Original Article

Segond Fractures Can Be Identified With Excellent Accuracy Utilizing Deep Learning on Anteroposterior Knee Radiographs

Jacob F. Oeding, M.S., Ayoosh Pareek, M.D., Kyle N. Kunze, M.D., Benedict U. Nwachukwu, M.D., M.B.A., Harry G. Greditzer IV, M.D., Christopher L. Camp, M.D., Bryan T. Kelly, M.D., Andrew D. Pearle, M.D., Anil S. Ranawat, M.D., Riley J. Williams III, M.D., and HSS ACL Reconstruction Registry

Purpose: To develop a deep learning model for the detection of Segond fractures on anteroposterior (AP) knee radiographs and to compare model performance to that of trained human experts. **Methods:** AP knee radiographs were retrieved from the Hospital for Special Surgery ACL Registry, which enrolled patients between 2009 and 2013. All images corresponded to patients who underwent anterior cruciate ligament reconstruction by 1 of 23 surgeons included in the registry data. Images were categorized into 1 of 2 classes based on radiographic evidence of a Segond fracture and manually annotated. Seventy percent of the images were used to populate the training set, while 20% and 10% were reserved for the validation and test sets, respectively. Images from the test set were used to compare model performance to that of expert human observers, including an orthopaedic surgery sports medicine fellow and a fellowship-trained orthopaedic sports medicine surgeon with over 10 years of experience. **Results:** A total of 324 AP knee radiographs were retrieved, of which 34 (10.4%) images demonstrated evidence of a Segond fracture. The overall mean average precision (mAP) was 0.985, and this was maintained on the Segond fracture class (mAP = 0.978, precision = 0.844, recall = 1). The model demonstrated 100% accuracy with perfect sensitivity and specificity when applied to the independent testing set and the ability to meet or exceed human sensitivity and specificity in all cases. Compared to an orthopaedic surgery sports medicine fellow, the model required 0.3% of the total time needed to evaluate and classify images in the independent test set. **Conclusions:** A deep learning model was developed and internally validated for Segond fracture detection on AP radiographs and demonstrated perfect accuracy, sensitivity, and specificity on a small test set of radiographs with and without Segond fractures. The model demonstrated superior performance compared with expert human observers. **Clinical Relevance:** Deep learning can be used for automated Segond fracture identification on radiographs, leading to improved diagnosis of easily missed concomitant injuries, including lateral meniscus tears. Automated identification of Segond fractures can also enable large-scale studies on the incidence and clinical significance of these fractures, which may lead to improved management and outcomes for patients with knee injuries.

From the School of Medicine, Mayo Clinic Alix School of Medicine, Rochester, Minnesota, U.S.A. (J.F.O.); Sports Medicine and Shoulder Service, Hospital for Special Surgery, New York, New York, U.S.A. (A.P., K.N.K., B.U.N., B.T.K., A.D.P., A.S.R. R.J.W.); Department of Radiology and Imaging, Hospital for Special Surgery, New York, New York, U.S.A. (H.G.G.); and Department of Orthopedic Surgery and Sports Medicine, Mayo Clinic, Rochester, Minnesota, U.S.A. (C.L.C.).

Received December 1, 2023; accepted March 25, 2024.

Address correspondence to Ayoosh Pareek, M.D., Sports Medicine and Shoulder Service, Hospital for Special Surgery, 535 East 70th Street, New York, NY 10021, U.S.A. E-mail: ayooshp@gmail.com

2666-061X/231672

https://doi.org/10.1016/j.asmr.2024.100940

Segond fractures are relatively rare yet serious osseous avulsion fractures of the anterolateral proximal tibia that occur in association with anterior cruciate ligament (ACL) tears.[1,2] The presence of a Segond fracture has been reported to indicate injury to the ACL up to 75% to 100% of the time, with an overall incidence of 7% to 10% in all ACL tears.[3-5] Due to the strong association with ACL tears, failure to identify Segond fractures on radiographs of the knee may result in functional limitations, concomitant injury, and increased long-term rates of osteoarthritis for patients in whom ligamentous instability is not initially suspected and advanced imaging not obtained.[6] Therefore, accurate identification of Segond fractures

when present is imperative to mitigate the risk of additional adverse and preventable consequences.

More recently, a strong link between Segond fractures and concomitant meniscal injury has been identified.[7-9] With almost 3 in 4 patients with Segond fractures found to have a concomitant lateral meniscus injury,[7] the presence of a Segond fracture should raise suspicion for injury to the lateral meniscus, and the surgeon should more carefully probe the lateral meniscus at the time of arthroscopy. Particularly when considering the role of the lateral meniscus in knee stability and the prevention of accelerated cartilage degeneration, as well as the fact that prior studies have shown that lateral meniscus tears associated with ACL injury can be particularly difficult to detect on magnetic resonance imaging (MRI) (with sensitivity dropping from 0.94 for patients with an intact ligament to 0.69 for patients with an ACL rupture),[10-12] improved identification of Segond fractures has substantial potential to improve the rate at which lateral meniscus tears are diagnosed and repaired in the setting of ACL injury, leading to enhanced joint preservation and improved long-term outcomes.

Deep learning is a subset of artificial intelligence that can analyze features present in unprocessed data, such as medical images, to make an informed prediction on its content.[13-16] It is thus poised to detect the presence of salient findings used to inform diagnosis or guide treatment.[13,14,17] Given that Segond fractures may be subtle or overlooked findings on anteroposterior (AP) knee radiographs, especially in hospital or health care settings with a paucity of resources or a dedicated musculoskeletal radiologist,[18] the development of a deep learning model that can accurately and efficiently detect these fractures may be of great clinical utility. By highlighting these salient findings when present on a knee radiograph, a deep learning model has the potential to assist health care providers in making more accurate and objective diagnoses, to triage and expedite treatment, and to improve patient outcomes.[19,20]

The purposes of the current study were to develop a deep learning model for detection of Segond fractures on AP knee radiographs and to compare model performance to that of trained human experts. The authors hypothesized that performance of the deep learning model would match human observer performance while being substantially faster.

Methods

Imaging Data

Following institutional review board approval, preoperative AP radiographs of patients treated for ACL tears at a large tertiary academic institution were retrospectively reviewed to identify images with and without evidence of a Segond fracture. Images were drawn from the Hospital for Special Surgery ACL Registry, which enrolled patients between 2009 and 2013. All images corresponded to patients who underwent ACL reconstruction by 1 of 23 surgeons included in the registry data. Segond fractures were identified through both manual examination and inspection of radiology reports by a board-certified musculoskeletal radiologist (HGG) with over 15 years of experience. Images were reviewed with particular attention placed on the presence or absence of Segond fractures. Images were deidentified and saved as Portable Network Graphics files in compliance with the Health Insurance Portability and Accountability Act.

Data Annotation

Following extraction and classification of the data set based on presence or absence of a Segond fracture, images were further annotated by manually drawing bounding boxes around the area of interest. Bounding boxes refer to rectangular frames or regions that enclose a specific object or region of interest within an image. They are commonly used in object detection tasks, where the goal is to locate and identify objects within images. In these tasks, deep learning models are trained to not only classify objects but also predict the coordinates of the bounding boxes that tightly enclose these objects. In the present study, these "objects" are the presence or absence of a Segond fracture. For the purpose of identifying Segond fractures, the area of interest was systematically defined as the area where a Segond fracture may occur on an AP radiograph (Fig 1):

(1) Medial boundary—lateral margin of the proximal tibia
(2) Lateral boundary—opacity change representing the edge of the subcutaneous tissue
(3) Superior boundary—lateral articular margin of the tibial plateau
(4) Inferior boundary—superior tip of the fibular head

Each bounding box was assigned a label corresponding to whether a Segond fracture was present within it (Segond = presence of Segond fracture, No Segond = no Segond fracture detected).

Model, Initialization, and Training

Seventy percent of the images were used to populate the training set, while 20% and 10% were reserved for the validation and test sets, respectively.

A convolutional neural network (CNN) object detection model (YOLOv5)[21] was trained to detect the area of the radiograph most likely to contain a Segond fracture, defined by the annotated bounding boxes as described above. As a single-stage object detector, the YOLOv5 model predicts both the location and contents of bounding boxes in a single step, bypassing the region

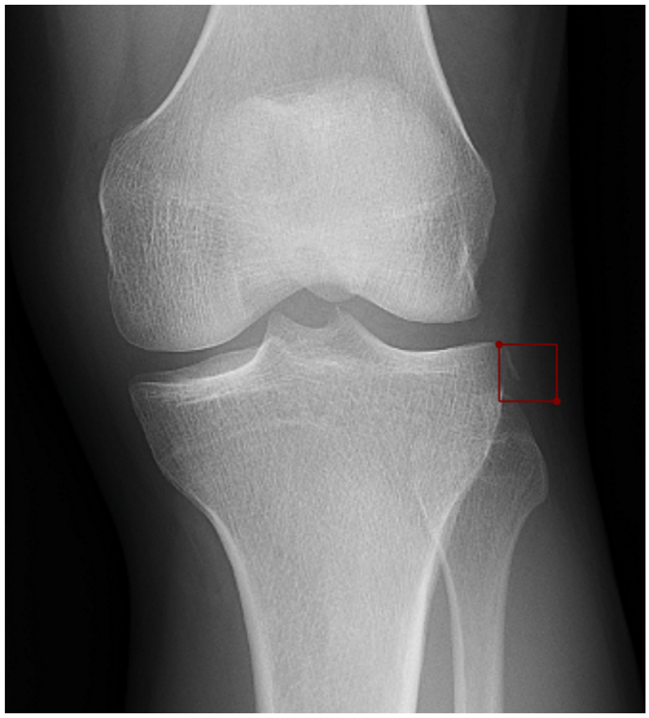

Fig 1. Anteroposterior radiograph of left knee demonstrating a Segond fracture with bounding box.

proposal stage of 2-stage models and thereby resulting in faster inference times.[22] Briefly, the YOLOv5 architecture consists of a "backbone," "neck," and "head." For its "backbone," YOLOv5 uses the cross-stage partial network (CSPNet) to extract features from the image.[23] Features are fused to create feature pyramids in the "neck" (PANet in YOLOv5), which are then used to generate model outputs in the "head."[24]

YOLOv5 was trained for 100 epochs, with a batch size of 16 and a starting learning rate of 0.01, which was reduced gradually. Adam (Adaptive Moment Estimation) was used as the optimizer, which is an optimization algorithm used to update the parameters of a neural network during the training process.[25] During training, the model that achieved the highest mean average precision (mAP) was selected as the final model. Briefly, mAP is a widely used metric for the evaluation of object detection models that calculates the average precision for recall values ranging from 0 to 1. Specifically, this metric compares the ground-truth bounding box to the detected bounding box, with higher scores suggesting more accuracy (range between 0 and 1). We trained our YOLOv5 model on an NVIDIA A100 Tensor Core GPU (Nvidia), utilizing the YOLOv5x configuration in PyTorch (The Linux Foundation).[21]

Model Performance

The model performance was assessed overall and on each individual class (Segond and No Segond). The precision-recall curve was utilized as the primary metric to determine model performance (standard in deep learning classification tasks).[26] Precision describes the success of a model at predicting the positive class (cases with Segond fractures) and is defined as the number of true positives divided by the number of cases the model predicted as positives (true positives plus false positives). A precision of 1 corresponds to a model with zero false positives (cases without a Segond fracture that the model predicted as cases with Segond fractures). Recall, however, is calculated by dividing the number of true positives by the number of cases the model should have predicted as positive (true positives plus false negatives). A perfect recall of 1 would therefore correspond to a model with zero false negatives (cases with a Segond fracture that the model did not identify).[27] We chose to utilize precision recall and the associated precision-recall curve to assess model performance as opposed to a receiver operator curve (ROC) as ROC may be predisposed to misrepresent model performance in theoretical cases of class imbalance and inadvertently lead to incorrect interpretations of the model results.[28] Finally, confidence scores reflecting how certain the model is that the predicted bounding box does or does not contain a Segond fracture were provided as part of the model output. Including confidence scores or local explanation metrics to explain the model's interpretations to humans will be particularly important for AI detection algorithms to have utility in the clinical setting.[29]

Model Attention

To visualize model attention, a class activation mapping (CAM) technique was applied.[30] Model attention mimics cognitive attention in that deep learning models learn to focus on some parts of the input data while ignoring other, less-important parts for optimal prediction. CAM techniques allow for visualization of what parts of the image the model is "looking at" to make its predictions. As CNNs consist of a substantial number of parameters and layers, model transparency is limited as the exact mechanisms through which the model learns are hidden. This results in outputs that may not be differentiable, limiting the use of commonly applied gradient-weighted class activation mapping (Grad-CAM) techniques.[31] Thus, an approach to class activation mapping was utilized that computes and visualizes the principal components of the learned features/representations from the convolutional layers and does not rely on the backpropagation of gradients (Eigen-CAM).[30] It is important to note that Eigen-CAM is an overlay of a low-resolution activation map from the second-to-last layer of the model (where the feature map resolution is low) on the high-resolution input image and thus does not necessarily provide a precise depiction of what aspect of the image the model is

Table 1. Model Overall and Class Specific Performance

Class	Precision	Recall	Mean Average Precision
Segond	0.844	1	0.978
No Segond	0.981	0.954	0.993
Overall	0.912	0.977	0.985

focusing on when making its decisions. Thus, the area in red on the heatmap is not necessarily the most important area in the model's final classification decision. However, CAM techniques still provide a good approximation and some utility in ensuring the interpretability of deep learning models.

Human Observer Comparison

Images from the test were used to compare model performance to 2 different levels of human observer expertise: (1) an orthopaedic surgery sports medicine fellow and (2) a fellowship-trained orthopaedic sports medicine surgeon with over 10 years of experience. The sensitivity and specificity of the model were compared to those of the analyses performed by the human observers, in addition to the time taken to identify the pathology/review the radiograph. To determine the average time spent reviewing each image, the sports medicine fellow was given the set of images from the test set and timed from the start of reviewing the first image to the end of reviewing the last image.

Results

In total, 324 AP knee radiographs were retrieved. Of these, 34 (10.4%) images demonstrated evidence of a Segond fracture, while 290 did not contain evidence of a Segond fracture. In total, 227 (70%) of the manually annotated images were randomly selected to create the training set, while 64 images (20%) were reserved for the validation set. The remaining 33 images (10%) were reserved for an independent test set to evaluate the model performance on data not previously used by the model during training or validation. Of the 227 images in the training set, 23 (10.1%) corresponded to patients with a Segond fracture. Of the 64 images in the validation set, 7 (10.9%) corresponded to patients with a Segond fracture. Of the 33 images in the test set, 4 (12.1%) corresponded to patients with a Segond fracture. The deep learning model demonstrated excellent performance in identifying the presence or absence of Segond fractures during training and validation, even on an imbalanced data set (i.e., substantially more images without Segond fractures than images with Segond fractures) (Table 1). The overall mAP was 0.985, and this performance was maintained on the Segond fracture class (mAP = 0.978, precision = 0.844, recall = 1). This corresponds to the precision-recall curve denoted in Figure 2.

Model Performance on the Independent Test Set

The model demonstrated excellent performance on the randomly selected set of 33 images held out from the training and validation process, suggesting the model can generalize predictions on new data. The model correctly classified all images in the test, demonstrating 100% accuracy on both classes, with perfect sensitivity (1.0) and specificity (1.0). An example of the test set with model predictions is displayed in Figure 3.

Model Versus Human Observer Performance

On the test set, the model conferred a sensitivity and specificity of 1.0, while the fellowship-trained sports medicine orthopaedic surgeon achieved a sensitivity of 1.0 and specificity of 0.83. The orthopaedic surgery sports medicine fellow demonstrated a sensitivity of 0.75 and a specificity of 1.0. The model analyzed all test set images with an average inference time of 21.0 ms per image, while the sports medicine fellow required an average of 8.1 seconds per image.

Model Visualization (Eigen-CAM)

The Eigen-CAM technique is demonstrated in Figure 4 using a representative set of 3 images. This analysis demonstrated that the model utilized the position of the distal femur and proximal tibia to locate the potential fracture site and subsequently conclude whether a Segond fracture was present in that area.

Discussion

The principal findings of the current study are as follows: (1) a deep learning algorithm developed on an institutional data set consisting of 324 AP knee radiographs from patients who underwent ACL reconstruction for acute ACL tears demonstrated excellent predictive ability for the automated detection of Segond fractures, and (2) when compared to expert human observers, the deep learning algorithm demonstrated superior performance while requiring 0.3% of the time to complete the task. Due to recent evidence suggesting that Segond fractures may serve as a "biomarker-like" indicator for lateral meniscus tears,[7] computer vision algorithms designed to detect Segond fractures may improve the rate at which these tears are diagnosed and repaired in the setting of ACL injury, particularly given the current difficulty in detecting concomitant lateral meniscal tears on MRI.

The performance of the deep learning model developed in the current study was perfect with an accuracy of 100% in detecting Segond fractures on the independent testing set of images. Overall and class-specific performance of the model during training and validation was also excellent, with a mAP ranging between 0.98 and 0.99, precision between 0.84 and 0.98, and recall from 0.95 to 1.0 (Table 1). High accuracy for

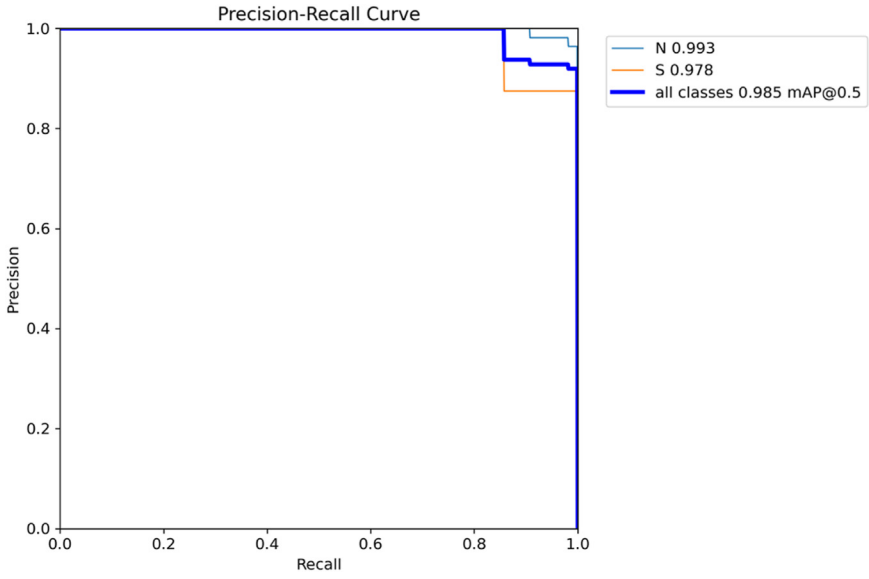

Fig 2. Precision-recall curve demonstrating mean average precision (mAP) of 0.985 for the overall image set with mAP of Segond (S) = 0.978 and mAP of No Segond (N) = 0.993.

detecting musculoskeletal pathology has been demonstrated in prior literature. For example, Zech et al.[32] developed a deep learning model for pediatric wrist fracture detection and reported that their model had an area under the receiver operating characteristic curve of 0.92, corresponding to an accuracy of 88%. Likewise, Liu et al.[33] developed a convolutional neural network deep learning model for intertrochanteric hip fractures and determined that their model predicted the presence or absence of these fractures with an accuracy of 88% and specificity of 87%. Ashkani-Esfahani et al.[34] trained a deep learning model to detect the presence of ankle fractures on 1,050 patients and controls. The authors reported that their model conferred a sensitivity of 98.7% and a specificity of 98.6% in detecting ankle fractures utilizing 3 radiographic views. Therefore, the high accuracy and performance of the current model are plausible, despite utilizing a small data set for training. The findings of the current study demonstrate that deep learning can be confidently applied to radiographs for the detection of Segond fractures, which may have important clinical and research implications. Indeed, utilization of automated computer-vision algorithms in resource-scarce health care settings or geographic areas where advanced imaging or availability of subspecialty trained radiologists remains limited has the potential to reduce the incidence of missed or delayed diagnoses. Anderson et al.[35] reported that clinicians (including radiologists, orthopaedic surgeons, physician assistants, primary care physicians, and emergency medicine physicians) aided by deep learning systems had higher accuracy (AUROC 0.94) compared to clinicians who were unaided (AUROC 0.90) for detecting fractures across 12 different anatomic regions ranging from the femur to the clavicle. The authors also found that artificial intelligence allowed clinicians with limited training in musculoskeletal imaging to reach performance close to expert

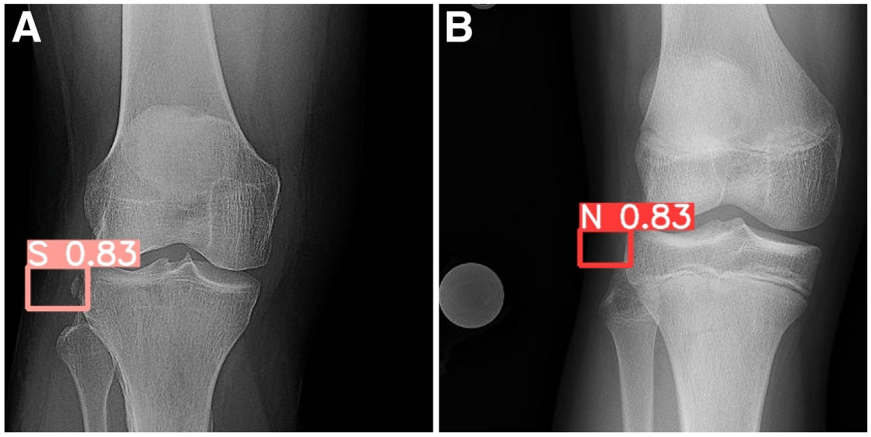

Fig 3. Images denoting bounding box generation and confidence scores of the model for presence of Segond fracture (S) or absence of Segond fracture (N). In this example, the model was 83% confident that each image did (A) or did not (B) contain a Segond fracture.

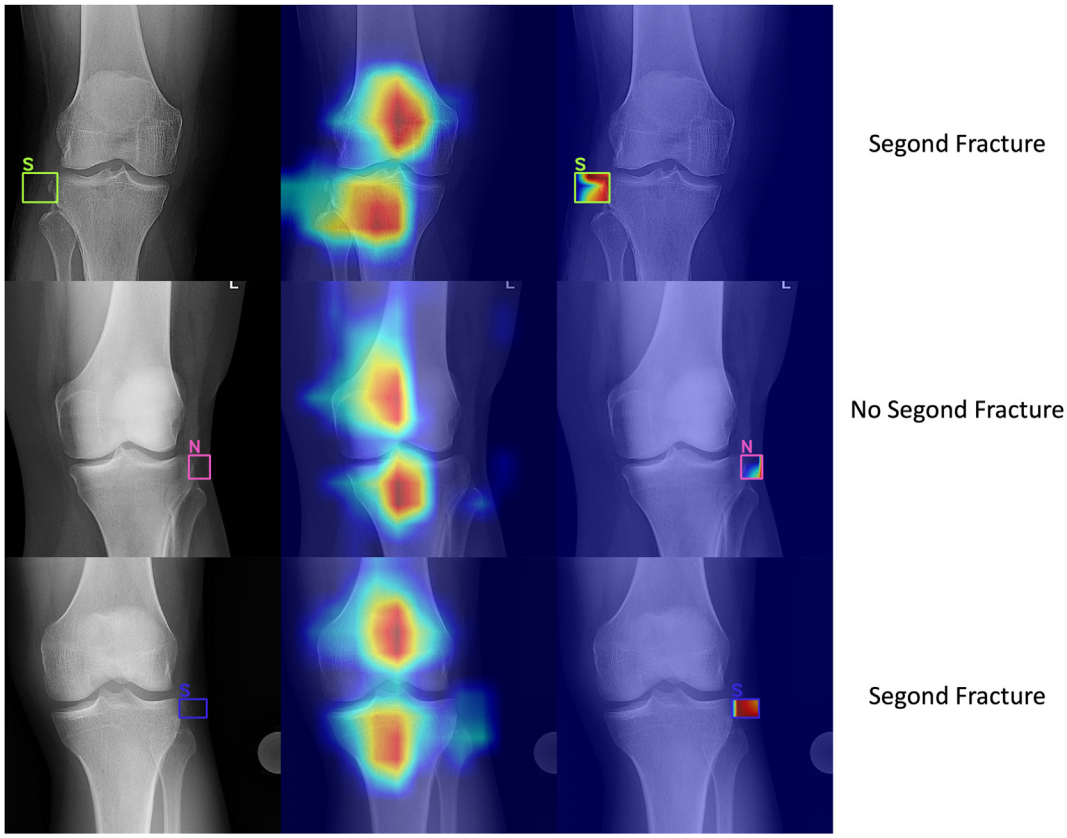

Fig 4. Eigen-class activation mapping (CAM) applied to 3 images from the test set. As an object detection algorithm, the model appears to utilize the position of the distal femur and proximal tibia to locate the potential fracture site and conclude whether a fracture is present in that area.

physicians and radiologists, closing the accuracy gap across clinician types and reducing morbidity for patients. Expanding on the current model for Segond fracture detection, rapid and accurate identification of pathology may subsequently allow for the expedited triage of patients to specialty centers where appropriate treatment or procurement of additional imaging can be performed in patients with acute knee pain with this radiographic finding.

Multiple studies have suggested a strong association between Segond fractures and concomitant meniscal injury.[7-9] In particular, Segond fractures have been demonstrated to indicate a significantly increased risk of lateral meniscus injury in patients with acute ACL ruptures, with an incidence of 72% among patients with Segond fractures, putting patients with Segond fractures at a nearly 3 times greater risk for lateral meniscus tears than patients without Segond fractures in the setting of acute ACL tears.[7] Furthermore, prior studies have shown that lateral meniscus tears can be particularly difficult to detect on MRI, especially tears associated with ACL injury, with sensitivity dropping from 0.94 for patients with an intact ligament to 0.69 for patients with an ACL rupture.[10-12] This sensitivity is even lower for patients with posterior and peripheral tears.[10,11] Given the important role of the lateral meniscus in knee stability and articular cartilage protection, it is important to identify and repair these lesions when they occur. Thus, the model developed in the present study has the potential for substantial improvement in the diagnosis and repair of lateral meniscus tears through enhanced preoperative planning prior to ACL reconstruction. Due to their low incidence, Segond fractures may not be routinely screened for by surgeons evaluating preoperative imaging, but through a deep learning model specifically trained to detect such fractures, surgeons can be alerted when a Segond fracture has occurred in the setting of an ACL rupture. This should raise suspicion of a lateral meniscal tear, and the surgeon should more carefully probe the lateral meniscus at the time of arthroscopy. Furthermore, an increased suspicion for lateral meniscus tears, even if not visible on MRI, allows the surgeon to plan for the necessary operative time and surgical equipment to repair the meniscus as well as ensure informed consent and discussion with the patient prior to the ACL reconstruction. Thus, even if a missed Segond fracture on radiograph does not necessarily lead to an ACL tear going undiagnosed, by facilitating improved identification and repair of meniscal

tears, the deep learning model developed in the current study has the potential to reduce cartilage degeneration over time and improve outcomes after ACL reconstruction by improving the rate at which lateral meniscal tears are diagnosed and repaired.

Automated fracture detection was substantially quicker with the application of a deep learning model compared with an expert human observer. Several prior studies have also corroborated this observation. Yao et al.[36] developed a deep learning model for rib fracture detection on computed tomography (CT) imaging, which not only outperformed 2 radiologists with regard to fracture detection accuracy but also required only 20 seconds as opposed to the greater than 150 seconds for each human observer to evaluate the CT. Interestingly, when collaborating with the model, 1 radiologist was able to decrease the time to evaluate CT scans from an average of 153 seconds per scan to 59 seconds per scan while increasing the accuracy at the same time.[36] In a musculoskeletal study evaluating pelvic fracture severity on CT scans, Dreizin et al.[37] analyzed 373 CT scans using deep learning. They noted the model to be as accurate as an expert radiologist in determining rotational instability and more accurate in determining translational instability with a mean inference time of <0.1 second per test image. This was a finding similar to one of the current study and also supports the plausibility of its high performance on a small data set. The clinical implications of more rapid and objective pathology detection that translate into reductions in time required for manual imaging review include the possibilities of allocating this additional time to shared decision-making, explanation of treatment, and related discussions with patients. Additionally, this may allow providers to use their saved time in more productive ways, such as responding to patient phone calls, completing administrative tasks, or having more patient-centered care.

Limitations

The current study is not without limitations. First, the deep learning model was developed and trained on radiographic data from a single institution with a relatively small test set. Further research is required to determine the external validity of this model and whether it can achieve high performance on larger, multi-institutional, or international data sets with varying quality of radiographs. With all images in the present study derived from a single institution, it is possible that overfitting occurred. For example, the model may depend on the quality of imaging data being analyzed, and performance may differ at facilities with lower-quality radiographs. Second, all patients in this study had confirmed ACL tears. As part of the external validation process, this algorithm should be tested on images of patients who do not have ACL tears but may demonstrate Segond fractures (which would be a rare event).

Conclusions

A deep learning model was developed and internally validated for Segond fracture detection on AP radiographs and demonstrated perfect accuracy, sensitivity, and specificity on a small test set of radiographs with and without Segond fractures. The model demonstrated superior performance compared with expert human observers.

Disclosures

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: J.F.O. is a consultant or advisor for Kaliber.ai. B.U.N. is a consultant or advisor for Stryker, has equity or stocks with Remote Health, and is a board member of Figur8 and BICMD. H.G.G. is a consultant or advisor for Motion Orthopaedics and Professional Imaging and has equity or stocks with Tate Greditzer, LLC. C.L.C. is a paid consultant for Arthrex, has received research support from Major League Baseball, and has received publishing royalties and financial or material support from Springer. B.T.K. is a consultant or advisor for Arthrex and BICMD. A.D.P. is a consultant or advisor for DePuy Synthes, Exactech, Smith & Nephew, and Zimmer and has equity or stocks with ACLIP, myHealthTrack, PerfectFict, and Vent Creativity. A.S.R. is a consultant or advisor for Anika Therapeutics, Arthrex, Bodycad USA, Marrow Cellution, Newclip Technics, Smith & Nephew, Stryker, and Xiros; has received speaking and lecture fees from Arthrex; and has equity or stocks with Conformis, Enhatch, and Ranfac. R.J.W. has equity or stocks with BICMD, CyMedica Orthopedics, Gramercy Extremity Orthopedics, Remote Health Ventures, Tonal, and Overture and is a board member of Pristine Surgical, Rehab Boost, and Tonal. All other authors (A.P., K.N.K.) declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors thank Sophia Madjarova, B.A., and Arjun Khorana, B.S., for their assistance with the preparation and submission of this manuscript.

References

1. Segond P. Recherces Cliniques et Expérimentales Sur Les Épanchements Sanguins Du Genou Par Entorse. *Progres Med* 1879.
2. Claes S, Luyckx T, Vereecke E, Bellemans J. The Segond fracture: A bony injury of the anterolateral ligament of the knee. *Arthroscopy* 2014;30:1475-1482.
3. Goldman AB, Pavlov H, Rubenstein D. The Segond fracture of the proximal tibia: A small avulsion that reflects

major ligamentous damage. *AJR Am J Roentgenol* 1988;151:1163-1167.
4. Hess T, Rupp S, Hopf T, Gleitz M, Liebler J. Lateral tibial avulsion fractures and disruptions to the anterior cruciate ligament. A clinical study of their incidence and correlation. *Clin Orthop Relat Res* 1994:193-197.
5. Yoon KH, Kim JS, Park SY, Park SE. The influence of Segond fracture on outcomes after anterior cruciate ligament reconstruction. *Arthroscopy* 2018;34:1900-1906.
6. Capps GW, Hayes CW. Easily missed injuries around the knee. *Radiographics* 1994;14:1191-1210.
7. Garra S, Moore MR, Li ZI, et al. Segond fracture: An indicator for increased risk of lateral meniscus injury in patients with acute anterior cruciate ligament ruptures. *Eur J Orthop Surg Traumatol* 2024;34:1883-1891.
8. Sulaiman Y, Li J, Chen G, Abudouaini H, Li Q, Tang X. The relationship between a Segond fracture and meniscus injury in patients with anterior cruciate ligament tears. *Knee* 2021;33:193-199.
9. Yeo PY, Seah AMJ, Visvalingam V, et al. Anterior cruciate ligament rupture and associated Segond fracture: Incidence and effect on associated ligamentous and meniscal injuries. *Asia Pac J Sports Med Arthrosc Rehabil Technol* 2022;30:36-40.
10. Krych AJ, Wu IT, Desai VS, et al. High rate of missed lateral meniscus posterior root tears on preoperative magnetic resonance imaging. *Orthop J Sports Med* 2018;6:2325967118765722.
11. De Smet AA, Graf BK. Meniscal tears missed on MR imaging: Relationship to meniscal tear patterns and anterior cruciate ligament tears. *AJR Am J Roentgenol* 1994;162:905-911.
12. Kim SH, Lee HJ, Jang YH, Chun KJ, Park YB. Diagnostic accuracy of magnetic resonance imaging in the detection of type and location of meniscus tears: Comparison with arthroscopic findings. *J Clin Med* 2021;10:606.
13. Currie G, Hawk KE, Rohren E, Vial A, Klein R. Machine learning and deep learning in medical imaging: Intelligent imaging. *J Med Imaging Radiat Sci* 2019;50:477-487.
14. Gore JC. Artificial intelligence in medical imaging. *Magn eason Imaging* 2020;68:A1-A4.
15. Oeding JF, Krych AJ, Pearle AD, Kelly BT, Kunze KN. Medical imaging applications developed using artificial intelligence demonstrate high internal validity yet are limited in scope and lack external validation. *Arthroscopy* 2024. doi: 10.1016/j.arthro.2024.01.043.
16. Oeding JF, Yang L, Sanchez-Sotelo J, et al. A practical guide to the development and deployment of deep learning models for the orthopaedic surgeon: Part III, focus on registry creation, diagnosis, and data privacy. *Knee Surg Sports Traumatol Arthrosc* 2024;32:518-528.
17. Wang DY, Liu SG, Ding J, et al. A deep learning model enhances clinicians' diagnostic accuracy to more than 96% for anterior cruciate ligament ruptures on magnetic resonance imaging. *Arthroscopy* 2024;40:1197-1205.
18. Ashkani-Esfahani S, Mojahed-Yazdi R, Bhimani R, et al. Deep learning algorithms improve the detection of subtle Lisfranc malalignments on weightbearing radiographs. *Foot Ankle Int* 2022;43:1118-1126.
19. Krogue JD, Cheng KV, Hwang KM, et al. Automatic hip fracture identification and functional subclassification with deep learning. *Radiol Artif Intell* 2020;2:e190023.
20. Skrede OJ, De Raedt S, Kleppe A, et al. Deep learning for prediction of colorectal cancer outcome: A discovery and validation study. *Lancet* 2020;395:350-360.
21. Jocher G, Chaurasia A, Stoken A, et al. ultralytics/yolov5: v7.0—YOLOv5 SOTA realtime instance segmentation. *Zenodo* 2022. https://zenodo.org/records/7347926. Accessed April 29, 2024.
22. Thuan D. *Evolution of YOLO Algorithm and YOLOv5: The State-of-the-Art Object Detection Algorithm*. Oulu, Finland: Oulu University of Applied Sciences, 2021.
23. Wang C-Y, Liao H-Y, Yeh I-H. CSPNet: A new backbone that can enhance learning capability of CNN. *CoRR* 2019. arXiv:1911.11929, https://arxiv.org/abs/1911.11929. Accessed April 29, 2024.
24. Wang K, Liew J, Zou Y, Zhou D, Feng J. PANet: Few-shot image semantic segmentation with prototype alignment. *CoRR* 2019. arXiv:1908.06391, https://arxiv.org/abs/1908.06391. Accessed April 29, 2024.
25. Kingma D, Ba J. Adam: A method for stochastic optimization. *CoRR* 2014. arXiv:1412.6980, https://arxiv.org/abs/1412.6980. Accessed April 29, 2024.
26. Precision-Recall. Scikit-learn developers website. https://scikit-learn.org/stable/auto_examples/model_selection/plot_precision_recall.html#:~:text=Precision%2DRecall%20is%20a%20useful,truly%20relevant%20results%20are%20returned. Accessed March 1, 2023.
27. Powers DMW. Evaluation: From precision, recall and F-Factor to ROC, informedness, markedness and correlation. *CoRR* 2020. arXiv:2010.16061, https://arxiv.org/abs/2010.16061. Accessed April 29, 2024.
28. Davis J, Goadrich M. The relationship between precision-recall and ROC curves In: *Proceedings of the 23rd International Conference on Machine learning—ICML '06*. Pittsburgh, PA: ACM Press, 2006;233-240.
29. Zhang Y, Liao Q, Bellamy R. Effect of confidence and explanation on accuracy and trust calibration in AI-assisted decision making In: *Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency—FAT '20* 2020. arXiv.2001.02114, https://arxiv.org/abs/2001.02114. Accessed April 29, 2024.
30. Muhammad M, Yeasin M. Eigen-CAM: Class activation map using principal components. In: *2020 International Joint Conference on Neural Networks (IJCNN) IEEE* 2020. https://arxiv.org/abs/2008.00299. Accessed April 29, 2024.
31. Selvaraju R, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: Visual explanations from deep networks via gradient-based localization. *Int J Comput Vis* 2020;128:336-359.
32. Zech JR, Carotenuto G, Igbinoba Z, et al. Detecting pediatric wrist fractures using deep-learning-based object detection. *Pediatr Radiol* 2023;53:1125-1134.
33. Liu P, Lu L, Chen Y, et al. Artificial intelligence to detect the femoral intertrochanteric fracture: The arrival of the intelligent-medicine era. *Front Bioeng Biotechnol* 2022;10:927926.
34. Ashkani-Esfahani S, Mojahed Yazdi R, Bhimani R, et al. Detection of ankle fractures using deep learning algorithms. *Foot Ankle Surg* 2022;28:1259-1265.
35. Anderson PG, Baum GL, Keathley N, et al. Deep learning assistance closes the accuracy gap in fracture detection

across clinician types. *Clin Orthop Relat Res* 2023;481:580-588.
36. Yao L, Guan X, Song X, et al. Rib fracture detection system based on deep learning. *Sci Rep* 2021;11:23513.
37. Dreizin D, Goldmann F, LeBedis C, 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.