Journal of Iranian Association of Electrical and Electronics Engineers
Vol. 22 No. 4, 2025
DOI: 10.61882/jiaeee.22.4.60

Lightweight Bone Fracture Detection in Radiographic Images Using an Optimized MobileNetV3 Approach

MohammadSam Andi[1], Payam Mahmoudi-Nasr[2]

[1] BSc. Student, Department of Computer Engineering, Faculty of Engineering and Technology, University of Mazandaran, Babolsar, Iran
m.andi05@umail.umz.ac.ir
[2]Associate Professor., Department of Computer Engineering, Faculty of Engineering and Technology, University of Mazandaran, Babolsar, Iran
p.mahmoudi@umz.ac.ir

Abstract:
Bone fractures are among the most common skeletal injuries, particularly in high-pressure sports environments and crowded emergency departments, where rapid and accurate diagnosis is of great importance. Delays or errors in fracture detection may lead to impaired healing, mobility limitations, and increased treatment costs. In this study, we propose a lightweight system based on an optimized version of MobileNetV3 to classify radiographic images into two categories: fractured and non-fractured. A dataset of 10,580 radiographs from different anatomical regions was used for training. The images were resized to 224×224 pixels, normalized, and fed into the network. The architecture was redesigned by incorporating global average pooling, fully connected layers, and regularization techniques to mitigate overfitting. During training, the AdamW optimizer with a learning rate of 0.0001 was employed, where the initial layers were frozen and the final layers were fine-tuned to improve performance. Evaluation on 506 test images demonstrated an accuracy of 99.21%, which is comparable to or even superior to some heavier architectures. Owing to its lightweight nature, the proposed system can be deployed on low-resource devices such as smartphones and medical tablets, making it a fast and reliable assistant for radiologists in emergency rooms, clinics, and crowded clinical settings.

Keywords: Fracture detection; Radiographic images; Deep learning; MobileNetV3; Binary classification

Article Type: Research

Received: 21. 08. 2025

Revised: 07. 09. 2025

Accepted: 03. 11. 2025

Corresponding author: Payam Mahmoudi-Nasr

Corresponding author's address: Department of Computer Engineering, Faculty of Engineering and Technology, University of Mazandaran, Babolsar, Iran

Copyright © 2025 The Authors. Published by Iranian Association of Electrical and Electronics Engineers.
This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International license (https://creativecommons.org/licenses/by-nc/4.0/). Non-commercial uses of the work are permitted, provided the original work is properly cited.

1. Motivation of the work

Bone fractures represent one of the most common musculoskeletal injuries, frequently occurring in high-impact sports and emergency settings. Accurate and timely fracture detection is critical for effective treatment, as delays or misdiagnoses can lead to complications such as delayed healing, reduced mobility, and increased healthcare costs. Despite radiologists' expertise, diagnostic errors remain prevalent due to fatigue and workload in busy clinical environments. In recent years, deep learning has demonstrated remarkable capabilities in medical image analysis, but many high-performing architectures such as DenseNet and ResNet are computationally expensive and unsuitable for devices with limited processing power. The motivation for this research stems from the need to develop a highly accurate yet lightweight system capable of real-time inference in clinical and mobile contexts. MobileNetV3-Small was selected for its balance between accuracy and efficiency, and the architecture was further optimized to adapt to the unique characteristics of radiographic data. The goal was to design a model that can maintain diagnostic accuracy comparable to complex architectures while being fast, portable, and feasible for deployment on low-resource devices. This approach can support radiologists in rapid decision-making, improve workflow efficiency, and enhance diagnostic reliability in emergency and outpatient settings.

2. Contributions

This research contributes to the field of medical image classification by optimizing a mobile-scale convolutional neural network for fracture detection. The proposed model improves MobileNetV3-Small through structural refinements and training enhancements specifically tailored to radiographic data. The Flatten layer was replaced by Global Average Pooling to reduce parameters and prevent overfitting, while two dense layers with 256 and 128 neurons enhanced feature abstraction. Batch Normalization ensured stable learning, and Dropout with a rate of 0.5 improved generalization. The AdamW optimizer, with controlled weight decay, promoted stable convergence and prevented excessive weight growth. The early layers were frozen to preserve general low-level features, and only higher layers were fine-tuned to capture fracture-specific characteristics. A diverse dataset of 10,580 labeled radiographs was utilized to ensure robustness and generalization. The system achieved a remarkable accuracy of 99.21%, exceeding or matching larger architectures such as DenseNet201 and YOLO-based models while requiring significantly fewer computational resources. This balance between performance and efficiency makes it suitable for real-time clinical deployment and edge computing applications. The research therefore bridges the gap between high diagnostic accuracy and practical deployability in resource-limited healthcare environments.

3. Procedures

The experimental process began with dataset preparation and preprocessing. The dataset contained 10,580 radiographs from different body parts including hands, feet, pelvis, lumbar, and knees. The data were divided into training, validation, and test subsets with 9,246, 828, and 506 samples, respectively. All images were resized to 224×224 pixels using bilinear interpolation and normalized to the range [-1,1] to match the expected input of ImageNet-pretrained networks. Data augmentation, including rotation, flipping, and scaling, was applied to increase diversity and improve model generalization. The base model used was MobileNetV3-Small, pretrained on ImageNet. Its early layers were frozen to retain general visual feature extraction capabilities, while upper layers were fine-tuned to specialize in detecting fracture patterns. The Flatten layer was replaced by Global Average Pooling to minimize parameters, followed by dense layers with ReLU activations, Batch Normalization, and a Dropout rate of 0.5 to regularize training. The network was trained using the AdamW optimizer with a learning rate of 0.0001 and weight decay of 0.0001, while the loss function used was Sparse Categorical Cross-Entropy. Early stopping was applied after 21 epochs when validation accuracy stabilized, preventing overfitting. The training was performed using two NVIDIA T4 GPUs in the Kaggle computational environment. Evaluation metrics included Accuracy, Precision, Recall, and F1-score, complemented by a confusion matrix to analyze misclassifications. Comparative analysis was conducted against existing studies utilizing architectures such as DenseNet121, DenseNet201, YOLO-v10, and MobLG-Net to position the proposed model's performance relative to contemporary benchmarks.

4. Findings

The optimized MobileNetV3-Small achieved 99.21% classification accuracy on the test dataset, correctly identifying 502 of 506 samples. The model's Recall of 99.63% and Precision of 98.74% reflect strong sensitivity to fracture detection and high specificity toward non-fractured samples. The F1-score of 99.26% indicates a balanced and robust performance. Training and validation curves demonstrated stable convergence without oscillations or overfitting, confirming the effectiveness of the optimization and regularization strategies. The confusion matrix analysis revealed minimal false negatives, an essential outcome in clinical scenarios where missing a fracture is more critical than false alarms. Compared to previous research, the model outperformed DenseNet201, which achieved 97.35% accuracy, and YOLO-v10, which achieved 96.4%. The improvements are attributed to architectural simplification, adaptive fine-tuning, and balanced regularization. The lightweight nature of the model enables its deployment on mobile devices without sacrificing diagnostic precision. This result confirms that a carefully tuned lightweight model can rival or exceed

the accuracy of deeper networks while remaining practical for real-time medical applications, particularly in emergency departments and low-resource clinics where computational capacity is limited.

5. Conclusion

This study developed an optimized MobileNetV3-Small-based system for accurate and efficient bone fracture detection in radiographic images. The method integrates structural, training, and regularization improvements that collectively enhance model stability, accuracy, and computational efficiency. Achieving an accuracy of 99.21%, the proposed approach demonstrates that lightweight convolutional architectures can perform competitively with deep, high-parameter models. Its low memory footprint and high inference speed make it a practical solution for on-device deployment in medical environments, providing radiologists with fast decision support tools in emergency rooms and outpatient clinics. Future extensions may focus on multi-task architectures for simultaneous localization and classification of fractures, validation on larger and more diverse clinical datasets, and improved interpretability mechanisms to strengthen clinical trust and integration into real-world diagnostic workflows.

نشریه مهندسی برق و الکترونیک ایران

نشریه مهندسی برق و الکترونیک ایران، سال بیست و دوم، شماره چهارم، زمستان ۱۴۰۴، صفحات ۶۰-۶۸

DOI: 10.61882/jiaeee.22.4.60

تشخیص خودکار شکستگی استخوان در تصاویر رادیوگرافی با رویکرد سبک‌وزن مبتنی بر MobileNetV3 بهینه‌شده

محمدسام اندی[۱]، پیام محمودی نصر[۲]

۱- دانشجوی کارشناسی- گروه مهندسی کامپیوتر- دانشکده مهندسی و فناوری ـ دانشگاه مازندران ـ بابلسر ـ ایران

m.andi05@umail.umz.ac.ir

۲- دانشیار- گروه مهندسی کامپیوتر- دانشکده مهندسی و فناوری ـ دانشگاه مازندران ـ بابلسر ـ ایران

p.mahmoudi@umz.ac.ir

چکیده: **شکستگی‌های استخوان از جمله آسیب‌های شایع اسکلتی هستند که در محیط‌های ورزشی پرفشار و اورژانس‌های پرمراجعه تشخیص سریع و دقیق آن‌ها اهمیت ویژه‌ای دارد؛ زیرا تأخیر یا اشتباه در شناسایی می‌تواند به تاخیر در التیام، ناتوانی حرکتی و افزایش هزینه‌های درمانی منجر شود. در این پژوهش سامانه‌ای سبک مبتنی بر نسخه اصلاح‌شده MobileNetV3 معرفی می‌شود که تصاویر رادیوگرافی را با دقت و سرعت بالا در دو گروه شکستگی و غیرشکستگی طبقه‌بندی می‌کند. برای آموزش از ۱۰۵۸۰ تصویر رادیوگرافی مربوط به نواحی مختلف بدن استفاده شد و داده‌ها پس از تغییر اندازه به ۲۲۴×۲۲۴ و نرمال‌سازی به شبکه داده شدند. معماری با به‌کارگیری میانگین‌گیری سراسری، لایه‌های تمام‌متصل و روش‌های کاهش بیش‌برازش بازطراحی شد و در آموزش از AdamW با نرخ یادگیری ۰.۰۰۰۱ استفاده گردید؛ لایه‌های ابتدایی فریز و لایه‌های انتهایی به‌صورت هدفمند تنظیم شدند تا عملکرد مدل بهبود یابد. ارزیابی بر روی ۵۰۶ تصویر آزمون نشان داد دقت ۹۹.۲۱٪ حاصل شده است و این مقدار با برخی معماری‌های سنگین‌تر قابل رقابت یا برتر است. ماهیت سبک سامانه، اجرای آن را بر روی دستگاه‌های کم‌منبع مانند تلفن هوشمند و تبلت پزشکی ممکن می‌سازد و می‌تواند به‌عنوان دستیار سریع رادیولوژیست‌ها در اورژانس، کلینیک‌ها و محیط‌های شلوغ بالینی به‌کار گرفته شود.**

کلمات کلیدی: **تشخیص شکستگی، تصاویر رادیوگرافی، یادگیری عمیق، MobileNetV3، طبقه‌بندی دودسته‌ای**

نوع مقاله: پژوهشی

دریافت: ۱۴۰۴/۰۵/۳۰

بازنگری: ۱۴۰۴/۰۶/۱۶

پذیرش: ۱۴۰۴/۰۸/۱۲

نام نویسنده‌ی مسئول: دکتر پیام محمودی نصر

نشانی نویسنده‌ی مسئول: ایران – مازندران – بابلسر – دانشگاه مازندران – دانشکده‌ی مهندسی و فناوری

حق انتشار این مستند، متعلق به نویسندگان آن است. ۱۴۰۴ ©. ناشر این مقاله، انجمن مهندسین برق و الکترونیک ایران است.

این مقاله تحت گواهی زیر منتشرشده و هر نوع استفاده غیرتجاری از آن مشروط بر استناد صحیح به مقاله و با رعایت شرایط مندرج در آدرس زیر مجاز است.

Creative Commons Attribution-NonCommercial 4.0 International license
(https://creativecommons.org/licenses/by-nc/4.0/)

۱- مقدمه

شکستگی استخوان، اعم از ترک‌خوردگی ناقص یا شکست کامل، یکی از شایع‌ترین آسیب‌های اسکلتی در سراسر جهان است و سالانه میلیون‌ها نفر را درگیر می‌کند [۱]. در حوزه‌ی ورزش، به‌ویژه طی فعالیت‌های پرفشار یا در اثر وارد شدن ضربات ناگهانی و شدید، احتمال وقوع شکستگی‌های استخوانی به‌طور محسوسی افزایش می‌یابد و این نوع آسیب‌ها می‌تواند به دوره‌های طولانی‌مدت بهبودی و کاهش توانایی عملکرد جسمانی منجر شود [۲]. تشخیص به‌موقع و دقیق این آسیب‌ها، برای تعیین استراتژی درمانی مناسب و جلوگیری از عوارضی نظیر تاخیر در جوش‌خوردن یا ناتوانی حرکتی ضروی است. اگرچه تفسیر تصاویر رادیوگرافی توسط رادیولوژیست‌ها استاندارد بالایی در این فرایند محسوب می‌شود، اما خستگی بالینی و افزایش حجم مراجعات می‌تواند منجر به بروز خطا شود؛ به‌طوری که در مطالعات چندمرکزی، شکستگی‌های تشخیص‌داده‌نشده به‌عنوان یکی از پرتکرارترین خطاهای تشخیصی در بخش اورژانس گزارش گردیده است [۳].

در سال‌های اخیر، پیشرفت‌های قابل‌ملاحظه در زمینه‌ی یادگیری عمیق چشم‌انداز جدیدی برای بهبود دقت و سرعت در تحلیل تصاویر پزشکی فراهم کرده است. مدل‌های مبتنی بر شبکه‌های عصبی پیچشی[۱]، با بهره‌گیری از مجموعه‌داده‌های گسترده و ظرفیت بالا در پردازش موازی، قادر به استخراج ویژگی‌های ظریف و پیچیده از تصاویر رادیوگرافی هستند. نتایج پژوهش‌های متعدد نشان می‌دهد که این سامانه‌ها می‌توانند با دستیابی به دقت تشخیصی بالاتر نسبت به روش‌های سنتی، نقش مؤثری در کاهش خطاهای تشخیصی و تسهیل روند درمانی ایفا نمایند [۴].

در این پژوهش از ۱۰۵۸۰ تصویر رادیوگرافی[۲] مربوط به وضعیت‌های شکستگی و غیرشکستگی استخوان‌های بدن مورد استفاده قرار گرفت. این تصاویر نواحی مختلفی از بدن مانند دست، پا، لگن، کمر و زانو را دربرمی‌گیرند. به منظور افزایش تنوع و حجم داده‌ها، از تکنیک‌های داده‌افزایی[۳] نظیر چرخش، قرینه‌سازی[۴] و سایر تبدیلات تصویری استفاده شد. به‌کارگیری داده‌افزایی با افزایش تنوع نمونه‌های آموزشی، خطر بیش‌برازش[۵] را کاهش داده و توانایی مدل برای تعمیم روی تصاویر ناشناخته را تقویت می‌کند. در ادامه نسخه اصلاح‌شده MobileNetV3 برای طبقه‌بندی دودسته‌ای به کار گرفته شد.

ساختار مقاله بدین ترتیب است: بخش ۲ پیشینه را مرور می‌کند، در بخش ۳ مجموعه‌داده معرفی می‌شود، بخش ۴ روش پیشنهادی را توضیح می‌دهد، بخش ۵ نتایج و تحلیل‌ها را ارائه می‌کند و در بخش ۶ جمع‌بندی می‌آید.

۲- پیشینه پژوهش

در سال‌های اخیر، مرورهای متعددی به بررسی کاربرد یادگیری ماشین و یادگیری عمیق در شناسایی شکستگی‌های استخوان پرداخته‌اند. به‌عنوان نمونه، کوتبی و همکاران (۲۰۲۴) در یک مرور نظام‌مند، پژوهش‌های مرتبط با به‌کارگیری هوش مصنوعی در تشخیص شکستگی استخوان را بررسی کردند. آنان نتیجه گرفتند که مدل‌های یادگیری عمیق نسبت به روش‌های سنتی عملکرد دقیق‌تری دارند و ترکیب آن‌ها با تصویربرداری سه‌بعدی می‌تواند دقت تشخیص را افزایش دهد. همچنین به نقش فناوری‌های نوینی مانند هوش مولد و مدل‌های زبانی بزرگ در بهبود فرایندهای بالینی اشاره کردند [۵].
سو و همکاران (۲۰۲۳) نیز با مرور ۴۰ مقاله‌ی منتخب، چارچوبی پردازشی برای تحلیل شکستگی‌ها ارائه داده و بر ضرورت تفسیرپذیری مدل‌ها، ادغام داده‌های بالینی و توسعه‌ی ابزارهای بصری در کاربردهای پزشکی تاکید داشتند [۶].

در زمینه‌ی روش‌های خاص تشخیص، آلام و همکاران (۲۰۲۵) مدل MobLG-Net را با ترکیب ساختار MobileNet و طبقه‌بندی LightGBM عرضه نمودند که دقت ۹۹٪ را در تفکیک تصاویر شکسته از سالم گزارش کرد [۷]. عبدالصمدو و همکاران (۲۰۲۵) چارچوبی سبک‌وزن مبتنی بر DenseNet121 در کنار آشکارساز کنای و پیش‌پردازش‌های تقویتی ارائه کردند که در دیتاست رادیوگرافی ورزشی به دقت ۹۰.۳٪ و حساسیت ۸۹٪ دست یافتند [۸]. سرینیواسو و همکاران (۲۰۲۵) با سفارشی‌سازی YOLOv10 و با به‌کارگیری ماسک تیزکننده و تطبیق کنتراست، دقت ۹۶.۴٪ را در تشخیص انواع شکستگی گزارش کردند [۹] و هوآ و همکاران (۲۰۲۵) با مقایسه بیست معماری مرسوم مانند VGG، ResNet، DenseNet و MobileNet دریافتند DenseNet با F1-Score برابر ۷۹٪ بهترین عملکرد را دارد [۱۰].

توسعه روش‌های داده‌افزایی نقش مهمی در بهبود اغلب این سامانه‌ها داشته است؛ برای مثال، عبدالصمدو و همکاران با اعمال چرخش، مقیاس‌دهی و قرینه‌سازی بر روی حدود ۵۰۰۰ تصویر رادیوگرافی ورزشی تنوع داده‌ها را افزایش دادند [۸].

علاوه بر ارتقای دقت خودکار، مطالعاتی مانند کوهن و همکاران (۲۰۲۳) نشان داده‌اند که ترکیب تحلیل یادگیری عمیق با ارزیابی رادیولوژیست‌ها می‌تواند حساسیت تشخیص را از ۷۶٪ (فقط رادیولوژیست) و ۸۳٪ (فقط سیستم هوش مصنوعی) به ۸۸٪ برساند [۱۱].

علاوه بر مطالعات فوق، آلدیانی و همکاران (۲۰۲۵) با بررسی مجموعه‌ای شامل ۱۰۵۸۰ تصویر رادیوگرافی از نواحی مختلف بدن، به مقایسه عملکرد چندین معماری مشهور از جمله VGG16 و DenseNet201 پرداختند و نشان دادند که مدل DenseNet201 با کسب دقت ۹۷.۳۵٪، حساسیت ۹۷.۷۸٪ و F1-Score برابر ۹۷.۴۱٪ بهترین کارایی را در تشخیص دودسته‌ای شکستگی/غیرشکستگی ارائه می‌دهد. نتایج این مطالعه تأکید می‌کند که استفاده از مدل‌های عمیق با ظرفیت بالای استخراج ویژگی، می‌تواند به طرز چشمگیری دقت تشخیص شکستگی را در کاربردهای بالینی بهبود بخشد [۱۲]. از آنجا

که مجموعه داده به‌کار گرفته‌شده در مطالعه‌ی آلدیانی و همکاران (۲۰۲۵) با مجموعه داده‌ی مورد استفاده در پژوهش حاضر یکسان است، نتایج آن‌ها می‌تواند به‌عنوان مبنای مناسبی برای مقایسه با نتایج حاصل از مدل پیشنهادی این پژوهش به کار رود.

بدین ترتیب، هرچند یادگیری عمیق ظرفیت افزایش قابل‌توجه دقت تشخیص شکستگی را دارد، هنوز بهینه‌سازی معماری‌ها، تضمین قابلیت تعمیم‌پذیری در محیط‌های بالینی گوناگون و ارتقای شفافیت تصمیم‌گیری مدل‌ها چالش‌های اصلی پیش‌رو باقی مانده‌اند.

۳- مجموعه داده

در این پژوهش، به منظور توسعه و ارزیابی مدل‌های یادگیری عمیق در تشخیص شکستگی‌های استخوانی از مجموعه‌داده‌ای ترکیبی و غنی بهره برده‌ایم. منظور از ۱۰۵۸۰ تصویر، نمونه‌های رادیوگرافی دوبعدی است که از مجموعه‌داده‌ی Fracture Multi-Region X-ray Data (Kaggle) گردآوری شده‌اند. هر تصویر نمایی از وضعیت شکستگی یا غیرشکستگی در یکی از نواحی بدن است. این تصاویر، که در حالت نمای قدامی پس‌سری[۶] و جانبی[۷] ثبت شده‌اند، از کیفیت و وضوح مناسبی برخوردار بوده و تنوع گسترده‌ای از انواع شکستگی‌ها (خطی، مغزی و پیچیده) و نمونه‌های بدون شکستگی را در بر می‌گیرد.

برای تأمین این حجم وسیع از داده و افزایش تنوع نمونه‌ها، سه مجموعه‌داده با یکدیگر ادغام شده‌اند. در نهایت، ۹۲۴۶ تصویر برای فاز آموزش[۸]، ۸۲۸ تصویر برای فاز اعتبارسنجی[۹] و ۵۰۶ تصویر برای مرحله‌ی آزمون[۱۰] انتخاب گردید تا علاوه بر حفظ تعادل کلاس‌ها، امکان برآورد دقیق عملکرد مدل و جلوگیری از بیش‌برازش فراهم آید. تمامی تصاویر پیش از ورود به شبکه‌های عصبی، تحت پردازش‌های پایه مانند تغییر اندازه[۱۱] و نرمال‌سازی[۱۲] قرار گرفتند تا ویژگی‌های بافتی و ساختمانی استخوان‌ها به‌خوبی حفظ شوند.

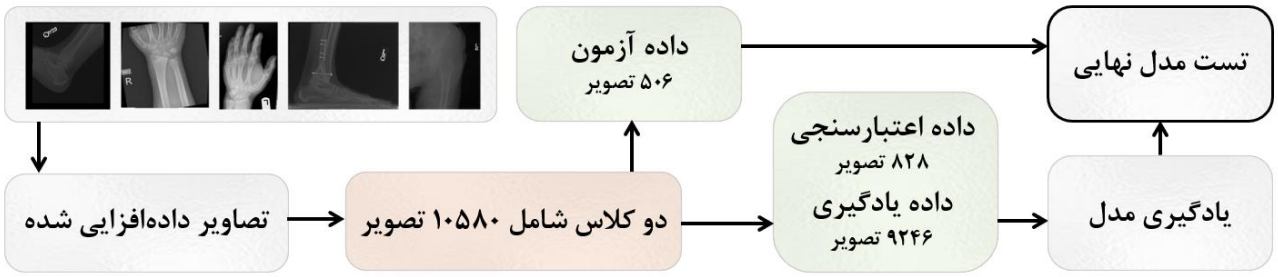
این مجموعه‌داده تحت مجوز Open Data Commons Attribution License (ODC-By) منتشر شده و استفاده، اشتراک‌گذاری و اصلاح آن، با ذکر منابع اصلی، برای پژوهشگران و توسعه‌دهندگان آزاد است. بدین‌ترتیب، با بهره‌گیری از این داده‌های جامع و متنوع، می‌توان گام مؤثری در جهت بهبود دقت و کارایی سامانه‌های تشخیص شکستگی در تصویربرداری پزشکی برداشت. چند نمونه از تصاویر این مجموعه داده در شکل ۱ نشان داده شده است.



شکل (۱): نمونه‌هایی از تصاویر مجموعه‌داده

۴- روش پیشنهادی

روش پیشنهادی با آماده‌سازی داده‌ها آغاز می‌شود؛ تصاویر پس از یکسان‌سازی اندازه و نرمال‌سازی شدت به سه بخش آموزش، اعتبارسنجی و آزمون تقسیم می‌شوند تا ارزیابی بی‌طرفانه‌تری به‌دست آید. سپس نسخه اصلاح‌شده MobileNetV3 به‌عنوان هسته استخراج ویژگی و طبقه‌بندی به کار می‌رود و خروجی شبکه برچسب شکستگی یا غیرشکستگی را بر اساس الگوهای آموخته‌شده تعیین می‌کند. نمای کلی این فرایند در شکل ۲ آمده است.



شکل (۲): نمای کلی پردازش از آماده‌سازی داده تا طبقه‌بندی شکستگی و غیرشکستگی

۴-۱- معماری شبکه عصبی موبایل‌نت

شبکه‌های عصبی کانولوشنی به دلیل توان استخراج الگوهای چندمقیاسی از تصاویر، در مسائل طبقه‌بندی پزشکی کارآمد هستند و در ترکیب با داده‌های کافی می‌توانند جزئیات کم‌کنتراست رادیوگرافی را نیز آشکار کنند. از میان معماری‌های سبک، خانواده MobileNet توازن مناسبی میان دقت و هزینه‌ی محاسباتی فراهم می‌کند و برای سناریوهای کم‌منبع گزینه‌ای عملی به شمار می‌آید. در این پژوهش، تمرکز بر سازگارسازی MobileNetV3 با ویژگی‌های خاص تصاویر رادیوگرافی است تا بدون اتکا به سخت‌افزار سنگین، مدلی دقیق و سریع برای تفکیک شکستگی و غیرشکستگی در اختیار باشد [۱۳]. MobileNetV3 با رویکرد جست‌وجوی خودکار معماری طراحی شده و نسخه Small آن برای کاربردهایی که منابع پردازشی و حافظه محدود است، مناسب‌تر است. در این کار از نسخه Small بهره گرفته‌ایم تا امکان استقرار در محیط‌های کلینیکی و روی دستگاه‌های قابل‌حمل فراهم شود. انتخاب این نسخه به دلیل برقراری تعادل میان سه مؤلفه کلیدی است؛ دقت قابل اتکا در طبقه‌بندی دودسته‌ای، زمان پاسخ کوتاه در هنگام استنتاج و اندازه‌ی مدل که برای اجرا بر بسترهای کم‌منبع قابل مدیریت باقی می‌ماند [۱۴].

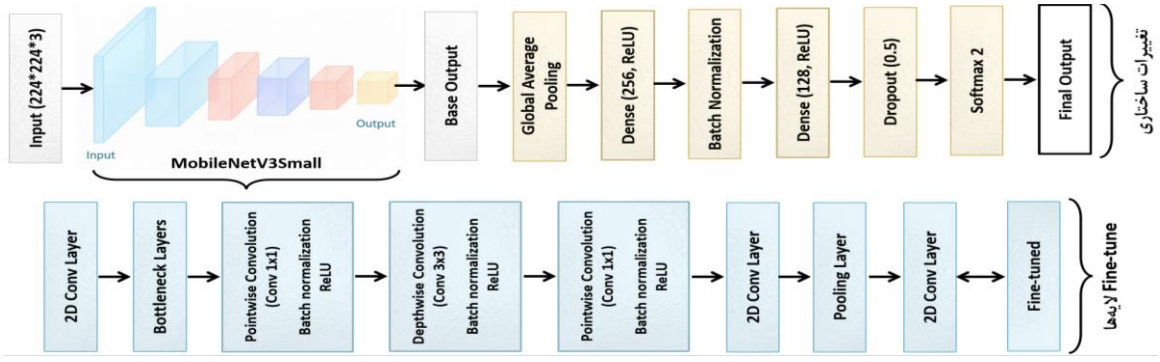
۴-۲- بهینه‌سازی ساختاری شبکه عصبی موبایل‌نت

با اعمال اصلاحات هدفمند بر معماری MobileNetV3، کارایی مدل در طبقه‌بندی تصاویر رادیوگرافی به‌طور معناداری بهبود یافت. مدل پایه MobileNetV3-Small از پیش روی ImageNet آموزش دیده است و ورودی‌های ۲۲۴×۲۲۴ را دریافت می‌کند. اصلاحات اعمال‌شده بر بخش‌های انتهایی و تنظیمات آموزشی در شکل ۳ نشان داده شده‌اند.

در این مدل، تغییرات مهمی اعمال شده است که کارایی آن را به‌ویژه در محیط‌های کم‌منبع مانند تلفن‌های همراه بهبود می‌بخشد. یکی از اصلاحات اساسی در ساختار مدل پیشنهادی، جایگزینی لایه‌ی Flatten با لایه‌ی Global Average Pooling بود. این تغییر موجب شد به جای تولید یک بردار بسیار بزرگ و پرپارامتر، تنها میانگین ویژگی‌های استخراج‌شده در هر کانال محاسبه شود. در نتیجه، هم تعداد پارامترهای شبکه کاهش یافت و هم بار محاسباتی سبک‌تر شد. مهم‌تر آن‌که این تغییر اجازه داد ویژگی‌های کلیدی و معنادار هر کانال حفظ شود، در حالی که پیچیدگی غیرضروری حذف گردید. این اصلاح علاوه بر صرفه‌جویی در منابع، موجب شد مدل در داده‌های آزمون عملکرد پایدارتر و تعمیم‌پذیرتری داشته باشد.

در گام بعدی، برای افزایش ظرفیت شبکه در یادگیری الگوهای پیچیده‌تر، دو لایه Dense به معماری افزوده شد؛ نخست یک لایه با ۲۵۶ نورون و سپس یک لایه با ۱۲۸ نورون، که هر دو با تابع فعال‌سازی ReLU عمل می‌کردند. قرار گرفتن این لایه‌ها باعث شد مدل بتواند روابط غیرخطی و ویژگی‌های ترکیبی را با عمق بیشتری در داده‌ها کشف کند. میان این دو لایه از Batch Normalization استفاده شد تا پایداری یادگیری تضمین گردد و تغییرات ناگهانی در توزیع ورودی‌ها کنترل شود. همچنین، برای جلوگیری از بیش‌برازش، یک لایه Dropout با نرخ ۰.۵ به‌کار گرفته شد که طی آموزش بخشی از نورون‌ها را به‌طور تصادفی غیرفعال می‌کرد. این ترکیب سبب شد شبکه هم قدرت یادگیری بالاتری پیدا کند و هم از افت عملکرد در مواجهه با داده‌های جدید جلوگیری شود.

اصلاح مهم دیگر، استفاده از راهبرد فریزکردن لایه‌های ابتدایی و انجام Fine-Tuning تنها بر روی لایه‌های بالایی بود. لایه‌های آغازین شبکه عمدتا ویژگی‌های عمومی نظیر لبه‌ها و بافت‌های ساده را استخراج می‌کنند که برای تمام تصاویر کاربرد دارند و نیازی به بازآموزی کامل آن‌ها وجود ندارد. در مقابل، لایه‌های بالایی مسئول استخراج ویژگی‌های اختصاصی‌تر و مرتبط با شکستگی در تصاویر رادیوگرافی هستند. در نتیجه‌ی این رویکرد، کاهش حجم محاسبات و صرفه‌جویی در زمان آموزش و افزایش دقت مدل در شناسایی الگوهای تخصصی مرتبط با شکستگی استخوان اتفاق افتاد. این ترکیب از دانش عمومی لایه‌های پایین و یادگیری اختصاصی لایه‌های بالا موجب بهبود دقت مدل و افزایش پایداری آن در داده‌های آزمون شد؛ این ویژگی می‌تواند در کاربردهای بالینی واقعی نیز مفید واقع شود. دسته‌بندی کرد.

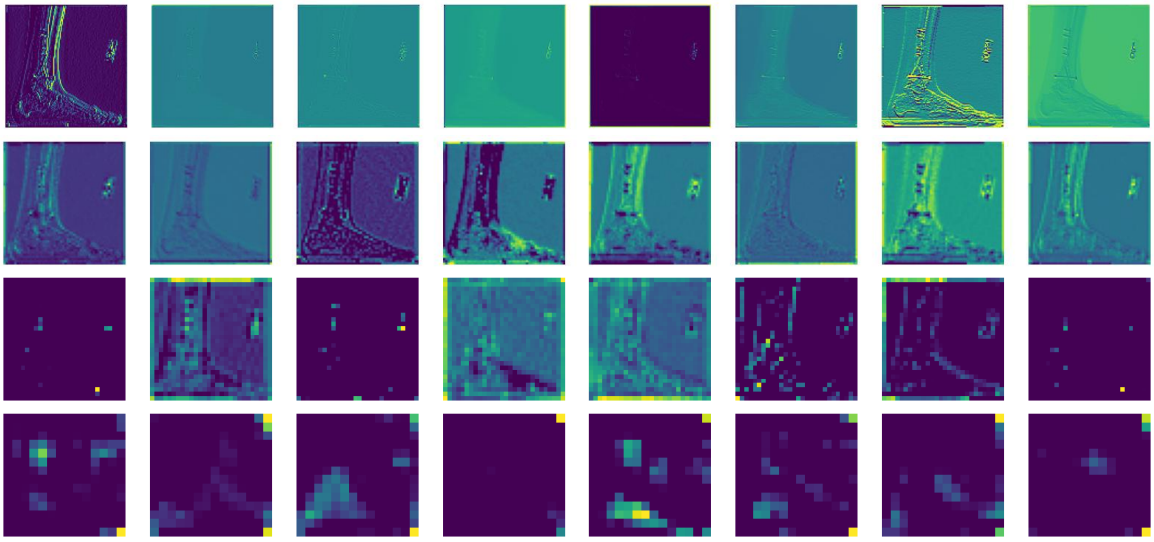


شکل (۳): معماری MobileNetV3-Small به همراه تغییرات اعمال‌شده در نسخه پیشنهادی

برای روشن‌تر شدن نحوه‌ی پردازش داده‌ها در شبکه، در شکل ۴ یک نمونه تصویر ورودی و در شکل ۵ خروجی برخی از لایه‌های اولیه نمایش داده شده است. همان‌گونه که مشاهده می‌شود، لایه‌های ابتدایی بیشتر الگوهای ساده مانند لبه‌ها و ساختارهای بافتی پایه را برجسته می‌کنند. با حرکت به سمت لایه‌های میانی و بالایی، ویژگی‌های پیچیده‌تر مرتبط با ساختار استخوان و شکستگی‌ها شناسایی می‌شوند. این روند به‌خوبی نشان می‌دهد که چگونه شبکه به‌صورت سلسله‌مراتبی از اطلاعات ساده به سمت الگوهای تخصصی حرکت کرده و به درک عمیق‌تری از محتوای تصویری دست می‌یابد.



شکل (۴): نمونه تصویر ورودی از مجموعه‌داده شکستگی استخوان



شکل (۵): خروجی لایه‌های شبکه

۴-۳- بهینه‌سازی شبکه عصبی برای یادگیری

برای پایداری فرایند یادگیری و کنترل پیچیدگی مدل، از بهینه‌ساز[۱۳] AdamW است که بهبودهایی نسبت به Adam دارد. AdamW با اضافه‌کردن کاهش وزن[۱۴]، از رشد بی‌رویه وزن‌های شبکه جلوگیری می‌کند و کمک می‌کند تا شبکه پایدارتر و با دقت بیشتری عمل کند. در این مدل، مقدار Weight Decay روی ۰.۰۰۰۱ تنظیم شده است که به اندازه کافی کوچک است تا وزن‌ها را به‌طور پیوسته کنترل کند و از انباشته شدن وزن‌های بزرگ جلوگیری کند. به‌طور پیش‌فرض، مقدار Weight Decay در بهینه‌ساز Adam برابر با صفر است، که ممکن است باعث رشد بی‌رویه وزن‌های شبکه و منجر به پیچیدگی بیش از حد مدل شود. با اعمال Weight Decay، مدل به‌طور مداوم وزن‌ها را کاهش می‌دهد و این کاهش به جلوگیری از پیچیدگی بیش از حد مدل کمک می‌کند. در نتیجه، شبکه به جای تمرکز بیش از حد روی داده‌های آموزشی، عملکرد بهتری روی داده‌های ناشناخته و جدید پیدا خواهد کرد. همچنین، نرخ یادگیری به جای مقدار پیش‌فرض ۰.۰۰۱ به ۰.۰۰۰۱ کاهش داده شده است. این کاهش باعث شد فرایند همگرایی با گام‌های کوچک‌تر و کنترل‌شده‌تری پیش برود. اگرچه نرخ‌های بالاتر می‌توانند در آغاز یادگیری شتاب ایجاد کنند، اما غالبا در مراحل پایانی موجب نوسان و بی‌ثباتی می‌شوند. بنابراین، با انتخاب نرخ یادگیری[۱۵] پایین‌تر، مدل در طول آموزش به‌صورت تدریجی و پایدار به نقطه بهینه نزدیک شد و در نهایت دقت بالاتری به دست آمد.

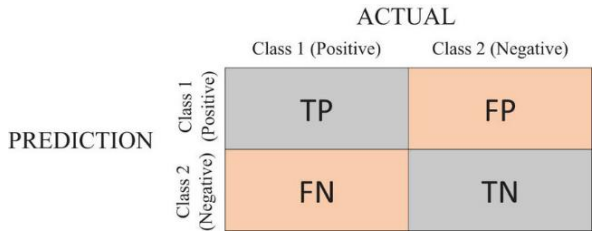
۵- نتایج پیاده‌سازی

به منظور ارزیابی کارایی مدل پیشنهادی، مجموعه‌ای از شاخص‌های استاندارد شامل Accuracy، Recall، Precision و F1-Score محاسبه شد. علاوه بر این، برای بررسی دقیق‌تر روند آموزش، نمودار تغییرات دقت و خطا ترسیم گردید و از ماتریس سردرگمی جهت تحلیل نحوه طبقه‌بندی نمونه‌ها استفاده شد. نتایج به‌دست‌آمده نه‌تنها کارایی مطلوب روش پیشنهادی را نشان داد، بلکه در مقایسه با پژوهش‌های شاخص گذشته نیز عملکرد بهتری ثبت گردید. این مقایسه جایگاه مدل را در میان رویکردهای موجود روشن‌تر ساخت و بر مزایای آن تاکید نمود.

۵-۱- نتایج پیاده‌سازی

ماتریس سردرگمی ابزاری پرکاربرد در ارزیابی مدل‌های طبقه‌بندی است که ارتباط میان پیش‌بینی‌های مدل و برچسب‌های واقعی داده‌ها را به‌صورت جدولی نمایش می‌دهد. در مسئله دودسته‌ای حاضر (شکستگی در برابر غیرشکستگی)، مقادیر روی قطر اصلی بیانگر

نمونه‌هایی هستند که به‌درستی تشخیص داده شده‌اند. در مقابل، مقادیر خارج از قطر نشان‌دهنده موارد خطای مدل در پیش‌بینی‌اند. این ساختار کمک می‌کند تا علاوه بر مشاهده دقت کلی، بتوان به‌طور مشخص نرخ خطاهای مثبت کاذب و منفی کاذب را نیز تحلیل کرد [۱۵]. همان‌طور که در شکل ۶ دیده می‌شود، ماتریس حاصل نشان‌دهنده عملکرد قوی مدل در شناسایی صحیح اکثریت نمونه‌ها است.



شکل (۶): نمایش ماتریس سردرگمی برای مسئله دو کلاسه [۱۶]

۵-۲- معیارهای ارزیابی

از مقادیر موجود در ماتریس سردرگمی می‌توان برای محاسبه شاخص‌های گوناگون بهره برد.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (۱)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (۲)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (۳)$$

$$\text{F1} - \text{Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (۴)$$

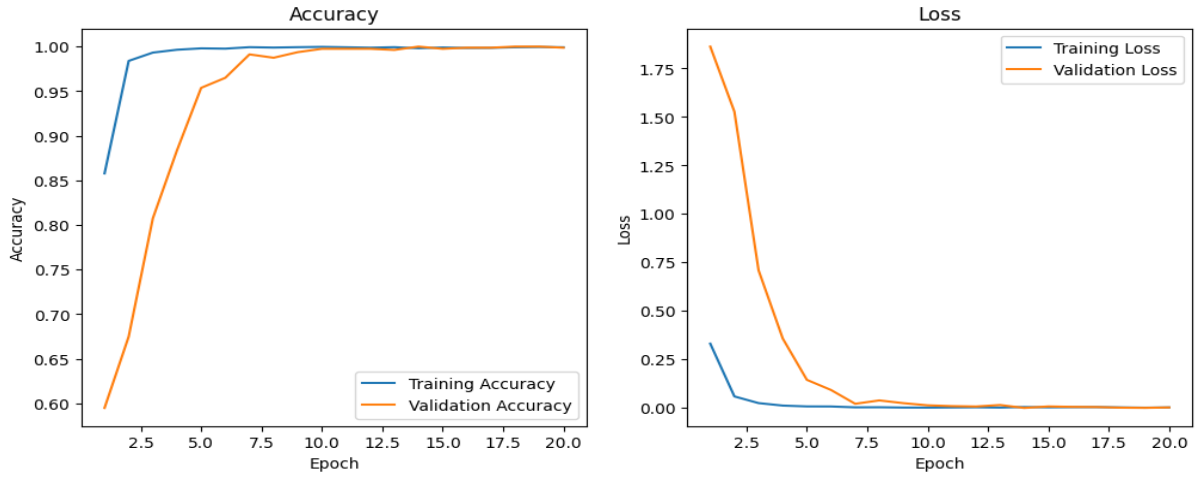
این شاخص‌ها ابزاری دقیق برای سنجش نقاط قوت و ضعف مدل در تشخیص هر دسته هستند [۱۷،۱۸].

۵-۳- نتایج روش پیشنهادی

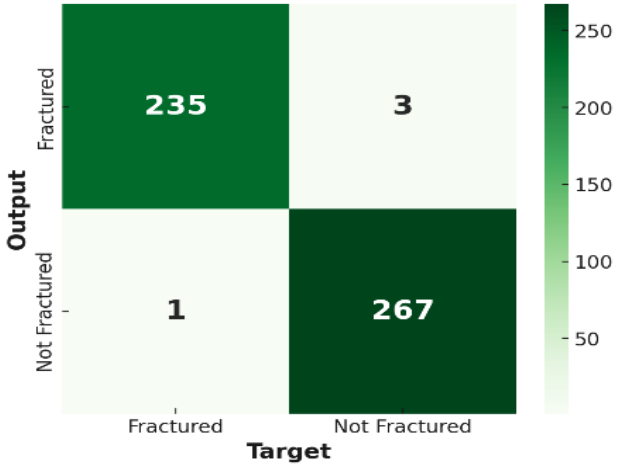
فرایند پیاده‌سازی با استفاده از زبان برنامه‌نویسی پایتون انجام شد. مدل پیشنهادی با بهینه‌ساز AdamW و تابع زیان Sparse Categorical Crossentropy آموزش داده شد. کلیه آزمایش‌ها در محیط محاسباتی Kaggle و با بهره‌گیری از دو پردازنده گرافیکی نوع T4 اجرا گردید. تصاویر ورودی پیش از ورود به مدل با روش درون‌یابی دوخطی به ابعاد ۲۲۴×۲۲۴ پیکسل تغییر اندازه داده شدند و شدت پیکسل‌های آن‌ها نیز مطابق استاندارد ورودی مدل‌های مبتنی بر ImageNet از بازه‌ی [۰,۲۵۵] به بازه‌ی [-۱,۱] نرمال گردید. سپس به کمک وزن‌های از پیش آموزش‌دیده فرایند یادگیری آغاز شد. این رویکرد، امکان بهره‌گیری از دانش عمومی استخراج‌شده از پایگاه‌های بزرگ داده را فراهم آورد و نیاز به آموزش کامل از ابتدا را کاهش داد. مدل طی ۲۱ دوره آموزشی مورد بهینه‌سازی قرار گرفت. شکل ۷ روند تغییرات دقت و خطا را در طول آموزش و اعتبارسنجی نشان می‌دهد. در آغاز آموزش، دقت مجموعه آموزش سریع‌تر از اعتبارسنجی افزایش یافت و در چند دوره نخست بالاتر باقی ماند؛ دقت اعتبارسنجی از مقدار پایین‌تر شروع شد و به‌تدریج به منحنی آموزش نزدیک شد و در ادامه تقریبا همگرا گردید. این الگو با توجه به استفاده از وزن‌های از پیش‌آموزش‌دیده، فریز بخشی از لایه‌ها در ابتدای آموزش و انتخاب نرخ یادگیری کوچک، طبیعی است و موجب شده مدل بر داده‌های آموزش سریع‌تر پایدار شود و نمودار اعتبارسنجی با اندکی تأخیر به آن برسد. با ادامه فرایند، دو منحنی به یکدیگر نزدیک شدند و در نهایت همگرا گردیدند. برای جلوگیری از بیش‌برازش و بهبود تعمیم‌پذیری، از Early Stopping استفاده شد که در پایان دوره بیست‌ویکم فعال گردید و آموزش متوقف شد.

تنظیم نرخ یادگیری بر مقدار ۰.۰۰۰۱ نقش تعیین‌کننده‌ای در دستیابی به همگرایی پایدار داشت. در حالی که نرخ‌های بالاتر می‌توانند در مراحل اولیه باعث افزایش سرعت یادگیری شوند، تجربه این پژوهش نشان داد که چنین رویکردی در فازهای پایانی منجر به نوسانات شدید می‌شود. با کاهش نرخ یادگیری به مقدار ۰.۰۰۰۱، مدل توانست با شیبی یکنواخت‌تر و بدون نوسان‌های بزرگ به سمت نقطه بهینه حرکت کند. نمودار خطا در هر دو مجموعه روندی یکنواخت و کاهشی داشت و واگرایی متاخر متداول پس از بیش‌برازش مشاهده نشد؛ توقف زودهنگام، وزن‌کاهی AdamW و Dropout از بروز آن جلوگیری کردند. این رفتار نشان می‌دهد که انتخاب صحیح نرخ یادگیری نقشی کلیدی در افزایش دقت و ثبات عملکرد مدل داشته است. این روند نشان‌دهنده تاثیر مثبت تنظیمات اعمال‌شده بر روی فرایند یادگیری مدل است.

ماتریس سردرگمی حاصل از روش پیشنهادی برای دو دسته تصویر در شکل ۸ نشان داده شده است. همان طور که از ماتریس سردرگمی مشخص است روش پیشنهادی به خوبی قادر است نمونه‌های هر دو کلاس را با دقت خوبی شناسایی کند.



شکل (۷): تغییرات دقت و خطا برای مجموعه‌های آموزش و اعتبارسنجی در طول دوره‌های آموزش؛ دقت آموزش در ابتدا بالاتر است و منحنی‌ها به‌تدریج همگرا می‌شوند، خطا نیز در هر دو مجموعه کاهشی و پایدار است.



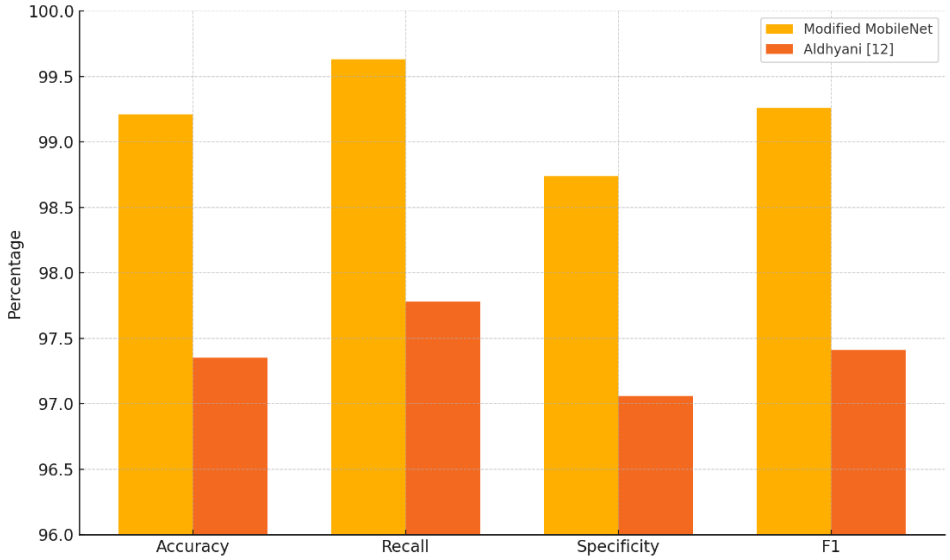
شکل (۸): ماتریس سردرگمی برای دو دسته شکستگی و غیرشکستگی

در جدول ۱، روش پیشنهادی با نتایج پژوهش پیشین مقایسه شده است. همان‌طور که از نتایج مشخص است، روش پیشنهادی عملکرد بهتری در تشخیص شکستگی دارد. حداکثر دقت روش پیشنهادی ٪۹۹.۲۱ می‌باشد که مدل به کمک آن توانست ۵۰۲ تصویر از ۵۰۶ تصویر را به درستی دسته‌بندی کند.

شکل ۹، نتیجه معیارهای ارزیابی روش پیشنهادی را نشان می‌دهد. در این شکل، روش پیشنهادی با بهترین نتیجه پژوهش‌های پیشین مقایسه شده است. همان‌طور که از نتایج مشخص است، روش پیشنهادی عملکرد بهتری در طبقه‌بندی تصاویر رادیوگرافی دارد.

جدول (۱): مقایسه روش پیشنهادی با پژوهش انجام شده برای طبقه‌بندی تصاویر رادیوگرافی

روش	دسته‌بند	Accuracy	Recall	Precision	F1	مجموعه‌داده
Alam [7]	MobLG-Net	۹۹	–	–	–	۱۷۷۳ تصویر رادیوگرافی
Abdusalomov [8]	DenseNet121	۹۰.۳	۸۹	۸۷.۵	۸۸.۲	۵۰۰۰ تصویر شکستگی ورزشی
Srinivasu [9]	YOLO-v10	۹۶.۴	–	–	–	مجموعه تصویر رادیوگرافی خصوصی
Aldhyani [12]	DenseNet201	۹۷.۳۵	۹۷.۷۸	۹۷.۰۶	۹۷.۴۱	مجموعه داده حاضر
مدل پیشنهادی	نسخه بهبودیافته MobileNetV3	۹۹.۲۱	۹۹.۶۳	۹۸.۷۴	۹۹.۲۶	مجموعه داده حاضر



شکل (۹): مقایسه نتیجه معیارهای ارزیابی حاصل از روش پیشنهادی با بهترین مدل پژوهش دیگر روی مجموعه داده یکسان

۶- نتیجه‌گیری

در این پژوهش، سامانه‌ای پیشرفته برای طبقه‌بندی تصاویر رادیوگرافی شکستگی و غیرشکستگی استخوان با استفاده از مدل بهینه‌شده MobileNetV3 Small معرفی شد. هدف اصلی افزایش دقت و پایداری مدل در تشخیص دودسته‌ای شکستگی بود. با بهره‌گیری از تکنیک‌های داده‌افزایی شامل چرخش، قرینه‌سازی و تبدیلات تصویری، جایگزینی لایه‌ی Flatten با Global Average Pooling، افزودن دو لایه‌ی Dense با ۲۵۶ و ۱۲۸ نورون، و استفاده از لایه‌های Batch Normalization و Dropout، قابلیت تعمیم‌پذیری شبکه تقویت گردید. به‌کارگیری بهینه‌ساز AdamW با مقدار Weight Decay معادل ۰.۰۰۰۱ و کاهش نرخ یادگیری به ۰.۰۰۰۱، ضمن تسریع همگرایی، از نوسانات بیش‌ازحد در فرایند آموزش جلوگیری کرد. همچنین، با فریز کردن لایه‌های ابتدایی تا لایه‌ی ۱۰۰ و اعمال Fine-Tuning بر لایه‌های بالایی، تعداد پارامترهای قابل آموزش کاهش یافت و منابع محاسباتی بهینه‌تر مصرف شد. نتایج ارزیابی بر مجموعه‌ی آزمون (۵۰۶ تصویر) نشان‌دهنده‌ی دستیابی به دقت ۹۹.۲۱٪ بود.

سبک‌وزن و کم‌مصرف بودن معماری MobileNetV3 امکان پیاده‌سازی سامانه بر روی دستگاه‌های کم‌منبع نظیر تلفن‌های همراه هوشمند و تبلت‌های پزشکی را فراهم می‌آورد و می‌تواند به‌عنوان ابزار کمکی در بخش‌های اورژانس و کلینیک‌های پرمراجعه جهت افزایش سرعت و دقت تشخیص مورد استفاده قرار گیرد. افزون بر این، این سامانه می‌تواند با کاهش خطاهای تشخیصی و تسهیل روند درمان، اثربخشی مراقبت‌های بالینی را ارتقا دهد.

برای مطالعات آتی، پیشنهاد می‌شود توسعه قابلیت‌های چندوظیفه‌ای شبکه شامل تشخیص، طبقه‌بندی و مکان‌یابی شکستگی، ارزیابی عملکرد مدل بر روی داده‌های بالینی متنوع‌تر و پیچیده‌تر، و

بهبود شفافیت فرآیند تصمیم‌گیری مدل در دستور کار قرار گیرد تا کاربرد بالینی آن به نحو چشمگیری گسترش یابد.

مراجع

[1] R. M. Jones, A. Sharma, R. Hotchkiss, J. W. Sperling, J. Hamburger, C. Ledig, … & R. V. Lindsey, (2020). Assessment of a deep-learning system for fracture detection in musculoskeletal radiographs. NPJ digital medicine, 3(1), 144.

[2] Y. D. Jeon, K. B. Park, S. H. Ko, J. M. Oh, & S. G. Kim, (2024). Sports-related fractures in the geriatric population at a level I trauma center. BMC geriatrics, 24(1), 464.
https://doi.org/10.1186/s12877-024-05095-x

[3] J. Crim, (2024). Bone radiographs: sometimes overlooked, often difficult to read, and still important. Skeletal Radiology, 53(9), 1687-1698.
https://doi.org/10.1007/s00256-023-04498-y

[4] N. Aghaei, G. Akbarizadeh, & A. Kosarian, (2022). Using ShuffleNet to design a deep semantic segmentation model for oil spill detection in synthetic aperture radar images. Journal of Iranian Association of Electrical and Electronics Engineers, 19(3), 131–144.

[5] M. Kutbi, (2024). Artificial intelligence-based applications for bone fracture detection using medical images: a systematic review. Diagnostics, 14(17), 1879.

[6] Z. Su, A. Adam, M. F. Nasrudin, M. Ayob, & G. Punganan, (2023). Skeletal fracture detection with deep learning: A comprehensive review. Diagnostics, 13(20), 3245.

[7] A. Alam, A. S. Al-Shamayleh, N. Thalji, A. Raza, E. A. Morales Barajas, E. B. Thompson,... & I. Ashraf, (2025). Novel transfer learning based bone fracture detection using radiographic images. BMC Medical Imaging, 25(1), 5.

[8] A. Abdusalomov, S. Mirzakhalilov, S. Umirzakova, O. Ismailov, D. Sultanov, R. Nasimov, & Y. I. Cho, (2025). Lightweight Deep Learning Framework for Accurate Detection of Sports-Related Bone Fractures. Diagnostics, 15(3), 271.

[9] P. N. Srinivasu, G. L. A. Kumari, S. C. Ahmed S. Narahari, & A. Alhumam, (2025). Exploring the impact of hyperparameter and data augmentation in YOLO V10 for accurate bone fracture detection from X-ray images. Scientific Reports, 15(1), 9828.

[10] T. N. Hoa, L. M. Pham, & Q. H. Vu, (2025). Classify bone fractures in X-ray images. In Lecture Notes in Computer Science (Vol. 15417, pp. 344–356). Springer.

[11] M. Cohen, J. Puntonet, J. Sanchez, E. Kierszbaum, M. Crema, P. Soyer, & E. Dion, (2023). Artificial intelligence vs. radiologist: accuracy of wrist fracture detection on radiographs. European radiology, 33(6), 3974-3983.

[12] T. Aldhyani, Z. A. Ahmed, B. M. Alsharbi, S. Ahmad, M. H. Al-Adhaileh, A. H. Kamal,... & J. Nazeer, (2025). Diagnosis and detection of bone fracture in radiographic images using deep learning approaches. Frontiers in Medicine, 11, 1506686.
https://doi.org/10.3389/fmed.2024.1506686

[13] S. Qian, C. Ning, & Y. Hu, (2021, March). MobileNetV3 for image classification. In 2021 IEEE 2nd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE) (pp. 490-497). IEEE.

[14] A. Howard, M. Sandler, G. Chu, L. C. Chen, B. Chen, M. Tan,... & H. Adam, (2019). Searching for mobilenetv3. In Proceedings of the IEEE/CVF international conference on computer vision (pp. 1314-1324).

[15] A. Vanacore, M. S. Pellegrino, & A. Ciardiello, (2024). Fair evaluation of classifier predictive performance based on binary confusion matrix. Computational Statistics, 39(1), 363–383. https://doi.org/10.1007/s00180-022-01301-9

[16] A. A. Mukhlif, B. Al-Khateeb, & M. A. Mohammed, (2022). An extensive review of state-of-the-art transfer learning techniques used in medical imaging: Open issues and challenges. Journal of Intelligent Systems, 31(1), 1085-1111.

[17] S. Asadi Amiri, & M. Andi, (2025). Classification of Pistachio Varieties Using MobileNet Deep Learning Model. Journal of Iranian Association of Electrical and Electronics Engineers, 22(1), 133–140.

[18] J. L. Speiser, M. E. Miller, J. Tooze, & E. Ip, (2019). A comparison of random forest variable selection methods for classification prediction modeling. Expert Systems with Applications, 134, 93–101. https://doi.org/10.1016/j.eswa.2019.05.028

زیرنویس‌ها

[1] Convolutional Neural Network
[2] X-ray
[3] Augmentation
[4] Flipping
[5] Overfitting
[6] AP
[7] Lateral
[8] Train
[9] Validation
[10] Test
[11] Resizing
[12] Normalization
[13] Optimizer
[14] Weight Decay
[15] Learning Rate