Journal Pre-proof

Applications of Deep Learning in Trauma Radiology: A Narrative Review

Chi-Tung Cheng, MD, Ph.D., Chun-Hsiang Ooyang, MD, Shih-Ching Kang, MD, Chien-Hung Liao, MD, FACS, FICS

PII: S2319-4170(24)00046-5

DOI: https://doi.org/10.1016/j.bj.2024.100743

Reference: BJ 100743

To appear in: *Biomedical Journal*

Received Date: 13 November 2023

Revised Date: 26 March 2024

Accepted Date: 24 April 2024

Please cite this article as: Cheng C-T, Ooyang C-H, Kang S-C, Liao C-H, Applications of Deep Learning in Trauma Radiology: A Narrative Review, *Biomedical Journal*, https://doi.org/10.1016/j.bj.2024.100743.

Applications of Deep Learning in Trauma Radiology: A Narrative Review

Chi-Tung Cheng, MD, Ph.D.; Chun-Hsiang Ooyang, MD; Shih-Ching Kang, MD; Chien-Hung Liao, MD, FACS, FICS

Department of Trauma and Emergency Surgery, Chang Gung Memorial Hospital, Linkou; Chang Gung University, Taoyuan Taiwan.

Correspondence to: Shih-Ching Kang MD. Department of Trauma and Emergency Surgery, Chang Gung Memorial Hospital, Linkou; Chang Gung University, Taoyuan Taiwan.

E-mail: b001089030@cgmh.org.tw

Tel: 886-3281200 Ext. 3651; Fax: 886-3-3285818

Abstract

Diagnostic imaging is essential in modern trauma care for initial evaluation and identifying injuries requiring intervention. Deep learning (DL) has become mainstream in medical image analysis and has shown promising efficacy for classification, segmentation, and lesion detection. This narrative review provides the fundamental concepts for developing DL algorithms in trauma imaging and presents an overview of current progress in each modality. DL has been applied to detect free fluid on Focused Assessment with Sonography for Trauma (FAST), traumatic findings on chest and pelvic X-rays, and computed tomography (CT) scans, identify intracranial hemorrhage on head CT, detect vertebral fractures, and identify injuries to organs like the spleen, liver, and lungs on abdominal and chest CT. Future directions involve expanding dataset size and diversity through federated learning, enhancing model explainability and transparency to build clinician trust, and integrating multimodal data to provide more meaningful insights into traumatic injuries. Though some commercial artificial intelligence products are Food and Drug Administration-approved for clinical use in the trauma field, adoption remains limited, highlighting the need for multi-disciplinary teams to engineer practical, real-world solutions. Overall, DL shows immense potential to improve the efficiency and accuracy of trauma imaging, but thoughtful development and validation are critical to ensure these technologies positively impact patient care.

Keywords

Deep learning, artificial intelligence, medical image, trauma, radiology, X-ray, CT, ultrasound

Introduction

In modern trauma care, various imaging modalities are essential, especially for patients with blunt trauma. Although there are situations in which patients require immediate surgical investigation, diagnostic imaging offers clinicians useful information that opens the door to endovascular and non-operative management and therapies. Key imaging modalities for initial evaluation include chest and pelvic X-rays, the Focused Assessment with Sonography for Trauma (FAST), and computed tomography (CT) scans focusing on the head and cervical spine. For areas of suspected injury, targeted CT imaging is frequently indispensable for making precise diagnoses. The use of whole-body CT scanning has grown increasingly common for evaluating trauma patients with polytrauma or high-risk injury mechanisms, offering comprehensive diagnostic insights quickly to inform management decisions. In the fast-paced environment of trauma care, clinicians are often tasked with interpreting numerous imaging results, sometimes in the absence of a radiologist's direct input. Moreover, emergency radiologists encounter distinct challenges in a high-stakes environment where time and safety are critical[1].

Deep learning (DL) is a machine learning algorithm based on artificial neural networks that has rapidly advanced in the past decade, emerging as a primary algorithm for tasks related to computer vision and natural language processing, which refer to artificial intelligence (AI) recently. Within the realm of medical imaging, DL has diverse applications. These encompass disease classification in chest radiography, lesion detection in chest CT scans, mammography interpretation, and analysis of brain MRI images[2]. Moreover, DL has been employed to segment lesions and organs in brain CT[3] and abdominal CT scans[4]. In the field of trauma imaging, a recent scoping review has highlighted the increasing prevalence of DL applications within computer-aided diagnosis (CAD) tools for trauma[5]. In this narrative review, we present a comprehensive overview of the foundational concepts of DL, which are crucial for guiding the development of medical DL projects. We explore the current applications of DL in trauma imaging and discuss potential avenues for future advancements.

Fundamental Concepts of Deep Learning

The basic concept of artificial neural networks was proposed as early as the 1960s, inspired by the biological nervous system as a machine learning algorithm. The architecture consists of multiple artificial neurons connected by weighted connections (perceptron), forming a hierarchical structure typically comprising input, hidden, and output layers. Each connection between neurons involves a mathematical computation that increases the computation requirement as neurons and layers increase, which limits its development and utilization in the early years. However, with advancements in computational capabilities, particularly the advent of graphic processing units (GPUs) utilized in the computation of neural networks in early 2010[6], developing deeper network structures of neural networks became possible. "Deep learning" is predominantly employed to describe such multilayered neural networks, thereby superseding the earlier nomenclature of multilayer neural networks.

Before the era of DL, the predominant approach to image analysis relied upon hand-crafted feature selection, feature extraction, and subsequent machine learning techniques. Developing a convolution neural network (CNN) breaks the concept and uses an end-to-end algorithm framework to learn the features automatically during the training process. This method rapidly became the principal image analysis technique after gaining achievements in the ImageNet competition[7] and became a cornerstone in DL.

While DL techniques have succeeded considerably in natural image classification, researchers have tried to adapt them to medical images. Training a robust DL model from scratch requires extensive volumes of image data, principally due to the high-dimensional feature space inherent to these algorithms. However, the collection of medical images is more challenging than natural images. Fortunately, the concept of "transfer learning" allows for fine-tuning pre-existing DL models on new image datasets, thereby enabling the application of models initially trained on one domain of images to be effectively re-purposed for a distinct target domain[8]. Early successful works in medical imaging include retina images, skin lesions classification, chest radiography interpretation, digital pathology, and lung nodule detection[2]. The DL algorithm in the image analysis field is mainly categorized as classification, segmentation, object detection, registration, and image generation.

The classification algorithms, such as AlexNet, ResNet series, Inception series, DenseNet, ResNeXt, and VisionTransformer, are the basis of DL in computer vision[9]. The trained model can recognize the input images as a specific class if the class is defined during the training process. For example, the collected ankle radiography can be defined as “fracture” or “normal” to train an ankle fracture detection algorithm. Multiple class identification is also possible if the image number of each class is sufficient for training. Segmentation algorithms such as Fully Convolutional Network, U-Net, DeepLab series, and SegNet[10] can partition an image into multiple parts and contour the target object by pixel-wise classification. The segmentation algorithms can precisely identify the target lesion, for example, hemorrhage in a brain CT[3]. However, intensive annotation effort is needed to draw the target lesion contour in each training dataset to train a segmentation algorithm, making high-quality datasets rare. Object detection networks such as R-CNN, Faster R-CNN, YOLO series, RetinaNet, and Single Shot MultiBox Detector[11] are designed to identify the target object in the image and place a bounding box to point out the location. The training data can be prepared with rectangle boxes to define the lesion area without precise contouring. An example of applying different algorithms for spleen injury detection in CT scans is shown in Figure 1. Additional techniques, such as image registration and generation algorithms, may enhance performance. Image registration is particularly useful for the precise alignment of different modalities, which is essential for accurate lesion detection. For example, Zhou et al. applied this technique to align the arterial phase and venous phase CT scans to proceed with vascular injury segmentation[12]. Image generation algorithms are commonly employed as training data augmentation methods in medical imaging.

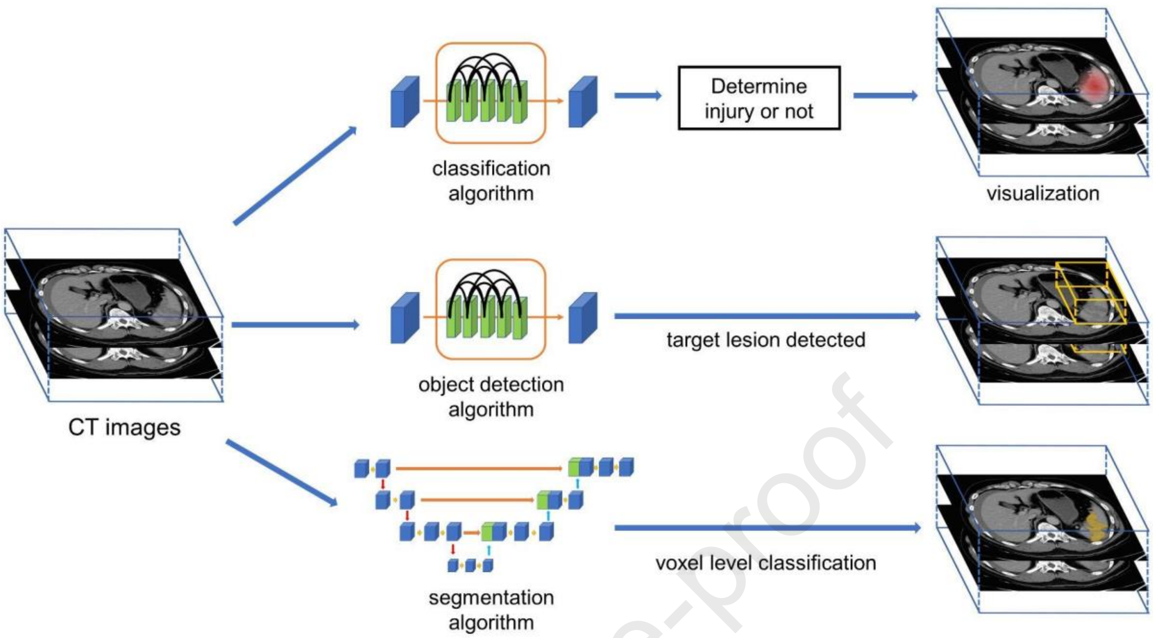

Figure 1. Applying different algorithms to detect spleen injury.

The most frequently used performance evaluation metric is the area under an operating characteristic curve (AUROC), which can be used to compare different models' performance in the same dataset because the model outputs are usually a probability value. After defining a cutoff value, the accuracy of diagnosing a specific injury by a binary class can be evaluated. However, the sensitivity (recall), specificity, positive predictive value (precision), and negative predictive value are much more critical to evaluating the true presentation of a model regarding the prevalence and the dataset imbalance[13]. The most commonly used evaluation indices in the segmentation algorithm include pixel accuracy, Intersection over union (IoU), and Dice similarity coefficient (DSC). Mean average precision (mAP) is specifically used in object detection algorithms to summarize prevision and recall across all object classes[14]. An independent or external test set is best for objectively evaluating the models because DL algorithms tend to overfit the training dataset. Researchers must avoid data leakage across training and test sets.

In medical imaging, some modalities, such as CT, should be interpreted through three-dimensional (3D) or dynamic approaches rather than relying solely on a single two-dimensional (2D) view. This process involves transforming the pixel-based 2D images into a 3D format and adapting the algorithm to analyze the scans on a voxel basis. Although 3D methods offer comprehensive insights, they demand considerable computational power. Furthermore, many DL algorithms, initially designed for 2D image analysis, are only partially transferable to 3D formats. To address this, some researchers have adopted a 2.5D approach, which incorporates additional spatial context from adjacent slices, thereby reducing computational demands while enhancing the depth of analysis[15,16]. An example of the difference between using 2D or 3D algorithms to develop splenic injury detection models is shown in Figure 2.

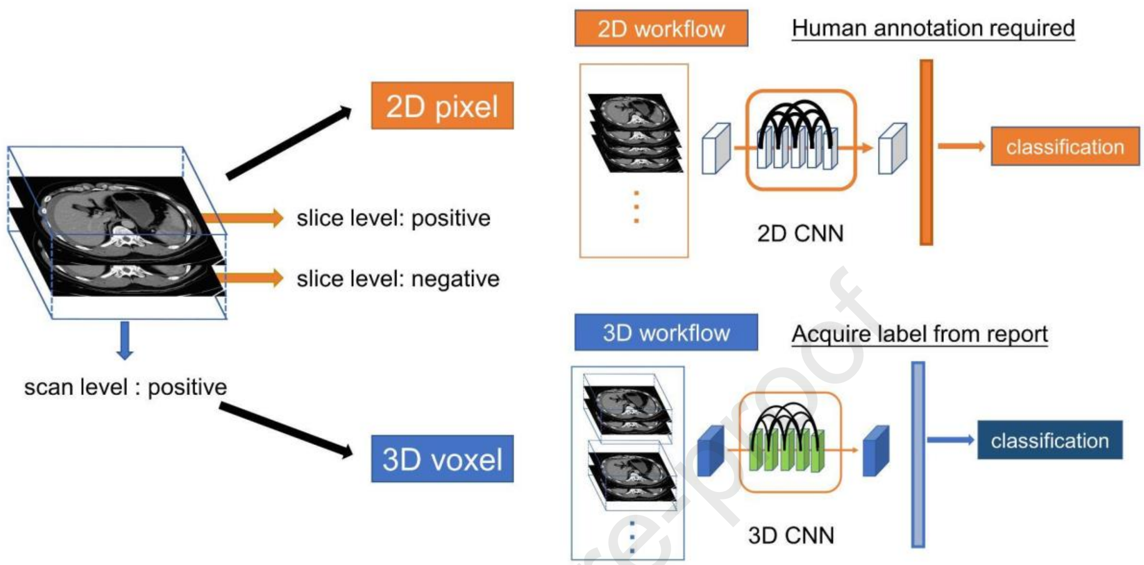

Figure 2. Example of classification algorithm workflow differences between 2D and 3D methods.

The development of DL models in medical institutions adheres to a technology readiness evaluation framework for machine learning projects[17]. This general prototyping workflow encompasses several key steps, as outlined in Figure 3. First, medical domain knowledge from experts should be the core of a specific project. Most projects fail because the solution provided is minimal or lacks clinical impact. Next, define the problem to solve from the clinical experience and understanding of specific image modality. For example, researchers may target a specific injury frequently misdiagnosed with severe consequences. The target application can be determined in the beginning as Computer-Aided Detection(CADe), Computer-Aided Diagnosis (CADx), Computer-Aided Triage (CADt), or image processing/quantification (IPQ) according to the FDA definition. All the clinical information and image acquisition must be approved by the Institutional Review Board (IRB) to evaluate medical data's ethical and privacy concerns. After the above preparation, researchers can start gathering the data from the source, mainly from the Picture Archiving and Communication System (PACS) repository in an institute or publicly released open datasets, followed by a de-identification process. The collaborating computer scientist must discuss with medical experts how to translate the question into a practical computer vision task, such as classification, segmentation, or object detection, and then design the algorithm further. Existing labels, including radiology reports and clinical diagnoses, can be obtained from hospital databases. However, the error rate in radiologists' reports for emergency plain radiographs may range from 3 to 6%[18], and segmentation masks are often absent.

Therefore, manual labeling is usually required[19,20]. Sometimes, the algorithm structure based on natural images can be directly applied to medical problems, but researchers can stack DL layers from scratch to suit their specific tasks. Finally, the iterations of model training and adjustment begin to enhance performance.

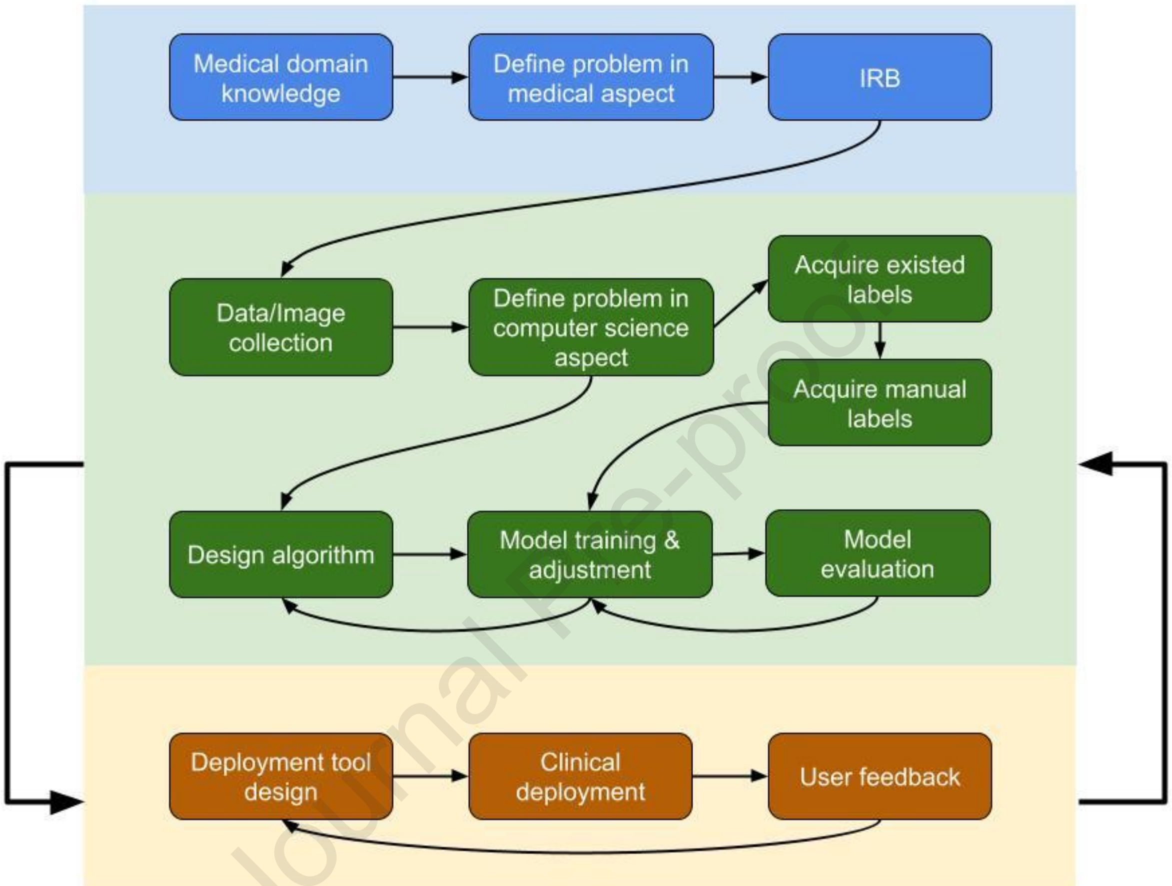

Figure 3. The prototyping workflow of developing a deep learning algorithm in trauma imaging. IRB: Institutional Review Board.

Applications of Deep Learning in Trauma Imaging

The trauma patient management workflow started with a primary survey to rapidly diagnose immediate life-threatening conditions. Figure 4 illustrates a clinical workflow of a patient with major abdominal bleeding. After the initial physical examination, the Advanced Trauma Life Support (ATLS) guidelines advocate for the focused assessment with sonography in trauma (FAST) or extended FAST (eFAST) to evaluate the patient. The FAST exam aims to detect hemoperitoneum or cardiac tamponade rapidly, and eFAST added chest exams to detect hemothorax or pneumothorax. In addition to the binary outcome of the exam, rough quantitative data can offer insights into the severity of each targeted body region, such as the volume of hemoperitoneum and hemothorax. Several research teams developed free fluid detection models[21–23] using ResNet50 and YoloV3, achieving over 95% accuracy in balanced datasets. Furthermore, Chiu et al. have experimented with the DL algorithm to assist novices in conducting FAST exams. This results in operators with AI guidance achieving a mean rate of acceptance score better than those without AI guidance(84% vs 68%; P =0.002) [24]. Kornblith et al. utilized ResNet-152 to automate the recognition of views in a pediatric cohort, while Taye et al. presented a CNN autoencoder for assessing the quality of ultrasound views[25]. In chest trauma, only Potter et al. applied YoloV3 to detect pericardial fluid, reaching 89% sensitivity in 78 cases[26]. However, there remains a lack of studies concerning the ultrasound detection of traumatic pneumothorax or hemothorax. Only Boice et al. tried to train a CNN pneumothorax classifier in a synthetic tissue phantom.

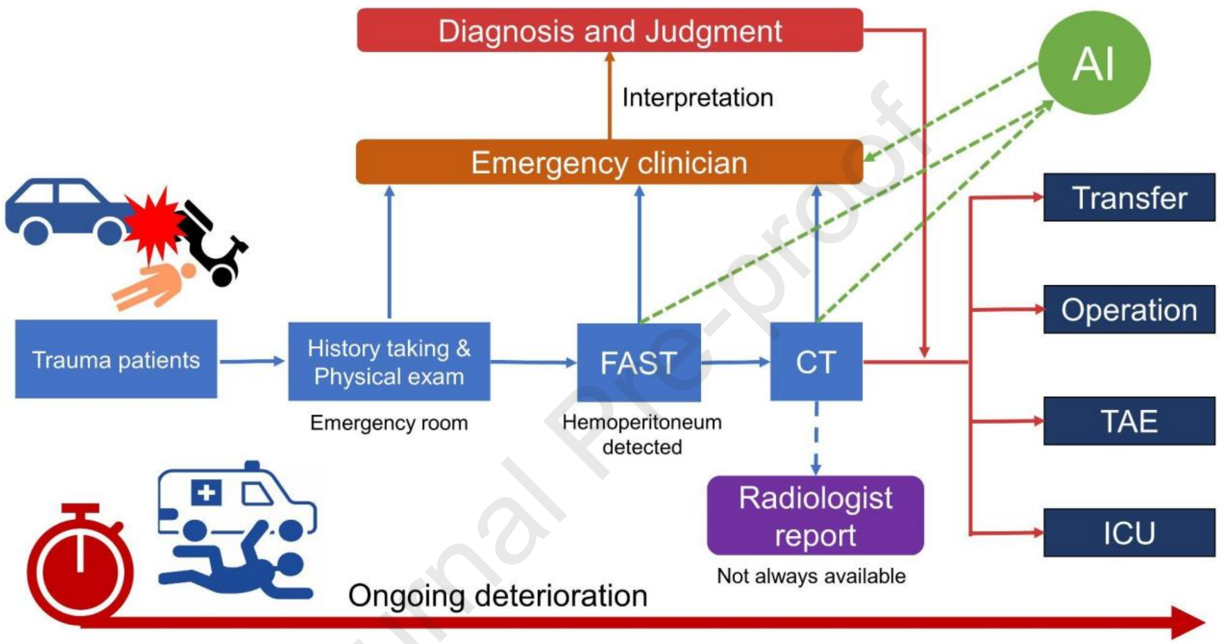

Figure 4. The clinical workflow for a patient with blunt abdominal trauma with abdominal internal bleeding. AI: artificial intelligence; CT: computed tomography; FAST: focused assessment with sonography in trauma; TAE: transarterial embolization; ICU: intensive care unit.

Plain chest radiographs (chest X-ray, CXR) are the most commonly used image modality in the medical domain. CXR can diagnose various medical conditions, such as cardiomegaly, pneumonia, spontaneous pneumothorax, pleural effusion, tuberculosis[27], lung nodule or lung mass, etc[28,29]. The CXR is also an initial workup for trauma patients to detect rib fractures, hemothorax, pneumothorax, aortic injury, or other skeletal fractures[30]. Patients with any thoracic symptom, major mechanism, or multiple trauma are indicated for CXR. Although many articles and even commercially available products regarding the detection of medical conditions in CXR[31], there are only a few focused on trauma findings. This might related to most of the released large-scale open CXR datasets such as NIH ChestX-ray14, CheXpert, MIMIC-CXR, PadChest, and VinDr-CXR lack of labels referring to trauma pathologies except for few rib fractures[32]. Since rib fractures are frequently missed condition, most studies are focused on detecting rib fractures using private datasets. Huang et al. manually crop the region of interest of ribs in CXR to classify rib fractures from normal using benchmark networks[33], while Ghosh et al. use a similar approach with a patch-based sliding window technique in the pediatric population with best AUROC of 0.89[34]. Sun et al. tried a simple framework for contrastive learning of visual representations on a small number of CXR test sets that achieved the highest sensitivity of 0.873 with five false positives per scan[35]. No convincing rib fracture detention model has been published except for the above experimental approach. A commercially available CXR algorithm was validated by Gipson et al. in trauma cases, resulting in better performance in identifying pneumothorax and segmental collapse compared to radiologists but poor performance in detecting fractures of the clavicle, humerus, and scapula[36]. Although hemothorax, lung contusion, and great vessel injuries are critical findings in CXR, none of the published articles focused on these topics. Figure 5 shows several important image modalities in trauma diagnosis.

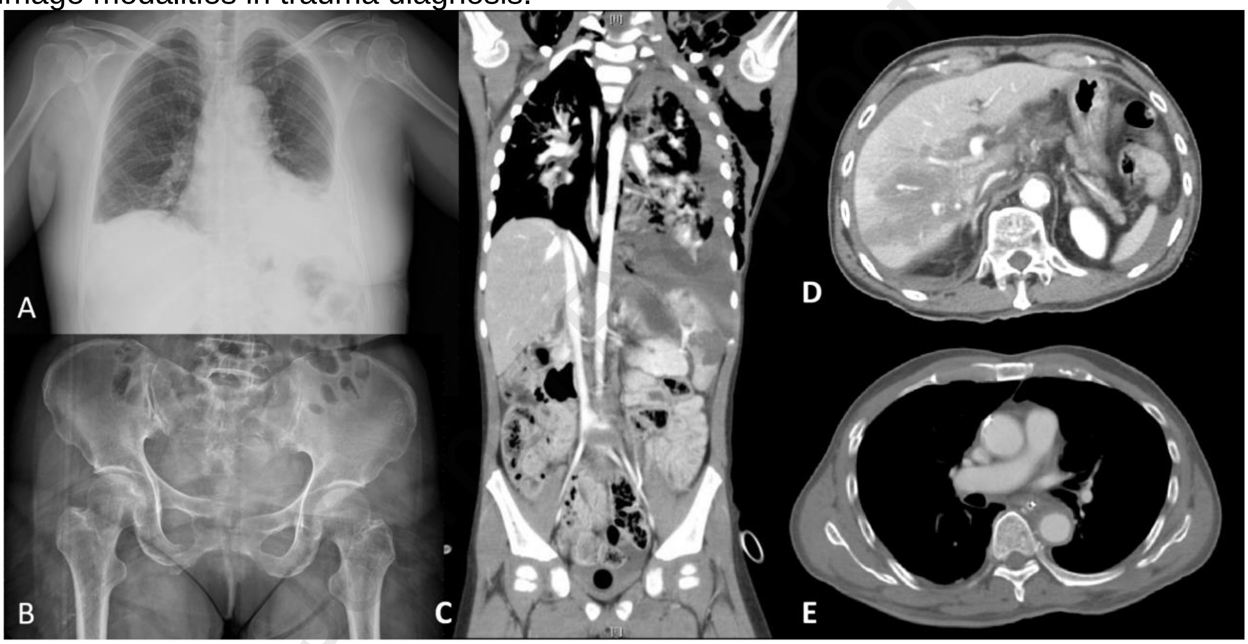

Figure 5. (A) A Chest X-ray shows left multiple rib fractures with hemothorax. (B) A pelvic plain film presented a vertically shearing type pelvic fracture. (C) A coronal view of a whole-body CT demonstrated multiple thoracoabdominal injuries. (D) An abdominal slice shows liver injury with active contrast extravasation. (E) A chest CT slice with a slightly displaced right rib fracture.

According to the ATLS guidelines, pelvic radiographs (PXRs) are another type of plain film recommended for initial evaluation to exclude pelvic fractures, which may cause massive bleeding. This modality can also evaluate the hip and proximal femur to diagnose fractures in these areas. Hip fracture detection is of great interest because timely intervention may reduce complications in older adults, which can be misdiagnosed in a chaotic emergency room. Several researchers have developed DL models for detecting hip fractures, including femoral neck and trochanteric fractures, achieving accuracy and sensitivity higher than 90%[37–43]. Furthermore, the subclassification of fracture types of the hip region has also been studied[37,41,44]. Not only comparing the algorithm with humans, Cheng et al. and Sato et al. also found the hip fracture detection model can assist inexperienced clinicians in improving diagnostic performance[45,46]. Pelvic fractures are presented as more complex in the PXR involving multiple areas. Cheng et al. developed a universal detection algorithm that combines DenseNet and feature pyramid networks, achieving 94.5% accuracy in detecting pelvic fractures. However, the study's heatmap only indicates the abnormal position without providing high-level diagnostic classification[47]. The accuracy of detecting certain types of fractures, including acetabular, sacral, and sacroiliac fractures, may be overestimated due to the inherent limitations of PXRs in accurately identifying injuries in these regions. An FDA-approved tool achieved a sensitivity of 93.0% and a specificity of 91.5% in detecting fractures on pelvic films. However, a detailed analysis comparing performance across different types of fractures is lacking[48].

Brain CT is the image of choice regarding head injuries. Timely detection of brain hemorrhage is the key to treating patients with traumatic brain injury. Several studies focused on stroke-related intracranial bleeding, which is sometimes hard to distinguish from traumatic brain injury. Other combining all intracranial hemorrhage regardless of the mechanisms. Most studies use a slice-based approach with classification and segmentation networks to detect bleeding[49–51]. The sensitivity and specificity performance ranged from 0.87 to 0.97 and 0.90 to 0.97, respectively. Chang et al. applied a hybrid 3D/2D method to hemorrhage detection and segmentation, achieving an AUROC of 0.98. Incorporating 3D-CNN could train models using examination labels from large-scale datasets with fewer labeling efforts, but those performances are lower than pixel or image label datasets[52–54]. The severity of different types of hemorrhage varies, such as subdural, epidural, intraparenchymal, and subarachnoid hemorrhage. Researchers also tried to subtype the hemorrhage at the same time of detection[55–57]. Although the performance of models cannot be compared in the same dataset, the best model achieved an AUC of 0.97 in a large-scale dataset[56], which might be ready for clinical use. The prediction of brain trauma outcomes using DL was also studied. Pease et al. developed a CNN and clinical model fusion model that achieved an AUC of 0.92 on mortality prediction on the internal data set but decreased on external Transforming Research and Clinical Knowledge in Traumatic Brain Injury dataset testing[58]. An open dataset CQ500[51], released in 2018, includes 491 head CT scans with scan-level classification labels annotated by three radiologists for five types of intracranial hemorrhage, the presence of calvarial fractures, midline shift, and mass effect. In 2019, the Radiological Society of North America (RSNA) launched a brain CT hemorrhage challenge, releasing a multi-institutional, multinational brain hemorrhage CT dataset. This dataset comprises 874,035 images from 25,312 examinations, each annotated with slice-level labels for five subtypes of hemorrhage[59]. These datasets facilitate benchmarking for classification tasks in newly developed algorithms.

The C-spine injury should be identified soon to prevent further deterioration. The preliminary diagnostic tool of choice shifted from plain film to C-spine CT since the 9th version of ATLS in 2012. Thoracolumbar fractures are checked with plain films first, then further confirmed by CT scan. Spinal cord injury should be further investigated with MRI if the patient's condition is feasible. Boonrod et al. investigated YOLO performance[60], and Naguib et al. applied classification algorithms with a saliency map in evaluating C-spine fracture on lateral view radiographs[61]. Rather than directly classify injury cases, several researchers start by performing the segmentation of vertebrae first on X-ray and CT images[62–65]. Voter et al. tested commercially available cervical spine fracture detection algorithms, resulting in a poor sensitivity of 54.9%[66]. Ma et al. use Faster R-CNN to detect spinal cord injury on MRI, achieving a mean average precision of 88.6[67].

Chest, abdominal, and pelvic CT scans are commonly utilized in the evaluation of specific organ injuries in relatively stable trauma patients. CT provides detailed information on body structures, organ integrity, and establishes itself as the diagnostic gold standard. Contrast-enhanced CT is crucial for identifying active bleeding, known as contrast extravasation. In the era of increasing reliance on non-operative management and endovascular interventions, CT delivers valuable information regarding the severity, characteristics, and location of detailed organ injuries, aiding clinicians in determining treatment options and prognosis[68]. Chest CT also assists in diagnosing rib fractures with precision, whereas CXR provides only approximate information due to the interference of overlapping organs[69], in addition to facilitating pneumohemothorax quantification. Furthermore, pelvic fractures require CT for accurate characterization, vascular injury detection, and to provide three-dimensional insights for orthopedic reconstruction[70]. In recent decades, WBCT scans have been increasingly advocated to rapidly evaluate patients with severe or multiple trauma to identify all potential life-threatening injuries, facilitating definite treatment despite increased radiation exposure. However, interpreting the WBCT in such cases is time-consuming and highly specialized, even for radiologists or trauma surgeons.

Chest CT is the gold standard for identifying lung, heart, great vessel, and rib cage injuries. The Medical Image Computing and Computer-Assisted Intervention (MICCAI) Society sparked researchers' interest by launching a RibFrac challenge in 2020, providing a dataset that includes around 5,000 rib fractures from 660 CT scans[71]. Several works followed by using private or multicenter datasets[72–75]. Choi et al. use 3D CNN to segment contused lung parenchyma and evaluate the severity among rib fracture patients, while Driezin et al. used nn-Unet to achieve a mean Dice score of 0.67. The quantification results were both correlated with clinical outcomes. Regarding major vessel injuries, Harris et al. developed a 5-layer CNN to binary classify scans for aortic dissection and rupture, including traumatic aortic injury, achieving sensitivities and specificities of 87.8% and 96.0% for aortic dissection and 100% and 96.0% for aortic rupture, respectively[76]. Abdominal CT contains numerous hollow digestive and solid organs, including the stomach, intestine, liver, spleen, pancreas, kidneys, etc. Cheng et al. utilize 3D CNN and a two-step algorithm to identify splenic injuries, resulting in an AUROC of 0.901 with accuracy, sensitivity, and specificity of 0.88, 0.81, and 0.92, respectively[77]. However, the annotations were performed by a single trauma surgeon, which may lead to bias due to the high inconsistency in AAST injury grading among clinicians. Chen et al. constructed a four-part algorithm pipeline to automate AAST splenic injury grading[78]. The classification agreement (weighted κ) between automated and consensus AAST grades was 0.79, which is a high performance regarding the discrepancies of blunt splenic injuries AAST grading agreement[79]. Dreizin et al. first utilized a combination of attentional networks and a dilated-CNN segmentation algorithm for CT volumetry on liver parenchyma, followed by decision tree analysis, to predict major arterial injury[80]. Farzaneh et al. also employed a simple U-Net approach to quantify liver injury[81].

CT evaluation for the pelvis accurately assesses stability, severity of pelvic fractures, and associated vascular injuries, which directly impact clinical outcomes[70]. Zapaishchykova et al. combined a fracture detection Faster-RCNN model with Bayesian tile grade refinement to automatically classify translational and rotational instability, achieving AUCs of 0.833 and 0.851, respectively[82]. Alternatively, Dreizin et al. utilized a ResNeXt-50 backbone combined with an LSTM RNN network to detect translational and rotational instability, performing comparably to radiologists. The translational instability detected by the model was further associated with the need for TAE and massive transfusion[83]. The application of these systems could enable rapid recognition of patients at risk.

Quantifying the hemorrhage amount and the extent of the injury can provide insights for the clinician regarding the severity of an injured patient's condition and potential blood loss. Furthermore, these quantifications are correlated with clinical outcomes and decisions. Dreizin et al. have published a series of works in this field. They utilized the Recurrent Saliency Transformation Network for hemorrhage quantification in pelvic fracture to automatically quantify the hematoma[84] and further predict the need for massive transfusion, angioembolization, or pelvic packing, achieving an AUC of 0.83[85]. A multiscale attentional network was applied to the hemoperitoneum, achieving a DSC of 0.61 and predicting composite outcomes, including hemostatic intervention, transfusion, and in-hospital mortality, achieving an AUC of 0.86. Regarding hemothorax, they applied an ensembled nn-Unet to achieve a mean DSC of 0.75, and an AUC of 0.74 in predicting composite outcomes[86,87]. Several groups have also worked on the quantification of pulmonary contusion and subsequent outcome prediction. Choi et al. used a pretrained lung segmentation U-net model to calculate the percentage of the contused area among rib fracture patients. They found the severity associated with mechanical ventilation and prolonged hospitalization[88]. Sarkar et al. trained a nnUnet on manually labeled pulmonary contusions to automate the quantification of the lung contusion index and subsequently predict the progression of patients to acute respiratory distress syndrome, achieving an AUC of 0.7[89].

The RSNA has hosted an annual AI competition since 2017, with the 2023 edition focusing on detecting abdominal trauma[90]. This iteration featured over 4,000 CT examinations, including various abdominal injuries. The dataset, collected from 23 sites across 14 countries, includes scan-level annotations for kidney, liver, spleen, bowel, and extravasation injuries. Notably, image-level labels were specially provided for bowel and extravasation injuries. The release of this dataset is anticipated to catalyze the development of high-performance algorithms.

Extremity fractures are the most common condition in trauma patients but are less emergent except for femoral shaft fractures, open fractures, and mangled extremities that require reperfusion procedures. Nevertheless, misdiagnosis of fractures can bring consequences such as malunion, nonunion, avascular necrosis, secondary injuries, and potential lawsuits for medical providers. The vast majority of fractures can be diagnosed initially using plain radiographs of each body part. Some minor or occult fractures can only be visible in CT or MRI under clinical suspicion. Among the limb fractures, wrist fracture is the most studied[91]. Some researchers tried to develop multiregion fracture detection algorithms in plain X-ray[48] and CT[92].

Challenges and Limitations

Several challenges restrict the development of DL models in applications in trauma imaging. First, the collection of medical image datasets is constrained by concerns surrounding patient privacy, which virtually limits the availability of data for researchers, which is the fuel for developing robust DL algorithms. Second, the accurate labels of medical images are expensive and can only rely on medical specialists in specific domains. The metadata of images stored in hospital information systems is typically optimized for clinical practice rather than algorithm training. Researchers are often compelled to re-label the images to acquire high-quality annotations. Lastly, there exists a pronounced class imbalance in diseases, as normal instances vastly outnumber the abnormalities in medical practice, especially for rare diseases.

The collection of high-quality trauma image datasets is challenging. Patient privacy and data security are crucial in trauma accidents and require a strict identification process to prevent information leakage. Medical institutes may also have concerns about obligated legal and ethical regulations like the Health Insurance Portability and Accountability Act (HIPAA) when considering releasing image datasets and labels. Additionally, the epidemiology of the trauma population differs widely across regions, and the imaging protocols are also diverse between institutes, complicating the standardizing of data[93].

Reannotating collected datasets is often necessary, as clinical labels typically only include radiology reports and Digital Imaging and Communications in Medicine (DICOM) headers. While some researchers utilize natural language processing (NLP) tools to analyze radiologist reports, this approach can generate inaccuracies and add noise to the dataset[94]. Report parsing generally yields only exam-level classification, not specific location details, which is particularly limiting for series of images like multiphase CT scans. For instance, in a whole-body CT of a patient with a minor splenic injury, most slices would appear normal, but the exam-level label would still indicate an abnormality. Manually annotating each image is labor-intensive and demands medical expertise. Additionally, segmentation of the injury site, especially for organ injuries, poses a unique challenge. Unlike tumor segmentation, where the target is more defined, the injured part of an organ is often irregularly shaped, with poorly defined boundaries, complicating the contouring task. To address these annotation challenges, some researchers attempted methods like point annotation[47] or generating pseudo labels[95], but a generalizable solution for these issues has yet to be established.

Most DL models are typically constructed using balanced datasets, meaning a nearly equal distribution of positive and negative cases, simplifying the model training and evaluation. In contrast, clinical scenarios often presented significant class imbalance, even in a targeted population[96]. Exams with specific positive findings are relatively rare, especially in some modalities used for screening. Consequently, even a classification model that achieved accuracy as high as 95% in a balanced dataset may still produce a large number of false positive cases when deployed in the real world, which causes alarm fatigue to clinicians if the specificity remains imperfect.

Models built on local datasets often suffer from overfitting to the training data, limiting their ability to generalize to images from other sources. While medical images of the same type are theoretically standardized under the DICOM format, variations in demographic distribution, equipment manufacturers, acquisition protocols, operators, and artifacts unique to different institutions can diminish model performance. Investigators have attempted to enhance the robustness of their models to new image sources by using generated images or applying various augmentations. Yet, the most effective solution to this issue relies on expanding the dataset scale to include diverse sources[20].

Several emergency-related DL-based products have received U.S. FDA approval for clinical use, but not all have peer-reviewed solid evidence[97]. Various multi-reader, multi-case studies with clinical benefit analysis are ongoing for commercially available AI tools. These tools include detection systems for brain hemorrhage, cervical spine fractures, compression fractures, and rib fractures in CT scans, pneumothorax in CXR, and extremity/junctional zone fracture identification in plain X-rays. A survey showed that more than half of American Society of Emergency Radiology members use AI tools in their practice[1]. While some products have been distributed regionally, global distribution remains limited due to legal constraints in each country and variations in medical practice. Additionally, the requirements for computational power and software engineering requirements pose significant challenges in transitioning from the laboratory to practical application. For data scientists working with medical images, the primary focus often lies in maximizing the model's performance, sometimes without regard to the computational resources required. However, designing a commercially available computer-assisted diagnosis product involves much more complex work and usually requires a team with diverse skill sets. Software engineers must prioritize data privacy, computational efficiency, a user-friendly interface, workflow integration, and cost-effectiveness in clinical settings.

Future Prospects and Trends

Even if the path ahead is filled with thorns, the road to AI remains bright. Several future directions may overcome those hardships in developing DL models in trauma imaging. Federated learning is one of the solutions that allows numerous collaborative teams to train the same global model across different servers without sharing their local data[93]. The advantage of this approach is that it can overcome the concern about data privacy and security and allows the pooling of diverse datasets for robust model training. Open datasets released from the institute may also stimulate the computer science society to focus on a specific topic and maximize its value. In our experience, most research advances are followed by the release of large-scale data on a specific topic.

Further, the explainability of DL models is essential for clinicians to build their trust while integrating AI tools into clinical workflow. The advancement of visualization techniques enhances the interpretability of AI decisions[98]. Some classification works can only present the prediction probability or result rather than correctly localize the lesion. Cheng et al. applied Grad-CAM to the hip fracture detection model and surprisingly found it could accurately localize the lesion site with heatmaps[42]. However, the heatmap can only highlight the area the algorithm relies on to determine the class without meaningful anatomical rationale. Although CAM-like approaches are practical in some topics as a weakly supervised approach, they often fail when the dataset is small, or the discrepancy between classes is less apparent.

Recent advances in developing multimodal DL models make integrating various data types possible[99]. Combining language and image models may enhance understanding of the correlation between radiologists' reports and images. Furthermore, it will be possible to incorporate additional clinical information, such as demographic data, laboratory results, physical examinations, and other modalities. Predicting patient outcomes using images represents another potential trend. The imaging itself reflects the severity of injury in anatomical structures. By predicting potential hemorrhage amounts and tissue damage, healthcare providers can make more informed decisions regarding treatment strategies and resource allocation, aligning with the modern concept of precision medicine[100]. Implementing DL in clinical practice as a computer-aided diagnosis system may change management paradigms by facilitating trauma diagnosis. Timely management of trauma patients begins with the identification of all lesions, which is crucial in a chaotic trauma scene. AI systems can assist clinicians in rapidly identifying injuries that require intervention, especially when specialist resources are scarce, thereby improving patient outcomes and healthcare quality.

Conclusions

The application of DL in trauma imaging has witnessed substantial advancements in recent years. Researchers achieved promising results across various domains, although certain areas remain in the developmental phase. Nonetheless, significant challenges persist, including data acquisition, privacy concerns, and the extensive efforts required for labeling. The cumulative effort in model development will ultimately enable AI products for trauma imaging to integrate into our clinical practice in the future, facilitating the rapid and accurate identification of life-threatening injuries. Healthcare staff should embrace this new technology, making diagnostics more precise, efficient, and patient-centric.

Conflicts of interest

The authors have no financial conflicts of interest.

Acknowledgments

This work was supported by Chang Gung Memorial Hospital under Grant CMRPG3K2141, CDRPG3M0031 and a grant from the National Science and Technology Council (Ministry of Science and Technology) MOST 110-2221-E-182A-005-, MOST-111-2221-E-182A-010-, and NSTC 112-2314-B-182A-053 -MY3

References

[1] Agrawal A, Khatri GD, Khurana B, Sodickson AD, Liang Y, Dreizin D. A survey of ASER members on artificial intelligence in emergency radiology: trends, perceptions, and expectations. Emerg Radiol 2023;30:267–77.
[2] Chen X, Wang X, Zhang K, Fung K-M, Thai TC, Moore K, et al. Recent advances and clinical applications of deep learning in medical image analysis. Med Image Anal 2022;79:102444.
[3] Inkeaw P, Angkurawaranon S, Khumrin P, Inmutto N, Traisathit P, Chaijaruwanich J, et al. Automatic hemorrhage segmentation on head CT scan for traumatic brain injury using 3D deep learning model. Comput Biol Med 2022;146:105530.
[4] Wasserthal J, Meyer M, Breit H-C, Cyriac J, Yang S, Segeroth M. TotalSegmentator: robust segmentation of 104 anatomical structures in CT images. arXiv [eessIV] 2022.
[5] Dreizin D, Staziaki PV, Khatri GD, Beckmann NM, Feng Z, Liang Y, et al. Artificial intelligence CAD tools in trauma imaging: a scoping review from the American Society of Emergency Radiology (ASER) AI/ML Expert Panel. Emerg Radiol 2023;30:251–65.
[6] Strigl D, Kofler K, Podlipnig S. Performance and Scalability of GPU-Based Convolutional Neural Networks. 2010 18th Euromicro Conference on Parallel, Distributed and Network-based Processing, IEEE; 2010, p. 317–24.
[7] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM 2017;60:84–90.
[8] Kim HE, Cosa-Linan A, Santhanam N, Jannesari M, Maros ME, Ganslandt T. Transfer learning for medical image classification: a literature review. BMC Med Imaging 2022;22:69.
[9] Chen L, Li S, Bai Q, Yang J, Jiang S, Miao Y. Review of Image Classification Algorithms Based on Convolutional Neural Networks. Remote Sensing 2021;13:4712.
[10] Liu X, Song L, Liu S, Zhang Y. A Review of Deep-Learning-Based Medical Image Segmentation Methods. Sustain Sci Pract Policy 2021;13:1224.
[11] Kaur R, Singh S. A comprehensive review of object detection with deep learning. Digit Signal Process 2023;132:103812.
[12] Zhou Y, Dreizin D, Wang Y, Liu F, Shen W, Yuille AL. External Attention Assisted Multi-Phase Splenic Vascular Injury Segmentation With Limited Data. IEEE Trans Med Imaging 2022;41:1346–57.
[13] Hicks SA, Strümke I, Thambawita V, Hammou M, Riegler MA, Halvorsen P, et al. On evaluation metrics for medical applications of artificial intelligence. Sci Rep 2022;12:5979.
[14] Padilla R, Passos WL, Dias TLB, Netto SL, da Silva EAB. A Comparative Analysis of Object Detection Metrics with a Companion Open-Source Toolkit. Electronics 2021;10:279.
[15] Avesta A, Hossain S, Lin M, Aboian M, Krumholz HM, Aneja S. Comparing 3D, 2.5D, and 2D Approaches to Brain Image Auto-Segmentation. Bioengineering (Basel) 2023;10.

https://doi.org/10.3390/bioengineering10020181.
[16] Cai J, Yan K, Cheng CT, Xiao J, Liao CH, Lu L. Deep volumetric universal lesion detection using light-weight pseudo 3d convolution and surface point regression. Conference on Medical … 2020.
[17] Lavin A, Gilligan-Lee CM, Visnjic A, Ganju S, Newman D, Ganguly S, et al. Technology readiness levels for machine learning systems. Nat Commun 2022;13:6039.
[18] Brady AP. Error and discrepancy in radiology: inevitable or avoidable? Insights Imaging 2017;8:171–82.
[19] Willemink MJ, Koszek WA, Hardell C, Wu J, Fleischmann D, Harvey H, et al. Preparing Medical Imaging Data for Machine Learning. Radiology 2020;295:4–15.
[20] Montagnon E, Cerny M, Cadrin-Chênevert A, Hamilton V, Derennes T, Ilinca A, et al. Deep learning workflow in radiology: a primer. Insights Imaging 2020;11:22.
[21] Cheng C-Y, Chiu I-M, Hsu M-Y, Pan H-Y, Tsai C-M, Lin C-HR. Deep Learning Assisted Detection of Abdominal Free Fluid in Morison's Pouch During Focused Assessment With Sonography in Trauma. Front Med 2021;8:707437.
[22] Lin Z, Li Z, Cao P, Lin Y, Liang F, He J, et al. Deep learning for emergency ascites diagnosis using ultrasonography images. J Appl Clin Med Phys 2022;23:e13695.
[23] Leo MM, Potter IY, Zahiri M, Vaziri A, Jung CF, Feldman JA. Using Deep Learning to Detect the Presence and Location of Hemoperitoneum on the Focused Assessment with Sonography in Trauma (FAST) Examination in Adults. J Digit Imaging 2023;36:2035–50.
[24] Chiu I-M, Lin C-HR, Yau F-FF, Cheng F-J, Pan H-Y, Lin X-H, et al. Use of a Deep-Learning Algorithm to Guide Novices in Performing Focused Assessment With Sonography in Trauma. JAMA Netw Open 2023;6:e235102.
[25] Taye M, Morrow D, Cull J, Smith DH, Hagan M. Deep Learning for FAST Quality Assessment. J Ultrasound Med 2023;42:71–9.
[26] Yıldız Potter İ, Leo MM, Vaziri A, Feldman JA. Automated detection and localization of pericardial effusion from point-of-care cardiac ultrasound examination. Med Biol Eng Comput 2023;61:1947–59.
[27] Kotei E, Thirunavukarasu R. Visual attention condenser model for multiple disease detection from heterogeneous medical image modalities. Multimed Tools Appl 2023. https://doi.org/10.1007/s11042-023-16625-x.
[28] Kotei E, Thirunavukarasu R. Ensemble Technique Coupled with Deep Transfer Learning Framework for Automatic Detection of Tuberculosis from Chest X-ray Radiographs. Healthcare (Basel) 2022;10. https://doi.org/10.3390/healthcare10112335.
[29] Kotei E, Thirunavukarasu R. A Comprehensive Review on Advancement in Deep Learning Techniques for Automatic Detection of Tuberculosis from Chest X-ray Images. Arch Comput Methods Eng 2024;31:455–74.
[30] Rodriguez RM, Hendey GW, Mower WR. Selective chest imaging for blunt trauma patients: The national emergency X-ray utilization studies (NEXUS-chest algorithm). Am J Emerg Med 2017;35:164–70.
[31] Meedeniya D, Kumarasinghe H, Kolonne S, Fernando C, Díez ID la T, Marques G. Chest X-ray analysis empowered with deep learning: A systematic review. Appl Soft Comput 2022;126:109319.
[32] Rajpurkar P, Lungren MP. The Current and Future State of AI Interpretation of Medical Images. N Engl J Med 2023;388:1981–90.
[33] Huang S-T, Liu L-R, Chiu H-W, Huang M-Y, Tsai M-F. Deep convolutional neural network for rib fracture recognition on chest radiographs. Front Med 2023;10:1178798.
[34] Ghosh A, Bose S, Patton D, Kumar I, Khalkhali V, Henry MK, et al. Deep learning-based prediction of rib fracture presence in frontal radiographs of children under two years of age: a proof-of-concept study. Br J Radiol 2023;96:20220778.
[35] Sun H, Wang X, Li Z, Liu A, Xu S, Jiang Q, et al. Automated Rib Fracture Detection on

Chest X-Ray Using Contrastive Learning. J Digit Imaging 2023;36:2138–47.
[36] Gipson J, Tang V, Seah J, Kavnoudias H, Zia A, Lee R, et al. Diagnostic accuracy of a commercially available deep-learning algorithm in supine chest radiographs following trauma. Br J Radiol 2022;95:20210979.
[37] Mutasa S, Varada S, Goel A, Wong TT, Rasiej MJ. Advanced Deep Learning Techniques Applied to Automated Femoral Neck Fracture Detection and Classification. J Digit Imaging 2020;33:1209–17.
[38] Adams M, Chen W, Holcdorf D, McCusker MW, Howe PD, Gaillard F. Computer vs human: Deep learning versus perceptual training for the detection of neck of femur fractures. J Med Imaging Radiat Oncol 2019;63:27–32.
[39] Twinprai N, Boonrod A, Boonrod A, Chindaprasirt J, Sirithanaphol W, Chindaprasirt P, et al. Artificial intelligence (AI) vs. human in hip fracture detection. Heliyon 2022;8:e11266.
[40] Bae J, Yu S, Oh J, Kim TH, Chung JH, Byun H, et al. External Validation of Deep Learning Algorithm for Detecting and Visualizing Femoral Neck Fracture Including Displaced and Non-displaced Fracture on Plain X-ray. J Digit Imaging 2021;34:1099–109.
[41] Yamada Y, Maki S, Kishida S, Nagai H, Arima J, Yamakawa N, et al. Automated classification of hip fractures using deep convolutional neural networks with orthopedic surgeon-level accuracy: ensemble decision-making with antero-posterior and lateral radiographs. Acta Orthop 2020;91:699–704.
[42] Cheng C-T, Ho T-Y, Lee T-Y, Chang C-C, Chou C-C, Chen C-C, et al. Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs. Eur Radiol 2019;29:5469–77.
[43] Urakawa T, Tanaka Y, Goto S, Matsuzawa H, Watanabe K, Endo N. Detecting intertrochanteric hip fractures with orthopedist-level accuracy using a deep convolutional neural network. Skeletal Radiol 2019;48:239–44.
[44] Krogue JD, Cheng KV, Hwang KM, Toogood P, Meinberg EG, Geiger EJ, et al. Automatic Hip Fracture Identification and Functional Subclassification with Deep Learning. Radiol Artif Intell 2020;2:e190023.
[45] Sato Y, Takegami Y, Asamoto T, Ono Y, Hidetoshi T, Goto R, et al. Artificial intelligence improves the accuracy of residents in the diagnosis of hip fractures: a multicenter study. BMC Musculoskelet Disord 2021;22:407.
[46] Cheng C-T, Chen C-C, Cheng F-J, Chen H-W, Su Y-S, Yeh C-N, et al. A Human-Algorithm Integration System for Hip Fracture Detection on Plain Radiography: System Development and Validation Study. JMIR Med Inform 2020;8:e19416.
[47] Cheng C-T, Wang Y, Chen H-W, Hsiao P-M, Yeh C-N, Hsieh C-H, et al. A scalable physician-level deep learning algorithm detects universal trauma on pelvic radiographs. Nat Commun 2021;12:1066.
[48] Jones RM, Sharma A, Hotchkiss R, Sperling JW, Hamburger J, Ledig C, et al. Assessment of a deep-learning system for fracture detection in musculoskeletal radiographs. NPJ Digit Med 2020;3:144.
[49] Patel A, Manniesing R. A convolutional neural network for intracranial hemorrhage detection in non-contrast CT. Medical Imaging 2018: Computer-Aided Diagnosis, vol. 10575, SPIE; 2018, p. 301–6.
[50] Cho J, Park K-S, Karki M, Lee E, Ko S, Kim JK, et al. Improving Sensitivity on Identification and Delineation of Intracranial Hemorrhage Lesion Using Cascaded Deep Learning Models. J Digit Imaging 2019;32:450–61.
[51] Chilamkurthy S, Ghosh R, Tanamala S, Biviji M, Campeau NG, Venugopal VK, et al. Deep learning algorithms for detection of critical findings in head CT scans: a retrospective study. Lancet 2018;392:2388–96.
[52] Jnawali K, Arbabshirani MR, Rao N, Patel AA. Deep 3D convolution neural network for CT brain hemorrhage classification. Medical Imaging 2018: Computer-Aided Diagnosis, vol.

10575, SPIE; 2018, p. 307–13.
[53] Ye H, Gao F, Yin Y, Guo D, Zhao P, Lu Y, 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.
[54] Titano JJ, Badgeley M, Schefflein J, Pain M, Su A, Cai M, et al. Automated deep-neural-network surveillance of cranial images for acute neurologic events. Nat Med 2018;24:1337–41.
[55] Lee JY, Kim JS, Kim TY, Kim YS. Detection and classification of intracranial haemorrhage on CT images using a novel deep-learning algorithm. Sci Rep 2020;10:20546.
[56] Burduja M, Ionescu RT, Verga N. Accurate and Efficient Intracranial Hemorrhage Detection and Subtype Classification in 3D CT Scans with Convolutional and Long Short-Term Memory Neural Networks. Sensors 2020;20. https://doi.org/10.3390/s20195611.
[57] Sharrock MF, Mould WA, Ali H, Hildreth M, Awad IA, Hanley DF, et al. 3D Deep Neural Network Segmentation of Intracerebral Hemorrhage: Development and Validation for Clinical Trials. Neuroinformatics 2021;19:403–15.
[58] 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;304:385–94.
[59] Flanders AE, Prevedello LM, Shih G, Halabi SS, Kalpathy-Cramer J, Ball R, et al. Erratum: Construction of a Machine Learning Dataset through Collaboration: The RSNA 2019 Brain CT Hemorrhage Challenge. Radiol Artif Intell 2020;2:e209002.
[60] Boonrod A, Boonrod A, Meethawolgul A, Twinprai P. Diagnostic accuracy of deep learning for evaluation of C-spine injury from lateral neck radiographs. Heliyon 2022;8:e10372.
[61] Naguib SM, Hamza HM, Hosny KM, Saleh MK, Kassem MA. Classification of Cervical Spine Fracture and Dislocation Using Refined Pre-Trained Deep Model and Saliency Map. Diagnostics (Basel) 2023;13. https://doi.org/10.3390/diagnostics13071273.
[62] Al Arif SMMR, Knapp K, Slabaugh G. Fully automatic cervical vertebrae segmentation framework for X-ray images. Comput Methods Programs Biomed 2018;157:95–111.
[63] Lessmann N, van Ginneken B, de Jong PA, Išgum I. Iterative fully convolutional neural networks for automatic vertebra segmentation and identification. Med Image Anal 2019;53:142–55.
[64] Cheng P, Yang Y, Yu H, He Y. Automatic vertebrae localization and segmentation in CT with a two-stage Dense-U-Net. Sci Rep 2021;11:22156.
[65] Mushtaq M, Akram MU, Alghamdi NS, Fatima J, Masood RF. Localization and Edge-Based Segmentation of Lumbar Spine Vertebrae to Identify the Deformities Using Deep Learning Models. Sensors 2022;22. https://doi.org/10.3390/s22041547.
[66] Voter AF, Larson ME, Garrett JW, Yu J-PJ. Diagnostic Accuracy and Failure Mode Analysis of a Deep Learning Algorithm for the Detection of Cervical Spine Fractures. AJNR Am J Neuroradiol 2021;42:1550–6.
[67] Ma S, Huang Y, Che X, Gu R. Faster RCNN-based detection of cervical spinal cord injury and disc degeneration. J Appl Clin Med Phys 2020;21:235–43.
[68] Brooks A, Reilly J-J, Hope C, Navarro A, Naess PA, Gaarder C. Evolution of non-operative management of liver trauma. Trauma Surg Acute Care Open 2020;5:e000551.
[69] Polireddy K, Hoff C, Kinger NP, Tran A, Maddu K. Blunt thoracic trauma: role of chest radiography and comparison with CT - findings and literature review. Emerg Radiol 2022;29:743–55.
[70] Raniga SB, Mittal AK, Bernstein M, Skalski MR, Al-Hadidi AM. Multidetector CT in Vascular Injuries Resulting from Pelvic Fractures: A Primer for Diagnostic Radiologists. Radiographics 2019;39:2111–29.
[71] Jin L, Yang J, Kuang K, Ni B, Gao Y, Sun Y, et al. Deep-learning-assisted detection and segmentation of rib fractures from CT scans: Development and validation of FracNet.

EBioMedicine 2020;62:103106.
[72] Meng XH, Wu DJ, Wang Z, Ma XL, Dong XM, Liu AE, et al. A fully automated rib fracture detection system on chest CT images and its impact on radiologist performance. Skeletal Radiol 2021;50:1821–8.
[73] Wang S, Wu D, Ye L, Chen Z, Zhan Y, Li Y. Assessment of automatic rib fracture detection on chest CT using a deep learning algorithm. Eur Radiol 2023;33:1824–34.
[74] Edamadaka S, Brown DJ, Swaroop R, Kolodner M, Spain DA, Forrester JD, et al. FasterRib: A deep learning algorithm to automate identification and characterization of rib fractures on chest computed tomography scans. J Trauma Acute Care Surg 2023;95:181–5.
[75] Tan H, Xu H, Yu N, Yu Y, Duan H, Fan Q, et al. The value of deep learning-based computer aided diagnostic system in improving diagnostic performance of rib fractures in acute blunt trauma. BMC Med Imaging 2023;23:55.
[76] Harris RJ, Kim S, Lohr J, Towey S, Velichkovich Z, Kabachenko T, et al. Classification of Aortic Dissection and Rupture on Post-contrast CT Images Using a Convolutional Neural Network. J Digit Imaging 2019;32:939–46.
[77] Cheng C-T, Lin H-S, Hsu C-P, Chen H-W, Huang J-F, Fu C-Y, et al. The three-dimensional weakly supervised deep learning algorithm for traumatic splenic injury detection and sequential localization: an experimental study. Int J Surg 2023;109:1115–24.
[78] Chen H, Unberath M, Dreizin D. Toward automated interpretable AAST grading for blunt splenic injury. Emerg Radiol 2023;30:41–50.
[79] Adams-McGavin RC, Tafur M, Vlachou PA, Wu M, Brassil M, Crivellaro P, et al. Interrater Agreement of CT Grading of Blunt Splenic Injuries: Does the AAST Grading Need to Be Reimagined? Can Assoc Radiol J 2023:8465371231184425.
[80] Dreizin D, Chen T, Liang Y, Zhou Y, Paes F, Wang Y, et al. Added value of deep learning-based liver parenchymal CT volumetry for predicting major arterial injury after blunt hepatic trauma: a decision tree analysis. Abdom Radiol (NY) 2021;46:2556–66.
[81] Farzaneh N, Stein EB, Soroushmehr R, Gryak J, Najarian K. A deep learning framework for automated detection and quantitative assessment of liver trauma. BMC Med Imaging 2022;22:39.
[82] Zapaishchykova A, Dreizin D, Li Z, Wu JY, Roohi SF, Unberath M. An Interpretable Approach to Automated Severity Scoring in Pelvic Trauma. Med Image Comput Comput Assist Interv 2021;12903:424–33.
[83] Dreizin D, Goldmann F, LeBedis C, Boscak A, Dattwyler M, Bodanapally U, et al. An Automated Deep Learning Method for Tile AO/OTA Pelvic Fracture Severity Grading from Trauma whole-Body CT. J Digit Imaging 2021;34:53–65.
[84] Dreizin D, Zhou Y, Zhang Y, Tirada N, Yuille AL. Performance of a Deep Learning Algorithm for Automated Segmentation and Quantification of Traumatic Pelvic Hematomas on CT. J Digit Imaging 2020;33:243–51.
[85] Dreizin D, Zhou Y, Chen T, Li G, Yuille AL, McLenithan A, et al. Deep learning-based quantitative visualization and measurement of extraperitoneal hematoma volumes in patients with pelvic fractures: Potential role in personalized forecasting and decision support. J Trauma Acute Care Surg 2020;88:425.
[86] Dreizin D, Nixon B, Hu J, Albert B, Yan C, Yang G, et al. A pilot study of deep learning-based CT volumetry for traumatic hemothorax. Emerg Radiol 2022. https://doi.org/10.1007/s10140-022-02087-5.
[87] Dreizin D, Zhou Y, Fu S, Wang Y, Li G, Champ K, et al. A Multiscale Deep Learning Method for Quantitative Visualization of Traumatic Hemoperitoneum at CT: Assessment of Feasibility and Comparison with Subjective Categorical Estimation. Radiol Artif Intell 2020;2:e190220.
[88] Choi J, Mavrommati K, Li NY, Patil A, Chen K, Hindin DI, et al. Scalable deep learning

algorithm to compute percent pulmonary contusion among patients with rib fractures. J Trauma Acute Care Surg 2022;93:461–6.
[89] Sarkar N, Zhang L, Campbell P, Liang Y, Li G, Khedr M, et al. Pulmonary contusion: automated deep learning-based quantitative visualization. Emerg Radiol 2023;30:435–41.
[90] Errol Colak, Hui-Ming Lin, Robyn Ball, Melissa Davis, Adam Flanders, Sabeena Jalal, Kirti Magudia, Brett Marinelli, Savvas Nicolaou, Luciano Prevedello, Jeff Rudie, George Shih, Maryam Vazirabad, John Mongan. RSNA 2023 Abdominal Trauma Detection 2023.
[91] Hansen V, Jensen J, Kusk MW, Gerke O, Tromborg HB, Lysdahlgaard S. Deep learning performance compared to healthcare experts in detecting wrist fractures from radiographs: A systematic review and meta-analysis. Eur J Radiol 2024;174:111399.
[92] Liu Y, Liu W, Chen H, Xie S, Wang C, Liang T, et al. Artificial intelligence versus radiologist in the accuracy of fracture detection based on computed tomography images: a multi-dimensional, multi-region analysis. Quant Imaging Med Surg 2023;13:6424–33.
[93] Kaissis GA, Makowski MR, Rückert D, Braren RF. Secure, privacy-preserving and federated machine learning in medical imaging. Nature Machine Intelligence 2020;2:305–11.
[94] Casey A, Davidson E, Poon M, Dong H, Duma D, Grivas A, et al. A systematic review of natural language processing applied to radiology reports. BMC Med Inform Decis Mak 2021;21:179.
[95] Mao J, Yin X, Zhang G, Chen B, Chang Y, Chen W, et al. Pseudo-labeling generative adversarial networks for medical image classification. Comput Biol Med 2022;147:105729.
[96] Gao L, Zhang L, Liu C, Wu S. Handling imbalanced medical image data: A deep-learning-based one-class classification approach. Artif Intell Med 2020;108:101935.
[97] van Leeuwen KG, Schalekamp S, Rutten MJCM, van Ginneken B, de Rooij M. Artificial intelligence in radiology: 100 commercially available products and their scientific evidence. Eur Radiol 2021;31:3797–804.
[98] Reyes M, Meier R, Pereira S, Silva CA, Dahlweid F-M, von Tengg-Kobligk H, et al. On the Interpretability of Artificial Intelligence in Radiology: Challenges and Opportunities. Radiol Artif Intell 2020;2:e190043.
[99] Thirunavukarasu R, C GPD, R G, Gopikrishnan M, Palanisamy V. Towards computational solutions for precision medicine based big data healthcare system using deep learning models: A review. Comput Biol Med 2022;149:106020.
[100] Collins FS, Varmus H. A new initiative on precision medicine. N Engl J Med 2015;372:793–5.