

Lung Disease Detection by Processing X-ray Image Database Using Deep Learning Techniques

Ahmed Khudhair Abbas¹, Ihsan Salman Jasim² and Adil Ibrahim Khalil¹

¹Department of Computer Science, College of Education for Pure Science, University of Diyala, 32001 Diyala, Iraq

²Department of Computer Science, College of Basic Education, University of Diyala, 32001 Diyala, Iraq
dr.ahmed.k.abbas@uodiyala.edu.iq, ihsan1975.83@gmail.com, dr.adil.khalil@uodiyala.edu.iq

Keywords: Chest X-Ray Images, Classifications, InceptionV3, Medical Imaging, Database.

Abstract: Integrating technology such as artificial intelligence in the medical field may contribute significantly to the early diagnosis of diseases. This study is conducted to achieve two goals: The first is to use three advanced deep learning models to identify tuberculosis and COVID-19 diseases from healthy cases by processing a chest X-ray image dataset. The second goal is to compare the three models and identify the most generalizable one based on accuracy and computational efficiency. The proposed models were applied to a large-scale chest X-ray of 6375 images. The findings show that InceptionV3 outperforms the CNNs and ResNet-50 algorithms by achieving high metrics values (Accuracy = 0.9636, Precision = 0.9752, Recall = 0.9554, and F1-Score = 0.9651). Furthermore, the results show that training neural networks across a wide range of epochs improves the model predictions of new data. The findings demonstrate InceptionV3's capability of providing a robust and automated method for classifying chest X-ray images. Apart from accuracy assessment, this computational analysis depicted marked variations in efficiency amongst the three models. The CNN architecture achieved the shortest training time of ≈ 12.4 minutes and attained the fastest inference speed of ≈ 4.1 ms per image, whereas ResNet-50 and InceptionV3 required longer training and inference times. This starkly indicates that there is a clear trade-off between computational cost and predictive performance across the evaluated architectures.

1 INTRODUCTION

Chest X-rays are a primary tool for symptomatic imaging that provide essential insights into the syndrome of thoracic depression [1].

Furthermore, Chest X-rays can provide important information about some diseases, such as COVID-19 and tuberculosis (see Fig.1).

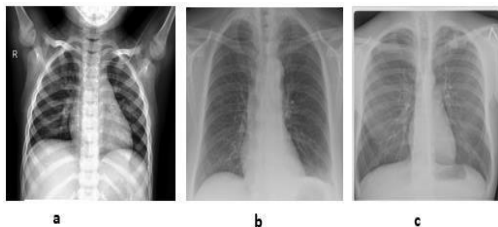


Figure 1: Three conditions of Chest X-ray images: a) Normal, b) COVID-19 and c) Tuberculosis.

Therefore, precisely and effectively interpreting these images is crucial for diagnosing, treating, and

controlling several cardiovascular disorders. Historically, this job has necessitated the advanced knowledge of experienced radiologists, a proficiency sometimes constrained by workload and geological irregularities[2].

Recent advancements in utilizing artificial intelligence techniques have emerged as a valuable tool in transforming the analysis of medical images. These models can enhance the capacities of radiologists, contributing to improvements in symptomatic accuracy and competence [3]. Therefore, efforts are ongoing to improve the quality of medical care by researching new algorithms and solutions to apply them in the medical field [4]. New techniques such as deep learning algorithms have attracted significant attention in many medical studies, especially pulmonary diseases. For example, the study of Nayak et al. [5] presented an automated early diagnosis system by processing X-ray images utilizing deep learning models. In their study, eight pre-trained CNNs models were investigated to

classify COVID-19 from healthy cases. The best model, ResNet-34, provided an accuracy of 98.33% on the tested chest X-ray images. The study of Singh et al. [6] used a CNNs algorithm to identify COVID-19 in chest X-ray images. The CNNs and VGG16 models were pretrained on the image dataset and used transfer learning and image processing techniques to increase detection accuracy. The dataset consists of normal and COVID-19 chest X-ray images. The results show that the CNN model successfully extracts useful features from the small dataset, suggesting that image processing and deep learning could be used for early and precise diagnosis. The CNNs model has the potential for medical professionals to diagnose COVID-19 and could help create automated screening systems for the pandemic's eradication. The Author found that the VGG16 model reduces overfitting and enhances model performance, making it particularly useful for smaller datasets like COVID-19 chest X-ray images. Six artificial deep neural networks were utilized to identify COVID-19 from chest X-rays images [7]. A total of 2905 COVID-19, viral pneumonia, and healthy chest X-rays were included in that study. With an overall accuracy of 96.91%, ResNet-50 was the most successful network. The model showed a high performance in the loss of detection of COVID-19 cases, indicating that it is a faster and more accurate method of detection.

Machine learning, especially deep learning, has been extensively explored in the context of corona detection, and a classifier has been trained to obtain positive corona or determine negative corona X-ray images as shown in [8]. This paper utilized Apache Spark and Deep Transfer Learning methods to create a classifier for COVID-19 chest X-ray images, achieving 100% accuracy in COVID/Normal/pneumonia, with 97 percent accuracy for InceptionV3, ResNet50, and VGG19 models. The study conducted by Mujahid et al. [9] detected pneumonia from X-ray images with pre-trained convolutional neural networks (CNNs) such as VGG16, Inception-v3, and ResNet50. The results indicate that Inception-V3 with CNNs had the best recall and accuracy rates, at 99.73% and 99.29%, respectively.

The goal of this paper is to propose a prediction model based on CNN algorithm. Therefore, the proposed model was applied to Chest image dataset to identify COVID-19 and Tuberculosis from healthy

ones. The performance of our proposed model was evaluated using accuracy, specificity, recall, and F1-score metrics.

2 MATERIAL AND THEORETICAL BACKGROUND

2.1 Model Architecture and Transfer Learning

Figure 2 presents the flowchart of the proposed methodology used in this study. Up to 6078 publicly available chest X-ray images were processed. The dataset is subdivided into normal, tuberculosis, and COVID-19 images. For preprocessing, the images are resized to 224×224 pixels. The images were subdivided into 3 groups: 60% training, 20% testing and 20% validation. The images were evaluated and classified using three proposed models of this study.

CNNs are widely used in medical imaging, such as image recognition and chest X-ray classification, due to their ability to learn spatial hierarchy and detect abnormalities with high accuracy. CNNs include an input layer, an output layer, and several hidden layers that learn new and complex objects and patterns. Convolution and pooling operations are then applied to a subsample of the provided input, which is then passed through an activation function.

In contrast, InceptionV3 is a CNN-based model and potentially a fundamental depth learning that is commonly used in various exploratory considerations for image classification tasks [8]. InceptionV3's victory lies in its architecture's capability to extricate complex features from pictures. It is a profitable tool for precise classification tasks in areas like mechanical apparatus support, medical diagnostics, and ecological observation [10].

In addition, ResNet-50 is a deep CNN architecture that is broadly utilized in different areas, including therapeutic picture investigation and plant illness discovery [11]. Moreover, ResNet-50 has been investigated for assessing optical properties in scrambling media, displaying comparable or way better-remaking precision with training on smaller datasets and potential for assist improvements [12].

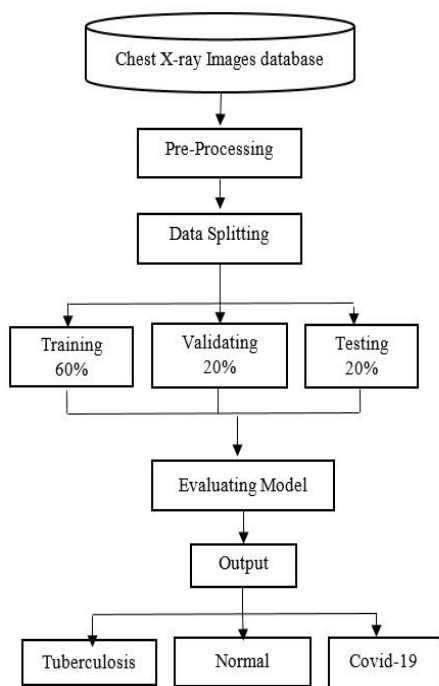


Figure 2: Flowchart of the proposed methodology.

The three models presented above achieved significant success in many computer vision tasks and have been adapted for the diagnostic analysis of medical images with promising results [13]. ResNet-50 has demonstrated exceptional performance in image classification tasks by addressing the paradox of the vanishing gradient problem[14]. In contrast, InceptionV3 employs a multi-scale preparation method, effectively capturing visual features at various degrees of reflection [15]. Therefore, this work investigates the suitability of three modern algorithms, CNNs, ResNet-50, and InceptionV3, for identifying tuberculosis and COVID-19 diseases from healthy cases by processing a chest X-ray image dataset. The study also aims to compare the three deep-learning architectures to determine which model provides the best balance of accuracy, generalization capability, and computational cost.

2.2 Data Splitting

The data is divided into training, validation, and testing sets, as shown in Table 1.

Table 1: Data splitting.

Class	Training	Validating	Testing	Total
Normal	1207	402	404	2013
Covid-19	1218	406	704	2328
Tuberculosis	1220	406	408	2034

2.2.1 Adjusting the Size of Image

All the used images of this study were resized to 224×224 pixels to be compatible with the requirement of deep learning models. This step enhances the process of extracting and selecting features.

2.2.2 Converting Image to Grayscale

The X-ray images were transformed into grayscale to emphasize details of structure essential for evaluation, thus minimizing the computational effort and removing unnecessary color data irrelevant to the purpose.

2.2.3 Normalization

Each pixel was normalized to the pixel intensity values range of [0, 1]. This step increases model training performance.

2.2.4 Noise Reduction

In this study, the quality of images was improved by reducing noise and artifacts through using median filter.

2.2.5 Contrast Adjustment

To improve the contrast of the images and accessing the small details such as changes in lung visibility, histogram equalization was applied.

2.2.6 Dataset Balancing

Class imbalance was assessed in the dataset, especially where one group (like COVID-19) can predominate over others. To guarantee an equal representation of all classes, synthetic oversampling

methods like SMOTE (Synthetic Minority Over-Sampling Technique) were investigated. An example of pre-processed lung image is shown in Figure 3, which presents the enhancements steps that applied to the images.



Figure 3: Example of image pre-processing step.

2.3 System Evaluation

The performance of a model is explained by evaluation metrics. Evaluation metrics are evaluated based on their ability to distinguish between model outcomes [16]. The classification process gives four conceivable results: false negative, true negative, true positive, and false positive. Accuracy, recall, F1 score, and precision are utilized as execution evaluation measurements in this consideration [17].

2.4 Model Architectures and Hyperparameters

This section discusses the three model's parameters and setting in this project. First model is (CNN):

- Structure: Three consecutive Convolution layers with 32, 64, and 128 filters (kernel size = 3x3) and a 2x2 MaxPooling after each layer.
- Flatten the layer, then create a Fully Connected layer with a size of 512 neurons and a Dropout of 0.5.
- An output layer with three nodes (softmax) for three categories (Normal, COVID-19, Tuberculosis).
- Training Parameters: Image size: 224x224, Batch size = 32, Optimizer: Adam, Learning rate: 0.001, Number of epochs (up to 50 as in experiments).
- This model represents a lighter and simpler baseline, providing a comparison with deep models (ResNet, Inception).

The second model is (InceptionV3):

- Structure: Use InceptionV3 pre-trained on ImageNet with `include_top = False`.
- Additions: `GlobalAveragePooling2D`, A 1024-sized Dense layer with ReLU, A 3-node softmax output layer, Freeze the base InceptionV3 layers during training (only the top layers are trained).
- Training Parameters: Image size 224x224, batch size = 32, `optimizer = Adam`, `loss = categorical_crossentropy`
- InceptionV3 is known for its ability to capture multi-scale features, making it suitable for detecting subtle lung patterns in X-rays.
- Using transfer learning reduces the need for large datasets and leverages the general features derived from ImageNet.

The third model (ResNet-50):

- Structure: Use pre-trained ResNet-50 with `include_top = False`.
- Additions: `GlobalAveragePooling2D`, `Dense 1024 with ReLU`, `Dense 3 with softmax`.
- Freeze the base ResNet layers (fine-tuning only for the top layers).
- Training Parameters: Image size 224x224, batch size = 32, Rescale images like Inception V3, `optimizer = Adam`, `loss = categorical_crossentropy`.
- Number of epochs (in the paper, ensure the number is consistent with what you actually implemented).
- ResNet-50 uses residual connections to address the vanishing gradient problem and allows for learning deep and accurate representations.
- It is a suitable middle ground between a simple CNN and Inception in terms of complexity and representational power.

3 RESULTS AND DISCUSSION

The proposed algorithms were used to explore the accuracy of diagnosing tuberculosis and COVID19 from the chest X-ray images. Table 2 contains the accuracy results of the proposed algorithms. The Inception-V3 model showed the highest accuracy when compared to the other models. Inception-V3 demonstrated effective performance in identifying X-ray images, with an accuracy of 0.9631, precision of 0.9752, recall of 0.9554, and an F1-score of 0.9651.

Table 2: Results of fine-tuned pre-trained models.

Model	Accuracy	Precision	Recall	F1-Score
CNNs	0.8613	0.8821	0.8116	0.8453
ResNet50	0.8975	0.9168	0.8963	0.9064
InceptionV3	0.9631	0.9752	0.9554	0.9651

Figure 4 presents the Accuracy results of the proposed algorithms versus epoch number. One can observe from Figure 4a, the validating and training accuracy of the ResNet50 model through 50 epochs. We notice that the highest precision is 89.75%. Figure 4.b shows the training and validation curves for Inception-V3.

The model reaches a training accuracy of 96.31%, and the validation accuracy remains consistently high across epochs, indicating stable learning behavior. All of the models are trained, approved and tested for normal, COVID-19, and Tuberculosis classes. Figure 4c shows the training and validation of the CNN model. We see that the accuracy is 86.13%. For 50 epochs, three models were trained on 3 classes, stability and performance enhancement registered in these models tune.

According to these curves, all models exhibit strong performance on both the training and validation datasets, achieving high levels of accuracy. The Inception-V3 shows (Fig. 4b) superior execution of the validation information with a precision of around 96%, making it the finest choice among the three models.

When using long periods of up to 50 epochs to train and test the data, we noticed that the results started to stabilize with little change after period 20. The same fluctuations appeared for the data during the training period, and the same thing was observed for the data's results during the test period.

The computational performance (training and inference time) for the three models shown in Table 3.

Table 3: Training and inference time in minutes.

Model	Training time	Inference time
CNNs	12.4	4.1
ResNet50	45.8	9.8
InceptionV3	59.6	16.3

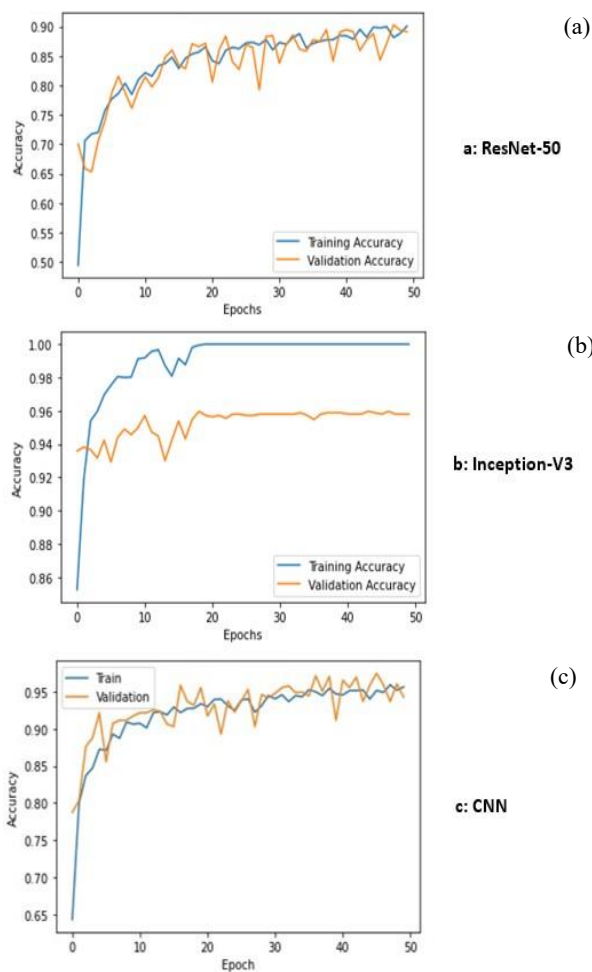


Figure 4: Accuracy results of the proposed algorithms: a) ResNet-50, b) Inception, c) CNN.

The computational analysis showed that the CNN model required the shortest training time (≈ 12.4 minutes), followed by ResNet-50 (≈ 45.8 minutes), while InceptionV3 needed the longest time (≈ 59.6 minutes) due to its deeper architecture.

For inference, CNN achieved the fastest prediction time per image (≈ 4.1 ms), whereas ResNet-50 and InceptionV3 required ≈ 9.8 ms and ≈ 16.3 ms respectively.

The bar chart in Figure 5 clearly demonstrates the computational cost across the three models.

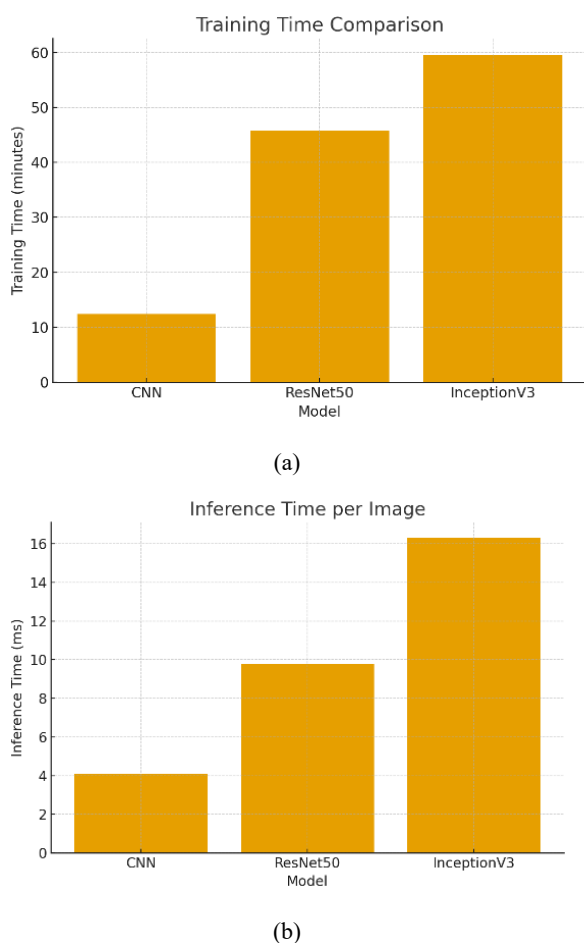


Figure 5: Proposed algorithms (ResNet-50, Inception, and CNN): a) Training time, b) Inference time.

4 CONCLUSIONS

Experimental results demonstrated the efficiency of the proposed system in achieving robust classification performance. Even though the models produced encouraging results, a further comparative analysis revealed that Inception-V3 had the best overall performance and the highest accuracy. This makes it reasonable for simple applications where high accuracy is essential, as therapeutic symptoms, and is particularly persuasive for this classification assignment. Inception V3 offers the highest accuracy and F1 score.

Therefore, it is medically the best choice for the early detection of COVID-19 and tuberculosis.

Measuring the results of training and inference time, CNN is the fastest in training and inference,

but the least accurate. ResNet-50 strikes a balance between accuracy and time. Inception V3 is the most accurate but the slowest in training and inference.

In conclusion, Inception V3 is suitable for systems where accuracy is a priority (such as central hospitals). CNN is suitable for systems requiring high speed (triage systems). ResNet-50 is a good compromise for applications with moderate resources. The importance of careful model selection and hyperparameter tuning for optimal performance in medical image investigation assignments has been highlighted by these findings. The effective implementation of deep learning in chest X-ray classification may enhance clinical treatment by increasing radiologists' proficiency and facilitating the identification of prior infection sites. Future studies should focus on helping to advance model execution, addressing issues like imbalanced datasets and clarifying capacity, and exploring how these models can be integrated into clinical workflows. The ultimate goal is to develop AI-based tools to improve patient care and advance diagnostic accuracy.

REFERENCES

- [1] E. L. Irede et al., "Medical imaging: a critical review on X-ray imaging for the detection of infection," *Biomed. Mater. Devices*, Jul. 2024, doi: 10.1007/s44174-024-00212-1.
- [2] G. Battineni, G. G. Sagaro, N. Chinatalapudi, and F. Amenta, "Applications of machine learning predictive models in the chronic disease diagnosis," *J. Pers. Med.*, vol. 10, no. 2, 2020, doi: 10.3390/jpm10020021.
- [3] X. Liu et al., "A comparison of deep learning performance against health-care professionals in detecting diseases from medical imaging: a systematic review and meta-analysis," *Lancet Digit. Health*, vol. 1, no. 6, pp. e271–e297, 2019, doi: 10.1016/S2589-7500(19)30123-2.
- [4] A. Khalil, A. Humeau-Heurtier, P. Abraham, and G. Mahe, "Comparative study to analyze the effect of aging on microvascular blood flow by processing laser speckle contrast images when Lorentzian and Gaussian velocity profiles are assumed for moving scatterers," in *Proc. Int. Conf. Image Process. Theory Tools Appl. (IPTA)*, 2014, pp. 1–6, doi: 10.1109/IPTA.2014.7001985.
- [5] S. R. Nayak, D. R. Nayak, U. Sinha, V. Arora, and R. B. Pachori, "Application of deep learning techniques for detection of COVID-19 cases using chest X-ray images: a comprehensive study," *Biomed. Signal Process. Control*, vol. 64, 2021, doi: 10.1016/j.bspc.2020.102365.

- [6] Y. Singh, N. Tripathi, S. Yadav, N. Gupta, A. U. Kumar, and J. V. N. Ramesh, "Transfer learning and chest X-ray-based image processing and modeling to detect COVID-19," in *Smart Technologies in Healthcare Management: Pioneering Trends and Applications*, CRC Press, 2024, pp. 240–263, doi: 10.1201/9781003330523-16.
- [7] M. M. R. Khan et al., "Automatic detection of COVID-19 disease in chest X-ray images using deep neural networks," in *IEEE Region 10 Humanitarian Technol. Conf. (R10-HTC)*, 2020, doi: 10.1109/R10-HTC49770.2020.9357034.
- [8] M. J. Awan, M. H. Bilal, A. Yasin, H. Nobanee, N. S. Khan, and A. M. Zain, "Detection of COVID-19 in chest X-ray images: a big data enabled deep learning approach," *Int. J. Environ. Res. Public Health*, vol. 18, no. 19, p. 10147, 2021, doi: 10.3390/ijerph181910147.
- [9] M. Mujahid et al., "Pneumonia classification from X-ray images with Inception-V3 and convolutional neural network," *Diagnostics*, vol. 12, no. 5, 2022, doi: 10.3390/diagnostics12051280.
- [10] S. Tang, S. Yuan, and Y. Zhu, "Convolutional neural network in intelligent fault diagnosis toward rotatory machinery," *IEEE Access*, vol. 8, pp. 86510–86519, 2020, doi: 10.1109/ACCESS.2020.2992692.
- [11] A. Rajbongshi et al., "Recognition of mango leaf disease using convolutional neural network models: a transfer learning approach," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 23, no. 3, pp. 1681–1688, 2021, doi: 10.11591/ijeecs.v23.i3.pp1681-1688.
- [12] S. Ganjei and D. K. Gunleiksrud, "Deep learning approach for binary classification of microscopic black holes and sphalerons: optimization by employing custom loss function," 2023. [Online]. Available: <https://hvlopen.brage.unit.no/hvlopen-xmlui/bitstream/handle/11250/3082038/Prosjekthandbok.pdf>
- [13] X. Chen et al., "Recent advances and clinical applications of deep learning in medical image analysis," *Med. Image Anal.*, vol. 79, 2022, doi: 10.1016/j.media.2022.102444.
- [14] B. Dong, X. Fu, and X. Kang, "SSGNet: semi-supervised multi-path grid network for diagnosing melanoma," *Pattern Anal. Appl.*, vol. 26, no. 1, pp. 357–366, Feb. 2023, doi: 10.1007/s10044-022-01100-4.
- [15] H. I. Hussein, A. O. Mohammed, M. M. Hassan, and R. J. Mstafa, "Lightweight deep CNN-based models for early detection of COVID-19 patients from chest X-ray images," *Expert Syst. Appl.*, vol. 223, p. 119900, Aug. 2023, doi: 10.1016/j.eswa.2023.119900.
- [16] G. Vilone and L. Longo, "Explainable artificial intelligence: a systematic review," 2020. [Online]. Available: <http://arxiv.org/abs/2006.00093>
- [17] Ž. Vujović, "Classification model evaluation metrics," *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 6, pp. 599–606, 2021, doi: 10.14569/IJACSA.2021.0120670.