© January 2026 | IJIRT | Volume 12 Issue 8 | ISSN: 2349-6002

DeepVision: An Interpretable Multi-Task Framework for Comprehensive Bone Fracture Diagnosis Using ResNet, DenseNet, YOLOv8, and Grad-CAM++

Para Rajesh[1], P Sushanth Chakravarthy[2], Neha Sheetal Kodakandla[3], B Siddhu[4], T Manoj Kumar[5]
[1,2,3,4,5]*Computer Science and Engineering, VNR Vignana Jyothi Institute of Engineering and Technology, Hyderabad-500090, Telangana, India*
doi.org/10.64643/IJIRTV12I8-191903-459

***Abstract*—Bones break X- rays In many things everyday medicine. Secure it fast helps patients Secure better faster. Most AI tools Just do it now one thing- Value spotting or tidying up- and often just for the sake of it one body part. It does them hard Implement as directed by physicians broad solutions.**

To fix these problems, this study brings out DeepVision- One step by step deep learning setup Built to place, sort, determine and pin broken bones I seven body areas: shoulders, upper arm, elbow, lower arm, Wrist, hand, fingers. It packs a punch four Tuned pieces: ResNet34 with a CBAM block This highlights the main features for the break. DenseNet121 Also CBAM to inform different break kinds In addition; YOLOv8n To fasten the breach without delay; And Grad- CAM++ To manifest exactly where the damage is.

Model power is extended using pre- trained networks, Additional synthetic images, With cleaning of input scans before processing. Performance checks depend on common measures- Favor correctness rate, Average detection score(map), and what the final shape is prefer the marked spots to do real ones- preserve results trustworthy. Administer tests known collections Esteem MURA and the Kaggle Bone Fracture set So we can understand how well it works new cases.

DeepVision Mix sharp detection power with clean visuals- Spit out label, box outline about box breaks, Plus shade- coded hot zone Shows Locations of damage. That way, Doctors can locate out what's causing it the system Something The flag instead of just zooming justice one area Favor older tools do, this setup Handle multiple regions At the same time minimize errors and perform well different bone types.

It is designed to be elementary to perform heavy computing needs, to create it handy to real clinics.

Sure, there are still some hiccups- Prefer to make fun of tiny cracks or uneven training data- But overall, This can increase how quickly and reliably it breaks down bones Procure Stuck in scans, for which no- nonsense helper offered skeletal imaging.

***Index Terms*—Bone fracture detection, Multi- level deep learning, X- ray image analysis, Resnett 34, DenseNet121, CBAM(Convolutional Block Attention Module), YOLOv8, Grade Chem++, Fracture classification, object recognition, Medical image localization, Multidimensional muscle imaging, computer- aided diagnosis(CAD), transfer learning, Medical image processing**

I. INTRODUCTION

The initial group of research Spotting watching the breach, with the aim To indicate that a X- ray The shows one. Instead of simply classifying images as standard or broken, Study[01] Utilize one CNN Setup is highlighted in How to data cleaning Extraction is very vital key symbols. Instead of building a building a model From the beginning,[02] taps I already trained networks- This shift helps catch tiny Indications Breaks More reliable While earlier efforts is used single angles,[03] connects several imaging methods shared prediction methods, It works well when data is limited. On top of that,[04] Adding attention mechanisms So the system In Zoom areas doctors Take care K most things, to make it accessible to create a deficient aspect unclear break patterns.

The next set Let's Dive into how fractures are classified by type. Instead, machines endeavor To ascertain out what kind of interval they are looking at. Study 5[05] sets over a system These are broken multiple groups- Still struggling some types Exhibition way less On top Of this, compared to others, better distinguishing traits Clearly necessary.

Moving forward, study 6[06] Deploy update versions of DenseNet and ResNet To dig up deep bone structure Details Because from this, these networks Manage subtle differences more effectively. I the meantime, study 7[07] checks How good is the computer? Match real doctors' calls. Seam the results The element of view expert level, Confusion arises when the fractures Check approx the same. Finally, study 8[08] Adding attention Tricks to explain classes faster. With both local and channel focus On, Accuracy is achieved a noticeable boost.

The survey Spotting Use boxes to mark where there will be cracks the damage is displayed.

Study instead 9[09] Check simple ways To ascertain errors, Measure how well the separate weights have the scales precision against recall- Why Is that an indication? Good boxes The Scenario Meanwhile, study 10[10] posted an anchor- Based on setup e. G Faster R- CNN To Improve positioning and range higher mAP results. On the contrary, Study 11[11] Test outside YOLO- type For the model the purpose speedier performance. These tools can be found broken bones Well, sometimes problems flag that aren't there.

I the end, Discern study 12[12]. Simpler setups made for clinics- Focusing quick results without losing health so they function smoothly actual healthcare settings.

The latest research On focus better ways For brake detection, shows exact spots through pixel highlights or area outlines By using heat- sensitive images or digital masks. Instead of relying on manually generated frames, Study 13[13] Applies to activation mapping minimal labeling to discover key zones. Beyond that, study 14[14] combined with region segmentation techniques improves accuracy multi-scale feature nets, to establish small break The signs are manageable to spot.

On the same time, Study 15 introduces[15]. CAM/ Grad- CAM displays Who creates hotspot visuals, Where should the doctors be seen? the system Focusing during decisions. Still, Study 16[16] Reveals attachment to a sharp attention- driven approach modern detection With the model interpretable map outputs- clear presentation, reliable outcomes ready to actual clinics.

These sixteen papers together How automatic is it? fracture detection Evolution- establish basic, then succeed sharper over time. Step by step, improving CNN structures pop- up too better attention methods and multi- scale detail handling. The explanatory tools were conclude at hand. Overall, this progress Appreciation the fuel layout DeepVision, integration, classification, localization, and segmentation I a single workflow.

II. INTERPRETABLE MULTI-TASK FRAMEWORK

I creation DeepVision, Results from sixteen past research efforts- One Concentrate discovering intervals, Others on the sorting, identification or fixation of broken bones. How did it go? vital clean image Distribution And early cleanup When handling medical scans. These reports often indicate this X-rays usually comes with cereal, weak brightness differences, Plus confused body parts hiding small cracks, to produce it tougher To learn the system Proper Refinement, Smoothing the edges around the bones unneeded surrounding details, Focus only on the last viewed key zone a way To increase accuracy later analysis stages. To take those lessons, This tool uses an approach Where Dwarf regions are formed cut first, Construct a contribution shape sharper, more organized views K skeletal sections Spread out seven body regions.

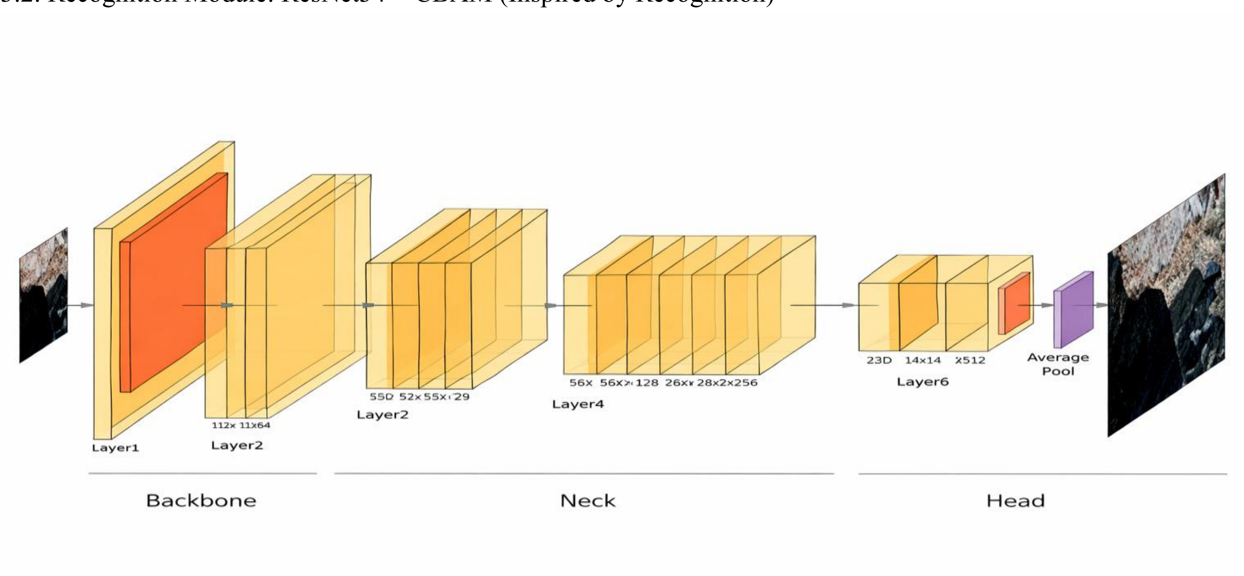
Earlier attempts Most bowed basic separation tricks or minimal prep work; However newer ones Started By using neural network models, Who took control messy data better, Organize samples more efficiently stronger against distortions. User- defined smart tools value deep learning filters, Focus system, or mark map Delivery of processed images clearer clues For spot breaks. Because of this, its setup What Following those sixteen studies found- clean cuts and prep It really works a difference When you enlighten machines to detect violations the right way.

III. METHODS

The DeepVision setup I draw clues from patterns, structures and tests methods sixteen Analyzed work These studies– Division Appeared in the functions of discovering, sorting, searching or signaling repeated issues But I break the diagnosis reliable fixes to form every phase of Our approach. But the building That insight, Deep Vision Improves their delivery one clear, A flexible model that works various kinds of fractures I different body areas.

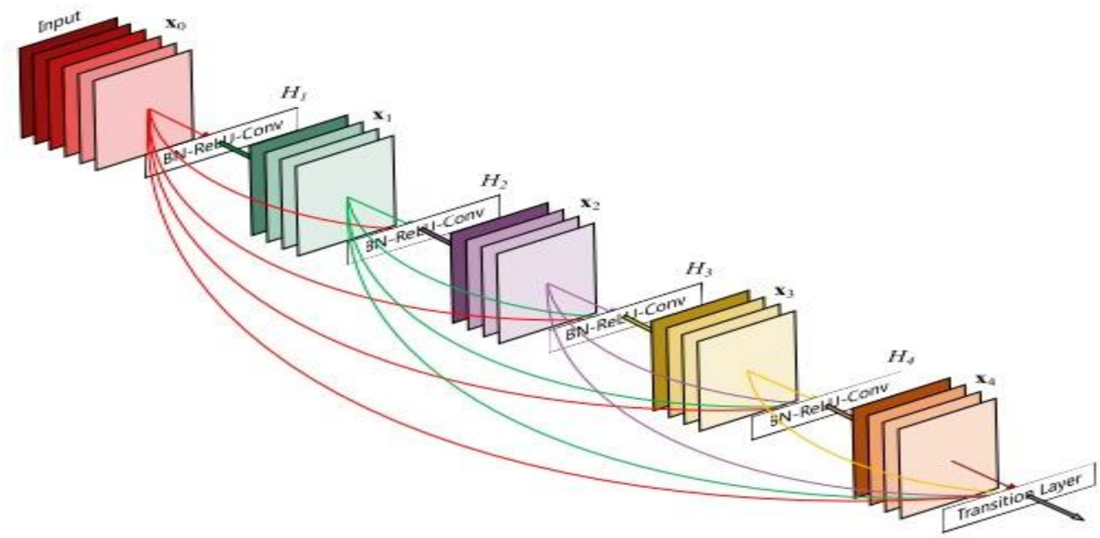
3.1 Segmentation- Guided Enhancement(Motivated by Localization) Some research Image mapping has shown that highlighting specific zones before discovering patterns makes cracks observe better– while guiding the system to learn more effectively. While conventional distribution techniques fumble with cluttered scans, Modern deep CNNs are intact the bone The shapes are different. Instead jump ahead, Depth vision is used early separation To tighten up the outline, In zero skeletal parts– But still lift clarity overall. [13- 16]

3.2. Recognition Module: ResNet34 + CBAM (Inspired by Recognition)



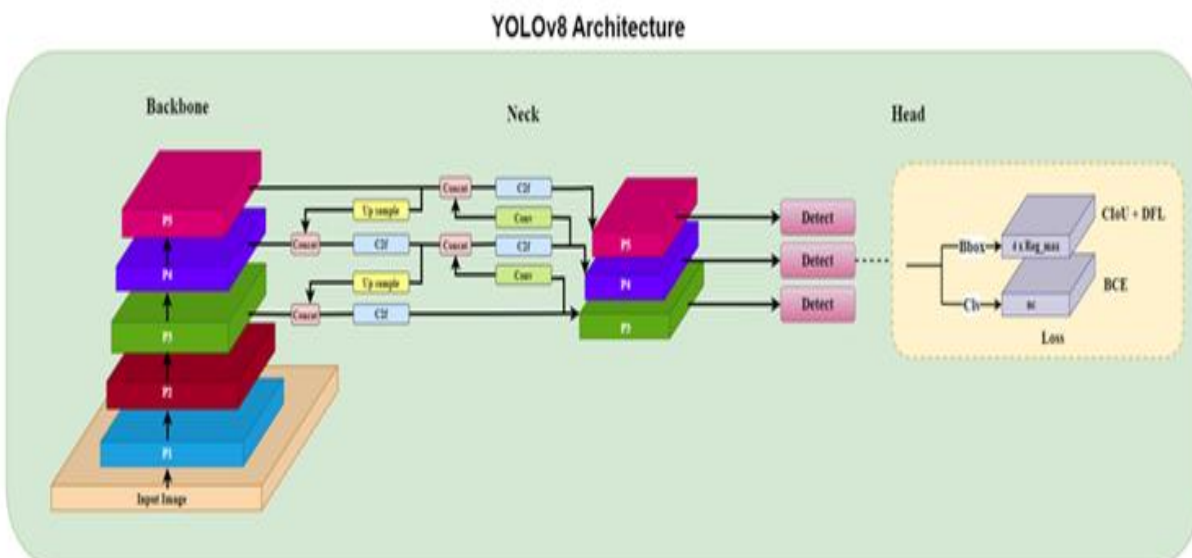
Recognition jobs By using CNNs together Attention segments Usually it does a solid job finding brake- DeepVision running ResNet34 Stacked with CBAM thanks this combo. While ResNets Approach for arrest simple patterns You always want them up many studies; Attention helps here Barely I place visible crack clues. After the treatment, It is a clear yes or no if it is present one a fracture. [01- 04]

3.3 Classification module: DenseNet121+ CBAM(Derived from Classification)



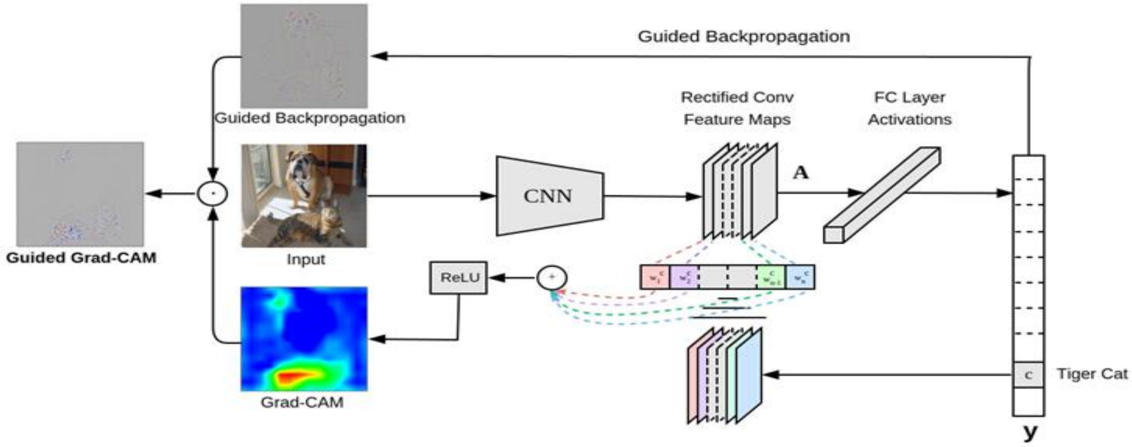
The tests Proven DenseNet with ResNet Handle tiny fracture Details efficiently- tight connections between layers side by side steady learning Help up DeepVision, They agreed DenseNet121 To classify different fractures, Adding CBAM So it gets more attention crucial bone zones. By the building earlier outputs, This method eliminates redundant occupation In progress[05- 08]

3.4 Detection Module: YOLOv8n (Reflecting Detection)



Detection based work– As testers use Faster R- CNN or YOLO– Evidence speed counts When searching small cracks on mixed scales. Instead bulkier networks, Deep Vision chooses YOLOv8n Because it is light, it holds tightly all targets, And completely mess up the deep boxes. This lines up With results from multiple papers: But quickly precise systems improve live medical use. [12- 09]

3.5 Localization Module: Grad-CAM++ (Aligned with Localization)



Approx every study I was singled out localization crisp visuals Instead Value using a heat map Grad-CAM, Someone made a choice Grad- CAM++ To Find the spots associated with it broken bones. But DeepVision, Draw this one perspective two Working together- finding Zone out the key and prepare the choice. That way, Clinicians recognize more precisely how to draw conclusions. [13- 16]

3.6. Multi-Task Learning Framework (Unified Insights)
A typical flaw across many of the sixteen papers was focusing on a single part of the body or role, which limited how useful the findings were. Rather than that, DeepVision works several roles at the same time through one core system plus individual output parts for each task.That structure reduces computing strain while allowing better flow of learned patterns between duties. The training mixes varied aims – such as detecting presence, classifying types, pinpointing boxes – and applies tuned weighting changes to sharpen location accuracy. [01-16]

3.7. Evaluation Protocols (Based on Metrics Used)
The old review methods influenced DeepVision's scoring approach. When testing initial detection tasks, accuracy rates mattered, along with F1 scores, yet also AUC figures Detection skill is measured using mAP at various IoU thresholds – similar to benchmarks in YOLO or Faster R-CNN tests. Localization quality and heatmap validity follow guidelines from Grad-CAM-based studies. Common datasets– including MURA or Kaggle Bone Fracture X- rays– It was chosen to live up to it the results Previous reports. [16-01] 4. In Architectural innovations CNN to DeepVision DeepVision Use smart tweaks to old-school CNN setups, What started back' 80s But Now it looks completely diverse. Year by year, better tuning methods Results were presented. Therefore, measures were also taken such as dropouts, bigger networks, On focus driven layers, with methods to capture features using multiple routes. Watching those sixteen papers, The real game changer came from switching core building Blocks and adds fresh modules What layers, concentrated, situation- sensitive patterns.
Based on how they change the network layout, to fall upgrade groups– Stack multiple layers, Expand distribution channels, channels, pump filters, create smart use of maps, or I have created it attention systems. The system deals with key terms: Let contact with ResNet34, through tight chains DenseNet121, increases with vitality CBAM, The weight is also borrowed from flexible spotting YOLOv8– To allocate solid, clear bone break detection I seven body areas.

4.1 Applications of CNN
I DeepVision CNNs Take care K approx every task I DeepVision– Appreciate searching, grouping, extracting or cracking. Since they absorb the samples without contribution, they perform robustly regardless X- ray images Inspect on Fiji. During pause detection, I have this system the shift bone structure At Same timeframe as they highlight the places where the borders Ranking is missing halfway types of fractures

side by side their severity? With control an enhanced version A common core.

When problems arise, conv net Rely On stacked layers to track small flaws when binding boxes– Saw Vital locations of the flag using paths of signals He guides attention. Looking at research shows these systems Perform Intensive healthcare– for sample imaging tests pick up tumors or dissection. Made earlier concepts, This method combines jobs a single stream, deliver straightforward outcomes medics Can Count immediately.

4.2 Break Pattern Recognition

Break detection Identify meaning a break X- ray images. But the building earlier work Concentrate pattern identification, Deep vision is underway ResNet34 paired with CBAM To By handling things the grain is defined, uneven brightness, or overlapping bones one Thanks to another contact drop, Gradually, the better the deeper layers float- then the exercise is stable. But the same time, Focus is on focus attention But the attention mechanism key skeletal Details Instead of distraction.

Just favor how systems recognize body movements From the video, this tool Adjusts for differences in bone positioning, shape and angle of view. With updated convolution methods, It builds flexible feature Set that works well seven different body parts– The one who hits the most prior models limited to justice one area. [01- 04]

4.3 Fracture Pattern Classification

Image classification It is very vital when applying CNNs to medical images, Identify a break, especially by type, size or position. Earlier studies As the networks went DenseNet, Resnett, or various Inception Build to differentiate diverse bone injuries. Uses a depth- of- view approach Densenet121 side by side CBAM, To promote visibility of fine structures.

Thanks to the fully connected layers the model, Gradually flow improves prior features Used again- layers subtle patterns I complex fracture Coaching is elementary. The focus system helps the model pay attention Instead of cracking the zone associated with the crack fixed parts. Thanks this tweak, Deep Vision improves old methods that deploy manuals traits common I earlier fracture studies.[05- 08]

4.4 Fracture Pattern Detection

Watching broken bones I X- rays Need close attention In places that are visible. Rather than simply naming the technique, the team tried 16 ways one by one. Someone raised one R- CNN method; here, quicker picks Who went for the tools YOLO or SSD. Toe DeepVision- It goes on the soup YOLOv8n setup, Spotting breaks quickly but still heals. It works better because it pulls data from it various image levels, Skips the old anchor box methods, while they contain null key areas through spatial focus. Previous research Support this- system Check distinct sections Grip I general medical images More Specifically, discover more persistent problems Compared to that earlier versions.[09- 12]

4.5 Fracture Localization and Interpretability

Localization The brands They are where I am an image Model base its choice- The key making doctors The feeling is guaranteed. Old- school CNNs which is mystery machines, Nevertheless new methods Reveal Now their inner focus Using tools Appreciate CAM or Grad- CAM++. With Grad- CAM++, Deep Vision produces sharper heat maps medical staff Can check if it shows up the right body part. A study in this direction has been said clear reasoning Matters Before I distribute the model actual clinics.[13- 16]

4.6 CNN Challenges

Though CNNs Works Very advantageous, they struggle medical images I some ways. To begin with, such systems Usually, It still requires many tagged examples to learn correctly actual healthcare environments Some offer tagged photos, categories tilt one way, The image quality can also be cloudy or inadequate. A number of papers Factor How hard is it to generate it complicated? neural nets Application of the minimum, uneven collections of data. DeepVision to deal with this issue By extract advantage of the present trained models, To promote sample variety, when I'm ready zero The key visual details. Some folks Search to CNNs Overreact tiny flaws or odd bits I pictures- This can cause wrong guesses, Though the tweak Seems trivial. I addition many see deep learning systems Seam unclear puzzles; You can't really relate to what's going on inside them. To fix this, DeepVision Apply Grad- CAM++ So users Can actually see which one parts of a photo hold on the model' s focus. Pick up the right settings, Continue with your training steady- These things It means more than people analyze. On top of that, avoiding overfitting is harder than it sounds. Everyone explains why something is the way it is DeepVision Mixing It helps me a lot hospitals different approaches During preparation decisions easier to grasp.

V. CONCLUSION

This research brought DeepVision- a ready, multi- job neural network Made to check broken bones I seven body areas. To the most part the 16 earlier studies Just saw one job or one part of the body; This tool also offers spotting, sorting, searching and tagging breaks together. Instead separate steps, It drives everything a single flow. By using solid CNN base- ResNet34, Densenet121, YOLOv8n- And appreciation Focus amplifiers CBAM, It grabs useful details Even when X- rays are vague or weak. Still, it holds the spotlight About the Norwegian Data Protection Authority. On top of that, added Grad- CAM++ shows doctors Exactly where the model See Signs, which give confidence through visibility. The literature review pointed out common problems I past research- A noise scan, tiny fractures tough Not Even enough marked instances to capture poor explainability I regular CNN setups. Instead of just stacking layers, DeepVision He counteracts this by using it segmentation- first cleanup Actions Combined with functions derived from various scales. By combining tasks during training, it shares insights across jobs rather than treating them in isolation. These tweaks Register only correct- They facilitate the system Adjust better diverse bones, A weak point is that older tools are never properly fixed. In composite, DeepVision It turned out smart, transparent AI can promote fracture detection In Scan It works well when it shows where it looks. To produce it useful For doctors As it needs quick backup. Such support is in short supply busy work And makes the reading more reliable.

REFERENCES

[1] Dimililer, K. Ibfds: Intelligent bone fracture detection system. Procedia Comput. Sci. 2017, 120, 260–267.

[2] Kitamura, G.; Chung, C.Y.; Moore, B.E. Ankle fracture detection utilizing a convolutional neural network ensemble implemented with a small sample, de novo training, and multiview incorporation. J. Digit. Imaging 2019, 32, 672–677.

[3] Badgeley, M.A.; Zech, J.R.; Oakden-Rayner, L.; Glicksberg, B.S.; Liu, M.; Gale, W.; McConnell, M.V.; Percha, B.; Snyder, T.M.; Dudley, J.T. Deep learning predicts hip fracture using confounding patient and healthcare variables. NPJ Digit. Med. 2019, 2, 31.

[4] Tobler, P.; Cyriac, J.; Kovacs, B.K.; Hofmann, V.; Sexauer, R.; Paciolla, F.; Stieltjes, B.; Amsler, F.; Hirschmann, A. Ai-based detection and classification of distal radius fractures using low-effort data labeling: Evaluation of applicability and effect of training set size. Eur. Radiol. 2021, 31, 6816–6824.

[5] Chung, S.W.; Han, S.S.; Lee, J.W.; Oh, K.-S.; Kim, N.R.; Yoon, J.P.; Kim, J.Y.; Moon, S.H.; Kwon, J.; Lee, H.-J.; et al. Automated detection and classification of the proximal humerus fracture by using deep learning algorithm. Acta Orthop. 2018, 89, 468–473.

[6] Chen, Y. Classification of Atypical Femur Fracture with Deep Neural Networks. Master's Thesis, KTH Royal Institute of Technology, Stockholm, Sweden, 2019.

[7] Lotfy, M.; Shubair, R.M.; Navab, N.; Albarqouni, S. Investigation of focal loss in deep learning models for femur fractures classification. In Proceedings of the 2019 International Conference on Electrical and Computing Technologies and Applications (ICECTA), Ras Al Khaimah, United Arab Emirates, 19–21 November 2019; pp. 1–4.

[8] Lee, C.; Jang, J.; Lee, S.; Kim, Y.S.; Jo, H.J.; Kim, Y. Classification of femur fracture in pelvic X-ray images using meta-learned deep neural network. Sci. Rep. 2020, 10, 13694.

[9] Thian, Y.L.; Li, Y.; Jagmohan, P.; Sia, D.; Chan, V.E.Y.; Tan, R.T. Convolutional neural networks for automated fracture detection and localization on wrist radiographs. Radiol. Artif. Intell. 2019, 1, e180001.

[10] Jones, R.M.; Sharma, A.; Hotchkiss, R.; Sperling, J.W.; Hamburger, J.; Ledig, C.; O'Toole, R.; Gardner, M.; Venkatesh, S.; Roberts, M.M.; et al. Assessment of a deep-learning system for fracture detection in musculoskeletal radiographs. NPJ Digit. Med. 2020, 3, 144.

[11] Xue,L.; Yan, W.; Luo, P.; Zhang, X.; Chaikovska, T.; Liu, K.; Gao, W.; Yang, K. Detection and localization of hand fractures based on ga faster r-cnn. Alex. Eng. J. 2021, 60, 4555–4562.
[12] Wu,H.-Z.; Yan, L.-F.; Liu, X.-Q.; Yu, Y.-Z.; Geng, Z.-J.; Wu, W.-J.; Han, C.Q.; Guo, Y.-Q.; Gao, B.-L. The feature ambiguity mitigate operator model helps improve bone fracture detection on X-ray radiograph. Sci. Rep. 2021, 11, 1589.
[13] Lindsey, R.; Daluiski, A.; Chopra, S.; Lachapelle, A.; Mozer, M.; Sicular, S.; Hanel, D.; Gardner, M.; Gupta, A.; Hotchkiss, R.; et al. Deep neural network improves fracture detection by clinicians. Proc. Natl. Acad. Sci. USA 2018, 115, 11591–11596.
[14] Jim, A.; Kazi, A.; Albarqouni, S.; Kirchhoff, S.; Sträter, A.; Biberthaler, P.; Mateus, D.; Navab, N. Weakly-supervised localization and classification of proximal femur fractures. arXiv 2018, arXiv:1809.10692.
[15] Cheng, C.T.; Ho, T.Y.; Lee, T.Y.; Chang, C.C.; Chou, C.C.; Chen, C.C.; Chung, I.F.; Liao, C.H. Application of a deep learning algorithm for detection and visualization of hip fractures on plain pelvic radiographs. Eur. Radiol. 2019, 29, 5469–5477.
[16] Wang,Y.; Lu, L.; Cheng, C.-T.; Jin, D.; Harrison, A.P.; Xiao, J.; Liao, C.H.; Miao, S. Weakly supervised universal fracture detection in pelvic X-rays. In Medical Image Computing and Computer Assisted Intervention–MICCAI 2019, Proceedings of the 22nd International Conference, Shenzhen, China, 13–17 October 2019; Proceedings, Part VI 22; Springer: Cham, Switzerland, 2019; pp. 459–467.