European Journal of Trauma and Emergency Surgery (2025) 51:119
https://doi.org/10.1007/s00068-025-02793-y

ORIGINAL ARTICLE

Integrating blockchain technology with artificial intelligence for the diagnosis of tibial plateau fractures

Yi Xie[1,2] · Xiaoliang Chen[3] · Huiwen Yang[2,4] · Honglin Wang[1,2] · Hong Zhou[1,2] · Lin Lu[5] · Jiayao Zhang[6] · Pengran Liu[1,2] · Zhewei Ye[1,2]

Received: 17 December 2024 / Accepted: 6 February 2025

Abstract

Purpose The application of artificial intelligence (AI) in healthcare has seen widespread implementation, with numerous studies highlighting the development of robust algorithms. However, limited attention has been given to the secure utilization of raw data for medical model training, and its subsequent impact on clinical decision-making and real-world applications. This study aims to assess the feasibility and effectiveness of an advanced diagnostic model that integrates blockchain technology and AI for the identification of tibial plateau fractures (TPFs) in emergency settings.

Method In this study, blockchain technology was utilized to construct a distributed database for trauma orthopedics, images collected from three independent hospitals for model training, testing, and internal validation. Then, a distributed network combining blockchain and deep learning was developed for the detection of TPFs, with model parameters aggregated across multiple nodes to enhance accuracy. The model's performance was comprehensively evaluated using metrics including accuracy, sensitivity, specificity, F1 score, and the area under the receiver operating characteristic curve (AUC). In addition, the performance of the centralized model, the distributed AI model, clinical orthopedic attending physicians, and AI-assisted attending physicians was tested on an external validation dataset.

Results In the testing set, the accuracy of our distributed model was 0.9603 [95% CI (0.9598, 0.9605)] and the AUC was 0.9911 [95% CI (0.9893, 0.9915)] for TPF detection. In the external validation set, the accuracy reached 0.9636 [95% CI (0.9388, 0.9762)], was slightly higher than that of the centralized YOLOv8n model at 0.9632 [95% CI (0.9387, 0.9755)] ($p > 0.05$), and exceeded the orthopedic physician at 0.9291 [95% CI (0.9002, 0.9482)] and radiology attending physician at 0.9175 [95% CI (0.8891, 0.9393)], with a statistically significant difference ($p < 0.05$). Additionally, the centralized model (4.99 ± 0.01 min) had shorter diagnosis times compared to the orthopedic attending physician (25.45 ± 1.92 min) and the radiology attending physician (26.21 ± 1.20 min), with a statistically significant difference ($p < 0.05$).

Conclusion The model based on the integration of blockchain technology and AI can realize safe, collaborative, and convenient assisted diagnosis of TPF. Through the aggregation of training parameters by decentralized algorithms, it can achieve model construction without data leaving the hospital and may exert clinical application value in the emergency settings.

Keywords Trauma · Tibial plateau fractures · Blockchain · Artificial intelligence

Abbreviations

AI	Artifcial intelligence
TPF	Tibial plateau fractures
CI	Confdence interval
CT	Computed tomography
MRI	Magnetic resonance imaging
DICOM	Digital imaging and communications in medicine
ROC	Receiver operating characteristic
AUC	Area under curve
YOLO	You Only Look Once
R-CNN	Region-based convolutional neural networks

Yi Xie, Xiaoliang Chen and Huiwen Yang: Equal contribution.

Extended author information available on the last page of the article

Introduction

Traumatic fractures, characterized by violent external trauma and involving bone joints, are highly disabling, difficult to diagnose and treat, and often have a poor long-term prognosis, posing a significant threat to both the physical

Published online: 21 February 2025

and mental health of patients [1]. Patients with traumatic fractures often present with joint dislocations, open fractures, multiple fractures, and may experience complications such as hemorrhagic shock, fat embolism syndrome, combined injuries to vital organs, spinal cord injury, osteofascial compartment syndrome, and infections [2]. Tibial plateau fracture (TPF), as a common type of lower limb traumatic fractures, can have varying degrees of compression and displacement of the joint surface. Inadequate treatment can result in lower limb dysfunction and disability. TPF accounts for approximately 1% of all fractures in human body and 8% of fractures in the elderly [1–3]. Generally, patients with knee joint trauma typically undergo the initial examination in the emergency department. During the diagnosis and treatment process, X-ray examination, compared to computed tomography (CT) and magnetic resonance imaging (MRI), offers advantages such as rapid localization of the fracture site and type, lower costs, and easier interpretation. X-ray plays an important role in the initial screening of trauma fracture patients. In the emergency departments of remote and underdeveloped regions, medical resources are often limited, and healthcare providers in communities may lack the experience and confidence to accurately identify and evaluate fractures based on X-ray images, particularly for occult fractures and fractures of the posterior column of the tibial plateau [3, 4]. The missed fracture diagnosis occurs when a fracture is not identified during the initial examination, delaying necessary treatment. This can hinder the healing process, extend recovery time, and potentially compromise the healing quality. Patients with undiagnosed fractures may continue to place weight on the affected area, which can lead to complications such as displacement, malunion, or osteoporosis. These issues can result in long-term functional impairments, including limited joint mobility, chronic pain, or reduced physical ability. Furthermore, delayed diagnosis increases the risk of serious complications. In cases of open fractures, untreated fractures can lead to infections or osteomyelitis, with delayed treatment heightening the likelihood of such risks. Therefore, prompt and accurate fracture identification is crucial for effective treatment and optimal recovery outcomes.

With the rapid advancement of information technology, artificial intelligence (AI), computer navigation, big data, and blockchain are continuously evolving, driving the digital transformation of medicine [5–7]. Among these technologies, AI, with its capabilities in feature engineering, artificial neural networks, and deep learning (DL), is increasingly being applied to the diagnosis of orthopedic diseases [8–10]. AI-powered imaging analysis tools have become instrumental in diagnosing conditions such as fractures, osteoarthritis, and spinal disorders. These tools assist radiologists and orthopedic surgeons by highlighting areas of concern and providing preliminary diagnostic suggestions, thereby improving decision-making and streamlining the treatment process. Lindsey and Russe reduced the risks of missed and misdiagnosed radial fractures by using deep learning model [11, 12]. Oakden et al. constructed a model based on 400 X-ray films to detect proximal femoral fractures, and its performance was comparable to that of radiologists [13]. As a critical load-bearing structure of the knee joint, the tibial plateau is susceptible to damage in traumatic incidents. Accurate and fast positioning diagnosis of subtle fractures, occult fractures, and the intercondylar eminence fractures of tibial plateau can be challenging for physicians with less experience. AI can assist clinicians in accurately locating fracture sites and performing precise image segmentation, thereby facilitating rapid preliminary screening and diagnosis of TPF, ultimately saving valuable time [14, 15]. The current training of AI models typically encounters problems such as the complexity of constructing specialized disease databases, small sample sizes, the susceptibility of privacy information leakage, and the difficulties in multi-center and cross-regional collaboration [16–18]. As a decentralized system, blockchain connects data blocks via a distributed ledger and utilizes consensus algorithms for ledger updates and encryption methods to ensure secure data transmission. Blockchain technology can effectively address data security challenges encountered during the establishment of orthopedic databases and model training. It functions without a central authority, enabling direct transactions between users in a peer-to-peer network. Once healthcare information is recorded, it becomes immutable and irreversible, ensuring the integrity of the data. In blockchain, individual distributed network are linked through a chain structure and broadcast business processes across the system. The nodes in blockchain can be thought of as a block on the blockchain network, and the interconnected blocks make up the nodes in the blockchain [19, 20]. Between hospital information systems, blockchain can help provide incentives to encourage devices to participate in model training, which is increasingly crucial for interactive decentralized deep learning platforms.

Blockchain offers several robust features that make it a compelling solution for safeguarding the privacy of sensitive healthcare information:

- Decentralization: Blockchain operates as a distributed ledger system, removing the need for a central authority to manage transactions. Unlike centralized systems, it enables medical data exchanges without intermediaries, thus reducing the risks associated with single points of failure common in traditional server-based systems.
- Traceability: By utilizing a time-stamped chain block structure, blockchain introduces a temporal aspect to data storage, significantly improving its verifiability and traceability. This allows for transparent tracking of data

history and access. The traceability of blockchain can help each participating node unit or researcher reward the contribution of local model updates based on incentive policies, such as online annotation of medical imaging data, data provision, and algorithm contributions.

- Immutability: Any attempt to alter a transaction would result in a different hash value, linking the current block to the next, making such changes detectable by other nodes running the same consensus algorithm. As blockchain operates as a distributed ledger, stored across thousands of nodes and synchronized in real-time, an attack would require more than 51% of the network's computing power to succeed, making tampering highly unlikely.
- Security: Blockchain incorporates economic incentives to ensure that all nodes in the distributed network participate in the verification of data blocks. New blocks are added through a consensus algorithm that selects specific nodes. Data is encrypted using asymmetric cryptography, and the robust computing power generated by consensus algorithms, such as proof of work, enhances the system's ability to resist external attacks. This makes blockchain highly secure, ensuring that data cannot be tampered [21].

Blockchain systems are built on key technologies such as smart contracts and Hyperledger, which are crucial for enabling a variety of applications. In the blockchain framework, smart contracts originate as patient-owned data elements, serving as the cornerstone for secure and transparent information exchange. Data owners have the capability to selectively grant or revoke access to individual data elements for various stakeholders. AI tools can interface with smart contracts, and these tools can also be integrated into the blockchain ecosystem. These analytical resources, including databases and clinical trials, are not confined to any single silo but are distributed across the network. Incentives can be allocated based on data ownership, the significance of the data to each process, or the value added by providers, healthcare entities, or consumer firms [22]. Hyperledger, a collection of open-source projects supported by the Linux Foundation, promotes collaboration between businesses and developers, creating an environment that fosters innovation. In blockchain networks, personal health records can be shared across multiple nodes within a peer-to-peer system. The source node generates a data file, which is broadcast across the network for validation. Validated transactions are then grouped into blocks, each updating the blockchain's state. Every node keeps a record of past transactions, preserving the historical integrity of the ledger. Once a block is validated, it is added to the chain, extending the sequence of previous blocks. To ensure privacy and security, blockchain utilizes two types of cryptography along with a hash function, maintaining data integrity. Distributed AI models further promote collaboration among healthcare institutions through parameter aggregation, enhancing both the scale of data and the stability of algorithms [23, 24].

Presently, several studies have investigated blockchain-based architectures for medical analysis, however, no research has yet examined the application of these technologies to fracture AI models. Therefore, we hypothesize that the integration of blockchain and AI could transform traditional methods of model training in the field of orthopedics and data sharing between hospitals. The purpose of this study was to propose the development of a specialized disease database and a fracture diagnosis algorithm for trauma orthopedics, leveraging blockchain technology. Additionally, we aimed to integrate the advanced distributed AI model to assist clinical emergency physicians in the rapid diagnosis of traumatic TPFs. This study not only compares the performance of the blockchain-based AI model with traditional diagnostic methods but also evaluates its impact on healthcare professionals' decision-making processes.

Material and methods

Ethical approval

This study has been approved by the Medical Ethics Committee of Union Hospital of Tongji Medical College of Huazhong University of Science and Technology (IRB No. 0840). Data collection conformed to clinical ethical practices and adhered to the ethical guidelines established in the Helsinki Declaration. The study was registered in the Chinese Clinical Trial Registry and obtained the registration number (ChiCTR2300070658). Relevant measures were taken to ensure that the study met ethical and regulatory standards and protected the privacy of the data.

Training data set

We have implemented the blockchain system across four distinct hospitals and integrated an interface within the client application of the imaging system to facilitate smart contract authorization for exporting orthopedic data. The collected preoperative and postoperative imaging data included X-ray films (both anteroposterior and lateral views of the knee joint), plain radiographs, and three-dimensional reconstructions. The inclusion criteria for tibial plateau fracture cases were as follows: (1) The patient's age was ≥ 18 years. (2) Patients with a clear diagnosis of TPFs, without fractures in other parts of the knee joint, such as distal femoral or proximal tibiofibular fractures, osteoarthritis, chronic osteomyelitis, and bone tumors. (3) Standard anteroposterior X-ray films without improper positioning, excessive exposure, artifacts, or obstructions. The inclusion criteria for non-fracture cases in the control group were standard anteroposterior X-ray films with no fractures in the knee joint. Additionally, some injuries of the tibial plateau in MRI manifested as bone marrow edema signals, which were also considered tibial plateau injuries [3]. In the image data preprocessing phase, we engaged four senior trauma physicians to independently evaluate and provide assessments of the images. In instances where discrepancies occurred in radiological interpretations, a detailed analysis of the diagnostic conclusions from two orthopedists, each with over 20 years of experience, was conducted to confirm the presence of TPFs. Subsequently, five senior attending physicians employed image annotation software to manually delineate the fracture areas of the tibial plateau in trauma patients, thereby generating regions of interest (as illustrated in Fig. 1) and extracting radiomics features from X-ray images.

Fracture detection model

This study employs the You Only Look Once (YOLO) algorithm, which features a simple network structure, rapid speed, and superior real-time performance in detecting fractures. The network structure consists of the input, the backbone, the neck, and the head network [25–27]. In this study, the original fracture imaging images was re-adjusted to the standard size before being input into the network. At the input end, the basic input is set to an image format of 640*640. Commonly, directly adopting the stretching method might potentially lead to imbalance and distortion of the target image proportion. This study utilizes the proportional scaling approach to effectively prevent detection image distortion and employs letterbox for scaling the input images to ensure the maximum restoration of the images. The algorithm addresses issues such as image distortion caused by region cropping and scaling, mitigates the redundancy in feature extraction by convolutional neural networks, significantly accelerates the generation of candidate boxes, and reduces computational costs. During training, the model's performance was evaluated on the validation set at the end of each epoch to monitor convergence and prevent overfitting. We utilized grid search or random search methods to explore the hyperparameter space, selecting the combination of hyperparameters that yielded the best performance on the validation set for final model training and evaluation. The fixed validation set strategy effectively reduces experimental complexity and minimizes computational resource consumption.

Distributed model based on the integration of blockchain and deep learning

In applying the AI model for traumatic orthopedics, blockchain technology can be leveraged to deploy AI algorithms. This approach integrates external expert knowledge and decision-making tools, thereby enhancing the accuracy and efficiency of diagnosing and treating traumatic orthopedic conditions. By leveraging a specialized disease database and a traceability system on the blockchain, an intelligent digital closed-loop treatment system can be established [28, 29]. This system assists physicians in medical history sharing, trauma identification, preoperative surgical design, intraoperative navigation, and postoperative risk prediction. In practice, key technologies such as secure sandbox environments, decryption, and summary generation are integrated into the modern hospital information system to support diagnostic and therapeutic decisions and quality evaluations [30]. This provides technical support for improving trauma treatment levels in multi-center collaborations and scenarios with controlled sensitive information. Trauma orthopedics and imaging professionals access patient health data through the blockchain system. After submitting a request, the smart contract verifies the node signature and forwards it. Upon authorization, the requested entity verifies the signature and returns the result and verification data via smart contract, ensuring traceability and security. When the AI-assisted diagnosis is required, the trauma orthopedic surgeon submits the request to the server through the embedded intelligent system, obtaining computational results and

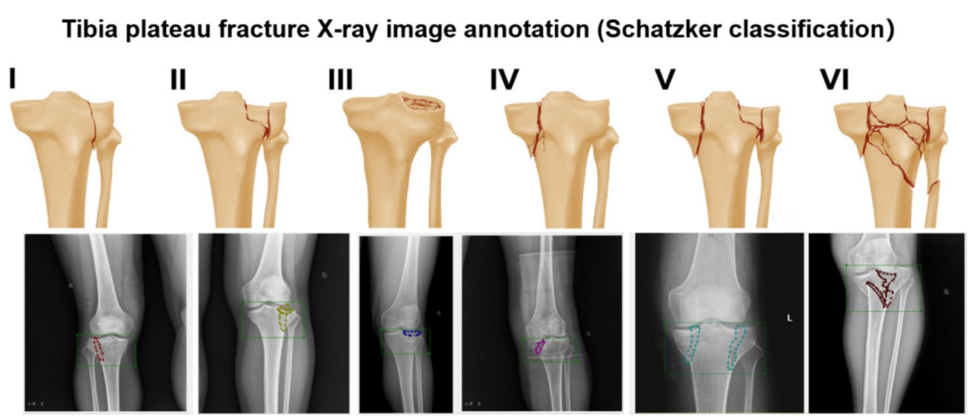


Fig. 1 Image annotation of TPFs

reference opinions. Regarding patient information security and privacy, blockchain-based AI model training adopts distributed data storage, point-to-point transmission, consensus mechanisms, and encryption algorithms, realizing a decentralized structure for medical data sharing, protecting patient data storage, transmission, and traceability [31, 32]. The AI model, collaboratively trained by the trauma orthopedics departments of various hospitals via a blockchain system, is deployed on cloud servers. When AI-assisted diagnosis are required, hospital physicians submit consultation requests. Permission control is executed through smart contracts. Upon authorization to access the model, the system forwards the node unit to the AI model for auxiliary decision-making and subsequently returns the clinical report results.

To validate the potential use of distributed AI model in fracture detection, we designed an experiment with a special focus on the clinic application in emergency care units. In this study, blockchain simulation nodes from three distinct hospitals in China were established. Independent servers were utilized to configure deep learning tasks through the platform, and key information such as participants, datasets, deep learning algorithms, and initial parameters was determined [26]. As depicted in Fig. 2, by employing privacy computing networks to distribute participating member nodes and uploading the basic information of the training task to the blockchain, the creation of the training task was accomplished. Once the participants obtained the task information, they could invoke the interface services of the DL module on the computing engines of each node. Regarding encrypted computation and parameter decryption, the DL module in this study parsed the task configuration, responded to requests, and asynchronously initiated the local computation tasks. The local training processes were launched based on the selected model, aggregation algorithm, and encryption scheme. The blockchain node was responsible for synchronizing the model's calculation status. Meanwhile, a smart contract was deployed on the consortium blockchain, randomly selecting one participant as the model aggregator for the current training round and disseminating this information to all member nodes. The model aggregator then initiated the aggregation process and made the aggregated model available to all participating nodes for further communication and scheduling [27]. Member nodes shared their local updated parameters with the model aggregator through the privacy computing network based on the single encoding aggregation algorithm. During the DL model training process, the task status and training metrics were synchronously and real-time updated to the blockchain. When the training ceased, the participating nodes shared the training parameters of the deep learning model. Throughout the entire training process, the member's data resources remained within their domains, reducing the risk of data leakage [28]. The jointly trained model exhibited higher accuracy compared to models trained with a single data source.

The blockchain-based AI model parameter updating and aggregation process is implemented through smart contracts. These contracts receive model parameters from the distributed ledger, aggregate them, and transmit the updated parameters back to the respective client-side ledgers. In this study, the distributed modeling process involves three trauma orthopaedics at different hospitals. Each hospital

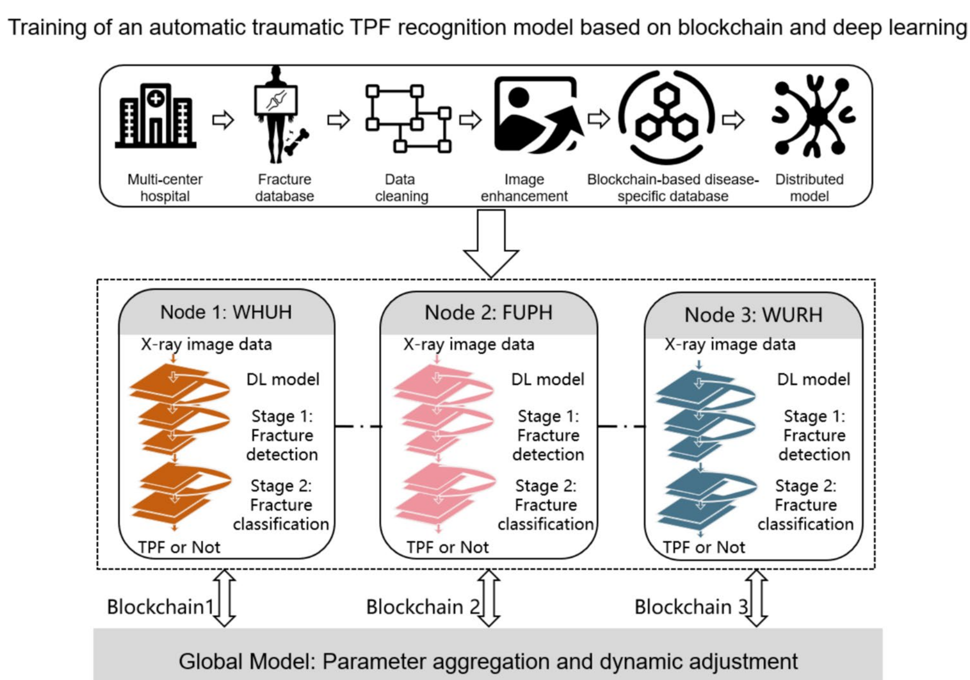


Fig. 2 The construction of the distributed model integrating blockchain and DL. *WHUH* Wuhan Union Hospital, *FUPH* Fuzhou University Affiliated Provincial Hospital, *WURH* Renmin Hospital of Wuhan University, *DL* deep learning

trains the model using its local fracture data and, at the end of each round, sends the updated model parameters, including weights and biases, back to the server for aggregation [30–32]. Once these parameters are aggregated, they are transmitted back to the nodes in different medical institution, where the models are updated with the new parameters. This iterative process continues until a predetermined number of rounds is reached. Smart contracts are utilized to store and aggregate both global and local updates [33, 34]. The smart contracts in this study were developed using the Solidity programming language and implement two primary processes: one for aggregating biases and another for aggregating weights.

The distributed DL model was executed on three AMAX GPU servers. Each server was configured with two Intel(R) Xeon(R) Gold 6226R CPUs (2.90 GHz), 24 TB HDD, 256 GB RAM, and four NVIDIA Tesla V100S GPUs (for detailed specifications, see: https://developer.nvidia.com/). The specific code and implementation procedures are provided in the Supplementary Material.

Test and diagnosis

In this study, we adjusted the scoring probability thresholds of the post-training deep learning model. The receiver operating characteristic (ROC) curve was plotted, and the area under the curve (AUC) was calculated to evaluate model performance. To further assess robustness, we calculated all performance indicators on the validation set and performed 100 inferential validations using the Bootstrap method to determine the 95% confidence intervals [30, 34]. After extracting the features from knee X-ray images in the training set, we randomly selected 60 X-ray images from an external validation set to compare the performance of several models and clinicians. We further compared the performance of the centralized model, the distributed AI model, and attending physicians (five orthopedic surgeons and five radiologists).

Statistical analysis

When analyzing the accuracy of the AI model and comparing the clinical effects, this study employed IBM SPSS 26.0 statistical software and R 4.2.2 software for data processing. For the measurement data, tests were initially conducted to determine whether they conformed to the normal distribution. The normality of the data was examined using the Kolmogorov–Smirnov test. The measurement data that conformed to the normal distribution were expressed as $x \pm s$, and the two-independent-sample t test was used for intergroup comparisons. Variables with $p < 0.05$ in the analysis were incorporated into the multivariate Logistic regression analysis. All tests were two-sided and $p < 0.05$ was considered statistically significant. Regarding the stability evaluation of the distributed AI model, the efficacy of the model was represented by combining the average training results and the 95% confidence interval.

Results

Demographics

In total, X-ray image data pertaining to 4181 radiographs were retrospectively collected from January 2016 and June 2024. The dataset was partitioned into a training set, testing set, and internal validation set at a ratio of 8:1:1. In the distributed database system, the demographic distribution of data for each node is elaborated in Table 1. In the development of YOLOv8n, the training set comprised 3026 images (1384 with TPF and 1642 without TPF), the testing set included 377 images (172 with TPF and 205 without TPF), the internal validation set consisted of 378 images (173 with TPF and 205 without TPF), and the external validation set had 400 images (200 with TPF and 200 without TPF).

Table 1 Distributed AI model training data partitioning

		Training dataset	Testing dataset	Internal validation dataset	External validation dataset	Total
WHUH	TPF	722	89	89	–	900
	Without TPF	791	100	100	–	991
FUPH	TPF	407	48	49	–	504
	Without TPF	500	66	66	–	632
WURH	TPF	255	35	35	–	325
	Without TPF	351	39	39	–	429
WHFH	TPF	–	–	–	200	200
	Without TPF	–	–	–	200	200

WHUH Wuhan Union Hospital, *FUPH* Fuzhou University Affiliated Provincial Hospital, *WURH* Renmin Hospital of Wuhan University, *WHFH* Wuhan Fourth Hospital

Performance of the centralized model

In the task of automatic recognition of traumatic TPFs, the AUC of YOLOv8n and YOLOv8x were 0.9884 [95% CI (0.9854, 0.9947)] and 0.9894 [95% CI (0.9864, 0.9897)], respectively. However, the Youden index of YOLOv8n was the highest when the score threshold was 0.5493, the corresponding accuracy, sensitivity, and specificity were 0.9754, 0.9613, and 0.9895, respectively. Therefore, YOLOv8n was selected as the baseline model for the recognition of TPF. The results of the binary classification task performed on the validation set are shown in Table 2. The confusion matrix and normalized confusion matrix of YOLOv8n for the binary classification are shown in Fig. 3. In the validation set, the detection accuracy of YOLOv8n reached 98%. By marking the identified results, this study presented the visualization effect of the centralized model (Fig. 4).

Table 2 The performance of DL model in distinguishing new TPF in a binary classification task

Evaluation index	Optimum value	Training plateau (starting from round 100)	
		x ± s	95% CI
Accuracy	0.9638	0.9622 ± 0.0016	(0.9628, 0.9667)
Precision	0.9464	0.9416 ± 0.0039	(0.9422, 0.9475)
Sensitivity	0.9875	0.9791 ± 0.0032	(0.9836, 0.9889)
Specificity	0.9378	0.9257 ± 0.0047	(0.9358, 0.9385)
F1-score	0.9252	0.9236 ± 0.0012	(0.9242, 0.9261)
AUC	0.9901	0.9899 ± 0.0001	(0.9891, 0.9902)

Evaluation of model efficacy based on the integration of blockchain and deep learning

The results indicated that the distributed AI model based on the integration of blockchain and deep learning achieved an accuracy of 0.9603 [95% CI (0.9598, 0.9605)], a precision of 0.9614 [95% CI (0.9612, 0.9616)], a sensitivity of 0.9354 [95% CI (0.9336, 0.9359)], a specificity of 0.9372 [95% CI (0.9358, 0.9375)], and an F1 score of 0.9253 [95% CI (0.9223, 0.9264)]. By drawing the ROC curve (as shown in Fig. 5), it can be observed that the AUC values of each node were 0.9792 [95% CI (0.9891, 0.9856)], 0.9776 [95% CI (0.9771, 0.9814)], and 0.9743 [95% CI (0.9725, 0.9832)], while the AUC value of the global model of the distributed AI was 0.9911 [95% CI (0.9893, 0.9915)], which was comparable to the AUC value of the centralized model of 0.990 ($p > 0.05$). The values of the experimental results are presented in Table 3.

Clinical effect assessment

In external clinical validation studies, distributed models demonstrate comparable accuracy to centralized models while also outperforming human physicians in terms of efficiency. The relevant data are presented in Tables 4 and 5. In the overall diagnosis of suspected traumatic TPF in the test set X-rays of knees, the accuracy of the distributed AI model was 0.9636 [95% CI (0.9388, 0.9762)], which was higher than that of the orthopedic attending physicians (0.9291 [95% CI (0.9002, 0.9482)]) and the radiology attending physicians (0.9175 [95% CI (0.8891, 0.9393)]), with statistically significant differences ($p < 0.05$). The

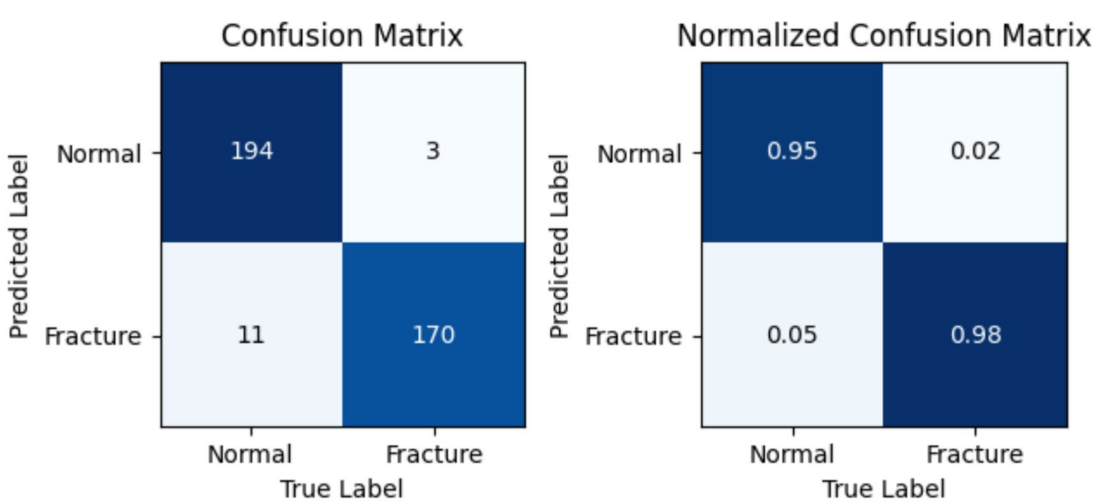
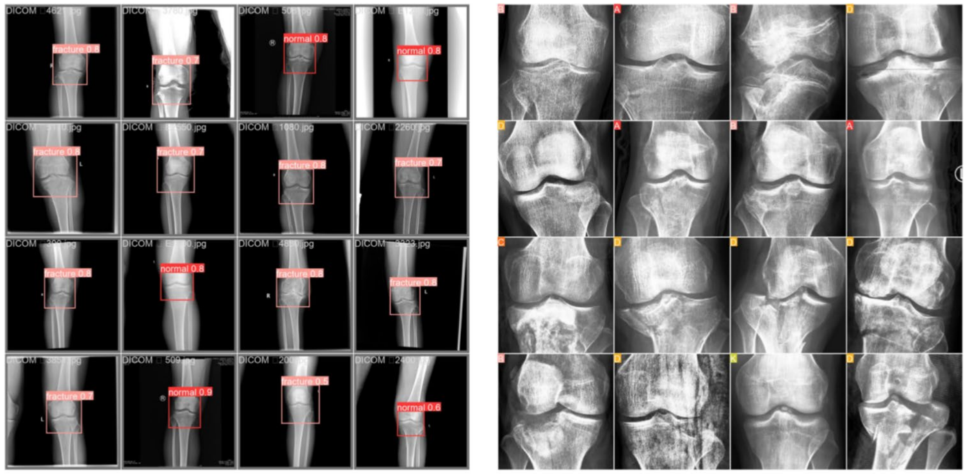


Fig. 3 Confusion and normalized confusion matrix for binary classification of TPFs using YOLOv8n



a. b.

Fig. 4 Output results of TPFs detection in validation dataset. a. Illustration of the predicted labels for the detection results. b. Visualization of TPF images output by the model

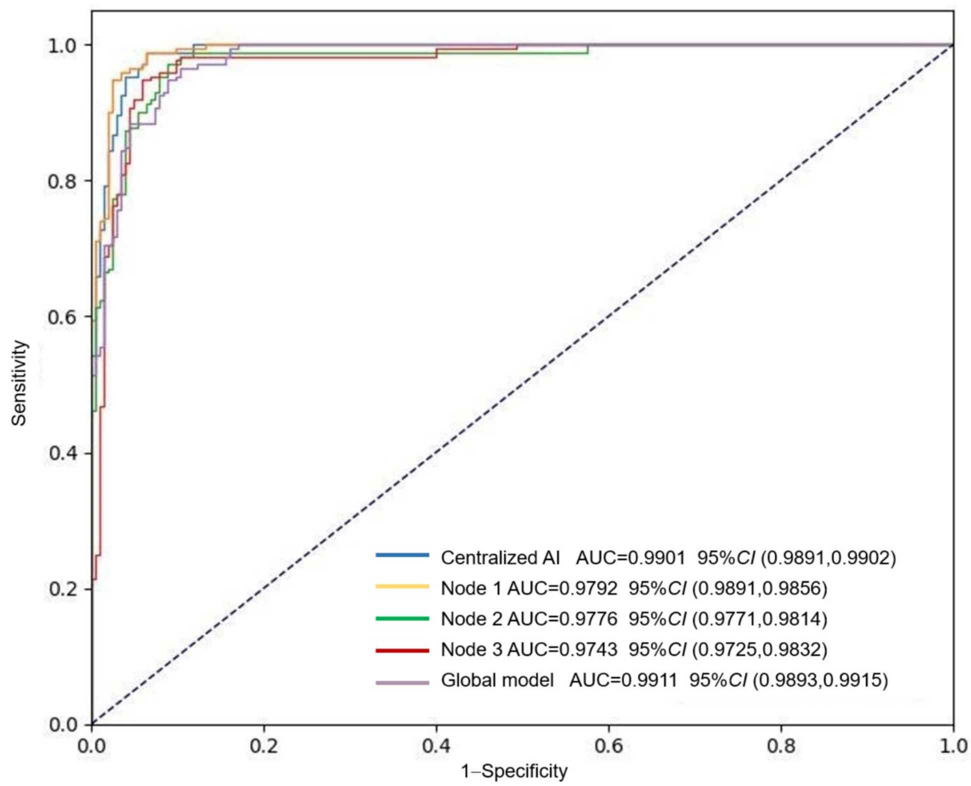


Fig. 5 Comparison of ROC curves between the blockchain and deep learning integrated model and the centralized AI model

Table 3 The performance of the integrated model in distinguishing TPF

Evaluation index	Node 1: WHUH ($n=1891$)	Node 2: FUPH ($n=1134$)	Node 3: WURH ($n=756$)	Global model: trained on all ($N=3781$)
Accuracy (95% CI)	0.9588 (0.9523, 0.9599)	0.9491 (0.9241, 0.9596)	0.9275 (0.9091, 0.9393)	0.9603 (0.9598, 0.9605)
Precision (95% CI)	0.9529 (0.9395, 0.9601)	0.9438 (0.9291, 0.9546)	0.9225 (0.8968, 0.9463)	0.9614 (0.9612, 0.9616)
Sensitivity (95% CI)	0.9494 (0.9221, 0.9592)	0.9135 (0.9039, 0.9240)	0.9155 (0.8497, 0.9276)	0.9354 (0.9336, 0.9359)
Specificity (95% CI)	0.9544 (0.9306, 0.9700)	0.9248 (0.9039, 0.9347)	0.9229 (0.9042, 0.9333)	0.9372 (0.9358, 0.9375)
F1-score (95% CI)	0.9452 (0.9242, 0.9561)	0.9363 (0.9289, 0.9402)	0.9219 (0.9122, 0.9243)	0.9253 (0.9223, 0.9264)
AUC (95% CI)	0.9792 (0.9891, 0.9856)	0.9776 (0.9771, 0.9814)	0.9743 (0.9725, 0.9832)	0.9911 (0.9893, 0.9915)

WHUH Wuhan Union Hospital, *FUPH* Fuzhou University Affiliated Provincial Hospital, *WURH* Renmin Hospital of Wuhan University

Table 4 Comparison of diagnostic performance between centralized AI and distributed AI model

Evaluation index	Centralized AI	Distributed AI model	*p*-value
Accuracy (95% CI)	0.9632 (0.9387, 0.9755)	0.9636 (0.9388, 0.9762)	> 0.05
Precision (95% CI)	0.9426 (0.9057, 0.9631)	0.9526 (0.9122, 0.9648)	> 0.05
Sensitivity (95% CI)	0.9742 (0.9556, 0.9898)	0.9837 (0.9510, 0.9918)	> 0.05
Time consumption (x ± s min)	4.99 ± 0.01	5.06 ± 0.02	> 0.05

Table 5 Comparison of diagnostic performance between orthopedic attending physicians and distributed AI model

Evaluation index	Distributed AI model	Orthopedists	Radiologists	*p*-value
Accuracy (95% CI)	0.9636 (0.9388, 0.9762)	0.9291 (0.9002, 0.9482)	0.9175 (0.8891, 0.9393)	0.004
Precision (95% CI)	0.9526 (0.9122, 0.9648)	0.9057 (0.8661, 0.9413)	0.8929 (0.8597, 0.9276)	0.004
Sensitivity (95% CI)	0.9837 (0.9510, 0.9918)	0.9523 (0.9159, 0.9755)	0.9431 (0.9142, 0.9632)	0.004
Time consumption (x ± s min)	5.06 ± 0.02	25.45 ± 1.92	26.21 ± 1.20	0.004

sensitivity of the distributed AI model was 0.9526 [95% CI (0.9122, 0.9648)], while the sensitivities of the orthopedic attending physicians and the radiology attending physicians were 0.9057 [95% CI (0.8661, 0.9413)] and 0.8929 [95% CI (0.8597, 0.9276)], respectively, with statistically significant differences ($p < 0.05$). The specificity of the distributed AI model was higher than the orthopedic attending physicians (0.9837 > 0.9523) and the radiology attending physicians (0.9837 > 0.9431). In terms of the misdiagnosis rate, the distributed AI model was lower than that of the orthopedic attending physicians (0.0163 < 0.0477) and the radiology attending physicians (0.0163 < 0.0569). In addition, the time taken by the distributed AI model (5.06 ± 0.02 min) was significantly less than that of the orthopedic attending physicians (25.45 ± 1.92 min) ($p < 0.05$) and the radiology attending physicians (26.21 ± 1.20) ($p < 0.05$).

We further examined the average diagnostic efficiency of five orthopedic physicians assisted by distributed AI in cases of traumatic TPF. The results revealed that, compared to orthopedic physicians without assistance, the accuracy of orthopedic physicians with distributed AI assistance increased to 0.9728 [95% CI (0.9602, 0.9882)], the sensitivity rose to 0.9548 [95% CI (0.9433, 0.9575)], the specificity improved to 0.9848 [95% CI (0.9723, 0.9975)], and the diagnosis time was 18.33 ± 1.25 min. The distributed AI global model exhibit significant potential in recognizing traumatic TPF, and the diagnostic efficacy might surpass that of orthopedic attending physicians in the emergency setting ($p < 0.05$). In the distributed training environment, the performance indicators such as the accuracy, sensitivity, and specificity of the AI model show minor differences from the centralized model, enabling the construction of the training model for traumatic TPFs while guaranteeing local data training and data privacy security. The diagnostic accuracy and efficiency of orthopedic attending physicians can be conspicuously enhanced with the assistance of distributed AI model.

Discussion

Blockchain technology offer a decentralized platform for the storage and sharing of images used in AI models, ensuring that no single entity has control over the data. By integrating smart contracts and encryption technologies, only authorized nodes and users are permitted to engage in computations or access the data, thereby safeguarding privacy and security. As the application of AI in the medical field continues to mature, the number of related applications is expanding. In the long term, ensuring the secure utilization and interoperability of medical data from diverse regions has become a critical prerequisite for the future development of AI in healthcare. This study focuses on the critical data requirements for AI model construction and proposes leveraging blockchain technology to integrate the YOLOv8 algorithm into a distributed AI collaborative training framework for TPF detection. In internal and external validation set, the performance of the decentralized model was optimized, demonstrating its potential in assisting orthopedic physicians with the initial screening of patients diagnosed with TPFs. By establishing a distributed blockchain-based specialized disease database, we conducted training of DL models across multiple nodes. The results indicate that the decentralized training environment offers superior security compared to the centralized approach. With the assistance of the distributed AI model, the diagnostic accuracy of orthopedic physicians was improved from 0.9291 to 0.9728. These experiments confirmed that the diagnostic speed and accuracy of orthopedic physicians in diagnosing TPF were enhanced, and the overall misdiagnosis rate was lower than the average level of orthopedic attending physicians without

AI assistance. In terms of the overall sensitivity the distributed AI model reached 0.9526, reaching the average level of orthopedic attending physicians. The results of this study suggest that the distributed AI model could serve as a method for multi-center training-related models and become an auxiliary tool for orthopedic physicians in diagnosing traumatic tibial plateau fractures.

In the data source selection and model construction, there have been studies on developing AI models based on X-ray images for the automatic detection of fractures. However, previous studies have primarily relied on centralized models for data storage and training, which introduces risks to data security. Furthermore, these approaches often neglect the critical role of data sharing in enhancing the performance of global models [8–16]. Our proposed blockchain framework addresses these issues by incorporating a privacy protection mechanism that allows for collaborative model training without compromising data confidentiality. Unlike traditional methods, our approach ensures strong performance in a distributed training environment where image data is not centralized [10, 12–16]. The results demonstrate that distributed method maintains high accuracy, comparable to scenarios using full datasets, by securely distributing data across multiple entities and aggregating model weights through our blockchain-based AI system. In terms of acquiring personal health data, medical institutions within the blockchain platform can verify data ownership, facilitating broader participation and data sharing. By introducing a more advanced convolutional neural network structure, we enhanced the feature extraction ability of the model, which enable it to identify the fracture regions more precisely. Additionally, the distributed AI model performed satisfactorily in the clinical diagnosis of TPF and even surpassed the level of orthopedic attending physicians, with the mean accuracy of 0.9636. In this study, our database originated from 4 different large-scale hospitals, the diversity of data further guarantees good data compatibility, applicability, and generalization of the algorithm after training. The results confirm that the model based on the blockchain can enhance the generalization ability and operational security, which represents a novel organizational modality. In the process of collaboration, the control rights of data and algorithms will be distributed among different computing entities, bringing a more equitable and inclusive training approach for AI models in trauma orthopedics. AI-based image diagnosis is expected to see significant advancements in the future.

To date, this is the first distributed AI system for fracture detection, and it has been externally validated using radiographs from more than one hospital. Comparative analysis reveals that our approach achieves greater robustness and generalization more quickly than centralized models, likely due to its effective management of statistical heterogeneity between datasets from different sources [29–31]. These improvements are due to the efficient aggregation of local model updates within a blockchain framework, which strengthens the model's ability to generalize across diverse datasets from various institutions. In orthopedic applications, our method excels at handling the complexity and variability of fracture detection tasks. Compared to other studies that created deep learning model for detecting TPFs, our algorithm performed with higher sensitivity and specificity [14, 15]. Current research encompasses blockchain-based federated learning for brain tumor image segmentation and blockchain-enabled swarm learning for pathology image analysis [31–34]. However, these models have predominantly relied on public databases and lack external validation. Notably, the field of orthopedics has yet to establish a distributed database framework, and there are no documented instances of targeted modeling utilizing blockchain technology in conjunction with machine learning methodologies [5, 6, 8, 10, 14, 16]. This study highlights the potential of blockchain-based distributed AI models to not only safeguard privacy but also promote collaborative learning in secure, decentralized environments. These findings confirm the feasibility of our approach for real-world medical applications but also illustrate its value in revolutionizing image analysis.

The study has several limitations. First, clinical diagnoses and annotations were made by orthopedic surgeons, with X-ray images serving as the gold standard for fracture identification. This approach focused primarily on visible fracture lines, which may have restricted the range of fracture types considered and simplified the deep learning task. Second, microfractures, which are typically detected through CT scans or other imaging techniques, were not included in the study. Moreover, the data set was limited to adults, all of whom had closed epiphyseal lines. While our distributed training framework provides significant advantages in handling large-scale data sets and complex models, it also faces challenges, including hospital system incompatibilities, high resource demands, and debugging complexities. Achieving optimal performance requires careful configuration of clusters, effective hardware management, and performance tuning—factors that could be challenging for smaller medical departments in clinical settings.

Future research should focus on the practical implementation of a multi-center blockchain system and expand participation to include more orthopedic surgeons and imaging specialists for online data annotation. This will significantly increase the volume of data available for AI training. To ensure a diverse range of patient data—such as 3D CT and MRI scans—are collected for training trauma fracture detection models, it is essential to develop targeted and fair incentive mechanisms for participants using distributed AI systems. Additionally, exploring more advanced fracture segmentation techniques could improve outcomes. Training algorithms on multiple radiograph views simultaneously may provide a more comprehensive understanding of fractures. On the policy front, medical administrators across different countries and regions should align on clinical data privacy sharing to support better communication and collaboration among trauma orthopedic professionals globally.

Conclusion

In conclusion, we propose a diagnostic model that integrates blockchain technology with the deep learning YOLOv8 algorithm. This model demonstrates excellent discriminatory power and clinical applicability for X-ray images of patients with traumatic tibial plateau fractures. It provides a convenient assessment tool for clinical practice. Through continuous optimization and expansion of this innovative approach, it is possible to classify and summarize fracture data and types, potentially guiding data management for specific diseases and AI model training in trauma orthopedics.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00068-025-02793-y.

Author contributions YX contributed to conceptualization, methodology, validation, resources, data curation, writing original draft preparation, XL C contributed to methodology, validation, formal analysis, resources, and writing review and editing, HW Y contributed to writing review and editing, and supervision, XD W, HL W, HZ, LL contributed to formal analysis, investigation, writing review and editing, resources, JY Z, PR L contributed to investigation, resources, writing review and editing, ZW Y contributed to funding, conceptualization, methodology, resources, writing review and editing, supervision. All authors read and approved the final manuscript.

Funding This study was supported by the National Natural Science Foundation of China (Grant no. 82172524 and 81974355), Major Technology Innovation of Hubei Province (Grant no. 2021BEA161), Establishing a national-level innovation platform cultivation plan (02.07.20030019), Open Project Funding of the Hubei Key Laboratory of Big Data Intelligent Analysis and Application, Hubei University (Grant no. 2024BDIAA03), and Wuhan Union Hospital Free Innovation Preliminary Research Fund (Grant no. 2024XHYN047).

Availability of data and materials The datasets used and analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors declare that they have no competing interests.

References

1. Reátiga Aguilar J, Rios X, González Edery E, et al. Epidemiological characterization of tibial plateau fractures. J Orthop Surg Res. 2022;17(1):106.
2. Khan K, Mushtaq M, Rashid M, et al. Management of tibial plateau fractures: a fresh review. Acta Orthop Belg. 2023;89(2):265–73.
3. Rudran B, Little C, Wiik A, Logishetty K. Tibial Plateau fracture: anatomy, diagnosis and management. Br J Hosp Med (Lond). 2020;81(10):1–9.
4. Kfuri M, Schatzker J. Revisiting the Schatzker classification of tibial plateau fractures. Injury. 2018;49(12):2252–63.
5. Rovere G, Bosco F, Miceli A, et al. Adoption of blockchain as a step forward in orthopedic practice. Eur J Transl Myol. 2024;24:12197.
6. Thomson C, Beale R. Is blockchain ready for orthopaedics? A systematic review. J Clin Orthop Trauma. 2021;23: 101615.
7. Allareddy V, Rampa S, Venugopalan SR, et al. Blockchain technology and federated machine learning for collaborative initiatives in orthodontics and craniofacial health. Orthod Craniofac Res. 2023;26(Suppl 1):118–23.
8. Hill BG, Krogue JD, Jevsevar DS, et al. Deep learning and imaging for the orthopaedic surgeon: how machines “read” radiographs. J Bone Jt Surg Am. 2022;104(18):1675–86.
9. Twinprai N, Boonrod A, Boonrod A, et al. Artificial intelligence (AI) vs. human in hip fracture detection. Heliyon. 2022;8(11):e11266.
10. Yoon AP, Lee YL, Kane RL, et al. Development and validation of a deep learning model using convolutional neural networks to identify scaphoid fractures in radiographs. JAMA Netw Open. 2021;4(5): e216096.
11. Lindsey R, Daluiski A, Chopra S, et al. Deep neural network improves fracture detection by clinicians. Proc Natl Acad SCI USA. 2018;115(45):11591–6.
12. Russe MF, Rebmann P, Tran PH, et al. AI-based X-ray fracture analysis of the distal radius: accuracy between representative classification, detection and segmentation deep learning models for clinical practice. BMJ Open. 2024;14(1): e076954.
13. Oakden-Rayner L, Gale W, Bonham TA, et al. Validation and algorithmic audit of a deep learning system for the detection of proximal femoral fractures in patients in the emergency department: a diagnostic accuracy study. Lancet Digit Health. 2022;4(5):e351–8.
14. Liu PR, Zhang JY, Xue MD, et al. Artificial intelligence to diagnose tibial plateau fractures: an intelligent assistant for orthopedic physicians. Curr Med Sci. 2021;41(6):1158–64.
15. Cai D, Zhou Y, He W, et al. Automatic segmentation of knee CT images of tibial plateau fractures based on three-dimensional U-Net: assisting junior physicians with Schatzker classification. Eur J Radiol. 2024;178: 111605.
16. Xu F, Xiong Y, Ye G, et al. Deep learning-based artificial intelligence model for classification of vertebral compression fractures: a multicenter diagnostic study. Front Endocrinol (Lausanne). 2023;14:1025749.
17. Xie Y, Zhang J, Wang H, et al. Applications of blockchain in the medical field: narrative review. J Med Internet Res. 2021;23(10): e28613.
18. Om Kumar CU, Gajendran S, Balaji V, et al. Securing health care data through blockchain enabled collaborative machine learning. Soft Comput. 2023;27(14):9941–54.
19. Rehman A, Abbas S, Khan MA, et al. A secure healthcare 5.0 system based on blockchain technology entangled with federated learning technique. Comput Biol Med. 2022;150:106019.
20. Shafay M, Ahmad RW, Salah K, et al. Blockchain for deep learning: review and open challenges. Cluster Comput. 2023;26(1):197–221.
21. Platt M, Hasselgren A, Román-Belmonte JM, et al. Test, trace, and put on the blockchain?: A viewpoint evaluating the use of decentralized systems for algorithmic contact tracing to combat a global pandemic. JMIR Public Health Surveill. 2021;7(4): e26460.

22. Kumar R, Kumar J, Khan AA, et al. Blockchain and homomorphic encryption based privacy-preserving model aggregation for medical images. Comput Med Imaging Graph. 2022;102: 102139.
23. Tagliafico AS, Campi C, Bianca B, et al. Blockchain in radiology research and clinical practice: current trends and future directions. Radiol Med. 2022;127(4):391–7.
24. Wu C, Tang YM, Kuo WT, et al. Healthcare 5.0: a secure and distributed network for system informatics in medical surgery. Int J Med Inform. 2024;186:105415.
25. Li J, Li S, Li X, et al. Primary bone tumor detection and classification in full-field bone radiographs via YOLO deep learning model. Eur Radiol. 2023;33(6):4237–4248.
26. Jeon YD, Kang MJ, Kuh SU, et al. Deep learning model based on You Only Look Once algorithm for detection and visualization of fracture areas in three-dimensional skeletal images. Diagnostics (Basel). 2023;14(1):11.
27. Ju RY, Cai W. Fracture detection in pediatric wrist trauma X-ray images using YOLOv8 algorithm. Sci Rep. 2023;13(1):20077.
28. Nath S, Rahimy E, Kras A, Korot E. Toward safer ophthalmic artificial intelligence via distributed validation on real-world data. Curr Opin Ophthalmol. 2023;34(5):459–63.
29. Tan TE, Anees A, Chen C, et al. Retinal photograph-based deep learning algorithms for myopia and a blockchain platform to facilitate artificial intelligence medical research: a retrospective multicohort study. Lancet Digit Health. 2021;3(5):e317–29.
30. Moulahi W, Jdey I, Moulahi T, et al. A blockchain-based federated learning mechanism for privacy preservation of healthcare IoT data. Comput Biol Med. 2023;167: 107630.
31. Warnat-Herresthal S, Schultze H, Shastry KL, et al. Swarm Learning for decentralized and confidential clinical machine learning. Nature. 2021;594(7862):265–70.
32. Moztarzadeh O, Jamshidi MB, Sargolzaei S, et al. Metaverse and medical diagnosis: a blockchain-based digital twinning approach based on MobileNetV2 algorithm for cervical vertebral maturation. Diagnostics (Basel). 2023;13(8):1485.
33. Saldanha OL, Quirke P, West NP, et al. Swarm learning for decentralized artificial intelligence in cancer histopathology. Nat Med. 2022;28(6):1232–9.
34. Lu MY, Chen RJ, Kong D, et al. Federated learning for computational pathology on gigapixel whole slide images. Med Image Anal. 2022;76: 102298.

Authors and Affiliations

Yi Xie[1,2] · Xiaoliang Chen[3] · Huiwen Yang[2,4] · Honglin Wang[1,2] · Hong Zhou[1,2] · Lin Lu[5] · Jiayao Zhang[6] · Pengran Liu[1,2] · Zhewei Ye[1,2]

✉ Jiayao Zhang
459882800@qq.com

✉ Pengran Liu
lprlprlprwd@163.com

✉ Zhewei Ye
yezhewei@hust.edu.cn

1 Department of Orthopedics Surgery, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

2 Laboratory of Intelligent Medicine, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

3 People's Hospital of Ningxia Hui Autonomous Region, Ningxia Medical University, Yinchuan, China

4 Department of Otorhinolaryngology, Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China

5 Department of Orthopedics Surgery, Renmin Hospital of Wuhan University, Wuhan, China

6 Department of Orthopedics Surgery, Fuzhou University Affiliated Provincial Hospital, Fuzhou, China