MULTIDISCIPLINARY
RESEARCH INSTITUTE
INDIA

Multidisciplinary Journal of Research in Engineering and Technology

ISSN: 2348-6953

Volume 13 Issue 2s (2026) | Pages: 218-223

Automated Skull Fracture Detection in CT Imaging: A Systematic Review of Computational Methods, Clinical Validation, and Future Directions

Vivek Bavaskar[1], Vedant Chaudhari[2], Tawheed Ahmed Ansari[3], Jayashree Khanapuri[4]

[1,2,3,4] Somaiya Institute of Technology, Mumbai, India

[1]vivek.bavaskar@somaiya.edu, [2]vedant.chaudhari@somaiya.edu, [3]tawheedahmed.a@somaiya.edu, [4]jayashreek@somaiya.edu

Peer Review Information

Type: *Article*
Received: *27 March 2026*
Revised: *12 April 2026*
Accepted: *26 May 2026*
Published: *16 June 2026*

Abstract

Skull fracture diagnosis is regarded as a significant diagnostic problem because it features a misdiagnosis rate up to 14.8% even among experienced radiologists. In this regard, the paper presents a systematic synthesis of 20 studies on skull fracture detection using conventional image-processing techniques, deep learning algorithms, and clinical validation. The best results obtained using the conventional technique are the recognition rates equal to 99-100%. In terms of deep learning models, their effectiveness is higher compared to others; for instance, CNN hybrid model reaches 97.6% accuracy, and YOLOv8 reaches 49% with recall 91.5%. Post-mortem skull fracture CT diagnosis features sensitivity 0.89 (0.80-0.94) with a base sensitivity of 0.87, and there are several visualization techniques such as azimuthal equidistant mapping that provide better skull fracture detection in less than one second of computational time. Important limitations are represented by the lack of dataset (CQ500 includes only 491 cases, including 84 patients with fractures), low inter-observer agreement (ICC < 0.4), and anatomical problems. Traumatic brain injury prevalence among maxillofacial fractures patients equals 51.04% (26,774).

Keywords: Skull Fracture Detection; Deep Learning; Computed Tomography (CT); Traumatic Brain Injury; Convolutional Neural Networks; Medical Image Analysis.

How to Cite This Article

Bavaskar, V., Chaudhari, V., Ansari, T. A., & Khanapuri, J. (2026). Automated skull fracture detection in CT imaging: A systematic review of computational methods, clinical validation, and future directions. Multidisciplinary Journal of Research in Engineering and Technology, 13(2s), 218–223.

Introduction

Skull fractures that lead to traumatic brain injuries occur in a population of 3.17 million people annually, and the total number of such fractures accounts for 23.4% of all head trauma incidents [6]. The possibility of subdural or epidural hematomas is increased due to such fractures, with the risk ranging between 10% and 30% in conjunction with CSF fistulae [8]. Misdiagnosis rates are currently at 14.8% and can be associated with negative outcomes in up to 41.1% of such incidents [4]. In the group of 84 patients who had a fracture analyzed based on the CQ500 data, only 14 subjects (16.7%) were positively identified by three radiologists with differing experience (8, 12, and 20 years), while the ICC was less than 0.4 [10].

In terms of computerized approaches, the detection process is based on a variety of techniques starting from histogram thresholding (0.99) [1] to multiple deep learning approaches such as CNNs, YOLOv8, and attention networks with sensitivities of 0.87 – 1.00 [3][5][14].

Traditional Image Processing

Histogram-Based Thresholding

Abubacker et al. [1] investigated 2,131 image slices of 10 subjects based on the characteristic intensity of skull bone (232-255 Hounsfield Units) to attain a sensitivity value of 1.00 and specificity of 0.995 (580 out of 590 fractured images recognized). Otsu multi-level thresholding gave an accuracy of 99% in a computation time of 1.7964 seconds, which is a 2.66 times enhancement compared to that of the k-means algorithm (4.7867 seconds) [2]. The performance evaluation of various edge detection algorithms showed Sobel as the best performer having 99% efficiency, then Canny (98%), Roberts (93%), and Prewitt (92%) [13][16][18].

Morphological Operations

Clinical validation of the technique was carried out using 213 abnormal cases based on multi-level segmentation using Fuzzy C-Means algorithm globally with two-level Otsu thresholding locally. The measures of performance revealed localization of ICH at 98% (65/65), calcification at 94% (82/87), ventricle misalignment at 94% (57/61), and an overall efficiency of 96.3% [19]. Further, black-hat operation achieved a minimum crack detection width of 0.35 mm with Jaccard Index of 80% for 1.05 mm crack; however, it wrongly detected sutures and vascular grooves as

Deep Learning Architectures

YOLOv8 Real-Time Detection

The performance of YOLOv8 was tested by Kumar and Agarwal [3] using the FracAtlas dataset, containing 880 triplets, for 100 epochs over a period of 0.881 hours with batch size 32, learning rate of 0.01, and Adam as the optimizer. The values for mAP50, precision, and recall were 0.49, 0.633, and 0.5, respectively. Performance by fracture type displayed significant variance (refer to Table I). Performance was highest in compression fractures while spiral/oblique fractures performed poorly due to small sample sizes (n=8, n=2).

Table 1. *YOLOv8 Fracture Type-Specific Performance on FracAtlas Dataset*

Fracture Type	n	mAP50	Precision	Recall
Compression	82	0.742	—	—
Spiral	8	0.264	—	—
Oblique	2	0.044	—	—
Overall	**880**	**0.49**	**0.633**	**0.500**

CNN-Based Hybrid Architecture

Kodavati and Kumar created a CNN hybrid architecture combining their classifier with DenseNet161 and U-Net algorithms to perform detection and segmentation. In particular, the training procedure involved the following parameters: batch size – 32, learning rate – 0.01, SGD optimizer, weight decay – 0.0001, and momentum – 0.9 for 40 epochs. Based on these settings, CNN achieved accuracy – 98.37%, AUC – 0.976, precision – 0.87, recall – 0.88, and F1 – 0.875, outperforming the baseline CNN model (AUC 0.90), SkullNetV1 (AUC 0.94), ResNet50+XGBoost (AUC 0.96), and DeepLab (AUC 0.9624).

Feature Pyramid Network with Attention

In turn, Liu et al. suggested a new CNN architecture based on the feature pyramid network with criss-cross attention (FPN+CCA) [5]. In terms of metrics, the model was tested on a database consisting of 19,222 training images and 1,922 test images and demonstrated an accuracy of 82.0%, recall of 91.5%, AUC of 0.972, and precision of 0.87. Compared with baseline architectures, the suggested method improved Faster R-CNN in terms of accuracy (+8.5%) and recall (+11.1%), as well as Cascade R-CNN in regard to accuracy (+11.8%) and recall (+15.7%). The contribution of each technique was analyzed through an ablation study, resulting in the following improvements: data augmentation (+3.2% in accuracy), criss-cross attention (+0.8% in accuracy and +1.2% in recall), ROI branch (+0.5% in accuracy), and Stretch IoU (+0.4%).

Weakly-Supervised Learning

Finally, Yang et al. applied the Grad-CAM algorithm to perform image localization with weak annotations [14]. In their experiment, the authors used 19,079 CT scans from 977 patients. In regard to accuracy, VGG-19 with full annotation reached 91.36%, while limited annotation did not affect its performance significantly, producing results of 89.37% at 10% annotation and 91.23% at 50% annotation (see Table 2).

Table 2. *Weakly-Supervised VGG-19 Accuracy vs. Annotation Level [14]*

Annotation Level	Accuracy (%)
Full (100%)	91.36
50%	91.23
10%	89.37

Alternative Visualization Techniques

Azimuthal Equidistant Mapping

A visualization technique for fractures was presented by Hadjittoouli et al. [10]. Applying cartographic principles, the researchers performed tests on CQ500 validation data, containing 491 CT scans, 193,317 images, and 84 positive cases. The inter-rater reliability index between radiologists' evaluations was below 0.4 (ICC < 0.4), demonstrating only 16.7% agreement. The algorithm produced Lower, Upper, Frontal, and Occipital (LUFO) projections, where fractures were depicted perpendicularly to the bone surface, with increased contrast. Linear computational complexity (O(n)) ensured that projection time per hemisphere remained less than 1 second on modern GPUs. In the case of patient CQ500-107, a temporal fracture was detected only by 2 out of 3 radiologists during traditional assessment; however, it was easily visible in LUFO projections. This approach allows for four LUFO projections compared to two standard reconstructions, providing better visual access to the basilar and frontal regions [9].

Disk Harmonic Maps

An approach to skull flattening based on differential geometry was suggested by Hadjittoouli et al. [18]. The segmentation of the skull involved Gaussian filtering (kernel size 5, $\sigma = 1$) followed by thresholding at 1200 HU and the use of alpha shape triangulation to address non-convex connectivity issues. The validation of the algorithm was carried out using data from 84 patients of CQ500, with successful flattening achieved, while distortion was restricted to featureless solid bone areas. The implementation of LUFO disk harmonic maps contributed to thorough visualization.

Biomechanical Modeling

Skull Implant Analysis

Shweta and Anburajan [9] examined implant material based on their analysis using the Finite Element Method for a defect in the front part, which was 34.07 mm long, 174.19 mm wide, and had a diameter of 38.87 mm. They had a mean fracture load of 1780 N and steel had the least displacement, 0.993×10^{-3} mm.

Table 3. *Detection Performance Metrics for Various Bone Fracture Types*

Type	n	Precision	Recall	mAP50	mAP50-95
Compression	82	0.754	0.748	0.742	0.353
Comminuted	137	0.666	0.596	0.613	0.225
Avulsion	57	0.651	0.639	0.592	0.180
Hairline	44	0.655	0.523	0.540	0.206
Spiral	8	0.400	0.250	0.264	0.068
Oblique	2	1.000	0.000	0.044	0.032

Traumatic Brain Injury Modeling

Khanuja and Unni [13] describe a three-dimensional finite element model of the head consisting of 288,962 tetrahedral elements, which are distributed as follows: cerebrum – 178,247; cerebellum – 17,936; brain stem – 10,383; cerebrospinal fluid (CSF) – 50,594; and skull – 31,802. The mechanical properties considered for each material are: skull (density $\rho = 3000$ kg/m³, Young's modulus $E = 7300$ MPa, Poisson's ratio ν = 0.22), CSF (ρ = 1000 kg/m³, E = 0.15 MPa, ν = 0.499), and soft tissues (ρ = 1040 kg/m³, shear modulus G∞ = 19.9 kPa, ν =0.499)

Table 4. *Impact Site vs. Injury Pattern from FE Head Model [13]*

Impact Site	Peak ICP (kPa)	Max Stress (kPa)	Key Finding
Frontal-top	322.4	—	Highest concussion risk
Occipital	—	76.6	Greatest neurological damage risk
Parietal	—	—	Strongest region (Pfrac = 0.83)

Clinical Validation

Post-Mortem CT Meta-Analysis

The systematic review by Henningsen et al. [4] included 18 eligible studies out of which 13 provided data used for meta-analysis. The total number of cases analyzed was 1,538. Standard scan parameters were set as the following: 120 kVp tube voltage, 110-250 mAs tube current, and 0.625 – 1.5 mm slice thickness. Results obtained after conducting a meta-analysis are presented below: the sensitivity for skull vault was 0.89 [95% CI: 0.80–0.94], for the skull base it was 0.87 [95% CI: 0.80–0.92]; the specificity for skull vault was 0.96 [0.91–0.98], and for skull base – 0.96 [0.90–0.98]; $I^2 = 45.6\%$ for skull vault and $I^2 = 0.0\%$ for skull base. Almost perfect agreement between the observers was reported by Le Blanc-Louvry et al. in 236 cases with Cohen's kappa values of 0.95 for skull base and 0.97 for skull vault [12]. Better accuracy of PMCT in detecting occipital condyle fractures not found in autopsy was established in the study by Daly et al. where 4 occipital condyle fractures and 10/14 of atlas-occipital fractures-dislocations were found via PMCT compared to 1 and 9/14 found by autopsy, respectively [11]. "False positive" values were obtained because the findings that could be seen on PMCT but were not visible during autopsy due to routine dissection restrictions. Acquisition parameters recommended by the authors are the following: 120 kVp tube voltage, 0.6 mm collimation, and the smallest possible slice thickness [4].

Maxillofacial Fracture–TBI Association

Othman et al. conducted a systematic review (PROSPERO CRD42020155912) which included 29 studies and 26,774 patients [7]. Road traffic accidents were reported to be the primary cause for 39.7-78.5% of patients. Maxillofacial fractures of the mandible occurred to be more common (22.8-70.4%), thus contributing to the higher rate of TBI. Severity of TBI was characterized by the following: concussions (47.27-100%), intracranial haemorrhage (28.6%), skull fractures (29.5%), brain contusion and laceration (16.9%). The age group where the highest prevalence was noted is 20-39 years. Glasgow Coma Scale ≤8 was found to be significantly associated with an increased risk of cervical spine injury [7].

Table 5. *Geographic Distribution and Prevalence of TBI in Maxillofacial Fracture Patients*

Region	Patients	MFF with TBI	Prevalence	M:F Ratio	Pooled Event Rate [95% CI]	I^2
Africa	3,714	1,684	45.3%	3.2:1	0.393[0.250-0.563]	95.9%
Asia	16,759	8,979	53.9%	5.9:1	0.556[0.345-0.748]	61.2%
Europe	4,095	1,908	46.5%	5.24:1	0.465[0.264-0.650]	98.4%
America	2,206	1,096	49.6%	4.75:1	0.480[0.315-0.650]	94.9%
Over all	**26,747**	**13,667**	**51.04%**	**4.8:1**	-	-

Research Gaps and Future Directions

Current Limitation

- Availability of Datasets: Limited access is only to the publicly available CQ500 dataset, containing 491 images of 193,317, and 84 patients with fractures [10]. The YOLOv8 algorithm was found to have highly variable performance due to insufficient samples: oblique fractures (n = 2, mAP50 0.044), spiral fractures (n = 8, 0.264), and compression fractures (n = 82, 0.742) [3]. Public annotated datasets include few calvarial fractures but lack annotations of basilar and facial fractures.
- Interobserver Reliability Issues: Poor interobserver agreement (ICC < 0.4) has been shown in pixel-level ground truth annotation, even in the case of three experienced radiologists with 8, 12, and 20 years of experience [4][10]. Annotation of fractures is hindered due to presence of sutures, vascular grooves, and anatomical variants that can be mistaken for fractures. Black-hat transformation

of the image has been applied to detect sutures and vascular grooves, leading to higher false positive rate [20].

- Clinical Utility: Timeliness of processing is vital for the emergency clinical setting. Lack of evaluation standards makes it difficult to compare results across different studies. Insufficient validation for diverse populations and trauma mechanisms is needed. Clinical validation is necessary to demonstrate non-inferiority compared to radiologist performance to obtain FDA/CE approval.

Future Directions

- Advanced architectures. ViT and Swin transformers showed better results in the medical image analysis because of self-attention mechanisms enabling long-range dependence. Three-dimensional CNN could also improve the fracture detection using its volumetric nature in the case of two-dimensional analyses missing some scan plane. Privacy-preserving federated learning is possible in order to train models without exchanging data.
- Multimodal fusion. Multimodal analysis using CT and MRI allows visualizing bone structures and analyzing soft tissues for more comprehensive diagnosis of TBI. Combined use of fMRI and diffusion tensor imaging is needed for the assessment of white matter tract injury. There is an opportunity for registration.
- Clinical decision support. Clinical decision support systems include automated detection, risk stratification for TBI, treatment plan and prognosis. Machine learning models are trained on huge clinical databases with correlations between clinical characteristics and image features. Optimization of automated triage identifies patients who should get urgent surgical care.
- Hierarchical classification. These systems classify the fracture according to their type (linear, diastatic, depressed, comminuted, compound) and analyze them based on their location (vault, base, and facial). The integration of classification with existing grading of the severity (Abbreviated Injury Scale) allows prognostic analysis. Type-based approach eliminates potential problems with YOLOv8.
- Standards. Consistent CT image acquisition protocol (tube voltage 120 kVp; tube current 240-420 mAs; slice thickness ≤0.625 mm). Consensus annotation guidelines for the extent of fractures, fracture localization, and classification according to severity. Multicenter research is performed in order to establish inter-rater reliability of annotation. Performance metrics include type-specific, location-based, and clinical impact analysis.
- Ethical considerations. Analysis of potential demographic bias including different age groups, male to female ratio being currently 4.8:1 and different ethnicities. Compliance with regulations regarding clinical validation studies. Development of liability allocation strategies. Reporting of limitations of the system and failure modes. Continuous post-market surveillance.

Conclusion

In this article, we provide a systematic evaluation of automated skull fracture detection models in twenty papers. In recent decades, there have been considerable advancements from classical image processing, namely the histogram thresholding technique (99%), to more sophisticated deep learning models (98.37%, AUC = 0.976). Deep learning models, including YOLOv8, a hybrid model based on the CNN network, and an attention-oriented model, performed clinically well in terms of sensitivity ranging from 0.87 to 0.98. Alternative visualization techniques, such as azimuthal equidistant projection and disk harmonic mapping, can detect all fractures using computations within less than one second. Clinical validation was conducted using post-mortem CT imaging meta-analysis of 1,538 images, showing 0.89 sensitivity and 0.87 sensitivity for cranial vault and base fractures, respectively, with a specificity of 0.96.

However, inconsistency in diagnosis remained significant, as 16.7% of complete agreement between radiologists and an ICC below 0.40 were observed. It has been revealed that the impact locations differ in biomechanical responses, where the impacts at the frontal-top area pose the maximum concussion probability (ICP = 322.4 kPa) and impacts on the occipital region represent high risks of neurological injuries (stress = 76.6 kPa). On a global level, the incidence of TBI was 51.04% (26,774 cases) in patients with maxillofacial fracture cases. Main problems identified are relatively small data size (CQ500 is the easiest access to), low interobserver agreement in establishing ground truth, challenges in distinguishing between fractures and anatomical regions, and under-representation of some fracture types (oblique n = 2, spiral n = 8).

Research directions discussed are: (1) building large-scale annotated data; (2) designing 3D convolutional networks and transformers utilizing continuity information in volume data; (3) introducing weakly-supervised learning methods; (4) multimodal fusion with additional imaging modalities, such as CT, MRI, and functional imaging; (5) clinical decision-support system combining automatic identification and predictive analytics; (6) standardized imaging, annotation, and evaluation protocols; and (7) prospective clinical validation to demonstrate non-inferiority to radiologists.

Automated systems have promising applications in emergency medicine, trauma surgery, neurosurgery, and forensic pathology. Real-time automatic detection may lead to better triage management, which will reduce time consumption. The clinical application is subject to careful validation and considerations regarding algorithmic biases and regulatory approval .

Acknowledgment

The authors acknowledge all contributing studies and institutions. This work was conducted as part of the undergraduate academic programme at Somaiya Institute of Technology, Mumbai.

References

1. Abubacker, S. N., Sulaiman, N., & Salleh, S. H. “An approach for automatic fracture detection of skull DICOM images based on neighboring pixels,” Proc. IEEE International Conference on Control System, Computing Engineering, pp. 558–563, Nov. 2015.
2. Zaki, W. M. H., Husain, M., Farok, M. L. A. M. R. M., & Sulaiman, N. A. “Automated method of fracture detection in CT brain images,” Proc. IEEE EMBS Conference on Biomedical Engineering Sciences, pp. 285–289, Dec. 2016.
3. Kumar, A., & Agarwal, D. “Automated bone fracture detection using YOLOv8,” International Journal of Research in Applied Science and Engineering Technology, vol. 11, no. 5, pp. 2998–3005, May 2023.
4. Henningsen, M. J., Larsen, S. T., Jacobsen, C., & Villa, C. “Sensitivity and specificity of post-mortem computed tomography in skull fracture detection—a systematic review and meta-analysis,” International Journal of Legal Medicine, vol. 136, pp. 1363–1377, Mar. 2022.
5. Liu, G., Wen, Y., Bai, X., Zhu, W., & Li, H. “Skull fracture detection method based on improved feature pyramid network,” Proc. IEEE 10th International Conference on CYBER Technology in Automation, Control, and Intelligent Systems, pp. 756–762, Oct. 2021.
6. Maas, A. I., Menon, D. K., Adelson, P. D., et al. “Traumatic brain injury: Integrated approaches to improve prevention, clinical care, and research,” Lancet Neurology, vol. 16, no. 12, pp. 987–1048, Dec. 2017.
7. Othman, A., AlQahtani, F. A., AlQahtani, H., Jaber, M., Bishawi, K., Khamis, A. H., & AlShanably, A. “Traumatic brain injuries and maxillofacial fractures: A systematic review and meta-analysis,” Oral and Maxillofacial Surgery, vol. 27, pp. 373–385, Jun. 2023.
8. Aarabi, T. M., Gerstein, B., & Harrop, C. “Skull base fractures and their complications,” Journal of Trauma, vol. 71, no. 6, pp. 1619–1625, Dec. 2011.
9. Ringl, C., Schernthaner, H., Schaefer-Prokop, R., et al. “Curved planar reformations of the skull: A new visualization technique for skull fractures,” Radiology, vol. 258, no. 2, pp. 604–610, Feb. 2011.
10. Hadjittoouli, N., Nikolaou, C., & Pitris, C. “Azimuthal equidistant mapping and projection of head CT scans for fracture visualization,” IEEE Access, vol. 13, pp. 78214–78220, May 2025.
11. Daly, B., Orave, P. J., & Cibelli, L. “Occipital condyle fractures in the era of multidetector CT,” Emergency Radiology, vol. 19, no. 5, pp. 407–413, Oct. 2012.
12. Le Blanc-Louvry, I., Thureau, A. C., Duval, E., Papin-Lefebvre, B., & Thiebot, G. “Postmortem computed tomography compared to forensic autopsy findings: A French experience,” European Radiology, vol. 23, no. 7, pp. 1829–1835, Jul. 2013.
13. Khanuja, T., & Unni, H. N. “Computational modeling of traumatic brain injury due to impact on different sides of human head,” Proc. IEEE 18th International Conference on Bioinformatics and Bioengineering, pp. 364–370, Oct. 2018.
14. Yang, C.-Y., Lo, C.-H., Wang, H.-C., Chou, J.-H., & Wang, Y.-C. F. “Weakly-supervised learning for attention-guided skull fracture classification in computed tomography imaging,” Proc. IEEE International Conference on Image Processing, pp. 1337–1341, Sep. 2019.
15. Kodavati, S. K., & Kumar, Y. C. A. P. “Detection and segmentation of skull fractures via CNN and U-Net hybrid model using computed tomography images,” SN Computer Science, vol. 3, no. 397, pp. 1–12, Sep. 2022.
16. Canny, J. “A computational approach to edge detection,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-8, no. 6, pp. 679–698, Nov. 1986.
17. Gonzalez, R. C., & Woods, R. E. Digital Image Processing, 3rd ed. Upper Saddle River, NJ: Prentice-Hall, 2008.
18. Hadjittoouli, N., & Pitris, C. “Skull fracture visualization using disk harmonic maps,” Proc. IEEE 22nd International Symposium on Biomedical Imaging, pp. 1–4, May 2025.
19. Zaki, Z., Rahiman, M. H. F., Hitam, M. S., & Sulaiman, N. “Multi-level segmentation method for serial computed tomography brain images,” Journal of Telecommunications, Electronic and Computer Engineering, vol. 9, no. 1–2, pp. 49–56, 2017.
20. Yamada, A., Teramoto, A., Otsuka, T., Kudo, K., Anno, H., & Fujita, H. “Preliminary study on the automated skull fracture detection in CT images using black-hat transform,” Proc. 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 6437–6440, Aug. 2016.