BENHA MEDICAL JOURNAL bmj

Print ISSN 1110-208X.

Online ISSN 2357-0016

Original article

The Value of Artificial Intelligence (AI) in Detection of Post Traumatic Brain Injury Using Non-Contrast CT Scans

Medhat M. Reffat [a], Waseem M. El Gendy [b], Ahmed E. Shalaan [a], Khaled E. Ahmed [a], Ahmed S.Raslan [a]

[a] Diagnostic and Interventional Radiology Department, Faculty of Medicine Benha University, Egypt.

[b] Diagnostic and Interventional Radiology Department, Military Medical Academy, Egypt.

Corresponding to:
Dr. Ahmed S. Raslan.
Diagnostic and Interventional Radiology Department, Faculty of Medicine Benha University, Egypt.

Email:
ahmed.s.raslan@gmail.com

Received:
Accepted:

Abstract:

Background: Traumatic brain injury is a leading cause of morbidity and mortality worldwide. Rapid and accurate assessment of TBI is crucial for timely intervention. Non-contrast computed tomography is the primary imaging modality for initial assessment. Aim of this study: To assess the role of artificial intelligence in the detection of post traumatic brain injury using non-contrast CT scans using a deep learning model from U-Net and ResNet50 architectures. **Methods:** A cross-sectional study was conducted using 628 patients, divided into training, testing and validation sets. A deep learning model, utilizing U-Net for semantic segmentation and ResNet-50 for feature extraction, was trained and validated for detection of PTBI, and was compared to two radiologists on the validation dataset of 1763 images. **Results:** The AI model demonstrated a sensitivity of 94.3% and a specificity of 84.6% in detecting PTBI when compared to the first radiologist's findings. Agreement between the first radiologist and AI in final diagnosis yielded a Kappa value of 0.796 ($p < 0.001$). The AI model achieved an AUC of 0.976 ($p < 0.001$) for detecting abnormalities compared with the first radiologist. Similarly, the comparison between the second radiologist and AI showed a sensitivity of 94.5%, specificity of 83.5%, and a kappa value of 0.787 ($p < 0.001$), with AUC of 0.968 ($p < 0.001$). **Conclusion:** This study demonstrates the potential of deep learning models to accurately detect PTBI on NCCT scans, with diagnostic performance comparable to expert radiologists. AI can aid radiologists in prioritizing critical cases and reducing diagnostic time.

Keywords: Artificial intelligence; Post traumatic brain injury; ResNet-50; U-Net; Deep learning.

Abbreviations

TBI: Traumatic Brain Injury
CT: Computed Tomography
AI: Artificial Intelligence
DL: Deep Learning
CNN: Convolutional Neural Network
AUC: Area Under the Curve
NCCT: Non-contrast computed tomography.

Introduction

Traumatic brain injury (TBI) significantly contributes to fatalities and long-term disabilities, affecting an estimated 55 million people globally [1, 2]. Patients with severe TBI, defined as a post-resuscitation Glasgow Coma Scale score of 8 or less, have mortality rates approaching 40% [3]. Timely neurosurgical intervention, particularly within 48 hours of injury, significantly improves outcomes [4,5].

Non-contrast computed tomography (NCCT) is the imaging modality of choice for the initial evaluation of TBI. It is fast, widely available, and sensitive for detecting acute intra-axial and extra-axial haemorrhages, mass effect, and skull fractures. However, NCCT has limitations in identifying non-haemorrhagic lesions, such as cortical contusions and diffuse axonal injuries, as well as in early detection of hypoxic-ischemic encephalopathy [6,7].

Artificial intelligence (AI), particularly deep learning, has demonstrated success in medical image analysis, with various applications in radiology, including classification, risk assessment, segmentation, diagnosis, and prognosis [8, 9]. AI-powered tools can assist in identifying abnormalities on imaging, including intracranial haemorrhage, acute infarction, and TBI on non-contrast head CT [10]. Integrating head CT images with clinical data may help predict long-term outcomes for patients with severe TBI [11].

The aim of this study is to assess the role of artificial intelligence in the detection of PTBI using non-contrast CT scans, utilizing U-Net and ResNet50 architectures.

Patients and methods

Data Collection and Study Design:

This cross-sectional study was conducted at the Radiology Department of Benha University Hospital between January 1st and December 31st, 2023. It included 628 patients with suspected post-traumatic brain injury (PTBI), imaged with NCCT. Data were collected from two main sources: the radiology department at Benha University and the CQ500 dataset. The study was approved by the ethical committee of Benha University Hospital, Approval code: MD 15-11-2022. An informed consent was obtained from all participants or their relatives. The patients' age ranged from 18 to 73 years with no gender preference. The dataset consisted of 249,475 CT cuts, with a subset of 1763 images selected for final testing.

Inclusion Criteria:

- Patients in post-traumatic status with suspected post-traumatic brain injury.

Exclusion Criteria:

- Patients with previous cerebral surgical intervention.

NCCT Technique and Data Acquisition:

Patients underwent NCCT using a third-generation dual-energy CT scanner. The acquisition parameters were set as follows: KV in the range of 80-100kv, a gantry speed of 0.35 s rotation, helical thickness of 0.2–0.4 mm, and prospective gating. Axial views were obtained with reconstructions in axial, coronal, and sagittal planes.

CT Image Interpretation:

All imaging data were anonymized and independently interpreted by two consultant radiologists, blinded to all patient information except for the reason for obtaining NCCT (suspected PTBI).

Building the AI Environment:

The AI model was developed using Python in a virtual environment managed by Anaconda. Key libraries included TensorFlow, Pydicom, Cv2, Matplotlib, Pandas, Keras, OS, and Scikit-learn.

Dataset Preparation and Pre-processing:

The dataset consisted of DICOM files, grouped by case. Pre-processing steps included:

- Dividing the DICOM files into normal and abnormal cases.
- Dividing abnormal files into major categories based on PTBI findings.

- Windowing to enhance specific regions, using Hounsfield Unit (HU) values.
- Resizing images to 512 x 512 pixels.
- Data augmentation, including rotation (0-20 degrees clockwise), vertical and horizontal shifts, zoom range (0-0.1), and horizontal/vertical flips.

Building the Deep Learning Model:

The deep learning model was constructed using U-Net for semantic segmentation and ResNet-50 for feature extraction. The U-Net architecture consisted of an encoder (feature extraction) path using ResNet-50 and a decoder (segmentation) path with skip connections between corresponding encoder and decoder layers.

Model Training:

The data were split into training (70%), testing (10%) and validation sets (20%). The model was trained using multiple trials with different numbers of epochs targeting achieving most optimal performance.

Statistical analysis

Analysis of data was performed using Statistical Package for Social Sciences (SPSS®) version 21 (IBM, Armonk, NY; United States of America). The radiologists' findings were compared using the Kappa measure of agreement. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), overall accuracy, and area under the ROC curve (AUC)- were calculated to assess the diagnostic performance of the AI model compared to the two radiologists in the validation dataset. Statistical tests were two-sided, with p-values < 0.05 considered significant.

Results

Patient Demographics:

The study included a total of 628 patients, with a mean age of 40.7 years (SD ± 14.3), with a range from 18 to 64 years. The study population was predominantly male, with 464 (73.9%) participants, while females comprised 164 (26.1%) **(Table 1)**.

Radiologist Assessments (Training Phase):

During the training phase, the first radiologist classified 309 cases (49.2%) as normal and 319 cases (50.8%) as abnormal, whereas, the second radiologist categorized 314 cases (50.0%) as normal and 314 cases (50.0%) as abnormal. The agreement between the two radiologists in their final diagnosis was exceptionally high, yielding a kappa value of 0.984 with a statistically significant p-value of less than 0.001, indicating a strong concordance **(Table 2)**.

The first radiologist also detailed the specific findings: parenchymal hemorrhage or contusion was found to be absent in 464 cases (73.9%) and present in 164 cases (26.1%); extra-axial hemorrhage was absent in 448 cases (71.3%) and present in 180 cases (28.7%); skull fractures were absent in 506 cases (80.6%) and present in 122 cases (19.4%); intraventricular hemorrhage was absent in 626 cases (99.7%) and present in 2 cases (0.3%); and finally, other post-traumatic brain injuries (PTBI) were absent in 612 cases (97.5%) and present in 16 cases (2.5%) **(Figure 1)**. These findings were nearly identical between the two radiologists. The second radiologist found the same distribution of parenchymal hemorrhage/contusion (26.1% present), extra-axial hemorrhage (28.7% present), and other PTBI (2.5% present), but detected skull fractures in slightly fewer cases (18.3%) **(Figure 2)**. The incidence of intraventricular hemorrhage (0.3%) was consistent across the two radiologists. The strong agreement between the two radiologists further supports their validity of diagnosis. Furthermore, the high agreement between radiologists' evaluations for each type of finding was supported by kappa coefficients all being above 0.95, and P values less than 0.001 **(Table 3)**.

Validation Phase Results

In the validation phase, 1763 images were independently assessed by the two radiologists and the AI model. The first radiologist categorized 754 cases (42.8%)

as normal and 1009 cases (57.2%) as abnormal. In contrast, the second radiologist categorized 769 cases (43.6%) as normal and 994 cases (56.4%) as abnormal **(Table 4)**.

AI Model Performance:

The AI model identified 696 cases (39.5%) as normal and 1067 cases (60.5%) as abnormal **(Figure 3)**. When compared to the final diagnoses of the first radiologist, the AI model achieved a kappa value of 0.796 with a p-value of less than 0.001, indicating a good level of agreement. Specifically, of the cases that were categorized as normal by the first radiologist, the AI model correctly identified 638 of them (91.7%) as normal, while classifying 58 of the abnormal cases (8.3%) as normal; on the other hand, 116 (10.9%) of the cases that were identified as normal by the radiologist were marked as abnormal by the AI and 951 (89.1%) cases identified as abnormal by the radiologist were classified as abnormal by the AI. When compared to the final diagnosis of the second radiologist, the kappa value was 0.787 ($p<0.001$), also demonstrating a good level of agreement. In this comparison, the AI correctly identified 641 of the normal cases (92.1%), misclassifying 55 of the abnormal cases (7.9%) as normal, while 128 of the cases marked as normal by the radiologist were considered abnormal by the AI. As well, of cases labeled as abnormal, 939 (88.0%) were also categorized as abnormal by AI. The AI model displayed an area under the ROC curve (AUC) of 0.976 ($p<0.001$) for detecting abnormalities when compared with the first radiologist, with a sensitivity of 94.3% and a specificity of 84.6%. Likewise, in comparison to the second radiologist, the AI yielded an AUC of 0.968 ($p<0.001$), with a sensitivity of 94.5% and a specificity of 83.5% **(Table 5), (Figure 3)**.

Case presentation (1): A 35-year-old male presented with a disturbed level of consciousness following head trauma; both radiologists diagnosed a right frontal hemorrhagic contusion. The AI model also classified the case as abnormal, with an abnormality prediction score of 76.48%, demonstrating complete agreement with the radiologists' findings **(Figure 4).**

Case presentation (2): A 31-year-old male presented with a disturbed level of consciousness following blunt head trauma. Both radiologists' diagnoses were in complete agreement: a left frontal fracture and left extra dural hematomas. The AI model agreed, classifying the case as abnormal with a prediction score of 80.09% **(Figure 5).**

Case presentation (3): A 33-year-old male presented after head trauma for assurance, and both radiologists deemed the findings normal. However, the AI model classified the case as abnormal, with an abnormality prediction score of 65.95%, thus disagreeing with the radiologists' interpretations. The final diagnosis of the case was normal **(Figure 6).**

Table (1): Characteristics of the study patients.

Demographic Characteristics	**(n=628)**	
Age		
Mean± S.D	40.7±14.3	
Range	18.0-64.0	
Sex		
Male	464	73.9%
Female	164	26.1%

This table shows that the mean age for the patients was 40.63±6.67 years old. Males represented 73.9%, while females represented 26.1%.

Table (2): Agreement between first and second radiologist in final diagnosis.

Final diagnosis By Second Radiologist	Final diagnosis By First Radiologist				Kappa	P
	Normal		Abnormal			
	n	%	n	%		
Normal	314	98.4%	0	0.0%		
Abnormal	5	1.6%	309	100.0%	0.984	0.001*
Total	319	100.0%	309	100.0%		

Table (3): Agreement between first and second radiologist in detection of TBI.

Second Radiologist	First Radiologist				Kappa	P
	Absent		Present			
	n	%	n	%		
Parenchymal Hemorrhage/Contusion						
Absent	464	100.0%	0	0.0%	1.000	0.001*
Present	0	0.0%	164	100.0%		
Extra Axial Hemorrhage						
Absent	448	100.0%	0	0.0%	1.000	0.001*
Present	0	0.0%	180	100.0%		
Skull Fracture						
Absent	506	100.0%	7	5.7%	0.964	0.001*
Present	0	0.0%	115	94.3%		
Intraventricular hemorrhage						
Absent	626	100.0%	0	0.0%	1.000	0.001*
Present	0	0.0%	٢	100.0%		
Other PTBi						
Absent	612	100.0%	0	0.0%	1.000	0.001*
Present	0	0.0%	16	100.0%		

Table (4): Final diagnosis by First Radiologist and second radiologist.

Final diagnosis	First radiologist		Second Radiologist	
	n	**%**	**n**	**%**
Normal	754	42.8%	769	43.6%
Abnormal	1009	57.2%	994	56.4%
Total	1763	100.0%	1763	100.0

Table (5): Agreement between first and second radiologist with AI in final diagnosis.

Final diagnosis	Final diagnosis by AI				Kappa	P
	Normal		Abnormal			
	n	%	n	%		
First radiologist						
Normal	638	91.7%	116	10.9%	0.796	0.001*
Abnormal	58	8.3%	951	89.1%		
Second Radiologist						
Normal	641	92.1%	128	12.0%	0.787	0.001*
Abnormal	55	7.9%	939	88.0%		

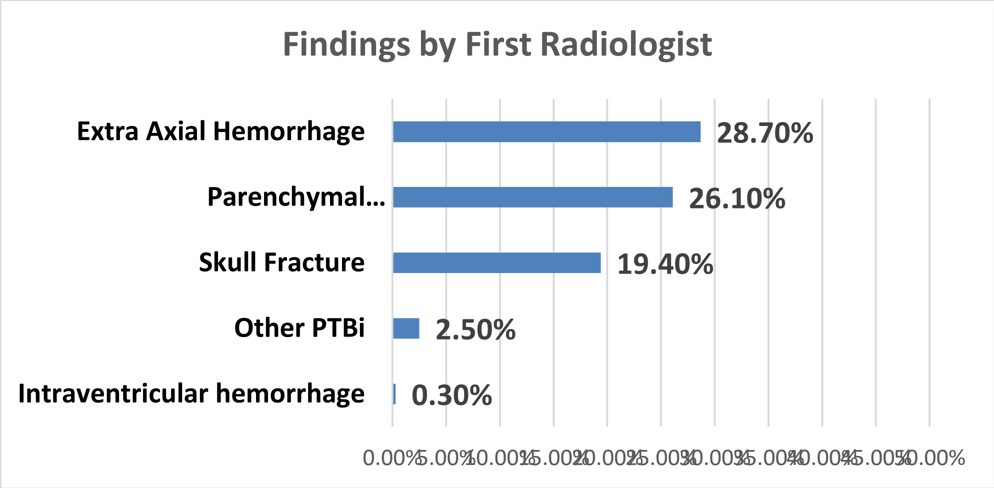

Figure (1): Findings by First Radiologist.

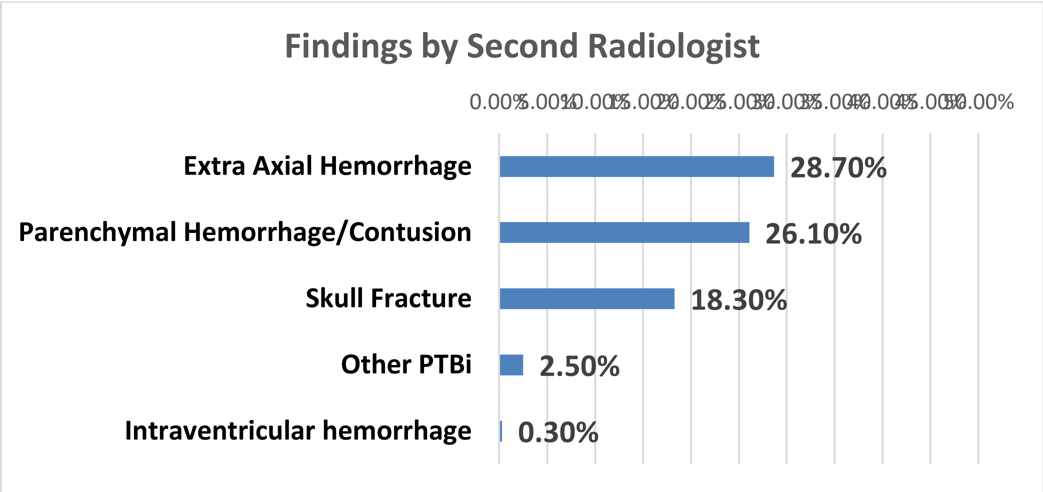

Figure (2): Findings by Second Radiologist.

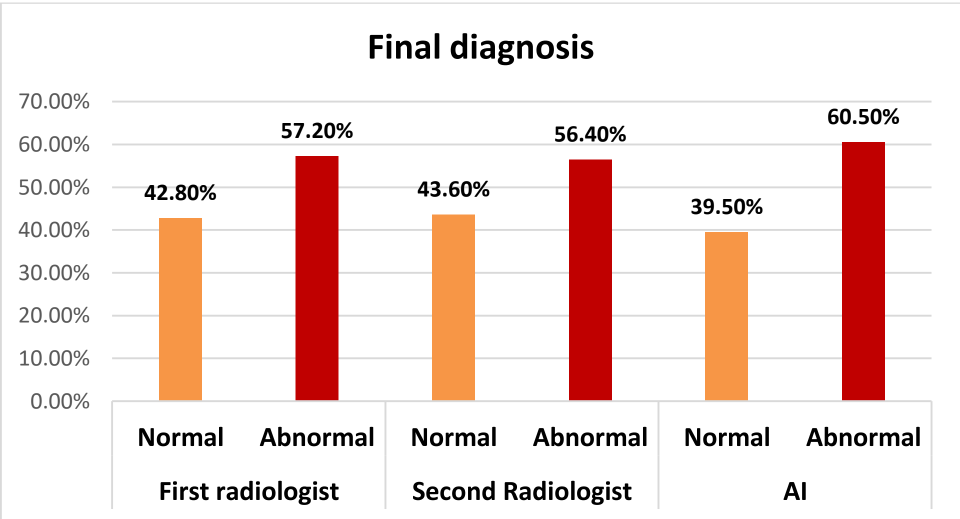

Figure (3): Agreement between first and second radiologist with AI in final diagnosis.

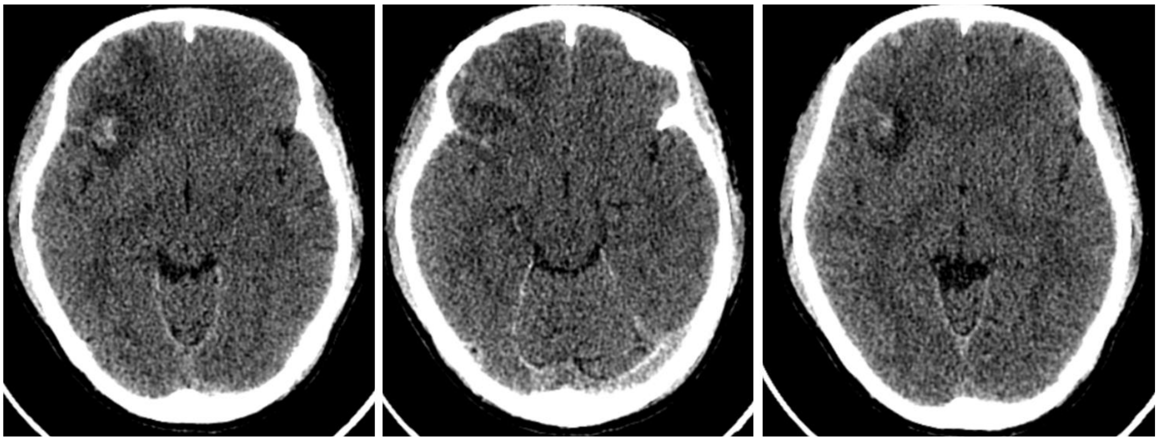

Figure (4): A 35-year-old male presented with a disturbed level of consciousness following head trauma; both radiologists diagnosed a right frontal hemorrhagic contusion. The AI model also classified the case as abnormal, with an abnormality prediction score of 76.48%, demonstrating complete agreement with the radiologists' findings.

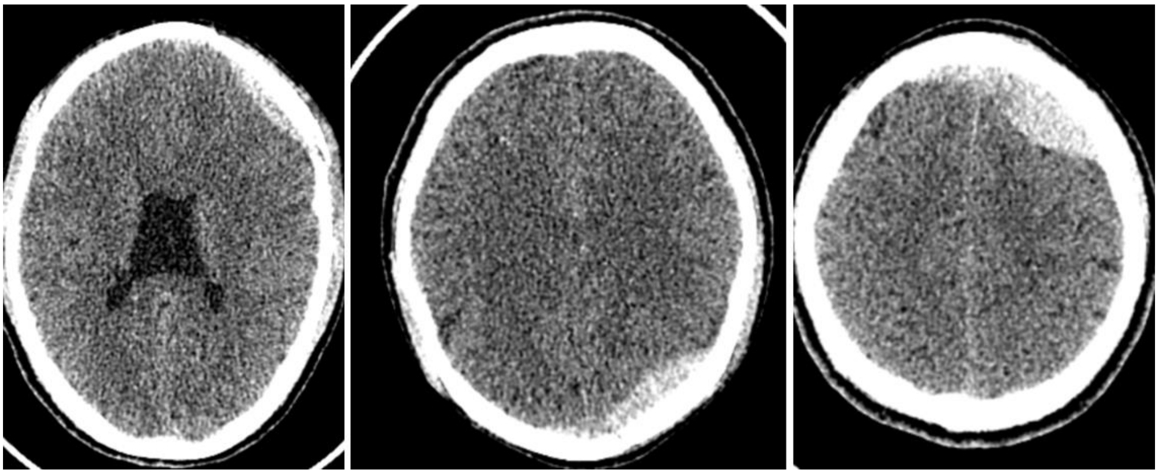

Figure (5): A 31-year-old male presented with a disturbed level of consciousness following blunt head trauma. Both radiologists' diagnoses were in complete agreement: a left frontal fracture and left extra dural hematomas. The AI model agreed, classifying the case as abnormal with a prediction score of 80.09%.

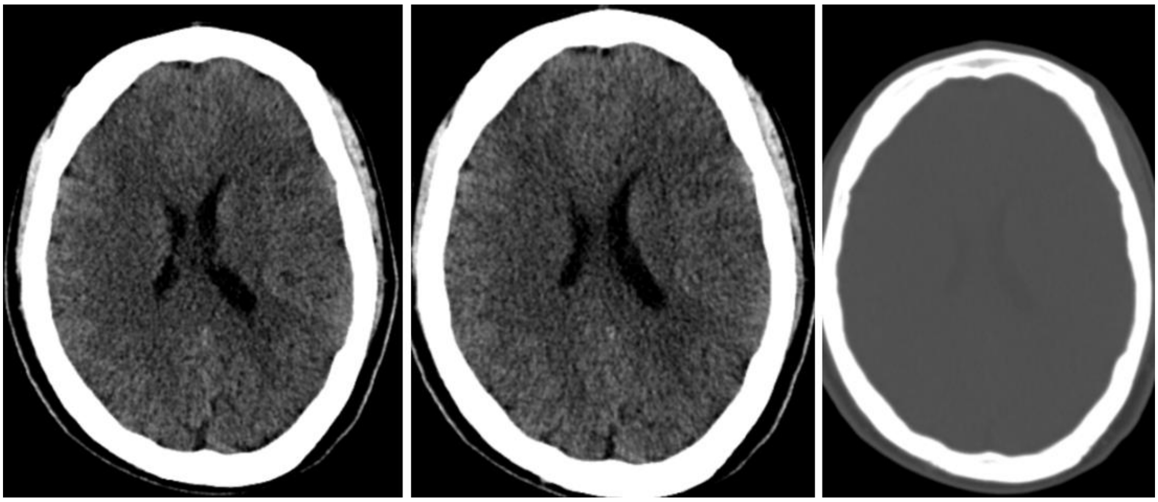

Figure (6): A 33-year-old male presented after head trauma for assurance, and both radiologists deemed the findings normal. However, the AI model classified the case as abnormal, with an abnormality prediction score of 65.95%, thus disagreeing with the radiologists' interpretations. The final diagnosis of the case was normal.

Discussion

Traumatic brain injury (TBI) remains a significant public health concern, marked by the complexity of its causes, pathophysiology, and outcomes [12]. Early detection and intervention are crucial for optimizing rehabilitation outcomes [13, 14]. NCCT is the primary imaging modality for TBI assessment however it is limited in detecting subtle injuries [6].

Deep learning models offer potential advantages in analyzing medical images, providing objective and quantitative evaluations [15]. Unlike traditional machine learning, deep learning can operate on raw data without human feature selection [16].

This study included 628 patients from the Benha University Hospital and CQ500 dataset, providing a diverse dataset for model training and validation. The agreement between the first and second radiologist during the training phase was strong. In the validation phase the AI model achieved high sensitivity (94.3% and 94.5% in comparison with the two radiologists) and specificity (84.6% and 83.5% in comparison with the two radiologists) demonstrating significant potential in detecting PTBI.

These findings align with other studies, although they utilized different deep learning architectures and outcomes, such as studies on long-term outcome prediction and lesion segmentation [11, 17]. Pease et al. [11] developed a prognostic model combining deep learning of head CT scans with clinical information to predict long-term outcomes. Their fusion model demonstrated higher accuracy than the IMPACT model, though it did not improve predictions in external testing. Jadon et al [17] demonstrated high accuracy of a U-Net++ 2D Architecture with Focal Tversky Loss Function in segmenting intraparenchymal hemorrhage, extra-axial bleeds, and traumatic contusions from non-contrast CT.

The observed false positive rate of 15.4% may be due to imaging artifacts, case complexity, or limitations in the training data quality, indicating areas for further model refinement.

Our results indicate that AI can be a vital tool for radiologists, allowing for faster prioritization of critical cases and potentially improving patient outcomes. However, the study was conducted at a single institution and the generalizability of the results may be limited. Further validation on larger, more diverse datasets and comparison with additional diagnostic tools- is necessary.

Conclusion

This study has demonstrated the feasibility and potential clinical utility of using a deep learning model utilizing U-Net for semantic segmentation and ResNet-50 for feature extraction in detecting post-traumatic brain injury on NCCT scans. The AI model achieved high sensitivity and specificity, comparable to experienced radiologists, offering the possibility to enhance workflow efficiency, potentially leading to earlier and more informed decisions for TBI patients. Future work will focus on multi-institutional validation and improving the model's performance in complex cases.

References

1. Control CfD, Prevention. Report to congress on traumatic brain injury in the United States: epidemiology and rehabilitation. National Center for Injury Prevention and Control. 2015;2:1-72.
2. Badhiwala JH, Wilson JR, Fehlings MG. Global burden of traumatic brain and spinal cord injury. The Lancet Neurology. 2019;18(1):24-5.
3. Jones DB, Higgins A, Bishop N, Chamberlain DJ, Venkatesh B, Myburgh J, et al. Epidemiology and 12-month outcomes from traumatic brain injury in Australia and New Zealand. 2008.
4. Hoyt DB, Holcomb J, Abraham E, Atkins J, Sopko G. Working Group on Trauma Research Program summary report: National Heart Lung Blood Institute (NHLBI), National Institute of General Medical Sciences (NIGMS), and National Institute of Neurological Disorders and Stroke (NINDS) of the National Institutes of Health (NIH), and the Department of Defense (DOD). Journal of Trauma and Acute Care Surgery. 2004;57(2):410-5.
5. Sahuquillo J. Decompressive craniectomy for the treatment of refractory high intracranial pressure in traumatic brain injury. Cochrane database of systematic reviews. 2006(1).
6. Schweitzer AD, Niogi SN, Whitlow CT, Tsiouris AJ. Traumatic brain injury: imaging patterns and complications. Radiographics. 2019;39(6):1571-95.
7. Davis PC, Expert Panel on Neurologic I. Head trauma. AJNR Am J Neuroradiol. 2007;28(8):1619-21.
8. Zhu G, Jiang B, Tong L, Xie Y, Zaharchuk G, Wintermark M. Applications of deep learning to neuro-imaging techniques. Frontiers in neurology. 2019;10:869.
9. Toruner MD, Wang Y, Jiao Z, Bai H. Artificial intelligence in radiology: where are we going? EBioMedicine. 2024;109:105435.
10. Lui Y, Chang P, Zaharchuk G, Barboriak D, Flanders A, Wintermark M, et al. Artificial intelligence in neuroradiology: current status and future directions. American Journal of Neuroradiology. 2020;41(8):E52-E9.
11. Pease M, Arefan D, Barber J, Yuh E, Puccio A, Hochberger K, et al. Outcome prediction in patients with severe traumatic brain injury using deep learning from head CT scans. Radiology. 2022;304(2):385-94.
12. OBASA AA, OLOPADE FE, JULIANO SL, OLOPADE JO. Traumatic Brain Injury or Traumatic Brain Disease: A Scientific Commentary. Brain Multiphysics. 2024:100092.
13. Buh FC, Hutchinson PJ, Anwar F. Early neuro-rehabilitation in traumatic brain injury: the need for an African perspective. BMC medicine. 2023;21(1):290.
14. Königs M, Beurskens EA, Snoep L, Scherder EJ, Oosterlaan J. Effects of timing and intensity of neurorehabilitation on functional outcome after traumatic brain injury: a systematic review and meta-analysis. Archives of physical medicine and rehabilitation. 2018;99(6):1149-59. e1.
15. Pacchiano F, Tortora M, Criscuolo S, Jaber K, Acierno P, De Simone M, et al. Artificial intelligence applied in acute ischemic stroke: from child to elderly. Radiol Med. 2024;129(1):83-92.
16. Alouani AT, Elfouly T. Traumatic Brain Injury (TBI) Detection: Past, Present, and Future. Biomedicines. 2022;10(10):2472.
17. Jadon S, Leary OP, Pan I, Harder TJ, Wright DW, Merck LH, et al., editors. A comparative

study of 2D image segmentation algorithms for traumatic brain lesions using CT data from the ProTECTIII multicenter clinical trial. Medical Imaging 2020: Imaging Informatics for Healthcare, Research, and Applications; 2020: SPIE.

To cite this article: Medhat M. Reffat, Waseem M. El Gendy, Ahmed E. Shalaan, Khaled E. Ahmed, Ahmed S.Raslan. The Value of Artificial Intelligence (AI) in Detection of Post Traumatic Brain Injury Using Non-Contrast CT Scans. BMFJ 2025;42(7):978-987.