IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15451

An EfficientNet-B4 Based Medical Deepfake Detection in Healthcare Image Analysis

Mr. A. Azeem[1], K. Gowthami[2], B. Indhu[3], K. Pavani[4]

Student, Department of Electronics and communication Engineering

Andhra Loyola Institute of Engineering and Technology, Vijayawada, A.P, India[1-4]

Abstract: Deepfake technology, powered by artificial intelligence and deep learning, can now create highly realistic fake images, audio, and videos. While this innovation has many uses, it also poses serious risks in healthcare, where medical images like X-rays and CT scans can be altered. Such manipulation may lead to wrong diagnoses, affecting patient safety and hospital operations. This study focuses on building a reliable deep learning approach to identify fake medical images. Two datasets—knee X-rays and lung CT scans—were prepared, preprocessed, and labeled as real or fake. The EfficientNet-B4 model was then applied to detect manipulations. Results show that the model performs very well, achieving high accuracy in both datasets, especially in knee X-ray images. It also maintains a good balance between speed and performance, making it suitable for real-time use. Overall, the study demonstrates that EfficientNet-B4 is an effective solution for detecting medical deepfakes quickly and accurately.

Index Terms: Medical deepfake image detection, deep learning, EfficientNet-B4, convolutional neural networks.

I. INTRODUCTION

In this project, the EfficientNet-B4 model is used as the core deep learning approach for detecting manipulated medical images. EfficientNet-B4 is a well-designed convolutional neural network that focuses on improving accuracy while keeping the model efficient and lightweight. It works by carefully balancing the size and depth of the network, allowing it to learn detailed patterns from images without requiring excessive computational power. This makes it highly suitable for analyzing complex medical images such as X-rays and CT scans. The model is capable of capturing subtle differences between real and altered images, which is essential in medical deepfake detection. Its strong performance and faster processing make it a reliable choice for practical healthcare applications where both accuracy and speed are important.

II. LITERATURE SURVEY

Recent advancements in artificial intelligence have made deepfake technology more powerful, raising concerns in the medical field where images like X-rays and CT scans can be manipulated. Many researchers have explored deep learning techniques, especially convolutional neural networks, to detect such fake images by identifying subtle patterns that are not visible to the human eye. Studies have shown that using diverse medical datasets along with data augmentation techniques improves detection performance. Modern models are designed to provide both high accuracy and faster processing, making them suitable for real-time healthcare applications. Overall, existing research highlights the need for reliable and efficient systems to ensure the authenticity of medical images and support accurate diagnosis.

III. PROPOSED SYSTEM

The proposed system is designed to detect whether a medical image is real or manipulated through a simple and user-friendly web interface. In this system, users can upload medical images such as X-rays or CT scans directly on the website. Once the image is uploaded, it is processed using a trained deep learning model, which analyzes the image and identifies hidden patterns or irregularities. Based on this analysis, the system quickly classifies the image as either real or fake. The result is then displayed on the webpage in an easy-to-understand format. This approach not only simplifies the detection process but also allows real-time usage, making it useful for medical professionals and researchers who need quick and reliable verification of medical images.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15451

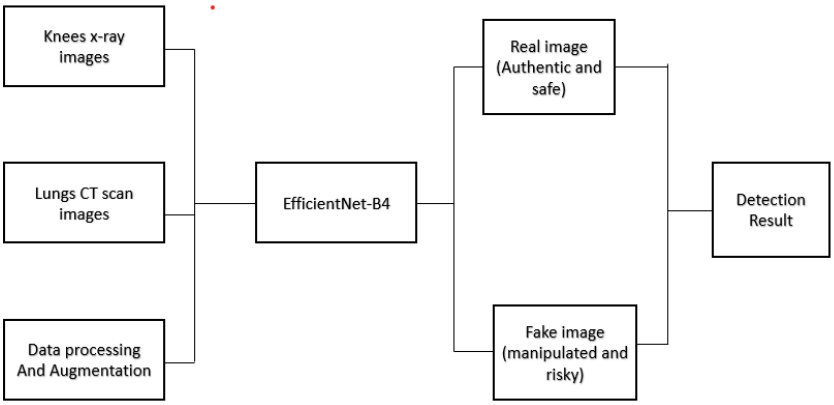


Figure.1 Block diagram

IV. METHODOLOGY

4.1 Data Pre-Processing Techniques:
In this step, the collected medical images are prepared for model training. Images are resized to a uniform size and converted into a consistent format. Noise and unwanted variations are reduced, and pixel values are normalized. This helps the model learn more effectively and ensures better accuracy during prediction.

4.2 Data Augmentation Techniques:
Data augmentation is used to increase the diversity of the dataset and improve model performance. Techniques such as rotation, flipping, zooming, brightness, and contrast adjustments are applied to X-ray and CT scan images. Small changes like slight rotations and contrast tuning help simulate real-world conditions without affecting medical meaning. This reduces overfitting and makes the model more reliable in detecting both real and fake images.

4.3 Splitting the Data:
The dataset is divided into three parts:

- Training – 60%
- Validation – 20%
- Testing – 20%
  This ensures proper learning, tuning, and evaluation of the model.

4.4 Algorithms Used in EfficientNet-B4:
Mobile Inverted Bottleneck Convolution (MBConv):
MBConv is the main building block of the model. It helps in extracting features efficiently while keeping the model lightweight and fast.
Squeeze and Excitation (SE) Module:
This module focuses on important features in the image by giving more weight to useful information and reducing less relevant details.
Swish Activation Function:
Swish improves learning by allowing smoother flow of information compared to traditional functions, leading to better accuracy.
FLOPs (Floating Point Operations):
FLOPs measure how efficient the model is. EfficientNet-B4 achieves high performance while using fewer computations, making it suitable for real-time applications.

V. EFFICIENTNET-B4

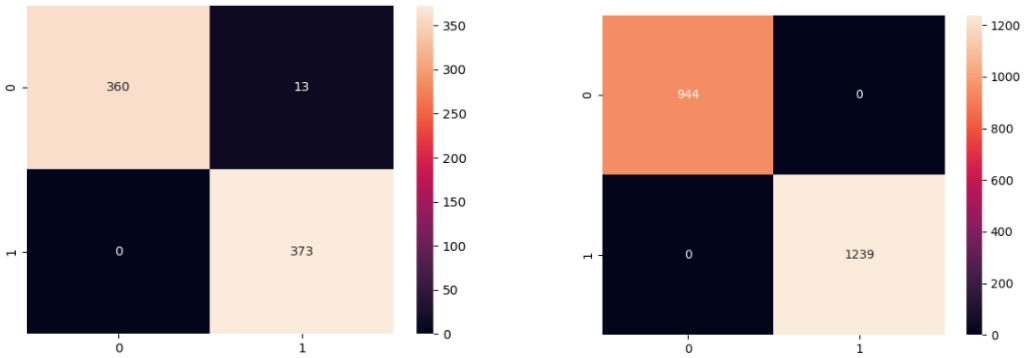
In this project, the EfficientNet-B4 model is used as the main deep learning approach for detecting fake medical images. It is a well-structured neural network that is designed to give high accuracy while keeping the model efficient and not too heavy. Instead of simply increasing the size of the network, it carefully balances different aspects like depth and resolution, which helps it learn detailed features from images. This makes it very effective in analyzing medical images such as X-rays and CT scans, where even small changes matter. The model is able to capture subtle differences between real and manipulated images, making the detection process more reliable and fast for practical use.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15451



(a) **(b)**

Figure 2. Confusion Matrix a)Lungs b)Knees

VI. DEEPFAKE DETECTION

Generative Adversarial Networks are widely used in image synthesis. With the latest developments in Generative Adver sarial Networks, we have entered a period where it is difficult for people to perceive whether the images are real or fake. Studies aimed at detecting this technology, which may pose great risks to humans, are also becoming widespread. Con volutional Neural Networks have also begun to be widely used in detecting fake image. Thanks to the stride parameter Convolution can be used in down-sampling images. With using them consecutive with pooling, CNN networks as feature extractors are obtained, where H is image height, W is image weight, pad is padding, K is convolution kernel, and S is stride .

$$Hout = Hin + (2\times pad) - Kheight/S$$

$$Wout = Win + (2\times pad) - Kwidth/S$$

VII. DATASETS FOR MEDICAL IMAGES

It is very difficult to access medical images because they are not often shared on the internet due to patients' personal rights. On the other hand, since studies on fraud detec tion generally focus on face replacement, the number of researchers working on producing deepfake medical datasets remains low. In this study, different data sets in the literature are used. The first dataset is the Osteoarthritis dataset. Prezja and his colleagues produced a total of 320,000 synthetic Osteoarthritis X-ray images with the GAN-based method they developed and made them available.

Figure 3.Dataset of Knees a)real image b)fake image

Data split	Real images	Fake images	total
Train	1500	1500	3000
Validation	1500	1500	3000
Test	1500	1500	3000

Table 1.The Number of instances in dataset of Knees (Osteoarthritis X-ray)

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15451

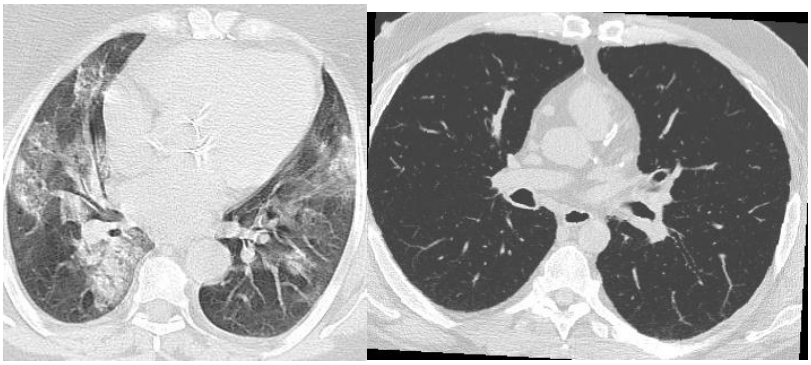


Figure 4. Dataset of Lungs a)real image b)fake image

Data split	Real Images	Fake images	total
Train	6504	531	7035
Validation	175	75	250
Test	175	75	250

Table 2.The number of instances in Dataset2 (lung CT scan)

VIII. EXTERIMENTAL RESULTS

This code builds a simple web application using Flask that allows users to register, log in, and upload medical images to check whether they are real or fake using deep learning models. It uses a lightweight database (TinyDB) to store user details like username, password, email, and mobile number. Once a user logs in, they can access a dashboard and choose between lung or knee image analysis. The system loads pre-trained EfficientNet-B4 models (for lung and knee) using PyTorch, and when a user uploads an image, it processes the image (resizing and converting it into a tensor) before passing it to the model for prediction. The model then classifies the image as either “real” or “fake,” and the result is displayed on the webpage along with the uploaded image. Overall, this project combines user authentication, image processing, and deep learning to create an interactive medical deepfake detection system.

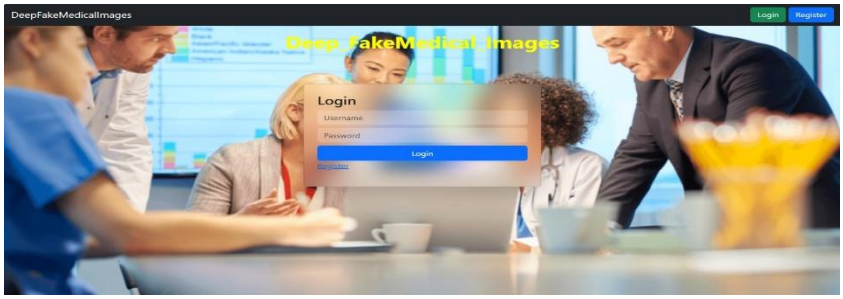


Figure 5. Login with username and password

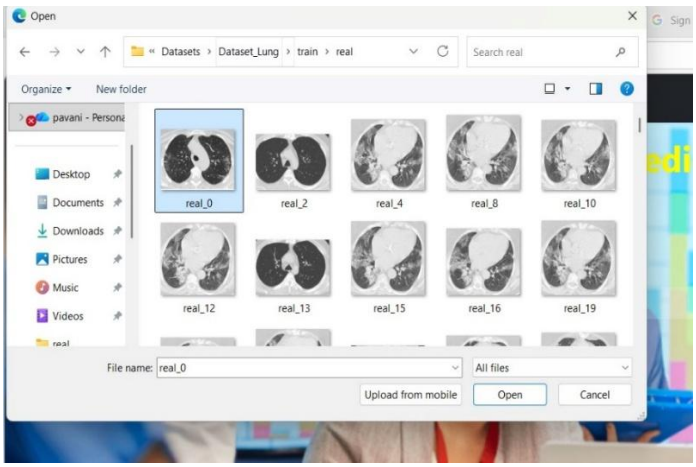


Figure 6. select a image

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15451

Figure 7. Output Real

IX. DISCUSSIONS AND ANALYSIS

This code is designed to build and test a smart system that can tell whether medical images (like lung and knee scans) are real or fake. It uses a powerful deep learning model called EfficientNet-B4 and improves it for this specific task. The images are first prepared by resizing and slightly modifying them (like flipping or adjusting brightness) so the model can learn better and become more accurate. The model is then trained step by step, where it keeps learning from the data and improving its predictions. At the same time, its performance is checked using validation data to make sure it is not overfitting. If the model stops improving after a few attempts, the training automatically stops to save time and avoid unnecessary learning. Once training is complete, the best version of the model is tested on new images it has never seen before. Finally, the code generates important results like accuracy, confusion matrix, and ROC curve, and saves them for analysis. In simple terms, this code handles the entire process—from learning to final evaluation—to create an effective medical deepfake detection system.

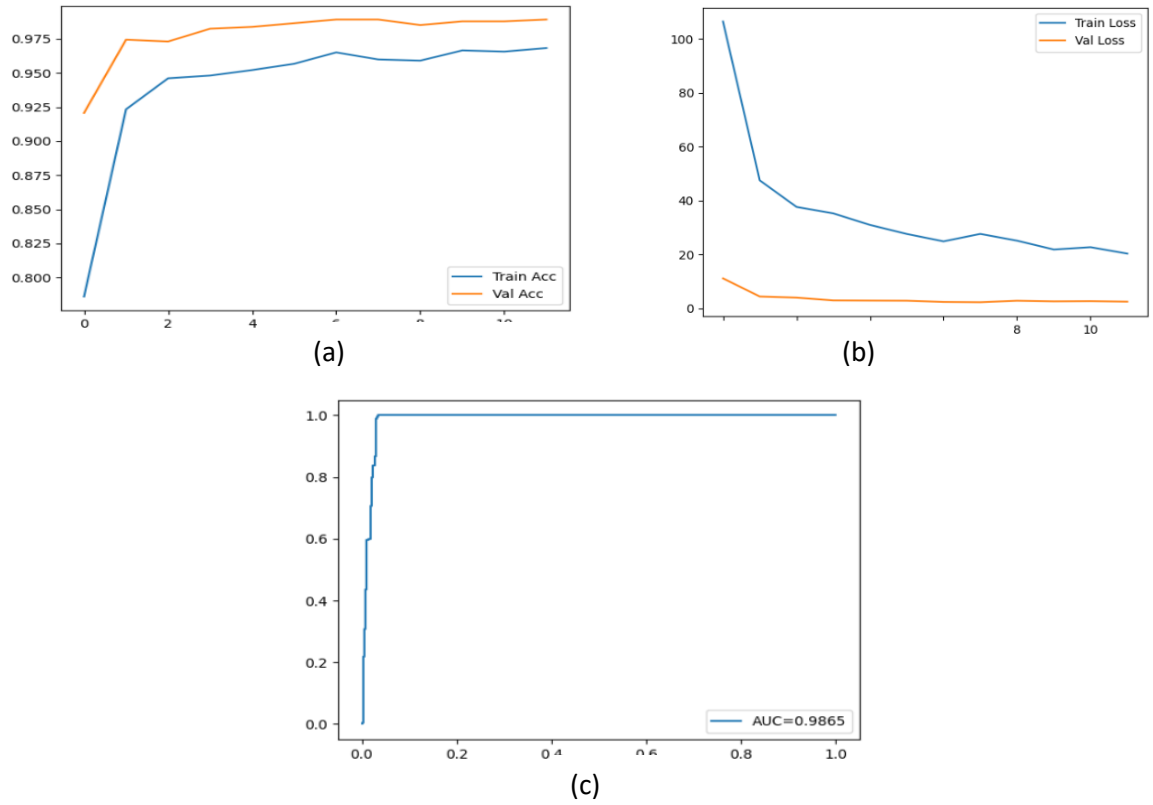


Figure 8. Lung Analysis a) accuracy b) Loss c)ROC curve

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15451

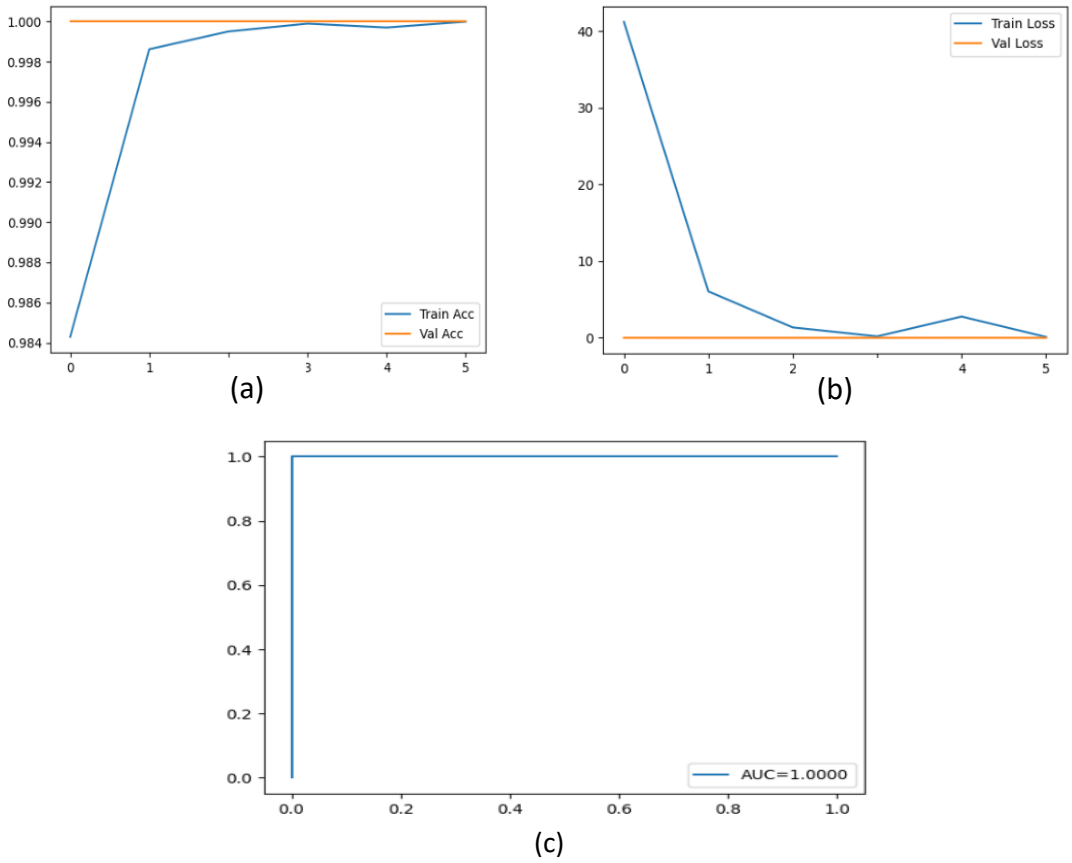


Figure 9. Knees Analysis a) accuracy b) Loss c)ROC curve

X. CONCLUSION

In this project, a reliable system has been developed to detect fake medical images using deep learning. With the growing risk of image manipulation in healthcare, ensuring the authenticity of medical data has become very important. The use of the EfficientNet-B4 model helped in accurately identifying whether an image is real or fake by learning detailed patterns from knee X-rays and lung CT scans.The system showed strong performance in terms of accuracy and consistency, proving that deep learning can effectively handle such critical tasks. By combining proper data processing, model training, and a simple web interface, the project provides a practical solution that can be used in real-world healthcare environments. Overall, this work contributes towards improving trust, safety, and reliability in medical image analysis.

XI. FUTURE SCOPE

This project can be further improved in several ways as technology continues to advance. One important direction is to enhance the model's accuracy by training it on larger and more diverse medical datasets. This will help the system perform better in real-world situations and handle different types of medical images more effectively.In the future, the system can also be extended to support additional imaging types such as MRI scans, ultrasound images, and even medical videos. Integrating the model into hospital systems or cloud platforms can make it more accessible for real-time use by doctors and healthcare professionals.Another possible improvement is to develop a mobile or user-friendly application so that the detection system can be easily used anywhere. The model can also be updated regularly to adapt to new and more advanced deepfake techniques. Overall, this project has strong potential to grow into a complete and widely used solution for ensuring the authenticity and security of medical data.

DOI: 10.17148/IJARCCE.2026.15451

REFERENCES

[1]. I. J. Goodfellow, ‘‘Generative adversarial nets,’’ Proc. Adv. Neural Inf. Process. Syst., vol. 27, pp. 2672–2680, 2014.
[2]. İ. İlhan and M. Karaköse, ‘‘Derin sahte videoların tespiti ve uygulamaları için bir karşılaştırma Çalışması,’’ Adıyaman Üniversitesi Mühendislik Bil imleri Dergisi, vol. 8, no. 14, pp. 47–60, Jun. 2021.
[3]. I. Ilhan, E. Bali, and M. Karaköse, ‘‘An improved DeepFake detection approach with NASNetLarge CNN,’’ in Proc. Int. Conf. Data Analytics Bus. Ind. (ICDABI), Oct. 2022, pp. 598–602.
[4]. S. Solaiyappan and Y. Wen, ‘‘Machine learning based medical image deepfake detection: A comparative study,’’ Mach. Learn. Appl., vol. 8, Jun. 2022, Art. no. 100298.
[5]. J. E. Dunn. (2018). Imagine You’re Having a CT Scan and Malware Alters the Radiation Levels—It’s Doable the Register. [Online]. Available: https://www.theregister.co.uk/2018/04/11/hacking_medical_devices/
[6]. Y. Mirsky, T. Mahler, I. Shelef, and Y. Elovici, ‘‘CT-GAN: Malicious tampering of 3D medical imagery using deep learning,’’ in Proc. 28th USENIX Secur. Symp., Jan. 2019, pp. 461–478.
[7]. D.Shen,G.Wu,andH.-I.Suk,‘‘Deep learning in medical image analysis,’’ Annu. Rev. Biomed. Eng., vol. 19, pp. 221–248, Jun. 2017.
[8]. Y. S. Kim, H. J. Song, and J. H. Han, ‘‘A study on the development of deepfake-based deep learning algorithm for the detection of medical data manipulation,’’ Webology, vol. 19, no. 1, pp. 4396–4409, Jan. 2022.
[9]. A. G. Eker and N. Duru, ‘‘Deep learning applications in medical image processing,’’ Acta Infologica, vol. 5, no. 2, pp. 459–474, 2021, doi: 10.26650/acin.927561.
[10]. H. MacMahon, D. P. Naidich, J. M. Goo, K. S. Lee, A. N. Leung, J. R. Mayo, A. C. Mehta, U. Ohno, C. A. Powell, M. Prokop, G. D. Rubin, C. M. Schaefer-Prokop, W. D. Travis, P. E. V. Schil, and A. A. Bankier, ‘‘Guidelines for management of incidental pulmonary nodules detected on Ct images: from the Fleischner society,’’ Radiology, vol. 284, no. 1, pp. 228–243, 2017.

