only external validation numbers is reported.)	LR: 83.0% RF: 69.8% MLP: 79.3% SVM: 67.9% (Only external validation numbers is reported.)	LR: 0.886 RF: 0.871 MLP: 0.851 SVM: 0.863 (Only external validation numbers is reported.)	-	Size: mean of 5.6 mm Location: PCoA, ACoA, ICA, MCA and Others
Detmer et al.[38]		0.770-0.925 (from lowest to highest ML)	0.348-0.758 (from lowest to highest ML)	MLP: 0.83 LRM: 0.82 (Best two MLs)	-	-

Table 2. Summary findings of studies included in this scoping review (Part II) (continue)

Author	Quality of evidence	Specificity	Sensitivity	AUC	Preprocessing algorithm	Aneurysm details
Chen et al.[39]	Low risk	-	Internal test dataset: 94.4% External test dataset: 82.9%	-	-	Unruptured cystic aneurysm Size: Training dataset 6.86 mm, Internal test dataset 6.30 mm, External test dataset 6.48 mm Locations: ICA, MCA, ACA, PCA, Basilar A, Vertebral A
Duan et al.[40]		-	-	-	-	Size: 3.7 mm mean Location: ACA, MCA, ICA
Jin et al.[41]	Moderate risk	-	89.3%	-	-	Size: - Location: Sidewall aneurysms, Bifurcation aneurysm
Lv et al.[42]	Moderate risk	Values ranged from 0.50 (knn) to 0.75 (gbm and glm)	Values ranged from 0.73 (mda and svmRadial) to 0.91 (rf, lda and knn)	Values ranged from 0.68 (pls) to 0.98 (gbm)	R-project library ‘caret’ was used to perform the preprocessing steps to center (subtracting the mean) and scale (divided by the standard deviation) the data.	-
Ou et al.[43]	Moderate risk	XGBoost: 77.0% ANN: 78.0% SVM: 81.0% LR: 83.0% PHASES: 64.0%	XGBoost: 90.9% ANN: 74.0% SVM: 72.6% LR: 72.0% PHASES: 79.7%	XGBoost: 0.881 ANN: 0.837 SVM: 0.838 LR: 0.801 PHASES: 0.758	-	Size: Unruptured group 3.63 mm mean Ruptured group 4.33 mm mean Location: ICA, MCA, ACA, PCA, BA, VA, AComA, PComA
Podgorsak et al.[44]	Moderate risk	-	-	-	-	-
Poppenberg et al.[45]		-	-	-	-	-
Rajabzadeh-Oghaz et al.[46]	Moderate risk	-	-	-	-	Size: 3.95 mm mean Location: ACA, AComA, ICA, MCA, PComA, Posterior circulation
Shi et al.[47]	Moderate risk	Internal cohort 1: 74.7% Internal cohort 2: 83.9% Internal cohort 3: 85.5% Internal cohort 4: 87.9% Internal cohort 5: 89.7% NBH cohort: 71.1% TJ cohort: 71.1% LYG cohort: 74.6%	-	-	-	Size: Internal cohort1: 4.3 mm Internal cohort 2: 4.8 mm Internal cohort 3: 4.2 mm Internal cohort 4: 3.5 mm Internal cohort 5: 5.1 mm NBH cohort: 4.4 mm TJ cohort: 5.3 mm LYG cohort: 4.4 mm Location: MCA, ACoA, ICA, PCoA, VBA, CA, ACA, PCA

Table 2. Summary findings of studies included in this scoping review (Part II) (continue)

Author	Quality of evidence	Specificity	Sensitivity	AUC	Preprocessing algorithm	Aneurysm details
Bhurwani et al.[48]	Low risk	0.57 (Only in Peak Height mode is reported)	0.92 (Only in Peak Height mode is reported)	-	-	Size: - Location: ICA, ACA, ACoA, MCA, PCA, PCoA, VA, BA
Wu et al.[49]	Moderate risk	0.564 (At 1 FPPV)	0.893 (At 1 FPPV)	0.873	-	Size: (Across all cohorts) 2.5-2.9 mm: 181 IAs 3-5 mm: 349 IAs 5-10 mm: 287 IAs ≥ 10 mm: 55 IAs Locations: ICA, MCA, ACoA, PCoA, BA, ACA, PCA
Xia et al.[50]	Moderate	Internal Test: 82.8% External Test: 83.1%	Internal Test: 78.3% External Test: 73.8%	-	-	Only Ruptured ACoA aneurysms were included. Size: mm in Good Outcome group 5.7 mm in Poor Outcome group
Yang et al.[51]	-	-	Sensitivities of the aneurysm class of five networks are 73.63%, 81.08%, 79.49%, 86.11%, and 80.77%,	-	Four preprocessing approaches A, B, C, and D were applied. A has only been applied as a necessary step in DeepMedic, while B,C, and D, were performed as additional masks for the skull-stripping of the TOF-MRA images.	-
Zeng et al.[52]	-	98.19%	99.38%	-	Gamma correction was performed on the original image and then its intensity stretched or shrank to the right levels. Digital subtraction was done to the original data by the pre-contrast sequences.	Size: 2-40 mm Location: -
Zhu et al.[53]	Low risk	RF: 90.9% SVM: 88.3% ANN: 92.9% Classic LR: 88.3% PHASES: 95.4%	RF: 54.4% SVM: 61.2% ANN: 51.5% Classic LR: 33.9% PHASES: 9.7%	RF: 0.850 SVM: 0.858 ANN: 0.867 Classic LR: 0.818 PHASES: 0.589	-	Size:- Location: ICA, MCA, ACA, AComA, Posterior circulation, PComA
Duan et al.[54]	-	-	-	0.942	-	Size: < 5.0 mm: 44 5.0–9.9 mm: 114 10.0–24.9 mm: 101 ≥ 25.0 mm: 2 Location: PCoA region of ICA

Table 2. Summary findings of studies included in this scoping review (Part II) (continue)

Author	Quality of evidence	Specificity	Sensitivity	AUC	Preprocessing algorithm	Aneurysm details
Hanaoka et al.[55]	-	-	80% when the number of false positives was three per case for both applications	-	-	-
Liu et al.[56]	-	-	-	General linear: 0.856 Ridge: 0.856 LASSO: 0.852	-	Size: 4 mm-8 mm Location: ACoA, PCoA, Posterior Circulation, MCA, ICA
Liu et al.[57]	Moderate risk	92.6%	95.0%	-	-	Size: 2.20 mm vessel size in unruptured group 1.94 mm vessel size in ruptured group Location: Anterior Communicating Artery
Castro et al.[58]	Low risk	-	0.78	0.946	-	-
Meuschke et al.[59]	-	-	-	-	-	-
Haraguchi et al.[60]	-	-	-	-	-	-
Johnson et al.[61]	-	-	-	-	-	-

IAs: Intracranial aneurysms; CNN: Convolutional neural network; SVM: Support vector machine; RF: Random forest; MLP: Multi-layer perceptron; TOF-MRA: Time-of-flight magnetic resonance angiography; OR: Operating room; HMD: Head-mounted display; VR: Virtual reality; CAV: Cerebral angiographic vasospasm; HU: Hounsfield units; TPOT: Tree-based Pipeline Optimization; UIAs: Unruptured intracranial aneurysms; ML: Machine learning; CAD: Computer-aided diagnosis; FCN: Fully convolutional network; DSA: Digital subtraction angiography; RRS: Rupture risk score; WSS: Wall shear stress; OSI: Oscillatory shear index; DNN: Deep neural network; RPN: Region proposal network; DPN: Dual-pass network; ACoA: Anterior communicating artery; WFNS: World Federation of Neurosurgical Societies; 3D-RA: 3D-Rotational angiography; SIF: Spatial information fusion; DIP: Digital image processing; CAD: Computer-aided diagnosis; ICA: Internal carotid artery; PCoA: Posterior communicating artery; NLP: Natural language processing; EMR: Electronic medical record; DL: Deep learning; MRA: Magnetic resonance angiography; ML: Machine learning

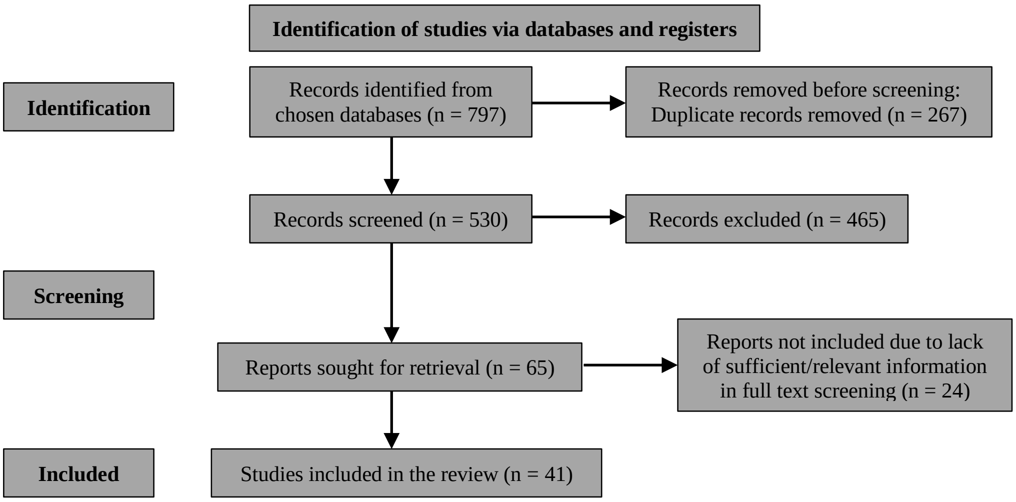

Figure 1. PRISMA flow diagram of the study selection procedure

Studies classified as "high risk of bias" were generally limited to single-center, retrospective designs with small samples or unclear inclusion criteria. The predominance of moderate to high risk of bias among outcome prediction studies in particular reduces the certainty of the evidence base for neurosurgical outcomes. As a result, conclusions regarding the clinical utility and generalizability of DL models for post-surgical or prognostic applications should be interpreted with caution. Risk of bias was less pronounced in larger, multi-center diagnostic studies, which contributed more robust evidence to the review's principal findings.

Imaging modalities and preprocessing approaches: The most frequently used imaging modality was CTA, followed by MRA, DSA, and 3D rotational angiography. Image preprocessing techniques were frequently applied to standardize inputs across varying acquisition protocols. These included skull stripping, vessel enhancement, signal normalization, DICOM-to-NIfTI conversion, and manual segmentation.

Figure 2. Global distribution of deep learning (DL) studies on intracranial aneurysms (IAs)
This map illustrates the geographic origins of studies included in the present scoping review. The size of each circle corresponds to the number of studies affiliated with each country, with text labels indicating the absolute count. Note that the sum of country counts exceeds the total number of unique studies (n = 42) because multi-country collaborations are credited for all contributing countries. The figure highlights the predominance of research output from China and the United States, and reveals areas with limited representation in the global literature.

Some studies employed multi-phase fusion techniques to improve input quality and reduce noise, particularly in models using volumetric CTA data.

DL architectures and analytical frameworks: CNNs were the most widely used architecture, applied in over half of the included studies. Three-dimensional CNN models such as DeepMedic, DAResUNet, and U-Net variants were commonly implemented for aneurysm detection and segmentation tasks, often achieving high levels of accuracy and sensitivity. Hybrid models combining CNNs with support vector machines (SVMs), random forests (RFs), or multilayer perceptrons were also explored to enhance classification and reduce overfitting. In several instances, these combinations demonstrated superior diagnostic performance compared to models relying solely on DL. A smaller subset of studies applied unsupervised learning, AutoML pipelines, or reinforcement learning-based frameworks, indicating a progressive trend toward fully automated and adaptive systems.

Clinical applications (detection, segmentation, and risk stratification): The most common clinical application was aneurysm detection and segmentation. DL-based models for this purpose were evaluated in 24 studies, with many reporting sensitivity rates above 90% and AUC values ranging from 0.85 to 0.98. Some models, such as the one developed by Shi et al., outperformed human readers in both patient-level and lesion-level sensitivity.[47] Segmentation performance was also strengthened in studies that employed 3D patch-based input, multi-scale feature extraction, and false-positive reduction modules.

Moreover, 13 studies focused on rupture risk prediction using DL, often incorporating geometric, hemodynamic, and clinical parameters. These models frequently outperformed conventional risk scoring systems, such as the PHASES score, in predicting rupture likelihood. For example, Liu et al.[57] developed a feedforward artificial neural network capable of predicting rupture risk in anterior communicating artery aneurysms with sensitivity and specificity above 90%. Other studies, such as those by Feng et al.[22] and Zhu et al.,[53] used radiomics-derived morphological features in conjunction with DL models to assess aneurysm instability, achieving comparable or improved predictive performance relative to logistic regression.

DL was applied to outcome prediction following surgical or endovascular treatment in 8 studies. These models aimed to predict periprocedural complications, long-term recanalization, and occlusion outcomes. Predictive variables included both patient-level features and procedure-specific parameters such as aneurysm morphology and blood flow dynamics. The study by Bhurwani et al. exemplified the use of intraoperative data for real-time outcome prediction using a deep neural network (DNN) trained on angiographic parametric imaging.[48]

Natural language processing (NLP) and clinical informatics integration: A subset of studies extended DL applications to broader clinical informatics by incorporating electronic medical record (EMR) data and NLP techniques. For instance, Castro et al. demonstrated that NLP applied to EMRs could effectively identify patients with cerebral aneurysms and their matched controls with a sensitivity of 94.6%, outperforming models based solely on coded variables.[58] These efforts highlight the potential of DL to support automated case detection, large-scale cohort construction, and integrated decision support in clinical environments.

Model validation and performance metrics: Despite generally strong performance across the included studies, validation approaches were often limited in rigor. While internal validation was conducted in most studies through test set separation or cross-validation, only 14 studies reported using external validation datasets. Fewer than 25% of studies employed independent control groups. Reported performance metrics varied but were generally favorable. Sensitivity values typically ranged between 73% and 99%, with specificity values between 54% and 98%. AUC values frequently exceeded 0.85, although methodological heterogeneity limited direct comparability.

Variability in aneurysm characteristics and predictive inputs: There was notable heterogeneity in how aneurysm characteristics were defined and incorporated into models. Some studies stratified by size, location, or aneurysm type (e.g., saccular vs. fusiform), while others included mixed cohorts without clear subcategorization. Aneurysm sizes across studies ranged from < 3 mm to > 25 mm, and many studies did not report rupture site, wall enhancement, or other pathophysiologically relevant features. This variability, along with inconsistent inclusion criteria and imaging protocols, underscores the challenge of synthesizing findings across studies and emphasizes the need for standardized definitions and reporting.

Synthesis and emerging trends: The cumulative evidence indicates that DL algorithms hold considerable promise in advancing the detection, classification, and management of IAs. CNN-based segmentation models consistently demonstrated excellent diagnostic performance, and emerging architectures, such as surface-based models and attention-enhanced networks, have begun to address challenges related to false positives and aneurysm localization. The integration of clinical, radiomic, and hemodynamic features into DL frameworks has further enhanced their predictive power, particularly in risk assessment and treatment outcome forecasting.

Nevertheless, critical limitations persist. The reliance on institution-specific, non-public datasets restricts reproducibility, while the lack of external and prospective validation weakens generalizability. Few studies reported on model explainability or integration within clinical workflows, issues that will be crucial for regulatory approval and adoption. Ethical considerations, including patient data privacy and the interpretability of model decisions, were also underreported in the current literature.

Overall, the included studies demonstrate that DL technologies are rapidly transforming the landscape of neurovascular imaging and risk stratification. While current models show high diagnostic and predictive accuracy, their translation into routine clinical practice will require methodological standardization, access to multicenter and open datasets, robust external validation, and continued development of interpretable, clinically aligned algorithms.

Thematic analysis of included studies: To synthesize the diverse applications and methodologies of DL across the included literature, we conducted a thematic analysis based on the extracted study characteristics. Seven key thematic categories emerged, reflecting both clinical relevance and methodological diversity (Figures 3 and 4).

1. Clinical application domains

The included studies addressed 4 primary clinical applications of DL in the context of IAs, detection and segmentation, rupture risk prediction, treatment outcome forecasting, and automated patient identification using NLP. Detection and segmentation were the most common focus, observed in more than half of the studies. Thirteen studies explored rupture risk stratification, while 8 evaluated DL models for forecasting outcomes such as procedural complications, occlusion success, or recanalization. A smaller subset leveraged NLP models in conjunction with electronic medical records to identify aneurysm cases at scale.

2. Model architectures and analytical frameworks

A variety of DL architectures were employed, including CNNs, ANNs, DNNs, and AutoML pipelines. CNNs were the predominant model type, particularly in 3D implementations such as DeepMedic, DAResUNet, and U-Net variants. Hybrid models that combined DL with traditional ML classifiers (e.g., SVMs, and RFs) were also common. A few studies explored unsupervised learning and reinforcement learning, indicating ongoing diversification in computational approaches.

3. Validation strategy and evidence rigor

Validation approaches varied widely. While most studies employed internal validation (e.g., hold-out test sets or cross-validation), only 14 conducted external validation using independent datasets. Furthermore, less than one-quarter of the studies incorporated control groups. This heterogeneity in methodological rigor reflects differing levels of evidence strength and reproducibility.

4. Imaging modalities and data inputs

The studies utilized a range of imaging inputs, including CTA, MRA, DSA, and 3D rotational angiography. CTA was the most commonly used modality, often preprocessed using normalization, skull stripping, and artifact reduction techniques. Some studies employed multi-phase imaging or computational fluid dynamics to enrich input features, particularly in rupture prediction tasks.

5. Geographic and institutional distribution

The geographic distribution of studies was heavily skewed toward East Asia, particularly China, followed by the United States, South Korea, and a few European countries. This concentration suggests a strong regional research interest but also raises concerns about potential population and data biases that may limit generalizability.

6. Performance reporting and metric completeness

There was considerable variation in how model performance was reported. Most studies included sensitivity, specificity, and AUC, but fewer provided Dice coefficients, false-positive rates, or confidence intervals. Reporting quality was highest in segmentation and detection studies and lower in outcome prediction or NLP-based investigations.

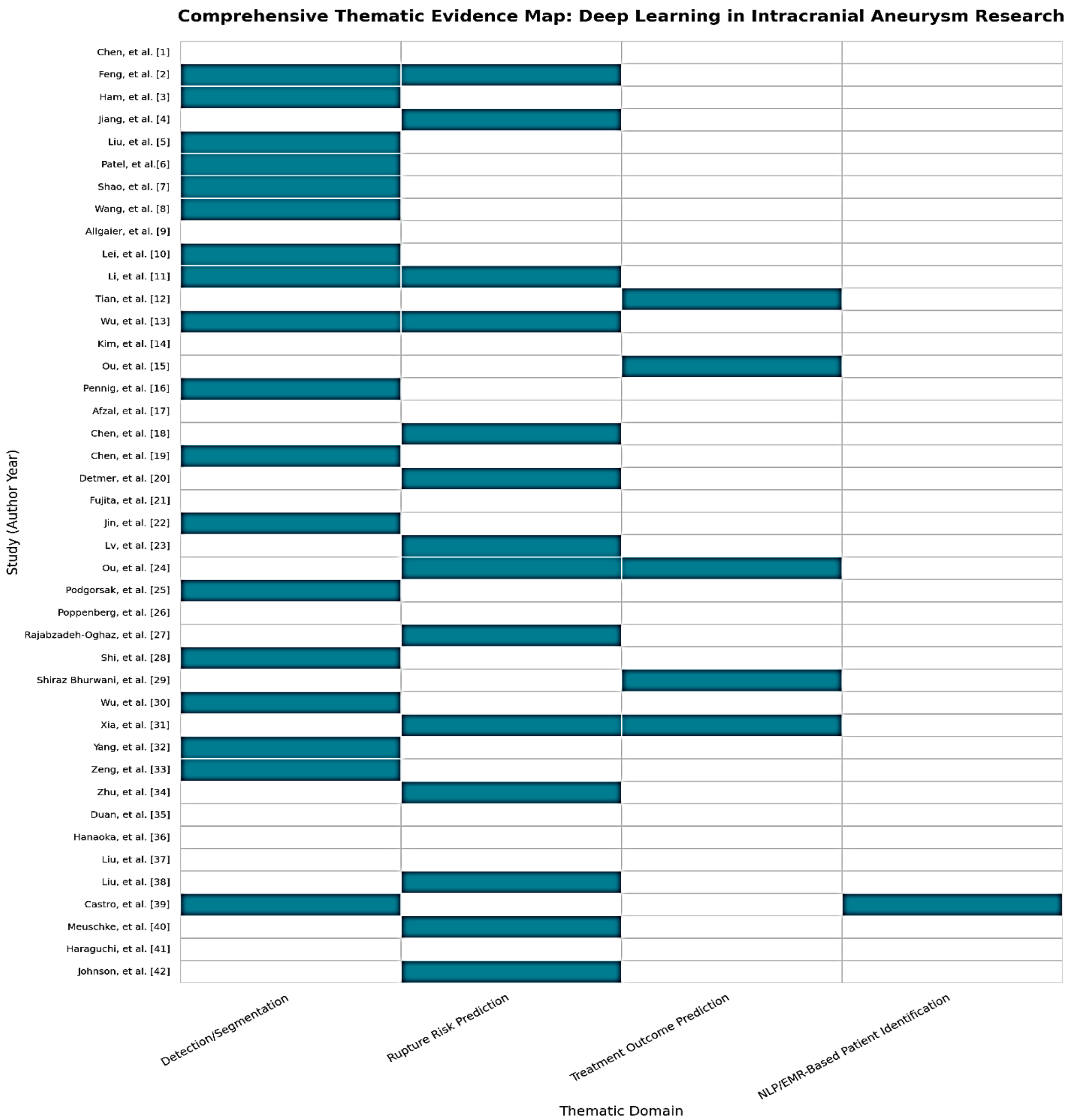

Figure 3. Thematic evidence map of deep learning (DL) applications in intracranial aneurysm research
This matrix visualizes the thematic classification of 42 studies included in the scoping review. Each row represents an individual study (abbreviated by first author and year), and each column denotes one of four primary domains of application: detection and segmentation, rupture risk prediction, treatment outcome prediction, and NLP/EMR-based patient identification. Shaded cells indicate that a study contributed substantively to the corresponding thematic area. This map highlights the concentration of research in detection-focused applications, while revealing gaps in external validation and clinical integration in underrepresented domains such as outcome forecasting and EMR-driven case identification.

7. Focus on ruptured vs. unruptured aneurysms

Several studies explicitly focused on unruptured aneurysms or specific subtypes, such as anterior communicating artery aneurysms. However, many studies did not clearly delineate between ruptured and unruptured lesions. This thematic ambiguity reflects a broader lack of consensus in the field regarding the most clinically actionable prediction targets.

Discussion

Several studies have recently used DL models to detect IAs using neuroimaging. Despite the challenges posed by size and location variability, image quality, imaging modality limitations, and artefacts in imaging, researchers have consistently reported high sensitivity, specificity, and accuracy. However, biases and concerns with the datasets restrict the overall diagnostic accuracy of this research.

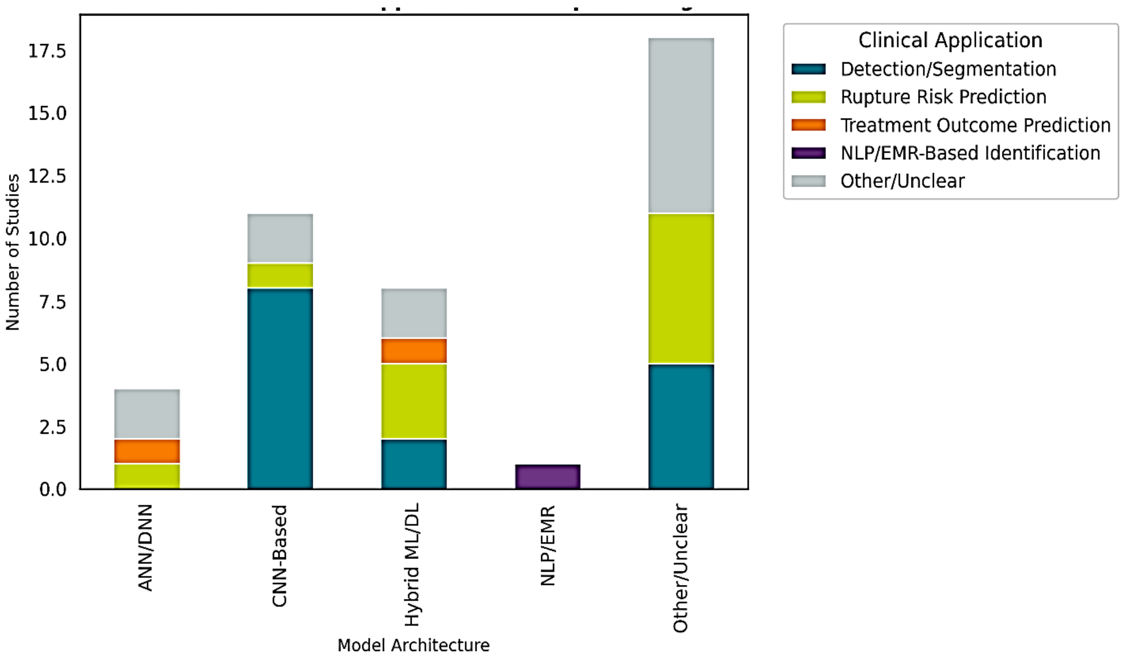

Figure 4. Model architecture versus clinical application in deep learning (DL) studies of intracranial aneurysms (IAs)

This stacked bar chart illustrates the distribution of DL model architectures (horizontal axis) mapped to their primary clinical applications (stacked colors) across the included studies. CNN-based models are most frequently employed for detection and segmentation tasks, while hybrid ML/DL approaches and ANN/DNN architectures demonstrate broader use in rupture risk and outcome prediction. The chart reveals a relative underrepresentation of NLP/EMR-based applications, underscoring the need for further exploration of clinical informatics integration in this domain

For example, several studies were conducted in a single center, which could be a primary cause of bias in the outcome.[3,6,8,16,25] Additional constraints included short test dataset sizes, small IAs in the test dataset, type of studies, and reliance on internal rather than external datasets.

Given the diversity of article types, no consistency or uniformity was found in the inclusion and exclusion criteria. In some of the studies examined, the criteria regarding IAs were based on factors such as the size of the aneurysms,[62] while in others, criteria were based on the presence of certain types or specific locations of IAs.[25] This nuanced approach to criteria underscores the complexity and variability inherent in the characteristics of IAs and their potential impact on the outcomes of the DL methods being investigated.

Additionally, within the scope of the reviewed studies in this article, it is imperative to note that only 2 of them incorporated a control group in their methodology.[63,64] This deficiency in including control groups can potentially introduce bias and significantly influence the outcomes derived from the proposed DL methodologies. Control groups serve as essential benchmarks against which the effectiveness and efficacy of new interventions or techniques can be evaluated. Consequently, the absence of control groups compromises the studies' internal validity and undermines their findings' reliability and generalizability. In future research endeavors, investigators must incorporate control groups systematically to enhance the robustness and credibility of their conclusions in DL methodologies.

Several studies included in this review were conducted at a single center. A limitation of single-centered studies is their potential lack of generalizability to broader populations or varied clinical settings, as they often reflect the characteristics and practices unique to a specific institution or patient population. Additionally, the findings from single-centered studies may be influenced by local biases or confounding factors, necessitating validation across multiple centers to establish robustness and applicability.

Despite the extensive utilization of various DL models in multiple studies focusing on IA analysis, a critical gap remains in the absence of a comprehensive comparative study dedicated to assessing the variability of results concerning model architecture. Among the studies reviewed, CNN was the most used model.

A notable limitation across all the reviewed studies is the exclusive validation of their DL models on private datasets. This practice hinders the independent verification of method effectiveness and the assessment of result variability concerning the utilized DL model architecture. The inability to access code and datasets from authors, often due to patient privacy policies and regulatory constraints, further compounds this issue, preventing researchers from replicating findings or comparing methodologies across studies effectively. However, it is essential to note exceptions, such as the studies by Feng et al.,[22] Ham et al.,[23] Chen et al.,[21] and Shi et al.[47]

Aneurysmal subarachnoid hemorrhage is one of the complications of a ruptured aneurysm, and its mortality rate is significantly high. Furthermore, developing a robust prediction model is necessary to assess the rupture risk of aneurysms. On the one hand, several studies have performed this process using ML-based algorithms.[24,29,65] On the other hand, of the included studies in this review, the number of articles that aimed to investigate this section was insufficient.

Preprocessing was applied in approximately a quarter of the studies. Unfortunately, the number of these studies was not significant compared to the number of studies included in this review, so it might affect the outcome and inference of the studies.

Interpretability and the "Black-Box" challenge

Despite the impressive diagnostic and predictive performance demonstrated by DL models in IA research, their clinical integration is fundamentally limited by the persistent "black-box" problem. Most DL models, particularly deep CNNs, operate through highly non-linear, high-dimensional feature spaces, making it challenging—even for developers—to elucidate the underlying logic of their outputs. The opacity of these models undermines clinicians' trust, as critical decisions must be explainable and justifiable, particularly in high-stakes neurovascular care. This interpretability gap complicates error analysis, bias detection, and model calibration, thereby impeding regulatory acceptance and routine clinical adoption. The lack of visual or quantitative explanation tools—such as heatmaps, attention maps, or saliency analyses—in the reviewed literature further amplifies this challenge, highlighting an urgent need for investment in explainable AI (XAI) frameworks and clinician-in-the-loop validation studies.[63,66]

Limitations of internal-only datasets and lack of multicenter validation

A substantial proportion of the included studies relied exclusively on internal, single-institution datasets for both model development and validation. While this approach may be sufficient for technical proof-of-concept, it significantly limits the generalizability of findings. Models trained and tested on a single dataset are prone to overfitting, potentially capturing site-specific imaging artifacts, acquisition protocols, or population demographics that do not translate across settings. The absence of multicenter, external validation further restricts the credibility and clinical applicability of these models. Without rigorous testing on diverse, independent cohorts, it is difficult to ascertain whether the reported performance reflects genuine clinical utility or is simply an artifact of local data characteristics. This methodological limitation is a major barrier to both academic benchmarking and regulatory endorsement.[67]

Lack of benchmarking across models and imaging platforms

Another critical gap identified in the current literature is the lack of systematic benchmarking across DL models, architectures, or imaging platforms. Many studies introduce novel models or variations but evaluate them only in isolation, using proprietary datasets and differing performance metrics. This heterogeneity prevents meaningful cross-study comparisons and impedes consensus regarding best-in-class models or optimal imaging modalities. Few studies provide head-to-head comparisons between different DL architectures (e.g., CNN vs. hybrid models) or assess robustness across imaging types such as CTA versus MRA. The field would benefit from collaborative benchmarking initiatives and standardized public datasets, which would enable transparent evaluation, facilitate reproducibility, and drive collective progress toward clinically reliable AI tools.[68,69]

Regulatory, ethical, and data privacy concerns

Clinical translation of DL in neurovascular imaging is also hindered by significant regulatory, ethical, and data privacy challenges. Most of the reviewed studies do not address how their models comply with existing data protection standards (e.g., HIPAA, GDPR) or how they would manage patient consent and data anonymization at scale. The use of proprietary, locally stored imaging data raises concerns about data security and re-identification risks, especially as model complexity increases. Furthermore, the black-box nature of many DL models presents regulatory hurdles, as agencies such as the FDA increasingly require explainability and robust post-market surveillance for AI-driven devices. Without frameworks for ongoing monitoring, auditing, and real-world performance assessment, the path from research prototype to approved clinical device remains fraught with uncertainty.[70]

Deficit in outcome-based studies

Despite the manuscript's stated emphasis on "neurosurgical outcomes," a striking deficit of outcome-focused research was identified. Only a small minority of the included studies directly addressed post-surgical or prognostic applications, such as predicting occlusion rates, recurrence, or patient functional outcomes following intervention. Instead, the majority of DL applications to date have prioritized anatomical detection, segmentation, or rupture risk stratification, with limited extension to longitudinal or post-therapeutic endpoints. This deficit represents a critical gap in the current landscape, as outcome-based modeling is essential for demonstrating the real-world value of AI in neurosurgical decision-making and patient management. Future research must prioritize the integration of perioperative and follow-up data, the development of longitudinal prediction models, and the validation of these approaches in diverse, prospectively collected cohorts.[25,62,64,65,71-73]

Limitations: There were several limitations in this study. First, only papers published in English were included in this review, which might create bias. Second, our investigation was limited to reviewing papers available through open access, thus overlooking any research inaccessible to a broader audience. This approach might have implications for the comprehensiveness and representativeness of our findings. Third, the research methodology involved investigation across only 3 databases, with an additional scrutiny of targeted pages within Google Scholar. Fourth, the variation in the proposed and utilized ML and DL model is another issue in this field. Although numerous options exist for the use and development of various models and algorithms that might increase the accuracy and efficiency of detection of IAs, this variant might affect the precise outcome.

Limitations and challenges of DL models in detecting IAs

The "black-box" problem, in which the methods of data processing from the input to the output layers are not fully understood, is brought about by the complex structure of DNN algorithms. As a result, doctors could be hesitant to accept the results of a classifier system like this "black-box." Though guidelines and recommendations have recently been proposed, there is currently no clear legal consensus regarding the regulations for AI and ML models, unlike medical devices. Moreover, no clear legal guidelines are available regarding the independent mathematical interrogation and validation of outputs generated by ML systems.[74]

The legitimacy of the decision-making made possible by ML models is also questionable regarding who would be more responsible for these systems—the data scientists and programmers or the treating clinician?

In addition to security concerns when sharing data between institutions and AI systems, ethical considerations when using large-volume patient data include data ownership and consent for an individual's data to be captured in an ML system. Since applied ML in the healthcare industry is still in its infancy, it is expected that problems with permission and data management may come up as the field develops and will need ongoing evaluation as AI advances.

Because insufficient data supports its usage, AI CAD systems for aneurysm detection are not yet ready to be incorporated into standard clinical practice.

If AI CAD tools assessed using internal test sets are reevaluated in subsequent research with anticipated external data, they will add further evidence to the body of knowledge. Large and representative datasets should be employed in studies that build AI tools to assure clinical uptake; clinical validation should then be accomplished through prospective multicenter trials.

Conclusion

The findings of the investigation demonstrate that DL methodologies exhibit promise in the detection of IAs. However, to enhance the robustness and reliability of these findings, future research endeavors necessitate the utilization of larger datasets.

Such datasets must encompass a comprehensive representation of all types of aneurysms, regardless of size and location, to effectively capture the intricacies inherent in aneurysm detection.

Additionally, to fully explain the impact of DL techniques in this field, it is recommended that the design of some studies should be diversified. By implementing varied study methodologies, researchers can better explain the breadth and depth of DL's efficacy in detecting IAs, thereby advancing the field toward more comprehensive and clinically relevant insights.

Conflict of Interests

The authors declare no conflict of interest in this study.

Acknowledgments

None.

References

1. Guo Y, Liu Y, Oerlemans A, Lao S, Wu S, Lew MS. Deep learning for visual understanding: A review. Neurocomputing 2016; 187: 27-48.
2. Awuah WA, Adebusoye FT, Wellington J, David L, Salam A, Weng Yee AL, et al. Recent Outcomes and Challenges of Artificial Intelligence, Machine Learning, and Deep Learning in Neurosurgery. World Neurosurg X 2024; 23: 100301.
3. Kim M, Yun J, Cho Y, Shin K, Jang R, Bae HJ, et al. Deep Learning in Medical Imaging. Neurospine 2019; 16(4): 657-68.
4. Latif J, Xiao C, Imran A, Tu S. Medical imaging using machine learning and deep learning algorithms: a review. 2nd International conference on computing, mathematics and engineering technologies (iCoMET); Sukkur, Pakistan. New York, NY: IEEE; 2019. p. 1-5.
5. Huang J, Shlobin NA, DeCuypere M, Lam SK. Deep Learning for Outcome Prediction in Neurosurgery: A Systematic Review of Design, Reporting, and Reproducibility. Neurosurgery 2022; 90(1): 16-38.
6. Guarneri B, Bertolini G, Latronico N. Long-term outcome in patients with critical illness myopathy or neuropathy: the Italian multicentre CRIMYNE study. J Neurol Neurosurg Psychiatry 2008; 79(7): 838-41.
7. Lu SL, Xiao FR, Cheng JC, Yang WC, Cheng YH, Chang YC, et al. Randomized multi-reader evaluation of automated detection and segmentation of brain tumors in stereotactic radiosurgery with deep neural networks. Neuro Oncol 2021; 23(9): 1560-8.
8. Rudie JD, Rauschecker AM, Bryan RN, Davatzikos C, Mohan S. Emerging Applications of Artificial Intelligence in Neuro-Oncology. Radiology 2019; 290(3): 607-18.
9. Keedy A. An overview of intracranial aneurysms. Mcgill J Med 2006; 9(2): 141-6.
10. Bonneville F, Sourour N, Biondi A. Intracranial aneurysms: an overview. Neuroimaging Clin N Am 2006; 16(3): 371-82, vii.
11. Cianfoni A, Pravatà E, De Blasi R, Tschuor CS, Bonaldi G. Clinical presentation of cerebral aneurysms. Eur J Radiol 2013; 82(10): 1618-22.
12. Abdollahifard S, Farrokhi A, Kheshti F, Jalali M, Mowla A. Application of convolutional network models in detection of intracranial aneurysms: A systematic review and meta-analysis. Interv Neuroradiol 2023; 29(6): 738-47.
13. Wang J, Sun J, Xu J, Lu S, Wang H, Huang C, et al. Detection of intracranial aneurysms using multiphase CT angiography with a deep learning model. Acad Radiol 2023; 30(11): 2477-86.
14. International Study of Unruptured Intracranial Aneurysms Investigators. Unruptured intracranial aneurysms--risk of rupture and risks of surgical intervention. N Engl J Med 1998; 339(24): 1725-33.
15. Turner CL, Higgins JN, Kirkpatrick PJ. Assessment of transcranial color-coded duplex sonography for the surveillance of intracranial aneurysms treated with Guglielmi detachable coils. Neurosurgery 2003; 53(4): 866-71; discussion 71-2.
16. Laukka D, Kivelev J, Rahi M, Vahlberg T, Paturi J, Rinne J, et al. Detection Rates and Trends of Asymptomatic Unruptured Intracranial Aneurysms From 2005 to 2019. Neurosurgery 2024; 94(2): 297-306.
17. Timmins KM, Van der Schaaf IC, Vos IN, Ruigrok YM, Velthuis BK, Kuijf HJ. Geometric deep learning using vascular surface meshes for modality-independent unruptured intracranial aneurysm detection. IEEE Trans Med Imaging 2023; 42(11): 3451-60.
18. Bizjak Ž, Špiclin Ž. A Systematic Review of Deep-Learning Methods for Intracranial Aneurysm Detection in CT Angiography. Biomedicines 2023; 11(11): 2921.
19. Tricco AC, Lillie E, Zarin W, O'Brien KK, Colquhoun H, Levac D, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. Ann Intern Med 2018; 169(7): 467-73.
20. Levac D, Colquhoun H, O'Brien KK. Scoping studies: advancing the methodology. Implement Sci 2010; 5: 69.
21. Chen B, Xie K, Zhang J, Yang L, Zhou H, Zhang L, et al. Comprehensive analysis of mitochondrial dysfunction and necroptosis in intracranial aneurysms from the perspective of predictive, preventative, and personalized medicine. Apoptosis 2023; 28(9-10): 1452-68.
22. Feng J, Zeng R, Geng Y, Chen Q, Zheng Q, Yu F, et al. Automatic differentiation of ruptured and unruptured intracranial aneurysms on computed tomography angiography based on deep learning and radiomics. Insights Imaging 2023; 14(1): 76.
23. Ham S, Seo J, Yun J, Bae YJ, Kim T, Sunwoo L, et al. Automated detection of intracranial aneurysms using skeleton-based 3D patches, semantic segmentation, and auxiliary classification for overcoming data imbalance in brain TOF-MRA. Sci Rep 2023; 13(1): 12018.
24. Jiang J, Rezaeitaleshmahalleh M, Lyu Z, Mu N, Ahmed AS, Md CMS, et al. Augmenting Prediction of Intracranial Aneurysms' Risk Status Using Velocity-Informatics: Initial Experience. J Cardiovasc Transl Res 2023; 16(5): 1153-65.
25. Liu X, Mao J, Sun N, Yu X, Chai L, Tian Y, et al. Deep Learning for Detection of Intracranial Aneurysms from Computed Tomography Angiography Images. J Digit Imaging 2023; 36(1): 114-23.
26. Patel TR, Patel A, Veeturi SS, Shah M, Waqas M, Monteiro A, et al. Evaluating a 3D deep learning pipeline for cerebral vessel and intracranial aneurysm segmentation from computed tomography angiography-digital subtraction angiography image pairs. Neurosurg Focus 2023; 54(6): E13.
27. Shao D, Lu X, Liu X. 3D intracranial aneurysm classification and segmentation via unsupervised dual-branch learning. IEEE J Biomed Health Inform 2022; 27(4): 1770-9.
28. Allgaier M, Amini A, Neyazi B, Sandalcioglu IE, Preim B, Saalfeld S. VR-based training of craniotomy for intracranial aneurysm surgery. Int J Comput Assist Radiol Surg 2022; 17(3): 449-56.
29. Lei X, Yang Y. Deep Learning-Based Magnetic Resonance Imaging in Diagnosis and Treatment of Intracranial Aneurysm. Comput Math Methods Med 2022; 2022: 1683475.
30. Li R, Zhou P, Chen X, Mossa-Basha M, Zhu C, Wang Y. Construction and Evaluation of Multiple Radiomics Models for Identifying the Instability of Intracranial Aneurysms Based on CTA. Front Neurol 2022; 13: 876238.
31. Tian Z, Li W, Feng X, Sun K, Duan C. Prediction and analysis of periprocedural complications associated with endovascular treatment for unruptured intracranial aneurysms using machine learning. Front Neurol 2022; 13: 1027557.
32. Wu K, Gu D, Qi P, Cao X, Wu D, Chen L, et al. Evaluation of an automated intracranial aneurysm detection and rupture analysis approach using cascade detection and classification networks. Comput Med Imaging Graph 2022; 102: 102126.
33. Kim KH, Koo HW, Lee BJ, Sohn MJ. Analysis of risk factors correlated with angiographic vasospasm in patients with

aneurysmal subarachnoid hemorrhage using explainable predictive modeling. J Clin Neurosci 2021; 91: 334-42.
34. Ou C, Liu J, Qian Y, Chong W, Liu D, He X, et al. Automated Machine Learning Model Development for Intracranial Aneurysm Treatment Outcome Prediction: A Feasibility Study. Front Neurol 2021; 12: 735142.
35. Pennig L, Hoyer UCI, Krauskopf A, Shahzad R, Jünger ST, Thiele F, et al. Deep learning assistance increases the detection sensitivity of radiologists for secondary intracranial aneurysms in subarachnoid hemorrhage. Neuroradiology 2021; 63(12): 1985-94.
36. Afzal M, Alam F, Malik KM, Malik GM. Clinical Context-Aware Biomedical Text Summarization Using Deep Neural Network: Model Development and Validation. J Med Internet Res 2020; 22(10): e19810.
37. Chen G, Lu M, Shi Z, Xia S, Ren Y, Liu Z, et al. Development and validation of machine learning prediction model based on computed tomography angiography-derived hemodynamics for rupture status of intracranial aneurysms: a Chinese multicenter study. Eur Radiol 2020; 30(9): 5170-82.
38. Chen G, Wei X, Lei H, Liqin Y, Yuxin L, Yakang D, et al. Automated computer-assisted detection system for cerebral aneurysms in time-of-flight magnetic resonance angiography using fully convolutional network. Biomed Eng Online 2020; 19(1): 38.
39. Detmer FJ, Lückehe D, Mut F, Slawski M, Hirsch S, Bijlenga P, et al. Comparison of statistical learning approaches for cerebral aneurysm rupture assessment. Int J Comput Assist Radiol Surg 2020; 15(1): 141-50.
40. Duan Z, Montes D, Huang Y, Wu D, Romero J, Gonzalez R, et al. Deep Learning Based Detection and Localization of Cerebal Aneurysms in Computed Tomography Angiography 2020.
41. Jin H, Geng J, Yin Y, Hu M, Yang G, Xiang S, et al. Fully automated intracranial aneurysm detection and segmentation from digital subtraction angiography series using an end-to-end spatiotemporal deep neural network. J Neurointerv Surg 2020; 12(10): 1023-7.
42. Lv N, Karmonik C, Shi Z, Chen S, Wang X, Liu J, et al. A pilot study using a machine-learning approach of morphological and hemodynamic parameters for predicting aneurysms enhancement. Int J Comput Assist Radiol Surg 2020; 15(8): 1313-21.
43. Ou C, Liu J, Qian Y, Chong W, Zhang X, Liu W, et al. Rupture Risk Assessment for Cerebral Aneurysm Using Interpretable Machine Learning on Multidimensional Data. Front Neurol 2020; 11: 570181.
44. Podgorsak AR, Rava RA, Shiraz Bhurwani MM, Chandra AR, Davies JM, Siddiqui AH, et al. Automatic radiomic feature extraction using deep learning for angiographic parametric imaging of intracranial aneurysms. J Neurointerv Surg 2020; 12(4): 417-21.
45. Poppenberg KE, Tutino VM, Li L, Waqas M, June A, Chaves L, et al. Classification models using circulating neutrophil transcripts can detect unruptured intracranial aneurysm. J Transl Med 2020; 18(1): 392.
46. Rajabzadeh-Oghaz H, Waqas M, Veeturi SS, Vakharia K, Tso MK, Snyder KV, et al. A data-driven model to identify high-risk aneurysms and guide management decisions: the Rupture Resemblance Score. J Neurosurg 2021; 135(1): 9-16.
47. Shi Z, Miao C, Schoepf UJ, Savage RH, Dargis DM, Pan C, et al. A clinically applicable deep-learning model for detecting intracranial aneurysm in computed tomography angiography images. Nat Commun 2020; 11(1): 6090.
48. Bhurwani MMS, Waqas M, Podgorsak AR, Williams KA, Davies JM, Snyder K, et al. Feasibility study for use of angiographic parametric imaging and deep neural networks for intracranial aneurysm occlusion prediction. J Neurointerv Surg 2020; 12(7): 714-9.
49. Wu D, Montes D, Duan Z, Huang Y, Romero JM, Gonzalez RG, et al. Deep learning based detection and localization of intracranial aneurysms in computed tomography angiography. arXiv:2005.11098v2 2021. [Preprint].
50. Xia N, Chen J, Zhan C, Jia X, Xiang Y, Chen Y, et al. Prediction of clinical outcome at discharge after rupture of anterior communicating artery aneurysm using the random forest technique. Front Neurol 2020; 11: 538052.
51. Yang X, Xia D, Kin T, Igarashi T. Surface-based 3D deep learning framework for segmentation of intracranial aneurysms from TOF-MRA images. arXiv.2006.16161v1 2020. [Preprint].
52. Zeng Y, Liu X, Xiao N, Li Y, Jiang Y, Feng J, et al. Automatic Diagnosis Based on Spatial Information Fusion Feature for Intracranial Aneurysm. IEEE Trans Med Imaging 2020; 39(5): 1448-58.
53. Zhu W, Li W, Tian Z, Zhang Y, Wang K, Zhang Y, et al. Stability Assessment of Intracranial Aneurysms Using Machine Learning Based on Clinical and Morphological Features. Transl Stroke Res 2020; 11(6): 1287-95.
54. Duan H, Huang Y, Liu L, Dai H, Chen L, Zhou L. Automatic detection on intracranial aneurysm from digital subtraction angiography with cascade convolutional neural networks. BioMed Eng OnLine 2019; 18(1): 110.
55. Hanaoka S, Nomura Y, Takenaga T, Murata M, Nakao T, Miki S, et al. HoTPiG: a novel graph-based 3-D image feature set and its applications to computer-assisted detection of cerebral aneurysms and lung nodules. Int J Comput Assist Radiol Surg 2019; 14(12): 2095-107.
56. Liu Q, Jiang P, Jiang Y, Ge H, Li S, Jin H, et al. Prediction of Aneurysm Stability Using a Machine Learning Model Based on PyRadiomics-Derived Morphological Features. Stroke 2019; 50(9): 2314-21.
57. Liu J, Chen Y, Lan L, Lin B, Chen W, Wang M, et al. Prediction of rupture risk in anterior communicating artery aneurysms with a feed-forward artificial neural network. Eur Radiol 2018; 28(8): 3268-75.
58. Castro VM, Dligach D, Finan S, Yu S, Can A, Abd-El-Barr M, et al. Large-scale identification of patients with cerebral aneurysms using natural language processing. Neurology 2017; 88(2): 164-8.
59. Meuschke M, Voß S, Beuing O, Preim B, Lawonn K. Glyph-Based Comparative Stress Tensor Visualization in Cerebral Aneurysms. Comput Graph Forum 2017; 36(3): 99-108.
60. Haraguchi K, Miyachi S, Matsubara N, Nagano Y, Yamada H, Marui N, et al. A mechanical coil insertion system for endovascular coil embolization of intracranial aneurysms. Interv Neuroradiol 2013; 19(2): 159-66.
61. Johnson E, Zhang Y, Shimada K. Estimating an equivalent wall-thickness of a cerebral aneurysm through surface parameterization and a non-linear spring system. Int J Numer Methods Biomed Eng 2011; 27(7): 1054-72.
62. Noori Mirtaheri P, Akhbari M, Najafi F, Mehrabi H, Babapour A, Rahimian Z, et al. Performance of deep learning models for automatic histopathological grading of meningiomas: a systematic review and meta-analysis. Front Neurol 2025; 16: 1536751.
63. Basem J, Mani R, Sun S, Gilotra K, Dianati-Maleki N, Dashti R. Clinical applications of artificial intelligence and machine learning in neurocardiology: a comprehensive review. Front Cardiovasc Med 2025; 12: 1525966.
64. Sajjadi SM, Mohebbi A, Ehsani A, Marashi A, Azhdarimoghaddam A, Karami S, et al. Identifying abdominal aortic aneurysm size and presence using Natural Language Processing of radiology reports: a systematic review and meta-analysis. Abdom Radiol (NY) 2025; 50(8): 3885-99.
65. Nafees Ahmed S, Prakasam P. A systematic review on intracranial aneurysm and hemorrhage detection using machine learning and deep learning techniques. Prog Biophys Mol Biol 2023; 183: 1-16.
66. Hanna MG, Pantanowitz L, Jackson B, Palmer O, Visweswaran S, Pantanowitz J, et al. Ethical and Bias Considerations in Artificial Intelligence/Machine Learning. Mod Pathol 2025; 38(3): 100686.
67. Paullada A, Raji ID, Bender EM, Denton E, Hanna A. Data and its (dis)contents: A survey of dataset development and use in machine learning research. Patterns (N Y) 2021; 2(11): 100336.
68. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts H. Artificial intelligence in radiology. Nat Rev Cancer 2018; 18(8): 500-10.
69. Casey A, Davidson E, Poon M, Dong H, Duma D, Grivas A, et al. A systematic

review of natural language processing applied to radiology reports. BMC Med Inform Decis Mak 2021; 21(1): 179.
70. White T, Blok E, Calhoun VD. Data sharing and privacy issues in neuroimaging research: Opportunities, obstacles, challenges, and monsters under the bed. Hum Brain Mapp 2022; 43(1): 278-91.
71. Yousefi M, Akhbari M, Mohamadi Z, Karami S, Dasoomi H, Atabi A, et al. Machine learning based algorithms for virtual early detection and screening of neurodegenerative and neurocognitive disorders: a systematic-review. Front Neurol 2024; 15: 1413071.
72. Yoonesi S, Abedi Azar R, Arab Bafrani M, Yaghmayee S, Shahavand H, Mirmazloumi M, et al. Facial expression deep learning algorithms in the detection of neurological disorders: a systematic review and meta-analysis. Bio med Eng Online 2025; 24(1): 64.
73. Sharifi G, Hajibeygi R, Zamani SAM, Easa AM, Bahrami A, Eshraghi R, et al. Diagnostic performance of neural network algorithms in skull fracture detection on CT scans: a systematic review and meta-analysis. Emerg Radiol 2025; 32(1): 97-111.
74. Luxton DD. Recommendations for the ethical use and design of artificial intelligent care providers. Artif Intell Med 2014; 62(1): 1-10.