Systematic Review and Meta-Analysis

Medicine

OPEN

Automated identification and quantification of traumatic brain injury from CT scans

Are we there yet?

Atsuhiro Hibi, MSc[a], Majid Jaberipour, PhD[b], Michael D. Cusimano, MD, FRCSC, DABNS, FACS, PhD, MHPE[a,d], Alexander Bilbily, MD, FRCPC[b,e], Rahul G. Krishnan, PhD[f,g], Richard I. Aviv, MBChB, FRCR, FRCPC, dABR[h], Pascal N. Tyrrell, PhD[a,b,c,*]

Abstract

Background: The purpose of this study was to conduct a systematic review for understanding the availability and limitations of artificial intelligence (AI) approaches that could automatically identify and quantify computed tomography (CT) findings in traumatic brain injury (TBI).

Methods: Systematic review, in accordance with PRISMA 2020 and SPIRIT-AI extension guidelines, with a search of 4 databases (Medline, Embase, IEEE Xplore, and Web of Science) was performed to find AI studies that automated the clinical tasks for identifying and quantifying CT findings of TBI-related abnormalities.

Results: A total of 531 unique publications were reviewed, which resulted in 66 articles that met our inclusion criteria. The following components for identification and quantification regarding TBI were covered and automated by existing AI studies: identification of TBI-related abnormalities; classification of intracranial hemorrhage types; slice-, pixel-, and voxel-level localization of hemorrhage; measurement of midline shift; and measurement of hematoma volume. Automated identification of obliterated basal cisterns was not investigated in the existing AI studies. Most of the AI algorithms were based on deep neural networks that were trained on 2- or 3-dimensional CT imaging datasets.

Conclusion: We identified several important TBI-related CT findings that can be automatically identified and quantified with AI. A combination of these techniques may provide useful tools to enhance reproducibility of TBI identification and quantification by supporting radiologists and clinicians in their TBI assessments and reducing subjective human factors.

Abbreviations: 2D = 2-dimensional, 3D = 3-dimensional, AI = artificial intelligence, CNN = convolutional neural network, CT = computed tomography, EDH = epidural (extradural) hemorrhage, ICH = intracranial hemorrhage, ICP = intracranial pressure, IPH = intraparenchymal hemorrhage, IVH = intraventricular hemorrhage, LSTM = long short-term memory, ML = machine learning, MRI = magnetic resonance imaging, RNN = recurrent neural network, SAH = subarachnoid hemorrhage, SDH = subdural hemorrhage, SVM = support vector machine, TBI = traumatic brain injury.

Keywords: artificial intelligence, computed tomography, machine learning, medical imaging, traumatic brain injury

This work was supported by a research grant from Nippon Steel Corporation (Fund Number 509533). RA receives peer reviewed grant support from CIHR project grant (148762), Faculty of Medicine, University of Ottawa, Translational Research Grant, Canada Foundation for Innovation, John R. Evans Leaders Fund (41174), New Frontiers in Research Fund – Exploration (00521), Michael T Richards Fellowship Clinical Research Fellowship, UOttawa, Brain and Mind Research Institute.

The authors have no consent to disclose.

AH is supported by a PhD funding award from Nippon Steel Corporation. AB is an officer and shareholder of 16 Bit Inc., and a consultant for Roche. RGK is a Scientific Advisory Board of Iterative Scopes. PNT is a shareholder of Ace Age Inc., an investigator and consultant of Novo Nordisk, an officer, director and shareholder of SofTx Innovations Inc., an advisory board member of Demeter Innovation Lab, and an advisory board member of Pulsar Music Inc. MJ, MDC, and RA have nothing to disclose.

All data generated or analyzed during this study are included in this published article [and its supplementary information files].

The authors have no ethical statement to disclose.

Supplemental Digital Content is available for this article.

[a] Institute of Medical Science, University of Toronto, Toronto, Ontario, Canada, [b] Department of Medical Imaging, University of Toronto, Toronto, Ontario, Canada, [c] Department of Statistical Sciences, University of Toronto, Toronto, Ontario, Canada, [d] Division of Neurosurgery, St Michael's Hospital, University of Toronto, Toronto, Canada, [e] Sunnybrook Health Sciences Centre, Toronto, Canada, [f] Department of Computer Science, University of Toronto, Toronto, Ontario, Canada, [g] Department of Laboratory Medicine & Pathobiology, University of Toronto, Toronto, Ontario, Canada, [h] Department of Radiology, Radiation Oncology and Medical Physics, University of Ottawa, Ottawa, Ontario, Canada.

** Correspondence: Pascal N. Tyrrell, Department of Medical Imaging, University of Toronto, 263 McCaul Street 4th floor rm 409, Toronto, Ontario M5T 1W7, Canada (e-mail: pascal.tyrrell@utoronto.ca).*

How to cite this article: Hibi A, Jaberipour M, Cusimano MD, Bilbily A, Krishnan RG, Aviv RI, Tyrrell PN. Automated identification and quantification of traumatic brain injury from CT scans: Are we there yet? Medicine 2022;101:47(e31848).

Received: 13 September 2022 / Received in final form: 25 October 2022 / Accepted: 26 October 2022

http://dx.doi.org/10.1097/MD.0000000000031848

1. Introduction

Traumatic brain injury (TBI) is defined as a disruption of brain function by external forces to the head.[1] TBI is a major health concern, and concern and is associated with significant morbidity and mortality. It is estimated that 69 million individuals suffer from TBI worldwide every year,[2] and the medical cost of severe TBI ranges from $600,000 to $1.8 million per patient lifetime.[3] computed tomography (CT) scan findings delineate the structural effects of TBI and neuroimaging techniques play a crucial role in guiding therapy for acute TBI.[4] Although magnetic resonance imaging (MRI) may be more sensitive in detecting small white-matter lessions in the later phases of TBI, conventional CT is the imaging modality of choice during the first 24 hours following injury, due to its wide availability and speed.[4]

There are some key CT findings that are important in estimating TBI diagnosis and prognosis, such as types of hematomas, its locations, extent of midline shift, and hematoma volume. Currently, radiologists and clinicians rely on manual reading of CT images to identify and quantify these neurological findings. However, manual assessment is not always reproducible the measurement can be different from reader to reader. The outcome can even vary in every trial by the same reader. Therefore, TBI identification and quantification process is worth automating for reducing the human factor. It is beneficial not only to radiologists but also referring clinicians because more standardized radiological reports are provided.

Recently, many artificial intelligence (AI) and machine learning (ML) methods have been proposed in an attempt to automate radiological routines related to TBI. However, no study has systematically investigated and summarized these ML studies with respect to the identification and quantification of a wide range of TBI abnormalities. The purpose of this work was to conduct a systematic review of ML studies that describe a methodology for identifying and quantifying TBI-related abnormalities. The question we wanted to answer was whether an automated identification and quantification of TBI from CT scans was currently possible. Specifically, for each paper, we summarized the following: ML's predictions (e.g., types of hematoma or localizations of hematoma); learning strategy (e.g., supervised learning); algorithm design (e.g., architecture of deep neural network); and algorithm performance (e.g., area under the curve). We also discuss the limitations of current ML methods and highlight the future research directions for improving the automated CT identification and quantification of TBI-related lesions.

In the next subsection, we overview the basic concept and some terminology commonly used in ML studies regarding automated TBI identification and quantification process.

1.1. Overview of AI/ML terminology

Generally, the goal of ML is to create a mathematical model that can be trained to produce the expected outputs when new, unseen input data are provided. ML types used in identified articles are roughly divided into *supervised learning* and *unsupervised learning* algorithms. In the supervised learning of medical imaging, the training dataset consists of medical images and paired labels that specify the ground-truth annotations created by medical experts. *Unsupervised learning* processes data without relying on annotations and aims to find useful patterns embedded in the data.

From the point of view of ML, radiological routines for identification and quantification of TBI can be seen using either image-level recognition (*classification* or *object detection task*), pixel-level recognition (*2-dimensional [2D] segmentation task*), or voxel-level recognition (*3-dimensional [3D] segmentation task*). An image *classification task* associates 1 or more labels with a given image. The annotation is a scalar or vector that represents the label, example, normal or anomalous as a *binary-class classification problem*, or multiple types of TBI abnormalities as a *multi-class classification problem*. An *image segmentation task*, which can be seen as a "pixel-wise" classification, refers to the process of assigning each pixel to 1 of the labels. The annotations in segmentation tasks, therefore, are multi-dimensionally labeled images. The *object detection task* combines classification and localization to determine which objects are in the image and to determine where they are by using bounding boxes.

Most of the recent studies on CT image recognition have relied heavily on deep neural network (DNN) frameworks, especially those using a convolutional neural network (CNN) architecture.[5] Other important types of DNN-based architecture that are utilized in identifying TBI are recurrent neural network (RNN) and long short-term memory (LSTM). Both RNN and LSTM have recurrent internal connections that ensure that sequential information, such as text or audio data, is accepted as input data. Several studies have used RNN or LSTM because a CT scan, which consists of multiple CT slices, can be used as sequential data.

2. Methods

2.1. Literature selection

We conducted a systematic review of ML studies that identified and quantified TBI based on CT images from the 4 major medical and scientific databases (Ovid Medline, Ovid Embase, IEEE Xplore and Web of Science) that were published before April 28, 2022. We used a combination of keywords related to CT, ML, and TBI with the Medical Subject Heading (MeSH) queries. See Methods, Supplemental Content 1, http://links.lww.com/MD/H961, which illustrated systematic search strings, for further details. We did not limit our retrieval to papers written in specific languages or to those written within a specific period.

The articles were assessed by 2 researchers independently (AH and MJ), both of whom had had extensive experience in ML and medical imaging literature. The papers that were judged by both researchers to meet the inclusion criteria and to be eligible for inclusion were sent for a second full-text appraisal to identify the papers that met the inclusion criteria. Any disagreements between the 2 researchers were resolved by a third author (PNT).

We focused on ML studies that used CT imaging to analyze human patients, and we included studies that included at least 1 of the following findings for open (penetrating) or closed TBI: Glasgow Coma Scale ≤ 15; concussion; skull fracture and intracranial hemorrhage (ICH). We chose papers that dealt with clinical variables and non-contrast CT images captured by either single-slice or multi-slice CT scanners as input data. We included ML studies with any of the following output data: image-level or pixel-level findings of abnormalities on CT images; severity; risk of death, and future outcomes. We excluded studies that used data from non-human participants or patients with non-TBI caused by, for example, a stroke or brain tumor. We also excluded papers that did not include CT data or only focused on other modalities, such as MRI or electroencephalogram. Existing work that focused only on statistical analyses, treatment strategies, and pathological research were also excluded. The detailed inclusion and exclusion criteria are listed in Table 1.

2.2. Protocols

This review was conducted in accordance with the Preferred Reporting Items for Systematic reviews and Meta-Analyses (PRISMA) 2020[6] and Standard Protocol Items: Recommendations for Interventional Trials (SPIRIT)-AI extension.[7]

Table 1

Inclusion and exclusion criteria.

	Inclusion	Exclusion
Study participants	•Patients with at least one of the following open (penetrating) and closed traumatic brain injury (TBI)-related findings: o Glasgow Coma Scale ≦ 15 o Skull fracture o Brain contusion o Concussion o Diffuse axonal injury o Epidural/Extradural hematoma o Intracerebral/Intraparenchymal hematoma o Subdural hematoma o Subarachnoid hemorrhage o Midline shift o Increased intracranial pressure	•Patients with non-TBI findings caused by o Stroke o Brain tumor o Toxic injury o Anoxic injury o Drug abuse o Aneurysm o Cerebral edema o Encephalitis o Heart attack o Hydrocephalus o Hypoxia/anoxia o Meningitis •Animal subjects
Input data	•Non contrast CT images captured by either single-slice or multi-slice CT scanners •Clinical records attached to CT images •Clinical variables	•MRI •Electroencephalogram •Positron emission tomography •Single photon emission CT •Ultrasound imaging
Output data	•Image-level findings of abnormalities on CT images •Pixel-level findings of abnormalities on CT images •Severity •Future outcome •Risk of death	•Rules to decide if neuroimaging is required
Methodology	•Machine learning (ML) approach for image recognition	•Statistical analysis •Treatment strategies •Pathological research
Publication type	•Peer reviewed journals •Conference proceedings	•Articles without full text

CT = computed tomography, ML = machine learning, MRI = magnetic resonance imaging, TBI = traumatic brain injury.

3. Results

A total of 759 papers were retrieved from Medline and Embase (n = 488), IEEE Explore (n = 24), and Web of Science (n = 247). After duplicates were removed and the initial screen conducted, 66 papers remained for further review (Fig. 1). The distribution of the included 66 articles according to year of publication is shown in Figure 2. It demonstrates that the number of ML studies on identifying and quantifying TBI has risen rapidly in recent years. The rest of this section discusses each article focusing on the algorithm design and relevance to clinical practice according to the learning types and the ML tasks listed in Table 2. An input and output of these studies are shown in a data extraction sheet (Table S1, Supplemental Digital Content, http://links.lww.com/MD/H962, which demonstrates details of all identified articles multi-class classification task) (Table S2, Supplemental Digital Content, http://links.lww.com/MD/H963, which demonstrates binary-class classification task) (Table S3, Supplemental Digital Content, http://links.lww.com/MD/H964, which demonstrates multi-class object detection task) (Table S4, Supplemental Digital Content, http://links.lww.com/MD/H965, which demonstrates multi-class 2D-segmentation task) (Table S5, Supplemental Digital Content, http://links.lww.com/MD/H966, which demonstrates binary-class 2D-segmentation task) (Table S6, Supplemental Digital Content, http://links.lww.com/MD/H967, which demonstrates multi-class 3D-segmentation task) (Table S7, Supplemental Digital Content, http://links.lww.com/MD/H968, which demonstrates binary-class 3D-segmentation task).

3.1. Presence or absence of TBI-related abnormalities

The first task for clinicians when they look at CT images of a potential TBI patient is to identify any TBI abnormalities. From the ML viewpoint, this process can be seen as either a 2-class (binary-class) or multi-class classification problem, depending on the clinical purpose and situation. For classifying CT slices as either ICH or non-ICH, some studies[34–37] assumed a binary-class classification problem. In the algorithm proposed by Patel et al,[34] the CNN was used as a feature extractor for each CT slice, and extracted features that represented multiple slices were stacked and fed into the LSTM model. By introducing the idea of LSTM, the algorithm was expected to acquire spatial relations, which can be more informative for model training than using individual slices.

Several previous studies[32,33,38–40] also employed binary-class classification problems for predicting other targets related to TBI lesions. The algorithm proposed in by Liu et al[40] aimed to distinguish normal CT slices from abnormal ones, including 5 types of hemorrhage (epidural/extradural hemorrhage [EDH], subdural hemorrhage [SDH], subarachnoid hemorrhage [SAH], intraparenchymal hemorrhage [IPH], and intraventricular hemorrhage [IVH]). The researchers classified vectors made up of 12 hand-crafted features using a support vector machine (SVM) to detect abnormal CT slices.

Another problem setting within binary-class classification is the estimation of intracranial pressure (ICP), which is considered an important indicator of TBI severity, as the ICP level is frequently elevated in patients after brain injury due to the mass effect of ICH.[1] Some articles[41–43] proposed methods for predicting the ICP level in the form of a binary classification, elevated ICP (ICP > 12 mm Hg), or normal ICP (ICP ≤ 12mm Hg), although the choice of threshold was not well justified in those papers. They used SVM models trained on clinical variables and features extracted from texture patterns embedded in CT images.

Although all of these papers only described models that accepted CT images, we also identified some algorithms[38,39]

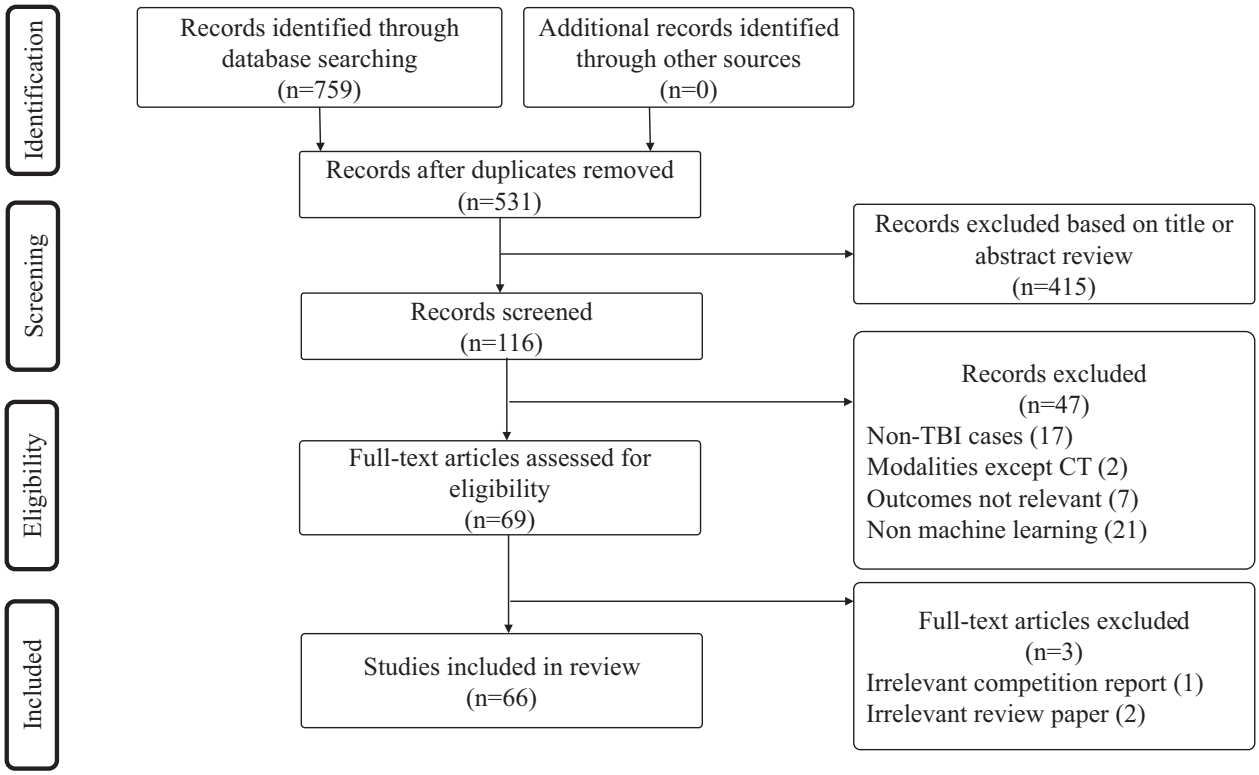

Figure 1. Flowchart of literature search.

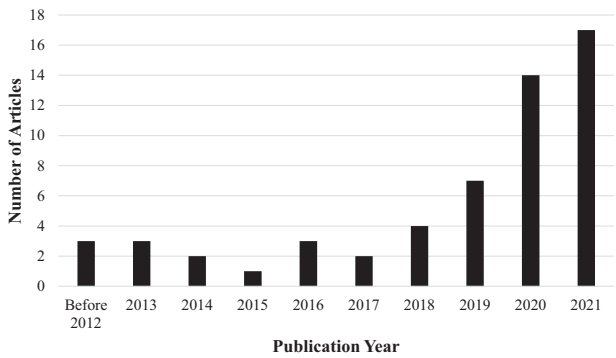

Figure 2. Papers included in this survey by year of publication.

that allowed clinical records attached to CT images to be used as input and predicted whether the corresponding CT images were normal (non-TBI) or abnormal (TBI) using natural language processing techniques.

3.2. Classification of ICH types

From the ML perspective, the multi-class classification task is the best strategy for automating categorization of ICH types. The 14 articles we identified that categorized different types of ICH were divided into 2 groups: 1 using publicly available datasets (n = 10), and the other using private clinical datasets (n = 4).

The publicly available CT imaging datasets containing TBI abnormalities that are currently available are CQ500,[68] the RSNA dataset,[69] and the Physio Net ICH dataset[70] (Table 3). The RSNA dataset was used in a competition,[71] and its first prize winner described their proposed model in.[8] They developed a primary CNN model followed by a 2-sequences CNN-based architecture to classify ICH, EDH, IPH, IVH, SAH, and SDH. Recent work[9,16] also aimed to build a CNN-based model using the RSNA dataset. The models that were developed by Sage and Badura in[16] consisted of a feature-extractable CNN-based architecture (ResNet-50) followed by a classifier (SVM or random forest). Several studies using the CQ500 proposed different types of ML models to classify TBI abnormalities, such as hemorrhages observed in various parts of a brain, fractures, or midline shift.[17–20]

Compared with the RSNA and CQ500 datasets, which contain hundreds of 1000s of CT scans, private or internal datasets were used in other studies on brain hematoma classification,[10–12,21–23] and most of these datasets were relatively small (150–2000 scans). The models used in these studies were trained with sophisticated ML pipelines, but there may have been limitations in the reproducibility and extendibility of the models, considering the wide variety of TBI abnormalities.

3.3. Localization of ICH

The identification of where a brain hematoma is located is important for determining a TBI treatment strategy in clinical practice. We identified 2 ML papers[24,25] which focuses on the object detection task, 1 of the most important ML tasks. The first study involved an object detection task for localizing brain hemorrhages,[25] where the model allowed for not only the prediction but also the localization of several types of hematoma (IPH, EDH, SDH, and SAH) by providing bounding boxes. Recent improvements in object detection were employed in by Ertuğrul and Akıl,[24] which who achieved accurate TBI localization by training YOLO-v4[72] architecture using the CQ500 dataset extension,[73] which contains additional bounding box annotations on the CQ500 dataset.

Table 2

Summary of reviewed articles.

Learning type	# Class	AI Task			
		Classification	Object detection	2D segmentation	3D segmentation
Supervised	Multi-class Binary-class	•Multiple TBI types[8–23] •Any hematoma[32,33] •ICH[34–37] •Normal/abnormal[38–40] •ICP level (high/low)[41–43] •Hematoma expansion[44]	•Multiple TBI types[24,25]	•Multiple hematoma types[10,26,27] •Any hematoma [45–48] •ICH[49–55] •SDH[56,57] •Normal/abnormal[58]	•Multiple hematoma types[28–31] •ICH[59] •SDH[60] •Normal/Abnormal[61]
	Others			•Midline delineation[62–64]	•Cisterns[59] •Midline[59]
Unsupervised	Binary-class			•Normal/abnormal[65–67]	

2D = 2-dimensional, 3D = 3-dimensional, ICH = intracranial hemorrhage, ICP = intracranial pressure, SDH = subdural hemorrhage, TBI = traumatic brain injury.

3.4. Pixel-level identification of hematoma

The pixel-level identification of various kinds of intracranial abnormalities is an important application of CT imaging in an acute setting. Although several attempts have been made to design an original CNN-based architecture,[49–53] the most common architecture used in this problem setting was U-net.[26,27,45,46,54,55,58] These studies employed the original U-net[74] or various kinds of U-net modifications. For instance, the authors in[54] collected 82 CT scans of subjects with ICH and achieved a Dice coefficient of 0.31 with the U-Net model. Their dataset was made publicly available; this is the only public dataset that contains pixel-level annotations of ICH, and it is known as the Physio Net ICH dataset.[70]

3.5. Measurement of midline shift

One important quantitative CT finding among TBI patients is the amount of midline shift or herniation, because the extent of midline displacement is a factor in predicting mortality.[1,75] Manually measuring the change in midline shift may also introduce inter- and intra-observer reliability concerns. We identified 2 ML studies to automatically measure the extent of midline shift to reduce the human factor. The authors in Nag et al[62] constructed a U-net model to predict the deformed boundaries between the left and right hemispheres followed by an estimation of midline shift. The authors validated their algorithm with private CT datasets and confirmed that the midline shift could be estimated with an average distance error of 1.29 ± 0.60mm. Another study that aimed to measure the midline shift in TBI patients was demonstrated by Wei et al,[63] where the proposed CNN-based model estimated the extent of midline shift with average distance errors of 1.1 ± 70.72 mm and 4.15 ± 3.97 mm on CQ500 and the internal dataset, respectively. Another recent study[64] utilized CNN-based architecture to predict several imaging landmarks to predict the extent of midline. We noted that no ML approaches existed that identified cerebellar tonsillar herniation.

Table 3

Publicly available dataset of CT images with TBI abnormalities.

	RSNA[69]	CQ500[68]	PhysioNet ICH dataset[70]
# CT slices (images)	674,257 (train) 78,545 (test)	171,390	2814
# CT scans or patients	19,530 scans (train) 2214 scans (test)	491 scans	82 patients
Annotated TBI lesions (H: hemorrhage or hematoma)	•Epidural H •Intraparenchymal H •Intraventricular H •Subarachnoid H •Subdural H •Any hemorrhage	•Epidural H •Intraparenchymal H •Intraventricular H •Subarachnoid H •Subdural H •Any fractures •Calvarial fractures •Midline Shift •Mass effect	•Epidural H •Intraparenchymal H •Intraventricular H •Subarachnoid H •Subdural H
Annotation level	Per slice	Per scan (some slices have bounding boxes)	Per pixel
# Annotators	60	3	2
Data source and period	Stanford University (1999–2014), Universidade Federal de São Paulo (2018), and Thomas Jefferson University Hospital (N/A)	Centre for Advanced Research in Imaging, Neurosciences and Genomics (2017)	Al Hilla Teaching Hospital, Iraq (2018)
CT scanner	N/A	GE BrightSpeed, GE Discovery CT750 HD, GE LightSpeed, GE Optima CT660, Philips MX 16-slice, Philips Access-32 CT	Siemens/ SOMATOM Definition AS
Slice thickness (mm)	3–5	5	5

ICH = intracranial hemorrhage, TBI = traumatic brain injury.

These algorithms would be a useful tool to support radiological reading and provide a standardized approach across specialties and institutions in assessing midline shift.

3.6. Measurement of hematoma volume

Another important quantitative CT description is the hematoma volume because it is a powerful prognostic predictor in moderate and severe TBI.[75] Estimating hematoma volume manually, however, is less reliable in an irregular shaped hematoma. Furthermore, lack of reproducibility still exists in the manual measurement process.

Some articles aim to estimate hematoma volume by stacking the prediction of 2D-segmented hematoma.[46,47,56] Although these 2D-segmentation methods are technically possible, 3D-hematoma recognition has recently attracted considerable attention as it enables end-to-end hematoma volume predictions. Furthermore, whereas a 2D-segmentation model is trained on each CT slice independently, a 3D-segmentation model accepts several CT slices as 3D volumetric data and is expected to learn the useful information in neighboring slices. Although using conventional texture-based imaging features is a feasible approach,[61] most recent studies relied on the CNN, which can easily deal with multi-dimensional images.

The authors of Monteiro et al[28] relied on Deep Medic,[76] a 3D-CNN architecture, for the 3D-segmentation of IPH, EDH, edema, and IVH. Phaphuangwittayakul et al The authors in[29] developed a 3D-segmentation model and trained it with the Physio Net dataset to detect EDH, SDH, and IPH. Some modifications of U-net were also implemented in other studies that focused on voxel-wise hematoma segmentations.[59,60] Jain et al,[59] the authors reported that the median volume difference between the 3D U-net prediction and the expert reference segmentations of ICH volume was 0.07 ml. They also attempted the delineation and volume estimation of the cisterns using the same 3D-architecture and found that the median volume difference was 0.01mL with a correlation coefficient of 0.94 between the proposed scheme and expert ground truth.

3.7. Unsupervised approach

We found several unsupervised approaches that did not require any training samples. A recent work led by Kärkkäinen et al[65] on segmenting ICH regions from CT images employed a clustering technique, which is a strategy commonly used in unsupervised methodologies. The proposed algorithm, based on the expectation-maximization process, adaptively determined the number of representative clusters, which are groups of pixels that have similar intensity values and are likely to be brain abnormalities. We noted that the clustering process does not require any ground truth, which implies that there is no need to prepare many annotations in advance.

Furthermore, there were 2 research articles proposing methods that used conventional image processing techniques in an unsupervised manner.[66,67] These studies focused on rule-based image processing that is applicable to ICH segmentation by taking full advantage of domain knowledge in clinical practice.

4. Discussion

In this section, we first discuss the difference between similar reviews focusing on TBI and our systematic review. Next, we discuss the limitations of existing ML studies and the possibility of fully automated TBI identification and quantification. We then outline the limitations of our systematic review and highlight future research directions.

4.1. Related work

A limitation of prior TBI -related reviews was the incomplete coverage of all components of CT findings and ML tasks (classification, object detection, and 2D- and 3D-segmentation. A literature review conducted by Vidhya et al[77] was limited to computer-aided systems that detected ICH and midline shift. Further, their review of midline shift considered only classification, and not segmentation tasks. A short review by Brossard et al[78] focused on recent ML developments in the automated determination of TBI lesions, but this was not a systematic review and many important papers were not included. Most notably, existing studies on 3D segmentation tasks that are applicable to TBI were not included, even though 3D recognition is a key factor in calculating the hematoma volume.

4.2. Limitations of ML studies and possibility of fully automated TBI identification and quantification

Most of the important components for identification and quantification of TBI-related abnormalities are covered by existing ML studies. Because ML algorithms are good at solving specific tasks, just 1 algorithm is not enough to cover a wide variety of TBI abnormalities and complications. However, as we demonstrated in this systematic review, each algorithm has a strong ability to automatically identify and quantify important CT findings caused by TBI. This suggests that a combination of existing ML algorithms can be a good supporting tool to alleviate the increased workload of radiologists and clinicians. Furthermore, the outcome predicted by ML is always the same as long as the input CT image is the same, indicating that the automated identification and quantification of TBI also contribute to the improvement of reproducibility, which is 1 of the inevitable problems in manual assessment.

As a limitation of current ML studies, 1 of the CT findings regarding TBI that were not completely covered was diffuse axonal injury. This is mainly due to the limited sensitivity of CT imaging compared to MRI. Another important CT findings that current ML studies did not investigate was identification of conditions in basal cisterns (normal, compressed or absent), which is a key component to estimate TBI prognosis.[79–81] This is possibly because the conditions are subjective interpretations and difficult to define to be used as a ML training dataset. However, the algorithm proposed in by Jain et al[59] enabled the delineation and volume estimation of the cisterns by using the 3D-segmentation, and this approach can be an initial approach to fill this gap. Therefore, a combination of ML algorithms identified in this work can be useful tools to enhance reproducibility and support radiologists and clinicians in their TBI identification and quantification process.

4.3. Recommendations

To maximize benefits that ML techniques provide to patients with TBI, it is important to develop ML algorithms that are generalizable to a wide range of TBI abnormalities and easily applicable to medical settings without imposing heavy burdens on clinicians. In this sub-section, we discuss several factors which prevent the development of generalizable and easily applicable ML models. Future research directions are proposed for overcoming these limitations from a ML perspective.

4.3.1. Large-scale segmentation dataset. Most of the studies we identified used private data that could only be accessed internally, especially for segmentation tasks. This makes it difficult to benchmark the segmentation algorithms using the same training and test dataset. Currently, the only public dataset for segmentation tasks is the Physio Net ICH dataset,

which contains CT scans of 82 TBI patients with pixel-level annotations. However, there is a risk of overfitting if we use this small-sized data for training recent CNN-based segmentation models, which contain millions of learning parameters.[5] Joint effort to coordinate creation of large-scale publicly available datasets for TBI segmentation tasks is in high demand to construct more generalizable segmentation models.

4.3.2. Improved learning strategy. The most used learning strategy for scoring TBI was supervised learning (n = 60 out of 66 articles), which generally requires a large number of training samples to achieve the expected performance. Unlike non-medical images that can be annotated by anyone, only clinicians with a strong background in brain trauma can annotate TBI lesions. Furthermore, because a CT scan comprises a 3D volumetric dataset, a heavy workload is inevitable in preparing enough annotations for the supervised ML models. To develop ML models that are smoothly applicable to medical settings, it is important to consider ways to reduce the annotation cost and workload required for constructing datasets. The study[48] is the only attempt that dealt with annotations in supervised learning for hematoma classification. The authors utilized an active learning framework, in which the algorithm "actively" chooses training samples that are likely to improve training performance and interactively asks human experts to annotate them, which allows for a reduction in the workload of human annotators without compromising accuracy.

There are learning techniques that alleviate annotation workload, such as semi-supervised learning[5] and self-supervised learning.[82] These approaches could be helpful in building ML models efficiently and accelerating the speed of technology translation from bench to bedside.

4.4. Limitations

We devised our search string to consider ML papers related only on CT images thereby excluding other imaging modalities, most notably MRI. As the included studies had different objectives, used different datasets, and reported different performance metrics, we were not able to statistically compare their performance as a meta-analysis.

5. Conclusion

A systematic review of published ML-based studies describing the identification and quantification of CT findings caused by TBI demonstrated that many TBI-related abnormalities could be automatically identified and quantified by AI studies at high resolutions. Combination of these studies can lead to useful tools to enhance reproducibility by reducing subjective human factors and to support radiologists and clinicians by providing guidance in their TBI identification and quantification assessment.

Author contributions

Conceptualization: Atsuhiro Hibi, Pascal N. Tyrrell.
Data curation: Atsuhiro Hibi, Majid Jaberipour.
Formal analysis: Atsuhiro Hibi, Majid Jaberipour.
Funding acquisition: Atsuhiro Hibi.
Investigation: Atsuhiro Hibi, Pascal N. Tyrrell.
Methodology: Atsuhiro Hibi, Pascal N. Tyrrell.
Project administration: Michael D. Cusimano, Pascal N. Tyrrell.
Supervision: Michael D. Cusimano, Alexander Bilbily, Rahul G. Krishnan, Pascal N. Tyrrell.
Validation: Michael D. Cusimano, Alexander Bilbily, Rahul G. Krishnan, Richard I. Aviv.
Writing – original draft: Atsuhiro Hibi.
Writing – review & editing: Michael D. Cusimano, Alexander Bilbily, Rahul G. Krishnan, Richard I. Aviv, Pascal N. Tyrrell.

References

[1] Silver JM, McAllister TW, Arciniegas DB. Textbook of Traumatic Brain Injury, Third Edition. Washington, DC, USA: American Psychiatric Association Publishing; 2018. Available at: https://books.google.ca/books?id=U60muQEACAAJ.
[2] Dewan MC, Rattani A, Gupta S, et al. Estimating the global incidence of traumatic brain injury. J Neurosurg. 2019;130:1080–97.
[3] Fu TS, Jing R, McFaull SR, et al. Health & economic burden of traumatic brain injury in the Emergency Department. Can J Neurol Sci. 2015;43:238–47.
[4] Lee B, Newberg A. Neuroimaging in traumatic brain imaging. NeuroRx. 2005;2:372–83.
[5] Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. Med Image Anal. 2017;42:60–88.
[6] Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021;372:n71.
[7] Cruz Rivera S, Liu X, Chan AW, et al. Guidelines for clinical trial protocols for interventions involving artificial intelligence: the SPIRIT-AI extension. Nat Med. 2020;26:1351–63.
[8] Wang X, Shen T, Yang S, et al. A deep learning algorithm for automatic detection and classification of acute intracranial hemorrhages in head CT scans. NeuroImage Clin. 2021;32:102785.
[9] Salehinejad H, Kitamura J, Ditkofsky N, et al. A real-world demonstration of machine learning generalizability in the detection of intracranial hemorrhage on head computerized tomography. Sci Rep. 2021;11:1–11.
[10] Gong T, Lim N, Cheng L, et al. Finding distinctive shape features for automatic hematoma classification in Head CT images from traumatic brain injuries. In: 2013 IEEE 25th International Conference on Tools with Artificial Intelligence. Herndon, VA, USA. 2013:242–249.
[11] Alis D, Alis C, Yergin M, et al. A joint convolutional-recurrent neural network with an attention mechanism for detecting intracranial hemorrhage on noncontrast head CT. Sci Rep. 2022;12:2084.
[12] Zhou Q, Zhu W, Li F, et al. Transfer learning of the ResNet-18 and DenseNet-121 model used to diagnose intracranial hemorrhage in CT Scanning. Curr Pharm Des. 2022;28:287–95.
[13] Mansour RF, Aljehane NO. An optimal segmentation with deep learning based inception network model for intracranial hemorrhage diagnosis. Neural Comput Appl. 2021;33:13831–43.
[14] Mansour RF, Escorcia-Gutierrez J, Gamarra M, et al. Artificial intelligence with big data analytics-based brain intracranial hemorrhage e-diagnosis using CT images. Neural Comput Appl. 2021.
[15] Ko H, Chung H, Lee H, et al. IEEE. Feasible study on intracranial hemorrhage detection and classification using a CNN-LSTM Network. In: 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC). Montreal, QC, Canada. 2020:1290–3.
[16] Sage A, Badura P. Intracranial hemorrhage detection in head CT using double-branch convolutional neural network, support vector machine, and random forest. Appl Sci. 2020;10:7577.
[17] Singh SP, Wang L, Gupta S, et al. Shallow 3D CNN for detecting acute brain hemorrhage from medical imaging sensors. IEEE Sens J. 2021;21:14290–9.
[18] Li JQ, Fu GH, Chen YD, et al. A multi-label classification model for full slice brain computerised tomography image. BMC Bioinformatics. 2020;21(15th International Symposium on Bioinformatics Research and Applications (ISBRA)-Bioinformatics).
[19] DelRocini M, Angelini C, Rasool G. Identification of abnormalities in head computerized tomography scans. In: 2020 IEEE Signal Processing in Medicine and Biology Symposium (SPMB). 2020:1-4.
[20] Shanna B, Venugopalan K. Classification of hematomas in brain CT images using neural network. In: Sharma A, Ahlawat A, Pandey A, Sharma V, (eds). Proc 2014 Int Conf ISSUES CHALLENGES Intell Comput Tech. Ghaziabad, India: (International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)). 2014:41-46.
[21] Kim KH, Koo H-W, Lee B-J, et al. Cerebral hemorrhage detection and localization with medical imaging for cerebrovascular disease diagnosis and treatment using explainable deep learning. J Korean Phys Soc. 2021;79:321–7.
[22] McLouth J, Elstrott S, Soun JE, et al. Validation of a deep learning tool in the detection of intracranial hemorrhage and large vessel occlusion. Front Neurol. 2021;12:656112.
[23] Ye H, Gao F, Yin YB, et al. Precise diagnosis of intracranial hemorrhage and subtypes using a three-dimensional joint convolutional and recurrent neural network. Eur Radiol. 2019;29:6191–201.

[24] Ertuğrul OF, Akıl MF. Detecting hemorrhage types and bounding box of hemorrhage by deep learning. Biomed Signal Process Control. 2022;71:103085.
[25] Chang PD, Kuoy E, Grinband J, et al. Hybrid 3D/2D convolutional neural network for hemorrhage evaluation on head CT. AJNR Am J Neuroradiol. 2018;39:1609–16.
[26] Jadon S, Leary OP, Pan I, et al. A comparative study of 2D image segmentation algorithms for traumatic brain lesions using CT data from the ProTECTIII multicenter clinical trial. In: Deserno TM, Chen P-H, (eds). Medical imaging 2020: imaging informatics for healthcare, research, and applications. Bellingham, Washington, USA: SPIE; 2020:48.
[27] Xu J, Zhang RG, Zhou ZJ, et al. Deep network for the automatic segmentation and quantification of intracranial hemorrhage on CT. Front Neurosci. 2020;14:541817.
[28] Monteiro M, Newcombe VFJJ, Mathieu F, et al. Multiclass semantic segmentation and quantification of traumatic brain injury lesions on head CT using deep learning: an algorithm development and multicentre validation study. Lancet Digit Heal. 2020;2:e314–22.
[29] Phaphuangwittayakul A, Guo Y, Ying FL, et al. An optimal deep learning framework for multi-type hemorrhagic lesions detection and quantification in head CT images for traumatic brain injury. Appl Intell. 2021;52:7320–38.
[30] Z RT, C Y, H R. Predicting hematoma expansion in intracerebral hemorrhage from brain CT scans via K-nearest neighbors matting and deep residual network. Biomed Signal Process Control. 2022;76:103656.
[31] Farzaneh N, Reza Soroushmehr SM, Williamson CA, et al. Automated subdural hematoma segmentation for traumatic brain injured (TBI) patients. In: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Vol 2017. Jeju Island, South Korea: IEEE; 2017:3069-3072.
[32] Mushtaq MF, Shahroz M, Aseere AM, et al. BHCNet: neural network-based brain hemorrhage classification using head CT scan. IEEE Access. 2021;9:113901–16.
[33] Garland J, Ondruschka B, Stables S, et al. Identifying fatal head injuries on postmortem computed tomography using convolutional neural network/deep learning: a feasibility study. J Forensic Sci. 2020;65:2019–22.
[34] Patel A, Leemput SC van de, Prokop M, et al. Image level training and prediction: intracranial hemorrhage identification in 3D non-contrast CT. IEEE Access. 2019;7:92355–64.
[35] Vidya MS, Mallya Y, Shastry A, et al. Recurrent sub-volume analysis of Head CT scans for the detection of intracranial hemorrhage. In: Shen D, Liu T, Peters TM, et al., (eds). Medical image computing and computer assisted intervention - MICCAI 2019, PT III. New York City, NY, USA: Springer, Cham. 2019:864-872.
[36] Arbabshirani MR, Fornwalt BK, Mongelluzzo GJ, et al. Advanced machine learning in action: identification of intracranial hemorrhage on computed tomography scans of the head with clinical workflow integration. NPJ Digit Med. 2018;1:9.
[37] Yang Y, Chen S, Tan D, et al. Fusion branch network with class learning difficulty loss function for recongnizition of haematoma expansion signs in intracerebral haemorrhage. In: 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Houston, TX, USA. 2021:3448-3455.
[38] Yadav K, Sarioglu E, Choi H-AA, et al. Automated outcome classification of computed tomography imaging reports for pediatric traumatic brain injury. Acad Emerg Med. 2016;23:171–8.
[39] Szlosek DA, Ferrett J. Using machine learning and natural language processing algorithms to automate the evaluation of clinical decision support in electronic medical record systems. EGEMS (Washington, DC). 2016;4:1222.
[40] Liu R, Tan CL, Leong TY, et al. Hemorrhage slices detection in brain CT images. In: 2008 19th International Conference on Pattern Recognition. Tampa, FL, USA. 2008:1-4.
[41] Chen W, Belle A, Cockrell C, et al. Automated midline shift and intracranial pressure estimation based on brain CT Images. J Vis Exp. 2013:3871.
[42] Chen W, Cockrell CH, Ward K, et al. Predictability of intracranial pressure level in traumatic brain injury: features extraction, statistical analysis and machine learning-based evaluation. Int J Data Min Bioinform. 2013;8:480–94.
[43] Chen W, Cockrell C, Ward KR, et al. Intracranial pressure level prediction in traumatic brain injury by extracting features from multiple sources and using machine learning methods. In: 2010 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Hong Kong, China. 2010:510-515.
[44] Inkeaw P, Angkurawaranon S, Khumrin P, et al. Automatic hemorrhage segmentation on head CT scan for traumatic brain injury using 3D deep learning model. Comput biol med. 2022;146:105530.
[45] Remedios SW, Roy S, Bermudez C, et al. Distributed deep learning across multisite datasets for generalized CT hemorrhage segmentation. Med Phys. 2020;47:89–98.
[46] Yao H, Williamson C, Soroushmehr R, et al. Hematoma segmentation using dilated convolutional neural network. Annu Int Conf IEEE Eng. 2018;2018:5902–5.
[47] Yao HM, Williamson C, Gryak J, et al. Automated hematoma segmentation and outcome prediction for patients with traumatic brain injury. Artif Intell Med. 2020;107:101910.
[48] Yao HM, Williamson C, Gryak J, et al. Brain hematoma segmentation using active learning and an active contour model. In: Rojas I, Valenzuela O, Rojas F, Ortuno F, (eds). Bioinforma Biomed Eng (IWBBIO 2019), PT II. New York City, NY, USA: Springer, Cham. 2019:11466.
[49] He X, Chen K, Hu K, et al. HMOE-Net: hybrid multi-scale object equalization network for intracerebral hemorrhage segmentation in CT Images. In: 2020 IEEE International Conference on Bioinformatics and Biomedicine (BIBM). Seoul, South Korea; IEEE. 2020:1006–1009.
[50] Islam M, Sanghani P, See AAQ, et al. ICHNet: Intracerebral Hemorrhage (ICH) segmentation using deep learning. In: Crimi A, Bakas S, Kuijf H, Keyvan F, Reyes M, VanWalsum T, (eds). Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries, Brainles 2018. New York City, NY, USA: Springer, Cham. 2019:456–463.
[51] Irene K, Masum MA, Yunus RE, et al. Segmentation and approximation of blood volume in intracranial hemorrhage patients based on computed tomography scan images using deep learning method. 2020 Int Work Big Data Inf Secur IWBIS 2020. Depok, Indonesia. 2020;65–72.
[52] Remedios S, Roy S, Blaber J, et al. Distributed deep learning for robust multi-site segmentation of CT imaging after traumatic brain injury. In: Angelini ED, Landman BA, (eds). Medical Imaging 2019: Image Processing. Bellingham, Washington, USA: SPIE; 2019:9
[53] Oka K, Fujita D, Nohara Y, et al. CNN Based hierarchical intracerebral hematoma region extraction method in head thick-slice CT Images. In: 2021 International Conference on Machine Learning and Cybernetics (ICMLC). Adelaide, Australia. 2021:1–5.
[54] Hssayeni MD, Croock MS, Salman AD, et al. Intracranial hemorrhage segmentation using a deep convolutional model. Data. 2020;5:141–18.
[55] Kuang Z, Deng X, Yu L, et al. Ψ-Net: focusing on the border areas of intracerebral hemorrhage on CT images. Comput Methods Programs Biomed. 2020;194:105546.
[56] Farzaneh N, Williamson CA, Jiang C, et al. Automated segmentation and severity analysis of subdural hematoma for patients with traumatic brain injuries. Diagnostics. 2020;10:773.
[57] Sharrock MF, Mould WA, Hildreth M, et al. Bayesian deep learning outperforms clinical trial estimators of intracerebral and intraventricular hemorrhage volume. J Neuroimag. 2022;32:968–76.
[58] Syeda-Mahmood TF, Karargyris A. Saliency U-Net: a regional saliency map-driven hybrid deep learning network for anomaly segmentation. In: Mori K, Petrick N, (eds). Medical Imaging 2018: Computer-Aided Diagnosis. SPIE; 2018:64.
[59] Jain S, Vyvere TV, Terzopoulos V, et al. Automatic quantification of computed tomography features in acute traumatic brain injury. J Neurotrauma. 2019;36:1794–803.
[60] Kellogg RT, Vargas J, Barros G, et al. Segmentation of chronic subdural hematomas using 3D convolutional neural networks. World Neurosurg. 2021;148:e58–65.
[61] Keshavamurthy KN, Leary OP, Merck LH, et al. Machine learning algorithm for automatic detection of CT-identifiable hyperdense lesions associated with traumatic brain injury. In: Armato SG, Petrick NA, (eds). Med IMAGING 2017 Comput DIAGNOSIS. Bellingham, Washington USA: SPIE. 2017;10134.
[62] Nag MK, Gupta A, Hariharasudhan AS, et al. Quantitative analysis of brain herniation from non-contrast CT images using deep learning. J Neurosci Methods. 2021;349:109033.
[63] Wei H, Tang X, Zhang M, et al. The delineation of largely deformed brain midline using regression-based line detection network. Med Phys. 2020;47:5531–42.
[64] Yan JL, Chen YL, Chen MY, et al. A Robust, fully automatic detection method and calculation technique of midline shift in intracranial hemorrhage and its clinical application. Diagnostics. 2022;12:693.
[65] Kärkkäinen K, Fazeli S, Sarrafzadeh M. Unsupervised acute intracranial hemorrhage segmentation with mixture models. In: 2021 IEEE 9th International Conference on Healthcare Informatics (ICHI). Victoria, BC, Canada. 2021:120–9.

[66] Qureshi AN; IEEE. A framework for segmentation and estimation of intracranial measurements in CT scans. In: 2014 Cairo International Biomedical Engineering Conference (CIBEC). Giza, Egypt. 2014:129–32.
[67] Cosic D, Loncaric S. Rule-based labeling of CT head image Art Intellig Med. In: Conference on Artificial Intelligence in Medicine in Europe. Berlin, Heidelberg: Springer. 1997:453–6.
[68] Chilamkurthy S, Ghosh R, Tanamala S, et al. Development and validation of deep learning algorithms for detection of critical findings in Head CT Scans. Available at: http://arxiv.org/abs/1803.05854 [access date January 18, 2018].
[69] Flanders AE, Prevedello LM, Shih G, et al. Construction of a machine learning dataset through collaboration: the RSNA 2019 Brain CT Hemorrhage Challenge. Radiol Artif Intell. 2020;2:e209002.
[70] Hssayeni MD, Croock MS, Salman AD, et al. Intracranial hemorrhage segmentation using a deep convolutional model. Data. 2020;5:14.
[71] Flanders AE, Prevedello LM, Shih G, et al. Construction of a machine learning dataset through collaboration: the RSNA 2019 Brain CT Hemorrhage Challenge. Radiol Artif Intell. 2020;2:e190211.
[72] Bochkovskiy A, Wang C-Y, Liao H-YM. YOLOv4: optimal speed and accuracy of object detection. Available at: http://arxiv.org/abs/2004.10934. [access date April 27, 2022].
[73] Reis EP, Nascimento F, Aranha M, et al. Brain Hemorrhage Extended (BHX): bounding box extrapolation from thick to thin slice CT images (version 1.1). PhysioNet. 2020.
[74] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation BT - Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. In: Navab N, Hornegger J, Wells WM, Frangi AF, (eds). New York, NY, USA: Springer International Publishing; 2015:234–41.
[75] Jacobs B, Beems T, van der Vliet TM, et al. Computed tomography and outcome in moderate and severe traumatic brain injury: hematoma volume and midline shift revisited. J Neurotrauma. 2011;28:203–15.
[76] Kamnitsas SK, Ferrante E, Parisot S, et al. DeepMedic for brain tumor segmentation. Lect Notes Comput Sci. 2016;10154:138–49.
[77] Vidhya V, Gudigar A, Raghavendra U, et al. Automated detection and screening of traumatic brain injury (Tbi) using computed tomography images: a comprehensive review and future perspectives. Int J Environ Res Public Health. 2021;18:6499.
[78] Brossard C, Lemasson B, Attye A, et al. Contribution of CT-Scan analysis by artificial intelligence to the clinical care of TBI Patients. Front Neurol. 2021;12:666875.
[79] Marshall LF, Marshall SB, Klauber MR, et al. A new classification of head injury based on computerized tomography. J Neurosurg. 1991;75:S14–20.
[80] Maas AIR, Hukkelhoven CWPM, Marshall LF, et al. Prediction of outcome in traumatic brain injury with computed tomographic characteristics: a comparison between the computed tomographic classification and combinations of computed tomographic predictors. Neurosurgery. 2005;57:1173–82.
[81] Raj R, Siironen J, Skrifvars MB, et al. Predicting outcome in traumatic brain injury. Neurosurgery. 2014;75:632–46.
[82] Xu J. A review of self-supervised learning methods in the field of medical image analysis. Int J Image Graph Signal Process. 2021;13:33–46.