RESNET101 AND GOOGLENET DEEP LEARNING MODELS: COMPARING SUCCESS LEVELS IN THE HEALTH SECTOR[1]

Kafkas University
Economics and Administrative Sciences Faculty
KAUJEASF
Vol. 15, Issue 29, 2024
ISSN: 1309 – 4289
E – ISSN: 2149-9136

Article Submission Date: 10.02.2024 *Accepted Date: 30.05.2024*

Muhammed Akif YENİKAYA
Asst. Prof. Dr.
Kafkas University
Faculty of Economics and Administrative Sciences,
Kars, Türkiye
akif.yenikaya@kafkas.edu.tr
ORCID ID: 0000-0002-3624-722X

ABSTRACT | Artificial intelligence (AI) applications in the healthcare sector have revolutionized medical diagnosis and treatment. Advances in this field provide many advantages such as early detection of diseases and increasing the efficiency of healthcare services. In this study, in order to investigate the usability of deep learning models for tuberculosis (TB) detection, the accuracy rates of deep learning models such as ResNet101 and GoogLeNet are compared in terms of TB detection potential in the healthcare sector. The results of the analyses revealed that deep learning networks are successful in classifying chest X-ray images with and without TB. In addition, when the success levels were analyzed, it was determined that the ResNet101 deep learning network, with a success rate of 99.3%, showed a higher score than the other deep learning model considered in the study, GoogLeNet (98.2%). These findings obtained within the scope of the research reveal the importance and functionality of AI applications in order to increase diagnostic accuracy rates.

Keywords: *Healthcare, Artificial intelligence, deep learning, Chest X-ray, disease detection.*
JEL Code: *I10, I20, M15*

Scope: *Management information systems*
Type: *Research*

DOI: 10.36543/kauiibfd.2024.015

Cite this article: Yenikaya M. A. (2024). ResNet101 and GoogLeNet deep learning models: Comparing success levels in the health sector. *KAUJEASF, 15*(29), 390-409.

[1] It has been declared that the relevant study complies with ethical rules.

RESNET101 VE GOOGLENET DERİN ÖĞRENME MODELLERİ: SAĞLIK SEKTÖRÜNDE BAŞARI DÜZEYLERİNİN KARŞILAŞTIRILMASI

Kafkas Üniversitesi
İktisadi ve İdari Bilimler Fakültesi
KAÜİİBFD
Cilt, 15, Sayı 29, 2024
ISSN: 1309 – 4289
E – ISSN: 2149-9136

Makale Gönderim Tarihi: 10.02.2024   *Yayına Kabul Tarihi: 30.05.2024*

Muhammed Akif YENİKAYA
Dr. Öğr. Üyesi,
Kafkas Üniversitesi,
İktisadi, İdari ve Bilimler Fakültesi,
Kars, Türkiye
akif.yenikaya@kafkas.edu.tr
ORCID ID: 0000-0002-3624-722X

ÖZ | Sağlık sektöründe yapay zekâ (YZ) uygulamaları, tıbbi teşhis ve tedavide önemli bir devrim niteliği taşımaktadır. Bu alandaki ilerlemeler, hastalıkların erken teşhis edilmesi ve sağlık hizmetlerinin verimliliğinin artırılması gibi birçok avantaj sağlamaktadır. Bu çalışmada, tüberküloz (TB) tespiti için derin öğrenme modellerinin kullanılabilirliğini araştırmak maksadıyla, ResNet101 ve GoogLeNet gibi derin öğrenme modellerinin sağlık sektöründe TB tespit potansiyeli bağlamında doğruluk oranları karşılaştırılmıştır. Yapılan analizlerden elde edilen bulgular, derin öğrenme ağlarının TB'li ve bu hastalığı bulundurmayan akciğer röntgen görüntüleri sınıflandırmasında başarılı olduğunu ortaya koymuştur. Ayrıca, başarı seviyeleri incelendiğinde ResNet101 derin öğrenme ağının %99.3 başarı oranı ile araştırmada ele alınan diğer derin öğrenme modeli olan GoogLeNet'e (%98.2) göre daha yüksek bir skor ortaya koyduğu tespit edilmiştir. Araştırma kapsamında elde edilen söz konusu bu bulgular, teşhis doğruluk oranlarının arttırılabilmesi için YZ uygulamalarının önem ve işlevselliğini ortaya koyar mahiyettedir.

Anahtar Kelimeler: *Sağlık hizmetleri, yapay zekâ, derin öğrenme, göğüs röntgeni, hastalık tespiti.*
JEL Kodları: *I10, I20, M15*

Alan: *Yönetim bilişim sistemleri*
Türü: *Araştırma*

1. INTRODUCTION

The healthcare sector plays a crucial role in protecting public health and preventing diseases. In this context, monitoring and controlling infectious diseases are critical elements that determine the effectiveness of health policies. Tuberculosis (TB) stands out among these diseases as a particularly notable and dangerous infection. Caused by a pathogen known as Mycobacterium tuberculosis, TB is more commonly seen in developing countries and poses a significant threat to public health (World Health Organization, 2018). Therefore, preventing the spread of TB and minimizing its impacts are among the primary goals of the healthcare sector. These countries are the biggest victims of this deadly disease due to weaknesses in their health systems. This is related to the large and persistent impact of latent TB infection, as well as biological factors such as underinvestment in the development of new medical tools, the adoption of standardized approaches that do not meet the different needs of individual patients, HIV (Human Immunodeficiency Virus) co-infection and the spread of drug resistance (Reid et al., 2019).

Pulmonary TB is the most common and most complicated subtype of TB and forms the basis of this study. This type mainly affects the chest cavity and more frequently the lungs. In most cases, prolonged chest pains, cough, fever and fatigue are among the many symptoms associated with this type of TB (World Health Organization, 2018). Chest X-rays are crucial for the diagnosis of TB as they show detailed anatomical information as well as specific radiologic patterns associated with the disease, including infiltrates, cavitations and nodules (Williams, 1907). However, human experts' assessment of chest X-ray images is subjective and prone to differences in experience and proficiency.

Recent advances in deep learning and machine learning algorithms offer a great chance to improve TB detection. By using artificial intelligence to scan chest X-ray images, these algorithms are able to understand the complex patterns and features associated with TB. This leads to a more accurate and consistent detection process.

The effective development of a uniform and impartial tool for medical practitioners, a TB detection model utilizing deep learning networks has the potential to completely transform TB diagnosis. These models can facilitate early diagnosis, quick treatment initiation, and the application of efficient TB control strategies (Khan et al., 2019). Additionally, areas with access to trained radiologists and particularly restricted health resources may benefit from the integration of artificial intelligence-based TB detection models into clinical processes. This could also aid international efforts to combat this fatal illness that keeps reoccurring.

The study highlights the importance of the role of artificial intelligence in improving diagnostic accuracy rates in healthcare. The use of artificial intelligence models, especially for the early diagnosis of diseases such as TB, accelerates patients' access to treatment and increases treatment success. This provides a significant improvement in healthcare services. Within the scope of this study, it was investigated whether ResNet101 and GoogLeNet, which are deep learning models, can be used in the health sector for TB detection and the accuracy rates obtained from these models were compared.

2. CONCEPTUAL FRAMEWORK

2.1. Deep Learning

In the field of science, deep learning initially gained prominence in 2012. A deep learning base architecture, which is regarded as a convolutional neural network (CNN), was used at the time to win the Large Scale Visual Recognition Competition (ImageNet), the biggest competition in the field of object recognition (Krizhevsky, Sutskever & Hinton, 2012). This incident marks the start of a phenomenal ascent in the deep learning sector. Using the CNN created by Krizhevsky et al. (2012), the top 5 error rate was lowered from 26.1% in the competition to 15.3%. The top 5 error rate has been further lowered to 3.6% by deep learning advancements. The imageNet competition's Top 5 error rate graph throughout time is displayed in Figure 1.

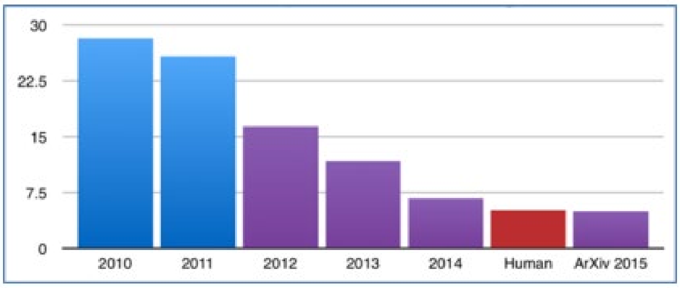


Figure 1: Top 5 Error Rates for The ImageNet Competition over The Years
Source: (Julia, 2016)

According to Deng and Yu (2014), deep learning is a type of machine learning that extracts and transforms features using multiple layers of nonlinear processing units. The output of the preceding layer serves as the input for each layer that follows.

The foundation of the deep learning methodology is the acquisition of several feature or data representation layers. A hierarchical representation is created by deriving higher-level features from lower-level features. According to Bengio (2009), this representation is capable of learning several levels of representation that correlate to various abstraction levels. One can consider features like clusters of edges or a vector of intensity values per pixel when representing an image. The features that most accurately reflect the data are selected from this group. At this point, deep learning techniques come into play, replacing manually derived features with effective hierarchical feature extraction algorithms that best describe the data (Song & Lee, 2013).

Even if there is a historical basis for early deep learning efforts, the large amount of data that is currently available is a major factor in their success. The learning algorithms used to tackle toy issues in the 1980s and the learning algorithms employed in today's demanding tasks share fundamental similarities in advanced deep learning models. Nevertheless, compared to earlier techniques, these models now include changes that make it easier to train deeper structures.

2.2. Convolutional Neural Network

Convolutional Neural Network (CNN) are a type of deep learning algorithm, often used in areas such as visual data analysis. CNN use a stack of layers that extract key features from the input data during training and make predictions close to human precision in specific problem domains. This architecture has been particularly successful in tasks such as image recognition, classification and object detection.

Deep learning algorithms often perform their tasks more efficiently by using complex architectures. This is based on features such as utilizing large data sets in the learning process, feature engineering and automatic updating of weights during learning. CNN has achieved great success, especially in areas such as computer vision and audio processing and has been used in many application areas (Krizhevsky et al., 2012).

The deep convolutional neural network (DCNN) offers an intuitive and powerful network architecture for deep learning. DCNN is a stack of CNN layers specially assembled to achieve a specific image classification, detection, or recognition problem. DCNN represents the application of deep learning techniques in CNN algorithms. In recent years, with the development of deep learning technology, more efficient DCNN models such as AlexNet, GoogLeNet, ResNet, VGG, DenseNet have become prominent. These DCNNs have demonstrated high performance, especially in image classification tasks, making it possible for computers to be more effective than humans in visual classification (Lopes & Valiati, 2017).

This study focuses on ResNet and GoogLeNet deep learning models, which are considered to be more successful in medical image analysis and are widely studied in the literature.

2.2.1. ResNet

ResNet is a deep learning model with more than 152 layers. In addition to this network, other versions with 34, 50 and 101 layers have been developed (He et al., 2016). The ResNet architecture has achieved great success, topping the 2015 ILSVRC competition and achieving accuracy approximating human precision (Targ, Almeida & Lyman, 2016). The first 34-layer network architecture of ResNet is given in Figure 2.

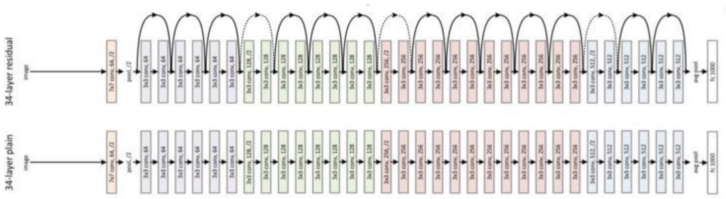


Figure 2: ResNet Architecture
Source: (He et al., 2016)

Figure 3 shows the Residual block structure (ReLu) of the ResNet architecture. In a normal situation, the process from input to output can be mapped by a nonlinear function H(x). However, in the ResNet architecture, instead of H(x), this process is mapped by another function defined as F(x) = H(x) - x.

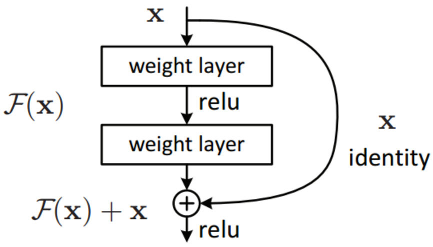


Figure 3: Residual Block Structure of the ResNet Architecture
Source: (He et al., 2016)

2.2.2. GoogLeNet

The winner of the 2014 ILSVRC competition, GoogLeNet, is also known as InceptionV1. This architecture consists of 22 layers, making it a very deep architecture. GoogLeNet has a fixed input size of 224 x 224 and uses ReLu activation to create linearity (Szegedy et al., 2015).

This architecture, in contrast to others, analyzes images in parallel as opposed to piling layers on top of one another. This is due to the fact that stacked operations take into account drawbacks such memory size increase and time loss. In Figure 4, the GoogLeNet network architecture is displayed.

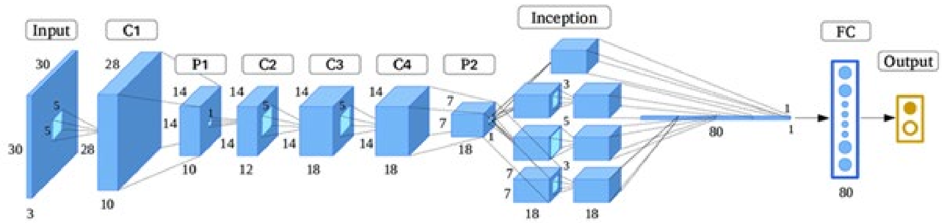


Figure 4. GoogLeNet Architecture
Source: (Guo et al., 2017)

In Figure 4, P1 and P2 are the pooling layer, C1, C2, C3, C4 are the convolution layer and FC is the fully connected layer.

3. STUDIES ON DEEP LEARNING

In the field of medical imaging, the potential of deep learning techniques offers a vast research area to improve computer-aided registration, detection, segmentation and finally diagnosis (Shen et al., 2017). Ideally, these computer-assisted capabilities would enable more effective analysis of images examined under a microscope. With the use of these techniques, diagnoses can be made more precisely and quickly on medical images, treatment plans can be created, and patient follow-up can be made more effective. Developments in this field have significant potential for early diagnosis and treatment of diseases in the medical field.

Hwang et al. (2016) introduced a 6-layer CNN architecture trained with full tuberculosis chest X-ray images. The obtained results provided values ranging from 0.96, 0.93 and 0.88 according to the area under the curve criteria.

The automatic diagnosis system proposed by Jaeger et al. (2013) provides 78.3% accuracy on the first dataset and 84.0% accuracy on the second dataset. These results were obtained using support vector machines (SVM) for classification after graph cutting for lung segmentation.

A report by Lakhani and Sundaram (2017) proposes the use of automation for tuberculosis diagnosis. The paper presents results from two different DCNN models, AlexNet and GoogLeNet. Both models include pre-trained and untrained cases. An ensemble model is also proposed that uses half of the weighted average of the probability scores obtained from AlexNet and GoogLeNet. In this context, the GoogLeNet, AlexNet and Ensemble models achieved high accuracy rates of 95.3%, 93.3% and 96.0% respectively.

Deep learning algorithms presented by Faust et al. (2018) were used to perform automatic diagnosis of tuberculosis. In this study, a neural network optimized using a genetic algorithm was developed and an average accuracy of 94.88% was achieved.

In a study by Bar et al. (2015), area under curve (AUC) values ranging from 0.87-0.94 were achieved for distinct diseases utilizing a dataset of at least 433 pictures. This study is the first of its type to demonstrate that domain-specific representations, which have not yet been used to general medical image identification tasks, may be effectively replaced by deep learning using ImageNet, a sizable non-medical image database.

Panicker, Kalmady, Rajan and Sabu (2018) used CNN for automatic TB detection. Using this technique, TB was identified in sputum smear images, with an F-score of 86.76% attained.

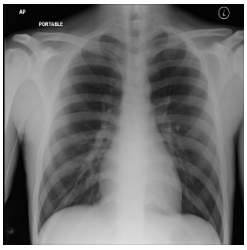
The questions to be answered within the scope of the research are as follows:

1. Can deep learning networks be used in image-based disease diagnosis?
2. At what training/test ratio do deep learning networks produce more successful results?
3. Which of the GoogLeNet and Resnet101 deep learning networks considered in the context of the research is more successful in disease detection by image parsing?

4. MATERIALS AND METHODS

For the training and testing of the deep learning models to be used for TB detection in this study, a dataset of 1400 chest X-rays (CXR), 700 of which are TB cases and the remaining 700 are normal cases, was obtained from the open source kaggle.com (Kaggle, 2021). Sample CXR images in this dataset are given in Table 1.

Table 1: Example of a CXR image from the Kaggle dataset

Without TB	With TB
	

To prepare the data, images with and without TB were organized into different folders. Each image was visually inspected to guarantee accurate identification and high quality. Before training the models, the images were subjected to a preprocessing procedure called scaling. Using the labeled CXR dataset, deep learning models were trained to identify patterns and attributes.

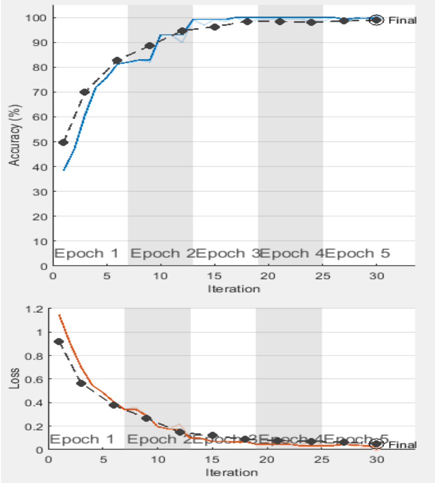
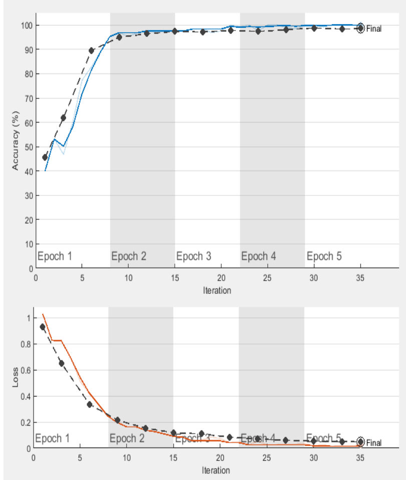
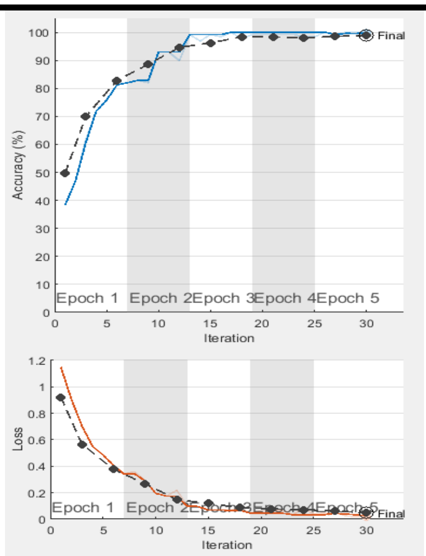
The computations for using the backpropagation algorithm during the training of a network can be performed faster with parallel processors. Therefore, specialized processors such as powerful graphics processing units (GPUs) should be preferred for training deep networks. For this reason, the models used in this study were run on a server with NVidia RTX 4000 Quadro GPU and 128 GB RAM.

The data preprocessing, model training and results evaluation processes of ResNet101 and GoogLeNet deep learning architectures used in this study were implemented and trained in Matlab environment. The dataset was split 60%-40%, 70%-30% and 80%-20% for training and testing, respectively. For 30 iterations with 0.001% learning rates, all models were executed for 5 epochs.

5. FINDINGS

Table 2 shows the accuracy and loss graphs for the dataset using the ResNet101 model, separated at different rates (60%-40%, 70%-30% and 80%-20%) for training and testing.

Table 2: Accuracy and Loss Graphs for ResNet101 Model

<table>
<tr><th>60%-40%</th><th>70%-30%</th></tr>
<tr><td>
</td><td>
</td></tr>
<tr><th>80%-20%</th><th></th></tr>
<tr><td>
</td><td>Accuracy
Training (smoothed)
Training
Validation
Loss
Training (smoothed)
Training
Validation</td></tr>
</table>

When the graphs are analyzed, it is seen that, as expected, the accuracy increases with each iteration and the loss decreases with each iteration.

Confusion matrix is a concept often used to evaluate the performance of a model. This matrix allows a model's predictions to be compared with actual values and shows how accurate or inaccurate the model is in making predictions.

The error matrix is widely used, especially for classification problems.

The confusion matrix usually contains four basic terms:

- True Positive (TP): When the model correctly predicts a positive sample.
- False Positive (FP): This is when the model incorrectly predicts a sample as positive when it is actually negative.
- True Negative (TN): When the model correctly predicts a negative sample.
- False Negative (FN): When the model incorrectly predicts a sample as negative when it is actually positive.

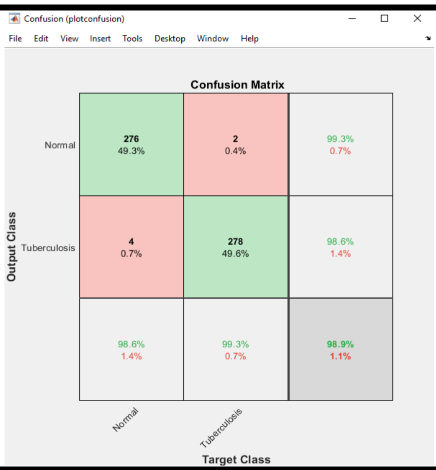
TN	FP
FN	TP

Interpreting the confusion matrix helps to better understand the performance of the model. For example, high TP values may indicate the model's ability to correctly recognize positive samples, while high FP values may indicate that the model tends to incorrectly recognize negative samples as positive.

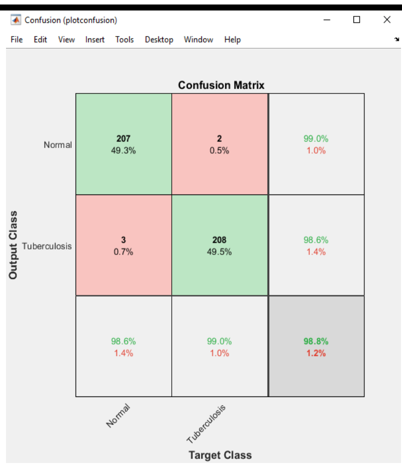
Table 3 contains the confusion matrix for training and test data separated by different percentages (60%-40%, 70%-30% and 80%-20%) in the dataset using the ResNet101 model.

Table 3: Confusion Matrix of ResNet101 Model

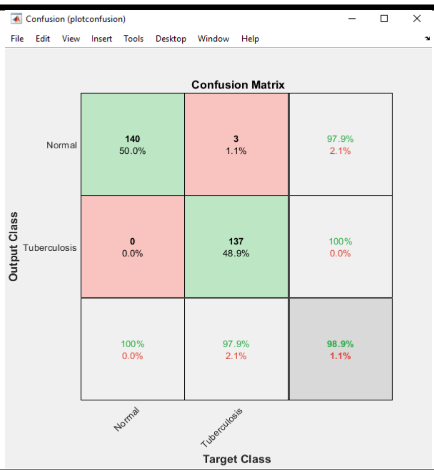
60%-40%



70%-30%



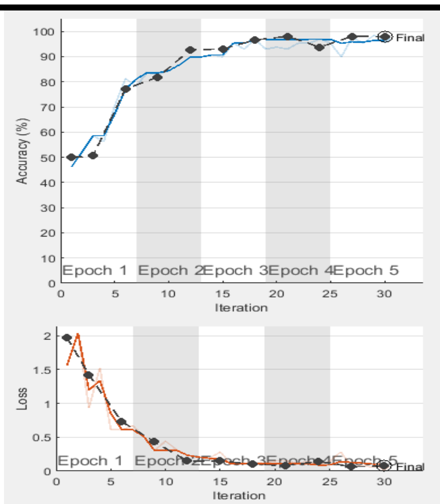
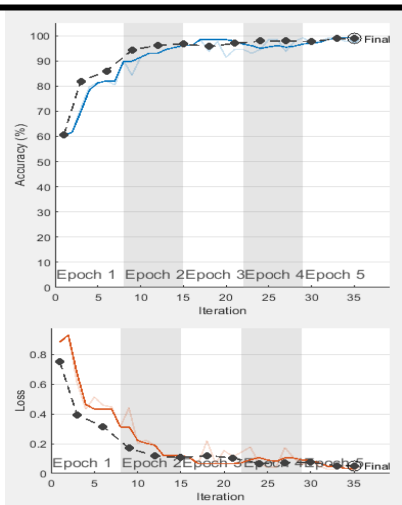
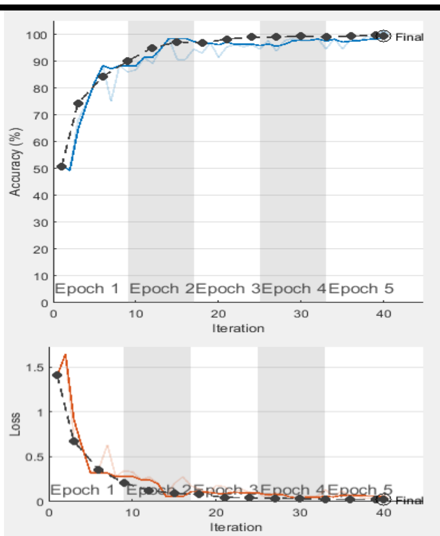
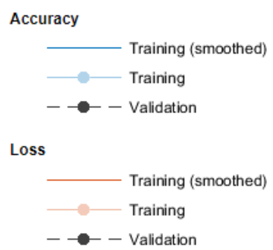
80%-20%



When the confusion matrix of the ResNet101 model is examined, an accuracy rate of 99.3% in normal cases (without TB), 98.6% in abnormal cases (with TB) and 98.9% in total was achieved for the data set reserved for 60% training and 40% testing. For the data set reserved for 70% training and 30% testing, an accuracy rate of 99.0% in normal cases (without TB), 98.6% in abnormal cases (with TB) and 98.8% in total was achieved. In the data set reserved for 80% training and 20% testing, an accuracy rate of 97.9% in normal cases (without TB), 100% in abnormal cases (with TB) and 98.9% in total was achieved.

Table 4 shows the accuracy and loss graphs for the dataset using the GoogLeNet model, separated by different percentages (60%-40%, 70%-30% and 80%-20%) for training and testing.

Table 4: Accuracy and Loss Graphs for GoogLeNet Model

<table>
<tr><th>%60-%40</th><th>%70-%30</th></tr>
<tr><td>
</td><td>
</td></tr>
<tr><th>%80-%20</th><th></th></tr>
<tr><td>
</td><td>
</td></tr>
</table>

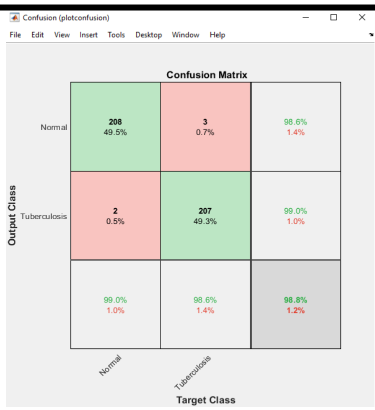
When the graphs are analyzed, it is seen that, as expected, the accuracy increases with each iteration and the loss decreases with each iteration.

Table 5 shows the confusion matrix for training and test data separated by different percentages (60%-40%, 70%-30% and 80%-20%) for the dataset using the GoogLeNet model.

Table 5: Confusion Matrix of GoogLeNet Model

60%-40%

70%-30%



80%-20%

When the confusion matrix of the GoogLeNet model is examined, an accuracy rate of 99.3% in normal cases (without TB), 96.5% in abnormal cases (with TB) and 97.9% in total was achieved for the data set reserved for 60%

training and 40% testing. For the data set reserved for 70% training and 30% testing, an accuracy rate of 98.6% in normal cases (without TB), 99.0% in abnormal cases (with TB) and 98.8% in total was achieved. In the data set reserved for 80% training and 20% testing, an accuracy rate of 99.3% in normal cases (without TB), 99.3% in abnormal cases (with TB) and 99.3% in total was achieved.

Considering these situations, the findings in Table 6 below were obtained.

Table 6: Accuracy Percentages Calculated According to the Training-Testing Ratios of the Models

Model	Training-Test Ratio	Accuracy
ResNet101	60%-40%	98.9%
ResNet101	70%-30%	98.8%
ResNet101	80%-20%	98.9%
GoogLeNet	60%-40%	97.9%
GoogLeNet	70%-30%	98.8%
GoogLeNet	80%-20%	99.3%

When the findings obtained in Table 6 are examined, it is seen that the highest accuracy rate, with 99.3%, belongs to the GoogLeNet model, whose data set is divided into 80% for training and 20% for testing.

6. CONCLUSION

In today's world, digital technologies take an active place in almost every field and significantly transform conventional practices and processes. One of the areas where this change is taking place is undoubtedly the health sector. Wide field of activity, need for qualified people, high workload, etc. In the healthcare sector, which has its own dynamics in terms of factors such as, technology is among the indispensable elements in terms of reducing errors, increasing quality, and making progress. In the context of these needs, in recent years, many academic studies have been conducted on the use of artificial intelligence in the healthcare sector, and with these studies, solutions are sought to some problems that have become evident in the healthcare sector (long waiting times, diagnosis and treatment errors, shortage of qualified people, etc.).

In this research, firstly, the general working logic of deep learning networks was explained, and then an experimental study was conducted on the usability of artificial intelligence in disease diagnosis, and the diagnostic accuracy rates of GoogLeNet and Resnet101 networks, which are important deep learning networks, were compared with the visual discrimination method. The findings and literary contributions obtained within the scope of the study are presented below.

In the research, 1400 chest x-ray images available in open sources (kaggle.com) were used, and these images were selected from 700 tuberculosis and 700 non-tuberculosis images. These selected images were defined to deep learning networks in three different ways: 60% training, 40% testing, 70% training, 30% testing and finally 80% training, 20% testing. After definition, both deep learning networks were tested with the same data groups and the results were compared separately. In line with the research findings, it has been determined that both deep learning architectures have successful accuracy rates (97.9% - 99.3%) and it has been demonstrated that deep learning can be used in disease diagnosis with the visual discrimination method. This finding is in parallel with studies (Onno, Khan, Daftary & David, 2023; David, Onno, Keshavjee & Khan, 2022; Codlin et al., 2021; Khan et al., 2020) showing that deep learning architecture is successful in detecting tuberculosis disease with the image parsing method. This finding is one of the questions to be answered in the research: "Can deep learning networks be used in image-based disease diagnosis?" It is an answer to the question.

Another issue emphasized within the scope of the research is whether the training and testing rates of deep learning networks affect the success achieved at the point of image parsing. At this point, both of the deep learning networks tested with training and testing performed with three different methods gave the highest score at a training/test ratio of 80/20, while GoogLeNet gave the lowest score at a training/test ratio of 60/40 and Resnet101 at a training/test ratio of 70/30. This finding is important in that it can create insight in future research on creating the optimum ratio for training network architectures. This finding in question is based on the second research question: "At what training/test ratio do deep learning networks provide more successful results?" It is an answer to the question.

As a result of the research, as mentioned before, Resnet101 and GoogLeNet deep learning networks were utilized and the success of these networks in diagnosing tuberculosis disease over radiology images was compared. The findings show that the GoogLeNet deep learning architecture has the highest success rate in terms of data identification, separation of patient/healthy images and diagnostic accuracy with 80/20 training/testing and 99.3% success rate in this study. This result answers the third research question, "Which of the GoogLeNet and Resnet101 deep learning networks considered in the context of the research is more successful in disease detection with image parsing?".

Yenikaya and Kerse's (2022) study on the detection of age-related macular degeneration types using deep learning models showed that GoogLeNet had higher success rates than AlexNet. Lee and Nam's (2021) study evaluating drug response in cancer treatment indicated that AlexNet and GoogLeNet deep learning models were more successful than the LASSO algorithm. Yenikaya and Oktaysoy's (2023) study on the identification of brain tumors revealed that ResNet101, one of the deep learning models used, was relatively more successful than GoogLeNet. According to the findings of the study by Yenikaya, Kerse and Oktaysoy (2024), ResNet101 was found to be more successful in detecting COVID-19, viral pneumonia, and normal (disease-free) conditions compared to the other models used in the study.

These results show that deep learning networks can be effective in different medical applications. For example, deep learning models such as GoogLeNet for eye disease identification and ResNet101 for brain tumor detection may be more successful. However, since each application has different requirements and data characteristics, it is important to choose the right model. Therefore, future research needs to more thoroughly evaluate the performance of different deep learning architectures and techniques in different medical applications. Moreover, larger-scale and multicenter clinical studies are needed to validate and generalize the results of these studies.

7. CONFLICT OF INTEREST STATEMENT

There is no conflict of interest between the authors.

8. FINANCIAL SUPPORT

This research received no specific grant from any funding agency.

9. AUTHOR CONTRIBUTIONS

The authors' contributions to the study are equal.

10. ETHICS COMMITTEE STATEMENT AND INTELLECTUAL PROPERTY COPYRIGHTS

The study does not require clearance from an ethics commission.

11. REFERENCE

Bar, Y., Diamant, I., Wolf, L., & Greenspan, H. (2015). *Deep learning with non-medical training used for chest pathology identification*. Paper presented at the Medical Imaging 2015: Computer-Aided Diagnosis.

Bengio, Y. (2009). Learning deep architectures for AI. *Foundations and Trends in Machine Learning, 2*(1), 1–127.

Codlin, A. J., Dao, T. P., Vo, L. N. Q., Forse, R. J., Van Truong, V., Dang, H. M., ... & Caws, M. (2021). Independent evaluation of 12 artificial intelligence solutions for the detection of tuberculosis. *Scientific Reports, 11*(1), 23895.

David, P. M., Onno, J., Keshavjee, S., & Khan, F. A. (2022). Conditions required for the artificial-intelligence-based computer-aided detection of tuberculosis to attain its global health potential. *The Lancet Digital Health, 4*(10), e702-e704.

Deng, L., & Yu, D. (2014). Deep learning: Methods and applications. *Foundations and Trends in Signal Processing, 7*(3–4), 197–387.

Faust, O., Hagiwara, Y., Hong, T. J., Lih, O. S., & Acharya, U. R. (2018). Deep learning for healthcare applications based on physiological signals: A review. *Computer Methods and Programs in Biomedicine, 161,* 1-13.

Guo Z., Chen Q., Wu G., Xu Y., Shibasaki R., & Shao X. (2017). Village building identification based on ensemble convolutional neural networks. *Sensors, 17*(11), 1-22.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). *Deep residual learning for image recognition.* Paper presented at the Proceedings of the IEEE conference on computer vision and pattern recognition.

Hwang, S., Kim, H.-E., Jeong, J., & Kim, H.-J. (2016). *A novel approach for tuberculosis screening based on deep convolutional neural networks.* Paper presented at the Medical Imaging 2016: Computer-Aided Diagnosis.

Jaeger, S., Karargyris, A., Candemir, S., Folio, L., Siegelman, J., Callaghan, F., Xue, Z., Palaniappan, K., Singh, R. K., & Antani, S. (2013). Automatic tuberculosis screening using chest radiographs. *IEEE Transactions on Medical Imaging, 33*(2), 233-245.

Julia, D. L. F. (2016). "Mocha.jl: Deep learning in Julia." Retrieved from https://devblogs.nvidia.com/parallelforall/mocha-jl-deep-learning-julia/

Kaggle (2021). Accessed: January 25, 2024: https://www.kaggle.com/datasets/tawsifurrahman/tuberculosis-tb-chest-xray-dataset

Khan, M. T., Kaushik, A. C., Ji, L., Malik, S. I., Ali, S., & Wei, D. Q. (2019). Artificial neural networks for prediction of tuberculosis disease. *Frontiers in Microbiology, 10,* 1-9.

Khan, F. A., Majidulla, A., Tavaziva, G., Nazish, A., Abidi, S. K., Benedetti, A., ... & Saeed, S. (2020). Chest x-ray analysis with deep learning-based software as a triage test for pulmonary tuberculosis: a prospective study of diagnostic accuracy for culture-confirmed disease. *The Lancet Digital Health, 2*(11), e573-e581.

Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). *ImageNet classification with deep convolutional neural networks.* In Advances in Neural Information Processing Systems (Vol. 25, pp. 1097–1105).

Lakhani, P., & Sundaram, B. (2017). Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. *Radiology, 284*(2), 574-582.

Lee, Y., & Nam, S. (2021). Performance comparisons of AlexNet and GoogLeNet in cell growth inhibition IC50 prediction. *International Journal of Molecular Sciences, 22*(14), 1-12.

Lopes, U. K., & Valiati, J. F. (2017). Pre-trained convolutional neural networks as feature extractors for tuberculosis detection. *Computational Biology and Medicine, 89,* 135–143.

Onno, J., Khan, F. A., Daftary, A., & David, P. M. (2023). Artificial intelligence-based computer aided detection (AI-CAD) in the fight against tuberculosis: Effects of moving health technologies in global health. *Social Science & Medicine, 327,* 115949.

Panicker, R. O., Kalmady, K. S., Rajan, J., & Sabu, M. K. (2018). Automatic detection of tuberculosis bacilli from microscopic sputum smear images using deep learning methods. *Biocybernetics and Biomedical Engineering, 38*(3), 691-699.

Reid, M. J., Arinaminpathy, N., Bloom, A., Bloom, B. R., Boehme, C., Chaisson, R., Chin, D. P., Churchyard, G., Cox, H., & Ditiu, L. (2019). Building a tuberculosis-free world: The Lancet Commission on tuberculosis. *The Lancet, 393*(10178), 1331-1384.

Shen, D., Wu, G., & Suk, H.-I. (2017). Deep learning in medical image analysis. *Annual Review of Biomedical Engineering, 19,* 221-248.

Song, H. A., & Lee, S.-Y. (2013). Hierarchical Representation Using NMF. In International Conference on Neural Information Processing (pp. 466–473).

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). *Going deeper with convolutions.* In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

Targ, S., Almeida, D., & Lyman, K. (2016). *Resnet in Resnet: Generalizing residual architectures.* arXiv preprint arXiv:1603.08029.

Williams, F. H. (1907). The use of X-ray examinations in pulmonary tuberculosis. *Boston Medical and Surgical Journal, 157,* 850–853.

World Health Organization (2018). *Global tuberculosis report 2018*. Geneva: World Health Organization.

Yenikaya, M. A., Kerse, G. (2022). *A comparison of accuracy rates of Alexnet and Googlenet deep learning models in image classification.* In Congress Book VII. International European Conference on Social Sciences, Antalya, Türkiye, pp. 713-719.

Yenikaya, M. A., & Oktaysoy, O. (2023). The use of artificial intelligence applications in the health sector: Preliminary diagnosis with deep learning method. *Sakarya University Graduate School of Business Journal, 5*(2), 127-131.

Yenikaya, M. A., Kerse, G., & Oktaysoy, O. (2024). Artificial intelligence in the healthcare sector: Comparison of deep learning networks using chest X-ray images. *Frontiers in Public Health, 12,* 1386110.