

Original Research

Cite this article: Liu C, Cai L-H, Shen Y-F, Li Z, He Z-J, Chen X-Y, Zhang L, Zhang Y, Xiao Y, Zeng F and Liu M (2025). Deep Learning-based Gait Recognition and Evaluation of the Wounded. *Disaster Medicine and Public Health Preparedness*, **19**, e272, 1–9
<https://doi.org/10.1017/dmp.2025.10179>

Received: 11 December 2024

Revised: 31 May 2025

Accepted: 24 July 2025

Keywords:

Gait recognition; assessment of the injuries; deep learning; YOLOv5

Corresponding author:

Minghua Liu;

Email: minghua_liu@tmmu.edu.cn

Deep Learning-based Gait Recognition and Evaluation of the Wounded

Chuanchuan Liu¹, Ling-Hu Cai¹, Yi-Fei Shen², Zhuo Li³, Zhi-Jian He⁴, Xiang-Yu Chen¹,
Liang Zhang¹, Yi Zhang¹, Yao Xiao¹, Feng Zeng¹ and Minghua Liu¹

¹Department of Emergency Medicine, The First Affiliated Hospital (Southwest Hospital) of Army Medical University, Chongqing, P.R. China; ²Department of Computing and Decision Science, Lingnan University, Hong Kong, P.R. China;

³School of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, P.R. China and ⁴Department of Electronic Computer Engineering, The Hong Kong University of Science and Technology, Hong Kong, P.R. China

Abstract

Objectives: Remote injury assessment during natural disasters poses major challenges for healthcare providers due to the inaccessibility of disaster sites. This study aimed to explore the feasibility of using artificial intelligence (AI) techniques for rapid assessment of traumatic injuries based on gait analysis.

Methods: We conducted an AI-based investigation using a dataset of 4500 gait images across 3 species: humans, dogs, and rabbits. Each image was categorized as either normal or limping. A deep learning model, YOLOv5—a state-of-the-art object detection algorithm—was trained to identify and classify limping gait patterns from normal ones. Model performance was evaluated through repeated experiments and statistical validation.

Results: The YOLOv5 model demonstrated high accuracy in distinguishing between normal and limp gaits across species. Quantitative performance metrics confirmed the model's reliability, and qualitative case studies highlighted its potential application in remote, fast traumatic assessment scenarios.

Conclusions: The use of AI, particularly deep convolutional neural networks like YOLOv5, shows promise in enabling fast, remote traumatic injury assessment during disaster response. This approach could assist healthcare professionals in identifying injury risks when physical access to patients is restricted, thereby improving triage efficiency and early intervention.

Introduction

Recently, humans have experienced frequent natural disasters worldwide.^{1–3} A few days ago, a massive 7.8 magnitude earthquake rocked Turkey and Syria, claiming the lives of over 36,000 people.⁴ Tens of thousands of people have been injured, with traumatic injuries,⁴ particularly fractures, being the most common.^{5–7} In this environment, due to the complexity of disaster scenes, rescue teams cannot reach and search every corner. To save the wounded as quickly as possible, the rescue team must perform a fast evaluation of the injuries at the scene of the disaster.¹ Moreover, collapsed buildings, unstable rubble, obscuring smoke or dust, poor lighting, and intermittent network connectivity commonly hinder clear visual access and reliable data transmission, making objective remote assessment of injuries extremely challenging. This relies heavily on the experience of rescue personnel because rescuers can only judge the condition of the injuries by their appearance.^{8,9} If not treated properly, the walking wounded will become seriously wounded, and even the lives of the badly injured will be endangered.

With the booming of deep learning technology, it has been introduced and widely used in the fast classification of the injured at the scene of disasters.^{10,11} As one of the biometric technologies, gait recognition aims to identify the wounded by analyzing their walking gaits.^{12,13} Compared to other biometrics, gait recognition shows the advantages of being contactless, nonintrusive, and easy to perceive. Gait recognition can be categorized into model-based and appearance-based methods. Generally speaking, the model-based approach is robust but less accurate. Some methods, represented by GaitGraph, the 2D/3D pose, and the SMPL model, take the estimated human skeleton as the model input, which is naturally robust against background noises (e.g., belongings and clothing) and prone to fail with low-resolution images, such that it lacks practicability.¹⁴ The appearance-based approach learns the shape features of objects directly from videos/images. It can work in low-resolution conditions and achieve high recognition accuracy but be sensitive to the object's appearance (e.g., belongings, postures, angle of view, etc.). Gait Set is one of the most influential gait recognition works in recent years, which innovatively takes a sequence of gaits as a set, then uses the maximum function to compress the sequence of spatial features at frame level, which

is simple and effective.¹⁴ GaitPart investigates the local information of the input silhouette and models the time-dependency through the micromotion capture module.¹⁴ Holding the point of view that the global spatial information of the gait representation usually neglects some details, meanwhile, using local information cannot describe the relationship across neighboring regions. GaitGL leverages both global and local convolutional layers to extract more comprehensive information.¹⁴

As deep learning has rapidly advanced, it has driven significant progress in healthcare and medicine—especially in medical image recognition and segmentation, disease diagnosis, and prognosis.^{15–18} Transformer architectures can further enhance medical-image understanding by modeling long-range dependencies.¹⁹ Deep learning-based gait recognition focuses on spatial feature extraction and gait time-dependency modeling. Several studies have been conducted on traumatic injury identification based on gait image recognition,^{20–23} and some of them have outperformed experienced physicians in terms of recognition accuracy.²⁴

To address the challenges in the fast identification and classification of injuries at the scene of disasters, in this work, we propose a deep learning model to recognize injuries with abnormal gait to improve the fast assessment accuracy of the wounded who should be treated immediately.^{25–27} It follows the architecture of the YOLOv5 model,^{28–29} a state-of-the-art deep neural network specifically for image classification and detection, which consists of 3 main pieces, namely, Backbone, Neck, and Head. The backbone is a convolutional neural network that aggregates and forms image features at different granularities; the neck is a series of layers to mix and combine image features to pass them forward to prediction; and the head consumes features from the neck and takes box and class prediction steps. We conducted extensive experiments on both experimental animals (dogs and rabbits) and humans. In the experiment, wood sticks and protectors are used to fix the arms and legs to simulate the injured at the scene of the disaster. We recorded more than 500 video clips and 4500 images for training the YOLOv5-based gait recognition model. Experimental results show that our model can achieve high accuracy in distinguishing the normal and the limped gaits, which validates the availability of our initial ideas that the pure vision-based deep learning model can be used to quickly identify the serious wounded.

Methods

Animals Source

We chose a Springer Spaniel as our pet due to its excellent adaptability and obedience. It is important to note that we treated the Springer Spaniel with utmost care and did not employ any harmful behavior. The simulated broken leg injury was created by gently immobilizing the joint on its foot for the study.

Datasets

The dataset consists of RGB video clips of humans, dogs, and rabbits recorded under disaster-simulation conditions. As illustrated in Figures 1(a, b), we recruited some healthy volunteers (ages 24–35, with no deformities in both lower limbs, and no previous injury) to be captured on video using an iPhone 13 under different environmental backgrounds (indoors, open grass fields, etc.). We divided all samples into 2 groups: the normal gait group and the limp gait group. In the normal gait group, the volunteers walked

back and forth in a playground at a speed of 1–2 m/s, whereas in the limp gait group, they imitated the fractured injured, armed and armored with the fracture protection equipment. At the same time, to enrich the variety and solidness of the study, we used the same method to collect dog and rabbit samples, as depicted in Figures 1(c–d). This yielded 6 distinct classes—human-normal, human-limp, dog-normal, dog-limp, rabbit-normal, and rabbit-limp. Altogether, we obtained >500 clips, each 3–5 s in duration. The recorded clips from the normal-gait and limping cohorts were decomposed into individual frames; any irrelevant or motion-blurred frames were discarded, yielding a curated set of 4500 clean images.

Images Preprocessing

In this study, we collected a total of 9 experimental subjects, 4500 normal gait images, as well as images with distinguished limp gait features. We manually annotated the normal gait and simulated limp gait of different experimental subjects, as displayed in Figure 2. Different colored rectangles are used to circle the regions of normal gait and limp gait of different experimental subjects. In machine learning and deep learning, these rectangles are called labels. Its role is to let the model know what the important part of the image is (category) and helps the computer obtain better meaning for later use of these labels to recognize different categories in new unseen images.

Model Construction

As demonstrated in Figure 3(a), we present the implementation of the YOLOv5 model for the classification of normal gait and simulated limp gait in humans, dogs, and rabbits. The model consists of 4 main components: Backbone, Neck, Dense Prediction, and Sparse Prediction.^{29–31} The backbone was responsible for extracting the original features of the input image. A neck network is used to enhance the feature fusion ability and diversity of the extracted features, which improves the performance of the detection network. Dense prediction and sparse prediction networks were used to obtain the output content and predict the position and category of the target using the previously extracted features. The dense prediction network predicts the position information of the target using the features obtained from the neck network and then concatenates the predicted position information with the features obtained from the neck network to obtain deep latent features. Finally, the model correctly identifies the target class.

Model Training

As depicted in Figure 3(b), the dataset was divided into a training set, test set, and validation set, with 70% of the data being the training set, 20% as the test set, and 10% as the validation set. To balance the class distribution, every limping frame was additionally subjected to small random rotations ($\pm 10^\circ$) and horizontal flips until parity with the normal-gait class was achieved. First, a convolutional neural network is used to extract the original features of the input image. The YOLOv5 model was then iteratively trained on the training set to fit the features of the training set images using the Adam adaptive estimation moving optimizer for iterative optimization. During the training process, a validation set was used to verify the accuracy of the model. After the model was trained, the test set was used to test the precision of the object recognition model in recognizing the gait in unknown scenarios. This method of data splitting is commonly

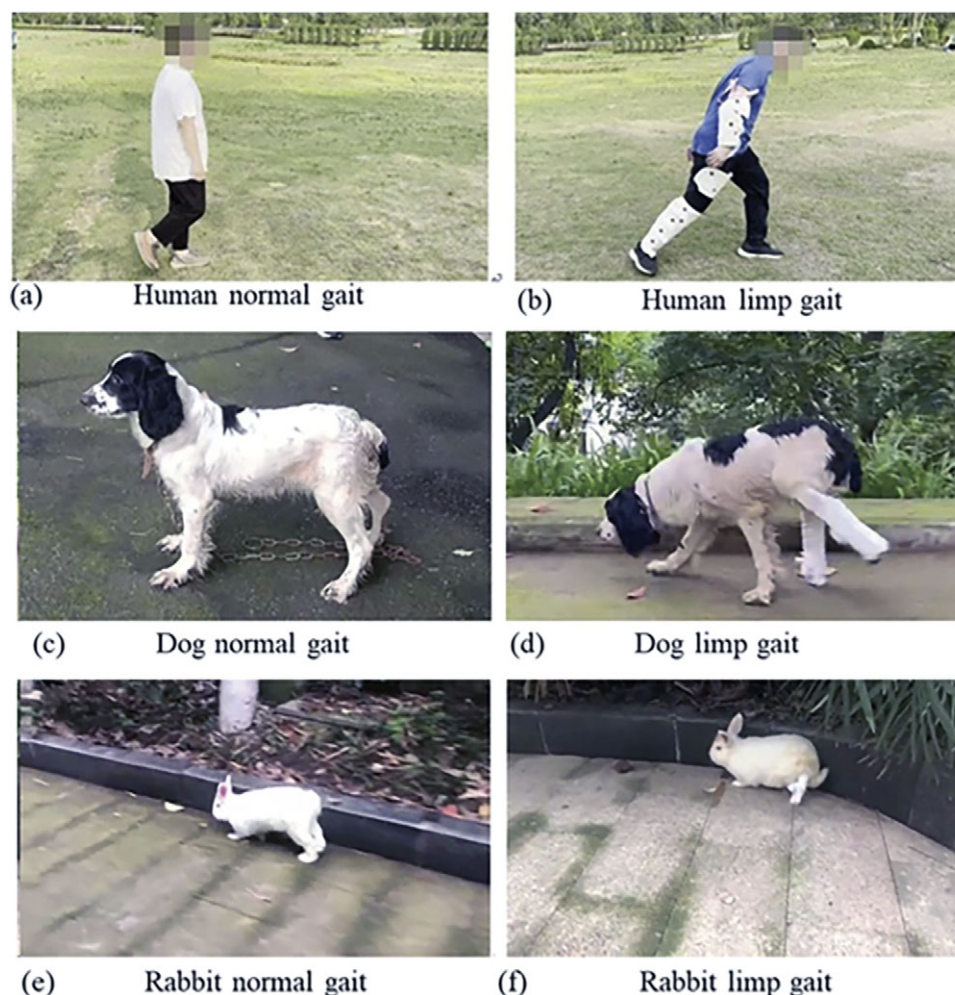


Figure 1. An illustration of the human images (a, b), dog (c, d) and rabbit (e, f) in both normal and limp groups.

used in machine learning/deep learning model development and evaluation. The training set was used to train the model, the validation set was used to tune the hyperparameters of the model, and the test set was used to evaluate the performance of the model. Splitting the data in this manner can help prevent overfitting, which occurs when the model is too complex and performs well on the training set but not on new data. Furthermore, it is crucial to ensure that the data split is random so that the model can be trained and evaluated on a representative sample of the data and that the data are balanced across classes, meaning that all classes have a similar number of samples.

Results

Evaluation Metrics

We used a group of evaluation metrics to evaluate the gait recognition performance by YOLOv5, including precision and recall, which are computed as follows:

$$\begin{cases} \text{Precision} = \frac{TP}{TP + FP} \\ \text{Recall} = \frac{TP}{TP + FN} \end{cases}$$

where TP, FP, TN, and FN represent true positive, false positive, true negative, and false negative samples, respectively. In our case, we used TP to represent the model that correctly predicts that a person walks in a normal state. FP indicates that the model predicts a person walking in a normal gait, but it is a limp gait. The higher the value obtained, the better the performance for all evaluation metrics.

Prediction Results

Figure 4(a) displays that a classification model was used to analyze gait recognition data, focusing on the classification of normal versus limping walking states. The results of the model, as illustrated in the figure, demonstrate that it is highly accurate in identifying normal walking states, with a classification accuracy of 97%. Additionally, the model had a relatively low rate of misclassifying a normal walking state as a limping state at 3%. However, the model is less accurate in identifying limping walking states, with an accuracy of only 70% and a higher rate of misclassifying a limping state as a normal state at 30%. Overall, the model demonstrated a high level of accuracy in identifying normal walking states but could benefit from further improvements in accurately identifying limping walking states.

An illustration of the precision-recall curve indicates that as recall increases, the precision decreases. As depicted in Figure 4(b),

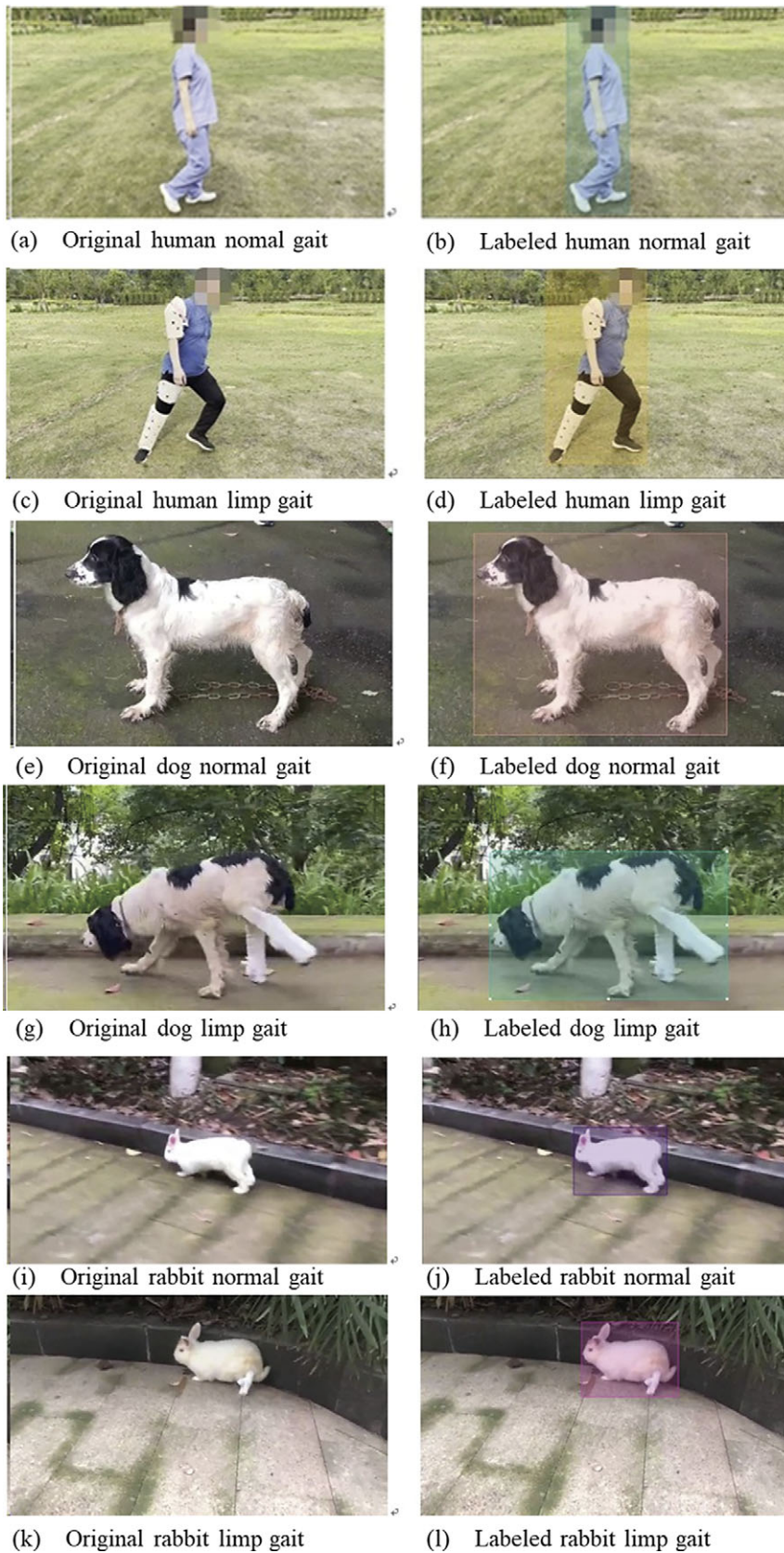


Figure 2. An illustration of the original and labeled human (a-d), dog (e-h) and rabbit (i-l) images in normal and limp groups.

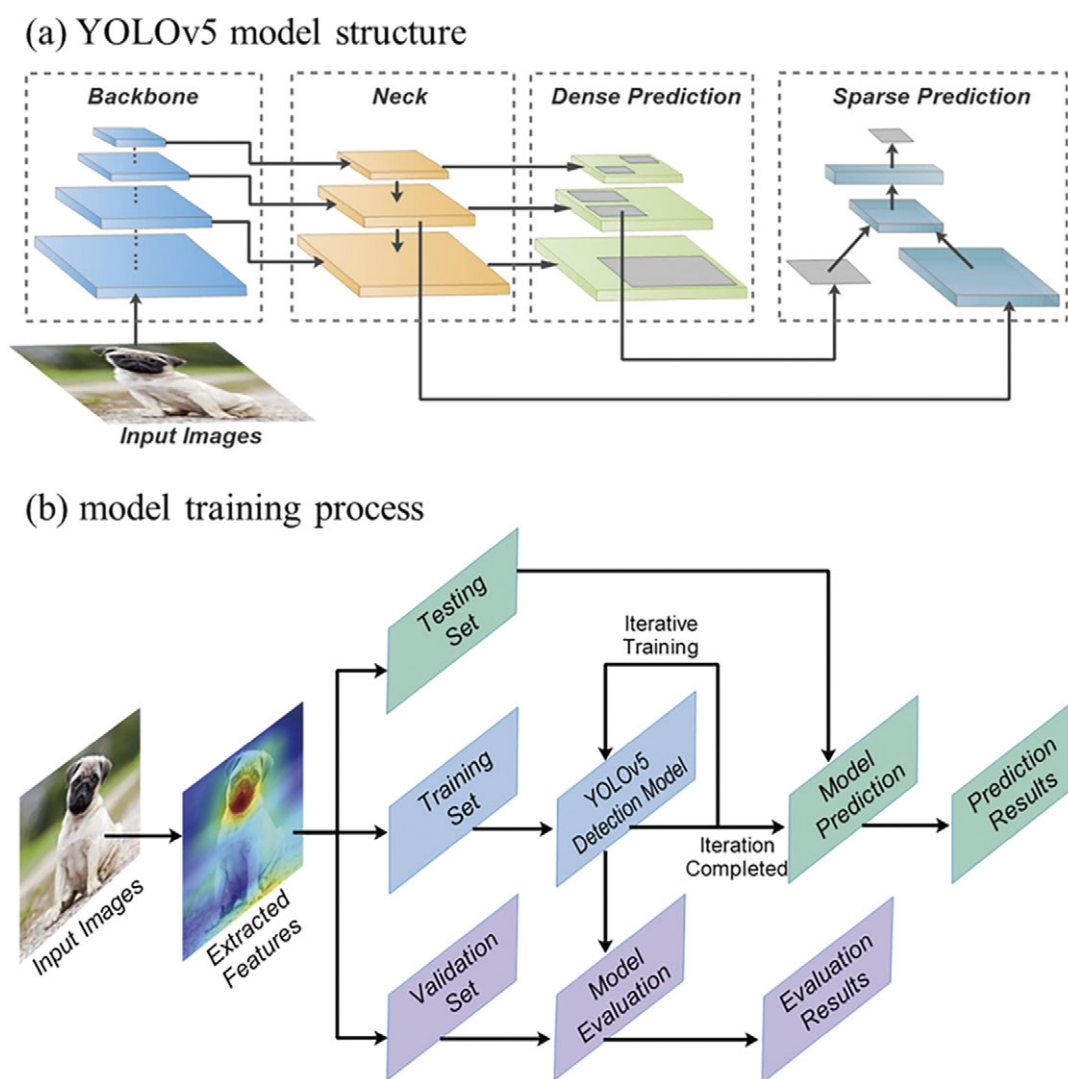


Figure 3. An illustrative diagram of the YOLOv5 model structure (a) and the model training process (b).

the model successfully predicted normal and limping gaits. However, it can be observed that the precision-recall curve for normal gait has a higher intersection point compared to that of limping gait, indicating that the model has slightly better performance in recognizing normal gait. In summary, the model demonstrated good overall performance in recognizing both normal and limping gait, with a slight inclination towards better recognition of normal gait.

Figure 4(c) illustrates the relationship between the classification loss on the training set and the number of training iterations. As the number of training iterations increases, the classification loss decreases accordingly. The graph shows that the decrease in classification loss rate is faster at the beginning of training but gradually slows as the number of iterations increases. This phenomenon is known as “convergence” and is a common characteristic of machine learning algorithms. The goal of training is to minimize the classification loss and achieve the lowest possible value, indicating that the model has learned the features of the data set. One could also use another measure, such as accuracy, to assess the performance of the model during the training process. Overall, these results depict the relationship between the classification loss and the number of training iterations and show how the model’s performance improves as training continues.

Figure 4(d) illustrates the relationship between the classification loss on the validation set and the number of training iterations. As the number of training iterations increased, the classification loss in the validation set decreased. The graph displays that the rate of decrease in classification loss is initially steep but gradually slows as the number of iterations increases. This phenomenon is known as “convergence” and is a common characteristic of machine-learning algorithms. The training aims to minimize the classification loss and achieve the lowest possible value, indicating that the model has learned the features of the validation set. Despite the decrease in the rate of descent, the model’s ability to recover from the classification losses demonstrates its robustness. These results depict that as the number of training iterations increases, the classification loss on the validation set decreases, indicating that the model learns the features of the validation set and is robust.

Figure 4(e) illustrates the relationship between the precision of the model and the number of training iterations. As can be seen, as the number of training iterations increases, the precision of the model also increases. The graph also shows that at the start of training, the increase in precision is more dramatic, but as the number of iterations increases, the rate of increase in precision slows down. This phenomenon is known as “convergence” and is a

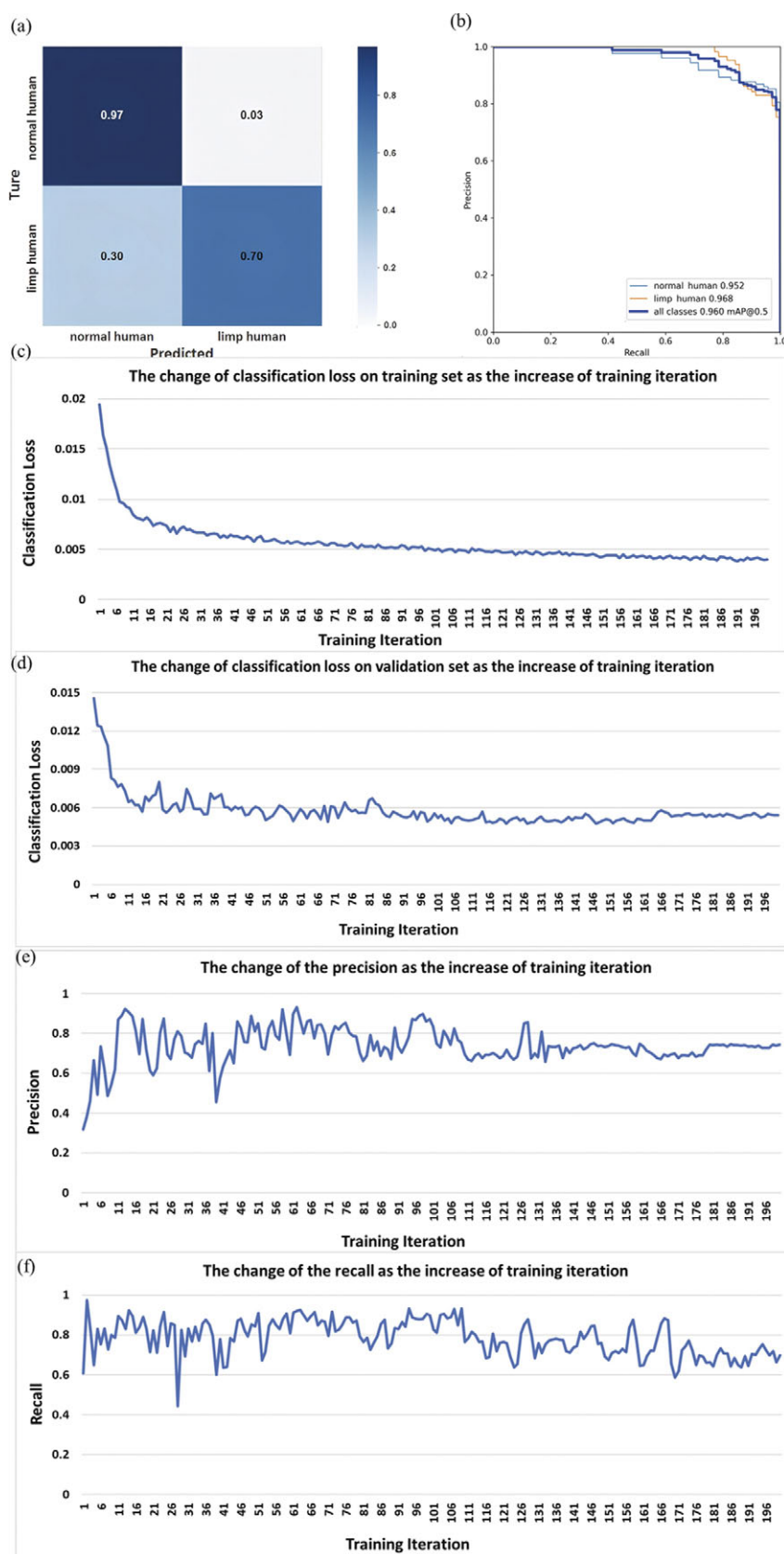


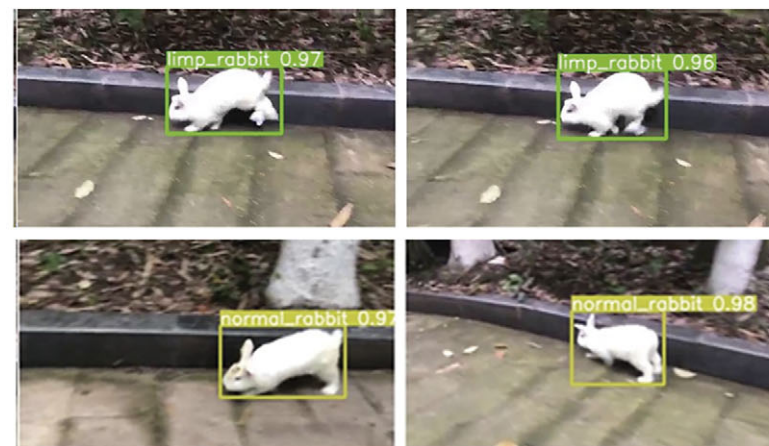
Figure 4. An illustration of the classification results of the gait recognition (a), the obtained precision-recall curve (PR curve) (b), the classification loss on the training set as the increase of training iteration (c), the classification loss on validation set as the increase of training iteration (d), the classification loss on precision as the increase of training iteration (e) and the classification loss on recall as the increase of training iteration (f).



(a) predicted normal and limp human gait



(b) predicted normal and limp dog gait



(c) predicted normal and limp rabbit gait

Figure 5. A case study of predicted normal and limp human gait (a), dog gait (b), and rabbit gait (c).

common characteristic of machine learning algorithms. The goal of training is to maximize the precision of the model, indicating that the model has learned the features of the data set effectively. In summary, the graph shows that as the number of training iterations increases, the precision of the model also increases.

Figure 4(f) illustrates the relationship between the recall metric and the number of training iterations for the model. As the number of training iterations increases, the recall fluctuates, with a consistent pattern around 0.8. The graph shows that initially, the fluctuation of recall is more pronounced, but as the number of iterations increases, the fluctuation gradually decreases. This suggests that the model is becoming more consistent in its ability to correctly identify relevant instances in the data set. In summary, the graph demonstrates how recall improves with an increase in the number of training iterations, with the rate of improvement slowing down as the model converges.

Finally, we tested the trained model on the test set and presented the case study results for gait recognition. As illustrated in Figure 5 (a), rectangles of different colors are used to circle the predicted regions of the model for the normal and limp gait of different subjects, and the prediction area includes the entire body of the experimental subject. We found that the accuracy of different gaits in the model predictions was maintained above 0.95. Concurrently, we also used the same method to predict the samples of dogs and rabbits, and the accuracy of recognition was also very good, even though some cases reached 0.98, as displayed in Figures 5(b, c). The experimental results depict that our model performs well in quickly recognizing normal and limp gaits.

Discussion

In traditional emergency disaster relief, rescuers are often confronted with the perilous task of entering intricate disaster zones, risking their lives. Through the deep learning model elucidated in this paper, we can adeptly differentiate between the normal and limping gaits of simulated injured individuals utilizing image recognition techniques. YOLOv5-s reaches ~30 FPS (FP32) and ~60 FPS (INT8) on a Jetson Xavier NX at 640 × 640 input; given our smaller 256 × 256 resolution, real-time inference on low-power edge devices is expected. This marks a pioneering endeavor in integrating cutting-edge AI technologies into disaster relief, offering a preliminary affirmation of the efficacy of our experimental design and methodology.

Although we have made some progress with our current experimental results, there are still some limitations compared to actual rescue practices. Firstly, there is the issue of the diversity of our experimental volunteers. The volunteers we currently recruit are relatively similar in terms of age, body shape, and appearance, which may affect the model's recognition accuracy in practical applications. To enhance the model's generalization ability, we plan to recruit a wider variety of volunteers in future studies. Secondly, we need to consider the timeliness of the assessment. Our ultimate goal is to achieve a real-time assessment of the victim's condition and provide corresponding emergency medical advice. Real-time recognition is crucial in rescue work because every second counts when a disaster occurs. Therefore, we will strengthen the improvement of this function in our subsequent work. In addition, in order to make the model more practical, we also plan to combine other AI technologies in future research to give the model more vital sign detection capabilities, such as respiration, heartbeat, and pulse, to more accurately assess the severity and urgency of the victim's

injuries in complex environments. Techniques ranging from self-supervised pretraining and regularization—effective with scarce annotations—to semisupervised consistency frameworks that leverage abundant unlabeled data will be explored in future work to further enhance the robustness of our system.^{32–33} In summary, we are steadily moving towards the goal of real-time, accurate rescue assessment and advice, and look forward to making greater breakthroughs in future research.

Conclusion

This paper proposed a deep convolutional neural network model based on YOLOv5 to recognize the limp and normal gait for injuries. Our experimental results validate that our model can distinguish injuries in complex environments using human walking images. With the assistance of this model, we can perform fast traumatic assessments in rescue, explore more artificial intelligence technologies, and apply them to disaster rescue.

Data availability statement. The code and trained model weights used in this study are publicly available at <https://github.com/chuangtouchu/Deep-Learning-based-Gait-Recognition-and-Evaluation-of-the-Wounded>. The datasets generated and analyzed during the current study are available from the corresponding author upon reasonable request.

Acknowledgments. We thank all the participating volunteers for their support in this experiment.

Author contribution. The authors declare that the research was conducted without commercial or financial relationships that could be construed as potential conflicts of interest. Chuanchuan Liu conceived the proposed idea, designed the experimental protocol, implemented the study, wrote the manuscript, and encouraged it. Chuanchuan Liu and Liang Zhang collected data from volunteers, and Chuanchuan Liu, Liang Zhang and Yi-Fei Shen processed, labeled, and supervised the results of this work and analyzed the results obtained. Yi Zhang, Feng Zeng, Yao Xiao, Ling-Hu Cai, Xiang-Yu Chen contributed to collecting animal data and daily animal husbandry and care. Zhi-Jian He provided technical guidance, and all authors discussed the results and contributed to the final manuscript. Ming-Hua Liu was primarily responsible for guiding the overall experimental design and critically reviewing the manuscript.

Funding statement. Not applicable.

Competing interests. The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Patient and public involvement. Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Ethics approval. I promise that the study was performed according to the international, national, and institutional rules considering animal experiments, clinical studies, and biodiversity rights. The study protocol was approved by the Army Medical University Ethics Committee. For the human participants, the study used anonymized data and did not involve any sensitive personal information or commercial interests. According to the *Measures for the Ethical Review of Biomedical Research Involving Humans* (2023), research using anonymized information data that does not cause harm to individuals or involve sensitive personal information or commercial interests can be exempt from ethics review. Therefore, this study did not require formal ethics committee approval. The authors attest that prior to the use of videotaping in the OR, all members of the experimental team at Southwest Hospital, including those pictured in Figures 1, 2, and 5, provided their consent to be taped. This consent included the use of the pictures for all academic purposes. Similarly, all members also provided consent. As indicated in the manuscript (Strategy section), the

consent form for each member included the following statement: "I consent to the recording, photography, closed circuit monitoring or filming for treatment or quality of care and academic."

References

- Rom A, Kelman I. Search without rescue? Evaluating the international search and rescue response to earthquake disasters. *BMJ Global Health*, 2020;5(12):e002398.
- Bradt DA. Site management of health issues in the 2001 World Trade Center disaster. *Academic Emergency Medicine: Official Journal of the Society for Academic Emergency Medicine*, 2003;10(6):650–660.
- Aitsi-Selmi A, Murray V. The Chernobyl disaster and beyond: Implications of the Sendai framework for disaster risk reduction 2015–2030. *PLoS Medicine*, 2016;13(4):e1002017.
- CNN. Over 36,000 dead from quake in Turkey and Syria. February 13, 2023. Available at: <https://www.cnn.com/middleeast/live-news/turkeysyria-earthquake-updates-2-13-23-intl>.
- Nohrstedt D, Hileman J, Mazzoleni M, et al. Exploring disaster impacts on adaptation actions in 549 cities worldwide. *Nature Communications*, 2022;13(1):3360.
- Xie J, Du L, Xia T, et al. Analysis of 1856 inpatients and 33 deaths in the West China Hospital of Sichuan University from the Wenchuan earthquake. *Journal of Evidence-based Medicine*, 2008;1(1):20–26.
- Bartels SA, VanRooyen MJ. Medical complications associated with earthquakes. *Lancet*, 2012;379(9817):748–757.
- Tang S, Ni F, Hu H, et al. Injury assessment of individuals wounded in the Lushan earthquake and the emergency department workload: A corresponding correlation study. *Disaster Medicine and Public Health Preparedness*, 2022;16(1):29–32.
- Shen J, Kang J, Shi Y, et al. Lessons learned from the Wenchuan earthquake. *Journal of Evidence-based Medicine*, 2012;5(2):75–88.
- Kaul V, Enslin S, Gross SA. History of artificial intelligence in medicine. *Gastrointestinal Endoscopy*, 2020;92(4):807–812.
- Yu KH, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nature Biomedical Engineering*, 2018;2(10):719–731.
- He J, Baxter SL, Xu J, et al. The practical implementation of artificial intelligence technologies in medicine. *Nature Medicine*, 2019;25(1):30–36.
- Schedl DC, Kurmi I, Bimber O. An autonomous drone for search and rescue in forests using airborne optical sectioning. *Science Robotics*, 2021;6(55):eabg1188.
- Mentec OL, Jocher G, Waxmann S. Comprehensive guide to ultralytics YOLOv5. 2023. <https://docs.ultralytics.com/yolov5/>.
- Al'Aref SJ, Anchouche K, Singh G, et al. Clinical applications of machine learning in cardiovascular disease and its relevance to cardiac imaging. *European Heart Journal*, 2019;40(24):1975–1986.
- Goto S, Mahara K, Beussink-Nelson L, et al. Artificial intelligence-enabled fully automated detection of cardiac amyloidosis using electrocardiograms and echocardiograms. *Nature Communications*, 2021;12(1): 2726.
- Libbrecht M, Noble WS. Machine learning applications in genetics and genomics. *Nature Reviews, Genetics*, 2015;16(6):321–332.
- Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 2017;42:60–88.
- Liu Z, Lv Q, Yang Z, et al. Recent progress in transformer-based medical image analysis. *Computers in Biology and Medicine*, 2023;164:107268.
- Matheny ME, Whicher D, Thadaney Israni S. Artificial intelligence in health care: A report from the National Academy of Medicine. *JAMA*, 2020;323(6):509–510.
- Hannun AY, Rajpurkar P, Haghighpanahi M, et al. Cardiologist-level arrhythmia detection and classification in ambulatory electrocardiograms using a deep neural network. *Nature Medicine*, 2019;25(1):65–69.
- Wagner N, Beuttenmueller F, Norlin N, et al. Deep learning-enhanced light-field imaging with continuous validation. *Nature Methods*, 2021;18(5):557–563.
- Ozturk T, Talo M, Yildirim EA, et al. Automated detection of COVID-19 cases using deep neural networks with X-ray images. *Computers in Biology and Medicine*, 2020;121:103792.
- Hosny A, Parmar C, Quackenbush J, et al. Artificial intelligence in radiology. *Nature Reviews, Cancer*, 2018;18(8):500–510.
- Jarchi D, Pope J, Lee TKM, et al. A review on accelerometry-based gait analysis and emerging clinical applications. *IEEE Reviews in Biomedical Engineering*, 2018;11:177–194.
- Zhang Z, Tran L, Liu F, et al. On learning disentangled representations for gait recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022;44(1):345–360.
- Hou SK, Lv Q, Ding H, et al. Disaster medicine in China: Present and future. *Disaster Medicine and Public Health Preparedness*, 2018;12(2):157–165.
- Zhu X, Lyu S, Wang X, et al. TPH-YOLOv5: improved YOLOv5 based on transformer prediction head for object detection on drone-captured scenarios. 2021 *IEEE/CVF International Conference on Computer Vision Workshops (ICCVW)*, 2021; 2778–2788.
- Yang G, Feng W, Jin J, et al. Face mask recognition system with YOLOv5 based on image recognition, 2020 *IEEE 6th International Conference on Computer and Communications (ICCC)*, 2020; 1398–1404.
- Bochkovskiy A, Wang CY, Mark-Liao HY. YOLOv4: Optimal speed and accuracy of object detection, arXiv preprint arXiv:2020.04.10934. <https://arxiv.org/abs/2004.10934>.
- Zahia S, Sierra-Sosa D, Garcia-Zapirain B, Elmaghraby A. Tissue classification and segmentation of pressure injuries using convolutional neural networks. *Computer Methods and Programs in Biomedicine*, 2018;159: 51–58.
- Liu Z, Lv Q, Lee C-H, et al. Segmenting medical images with limited data. *Neural Networks*, 2024;177:106367.
- Guo S, Liu Z, Lee C-H, et al. Multi-scale multi-object semi-supervised consistency learning for ultrasound image segmentation. *Neural Networks*, 2025;184:107–195.