➢ *ORAL PRESENTATION*

Automatic detection of Alzheimer's disease with transfer learning and image processing techniques

Atılay Yeşilada[1*] (*https://orcid.org/0000-0003-2845-5800*), Osman Gökalp[2] (*https://orcid.org/0000-0002-7604-8647*)

[*1]University of İzmir Kâtip Çelebi, Faculty of Natural and Applied Sciences, Department of Software Engineering, İzmir, Türkiye
[2]University of İzmir Kâtip Çelebi, Faculty of Engineering and Architecture, Department of Computer Engineering, İzmir, Türkiye

*Corresponding author e-mail: y210240039@ogr.ikcu.edu.tr

Abstract

Alzheimer's disease continues to be a serious public health issue. Alzheimer's disease affects generally old people, and it is a neurological disorder. The most significant sign of Alzheimer's disease is memory loss, which becomes worse over time. Therefore, early Alzheimer's disease identification can aid in patients getting the proper care. One of the aims of this study is to automate the early diagnosis of Alzheimer's disease. While detecting automatically, a system that can improve itself by benefiting from previous experiences each time can be established with artificial neural network models. In this way, in late diagnosis, the bad effects of the disease can be protected at an early stage. Thanks to artificial intelligence and image processing algorithms, successful results can be achieved with MRI data used by transfer learning and deep learning models. In this study, a deep learning method with transfer learning based on brain MRI images for the detection of Alzheimer's disease has been proposed. In the experimental work, different convolutional neural network architectures and transfer learning methods have been tried. The results show that the detection of Alzheimer's disease can be detected with remarkable accuracy with the transfer learning method.

Keywords: Alzheimer's Disease, Brain MRI, Deep Learning, Transfer Learning, Convolutional Neural Networks, InceptionV3

INTRODUCTION

Alzheimer's disease is a worldwide neurodegenerative disease that occurs in the human brain and significantly affects the patient's quality of life (Arendt, 2009). It is reported that the number of globally affected patients is expected to double in the next 20 years, and one in 85 people will be affected by 2050 (Jellinger, 2003). Each year, 7.7 million new cases of dementia are reported globally (Bhagyashree and Sheshadri, 2018). One of the most well-known cases of dementia among the elderly is Alzheimer's disease. The following are some of the key signs of this dementia: difficulty communicating, memory issues, and different disabilities that make the patient's everyday life exceedingly challenging (Burkea et al., 2016). Besides that, Alzheimer's sufferers are troubled by perplexity, difficulty learning new situations, speech and language challenges, poor motivation, and other complications (Shimizu et al., 2018). Early and accurate diagnosing of Alzheimer's disease can greatly benefit its treatment. Alzheimer's disease is typically diagnosed in clinics after a clinical evaluation and the completion of a questionnaire by patients and their loved ones. Due to the lack of knowledge regarding disease-related brain regions and known symptoms such as brain shrinkage, it is a challenging undertaking (Kazemi et al., 2018). For several types of brain scans, including magnetic resonance imaging (MRI), well-known machine learning classification techniques have been used. Due to the complexity of medical imaging, this topic is still challenging. With the aid of ongoing and upcoming research in this area, we can better understand the course of the disease and create new treatments. Deep learning, a subclass of machine learning, has been used for intelligent systems in many fields, particularly in the processing of medical pictures, and it can handle difficult decision-making tasks. Convolutional neural networks have shown promising results for illness diagnosis and organ segmentation (Litjens et al., 2017). CNNs combine the three key phases of a classification job, namely feature extraction, feature selection, and classification, in contrast to conventional machine learning approaches. In automated medical image analysis, convolutional neural networks have achieved major advancements. Transfer learning is one of the most effective methods for training deeper networks without overfitting when the amount of data is smaller (Xiao et al., 2018). A pre-trained network is the foundation of transfer learning. The transfer learning strategy can take advantage of the generic feature mappings and concentrate on learning the problem-specific ones for the problem of interest rather than training a specialized CNN from scratch.

In this work, we developed a deep learning method that automatically detects Alzheimer's disease by making tunings over a pre-trained model using the transfer learning method and applying data augmentation. We employed pre-trained networks that use different architectures such as InceptionV3, ResNet50V2, and Xception. An experimental work demonstrated that, as a base model, InceptionV3 can be used for the automatic detection of Alzheimer's disease and can be operated by the transfer learning method.

The rest of this work is organized in the following manner. The related work section reviews the existing methods and research in the literature related to Alzheimer's disease detection and classification. In the next section, it is an overview of the entire methodology of this study such as the used dataset and how it was obtained, deep learning, and transfer learning techniques and how they were used. Implementation details, results, and evaluation are discussed in the next section. Conclusions and final thoughts are mentioned in the last part of this paper.

RELATED WORK

In recent years, many different methods have been used for disease prediction in medical fields. Recent developments in the field of computer vision have focused on establishing models for the identification and categorization of Alzheimer's disease as well as extracting usable characteristics using a machine learning approach. Since the 1980s, brain imaging tools such as Magnetic Resonance Imaging (MRI), Positron Emission Tomography (PET) and Diffusion Tensor Image (DTI) have been used for the diagnosis and classification of Alzheimer's disease (McGeer, 1986).

For the segmentation of the gray matter, Nair and Mohan employed the Gaussian Mixing Model (Nahir and Mohan, 2017). Both the conventional classification problem and the difficulty of calculating the density are addressed by this paradigm. Then, features were retrieved using the Partial Least Squares approach, and the Support Vector Machine technique was employed for classification. In addition to these studies, a classification method built on level-set segmentation is offered (Rabeh et al., 2017). It begins by determining which of the four learning examples is closest to the new input using four distances in this investigation. A different approach has also been suggested, and they used 3D masking to isolate the Volume of Interest (VOI) region using voxel-based morphometry for feature extraction (Mufidah et al., 2017). Besides, the segmentation of images was used based on voxel morphometry to collect the 3 brain tissues (the white matter, the gray matter, and cerebrospinal fluid) (Angkoso et al., 2018). Different methods used deep learning architecture especially CNN to detect and classify Alzheimer's disease. A graph convolutional neural network (GCNN) classifier was employed by Song (Song et al., 2019). This network's eleven layers (nine convolutional layers and two fully connected layers) enable the division of Alzheimer's disease patients into four classes.

The work that used a 2D CNN with one convolutional layer, a max-pooling layer, and ultimately a neural network with a single hidden layer for classification is one of the earliest investigations of a transfer learning strategy for Alzheimer's disease diagnosis utilizing deep learning (Gupta et al., 2013). This study showed that the auto-classification encoder's performance in the subsequent layers can be improved by using real-world photos to train it. Later on, a different transfer learning technique was also put out, wherein three 2D CNNs with two convolutional layers were trained using only three slices from the middle of the MRI images of the hippocampal region (Aderghal et al., 2018). The impact of transfer learning when used for medical image classification was demonstrated in one of the subsequent studies (Shin et al., 2016). According to a study on several modalities, the results of the fine-tuning processes performed are superior. In addition, a transfer learning technique that uses ultrasound scans to locate plans and can transfer information on fewer layers is proposed (Chen et al., 2015). In recent years, as shown in Figure 1, interest in how artificial intelligence may be applied to medical imaging has increased exponentially. This interest is reflected in the rapid rise in publications in artificial intelligence from 200 peer-reviewed publications in 2010 to about 1000 in 2019 (Lui et al., 2020).

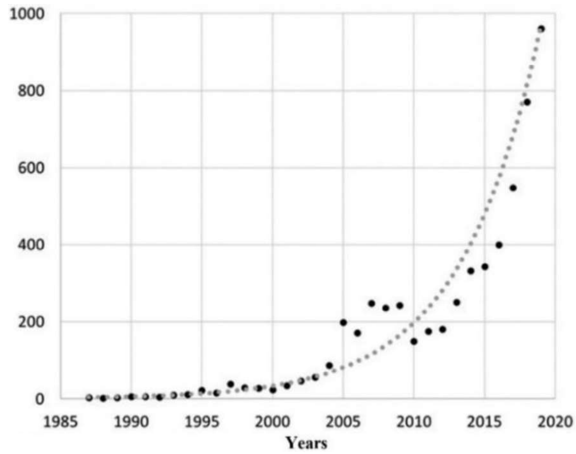

Figure 1. The number of neuroimaging artificial intelligence publications (Lui et al., 2020).

According to related works, transfer learning models can reach an accuracy equal to or greater than that of a 3D-CNN model trained from scratch. As in related previous studies, instead of creating something new from scratch, high accuracy values can be obtained and studies can become more efficient by transferring similar features with transfer learning, much easier.

MATERIALS AND METHODS

In this section, the dataset containing brain MRI images is prepared and described for Alzheimer's disease. In addition, deep learning models and transfer learning techniques are explained. After that, how Alzheimer's disease can automatically be classified in this study with artificial intelligence is explained with the dataset containing brain MRI images.

Dataset

The dataset used in this study is an open database on the Kaggle website [1]. The dataset consists of train and test folders containing a total of 6400 images. In this dataset, there are four classes of images: “Mild Demented”, “Moderate Demented”, “Non-Demented", and "Very Mild Demented". The main purpose of this dataset is to help researchers develop a model with high accuracy to detect and classify the stage of Alzheimer's disease.

Data Preparation

In the data preparation phase, libraries such as Tensorflow, Matplotlib, NumPy, and Pandas were installed and used. Afterward, paths to the images were identified and constants and seeds were determined for reproducibility. Then, the dataset was split into %80 train and %20 test batches consisting of 4 classes and determined the number of images for each class was. Images were copied and resampled from the train folder and split randomly with equal proportions in each class for the validation folder.

Data Preprocessing

Data were grayscale and they needed to be converted to RGB to use with pre-trained deep learning models. In the image transformation phase for data preprocessing, images were scaled. As shown in Figure 2, images are converted to RGB and set seed for reproducibility. Then, images were displayed in a grid of 9 images with their labels.

[1] Alzheimer's Dataset, https://www.kaggle.com/datasets/tourist55/alzheimers-dataset-4-class-of-images.

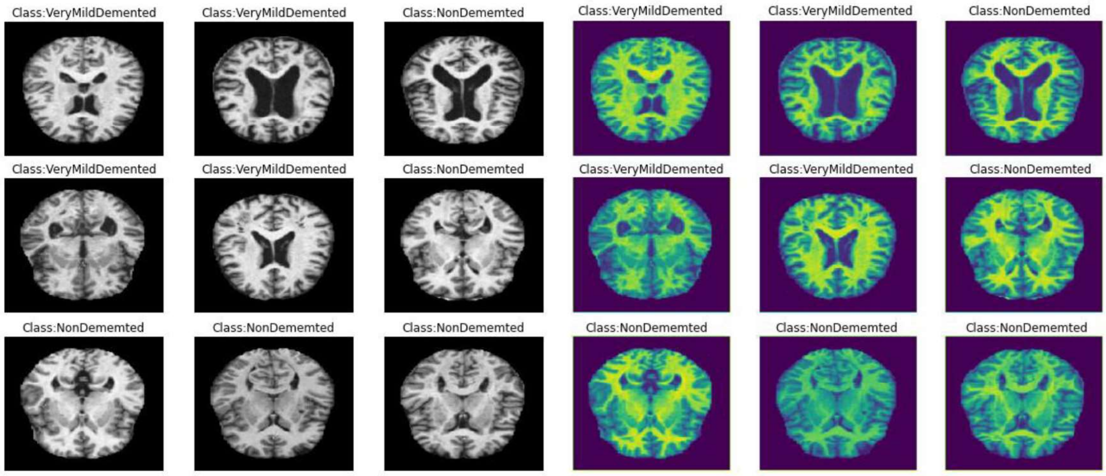

Figure 2. Brain MRI images with labels grayscale (left) and RGB (right).

Deep Learning

Deep learning is an interdisciplinary field of study developed on designing systems that require data, algorithms, and hardware knowledge, aiming to solve complex problems that humans can solve. Although there are different deep learning models used today, they vary according to the field of study and data. Convolutional neural networks are used for image data. Deep learning has approached the human level in image classification, voice recognition, and the ability to answer questions.

Convolutional Neural Network (CNN): CNN is a multilayered network structure that was developed using the multilayer perceptron, which is the standard neural network architecture. CNN designed for object detection finds its use in image classification, segmentation, pattern recognition, etc. Due to its autonomous style of operation, CNN has become a crucial tool for machine vision and artificial intelligence.

Pre-trained Models

InceptionV3: A module for GoogleNet, InceptionV3 is a convolutional neural network that facilitates object detection and picture analysis. The goal of the Inception V3 design is to support deeper neural networks without letting the number of parameters balloon out of control. Compared to AlexNet's 60 million parameters, InceptionV3 has fewer than 25 million (Szegedy et al., 2015).

ResNet50V2: The improved ResNet50V2 outperforms the original ResNet50 and ResNet101 on the ImageNet dataset.

Xception: The model is an extension of the Inception architecture, using depthwise separable convolutions in place of the normal Inception architecture modules.

Transfer Learning

One of the most well-known machine learning techniques for image categorization and detection is transfer learning. It enables the application of previously learned expertise from a highly accurate architecture to new issues (Mazo et al., 2018). Instead of rebuilding the models from scratch, transfer learning could use the model and the learned information to identify the images in the dataset containing brain MRI data. In our study, we use the InceptionV3 architecture of the convolutional neural network as a base model in our research. CNN is typically not optimized, despite the fact that training a CNN from scratch has been done in many experiments. In general, starting from scratch requires a lot of computing power, and there aren't always enough samples (Liu et al., 2018).

In this study, to reuse pre-trained model weights with the transfer learning method, it is necessary to freeze some model layers. For this process, a function is developed and it can freeze all, some, or none of the layers in a base model to freeze only some of the weights, layer names, and block numbers that need to be provided in this function.

Model Tuning: Model tuning is the process of choosing the model's parameters that produce the best outcomes for the research. The model's performance is evaluated in comparison to other builds after it has been trained numerous times using possible setups.

EXPERIMENTAL WORK

In this section, we evaluate the dataset containing brain MRI images. In addition, at this stage, the metrics used in this study and the performance obtained are mentioned.

Experimental Settings

To increase the test AUC of the model, some experimental settings have been made with the transfer learning method in this paper.

The function which builds the transfer model was tuned to add the additional layers. Also, additional convolution layers were added with dropout and batch normalization for regularization to increase the capacity of the model. In this way, it was given the additional capacity to the base model. At this stage, the epoch value was set to 50 and the learning rate was taken as 0.001.

Following this point, the training data revealed that the model was overfitting in the early epochs. Because of this, data augmentation was added to help generalize the model to new data. Changes have been made to the rescale, width shift range, rotation range, height shift range, and zoom range properties of images for data augmentation.

Finally, the model was re-trained with the tuned parameters using data augmentation and other techniques. In addition, early stopping was removed and the epochs were set to 150 to give the model as much time to train as possible without overfitting.

Metrics

The following metrics were used to evaluate the performance of the model described earlier.

1) Accuracy: It refers to the percentage of test samples that were successfully categorized.

2) F1-score: Calculating precision and recall is possible using the confusion matrix. After that, the F1-score is calculated as the harmonic mean of recall and precision.

3) Recall: Recall is a metric that shows how well we positively use our estimates.

4) Precision: Precision, on the other hand, shows how many of the values we estimated as positive are positive.

5) AUC: A popular metric for assessing the effectiveness of a multi-class classification system is the area under the receiver operator characteristic (ROC) curve. A ROC curve is a graph that shows the true and false positive rates.

Base Model Selection

In this section, InceptionV3, ResNet50V2, and Xception models are used. AUC, F1-Score, and Accuracy metrics of all models have been calculated. As a result of all these calculations, the InceptionV3 model got 85.00%, and the ResNet50V2 model and the Xception model got 81.00% AUC. The model with the best AUC value was selected from the 3 models. That model was InceptionV3, with an AUC value of 85.00%. After this section, some improvements were applied to the InceptionV3 base model.

Base Model Scores for Alzheimer's Disease	Accuracy	F1-Score	AUC
InceptionV3	0.65	0.49	%85.00
ResNet50V2	0.62	0.47	%81.00
Xception	0.57	0.49	%81.00

Table 1. Base model scores.

Performance Evaluation

In this study, after the base transfer learning model was trained, the base model was tuned and different improvements were made to the model to get better results. In this section, the changes made to the model and the results obtained are shown in Table 2.

Stages of the InceptionV3 model based on Transfer Learning	Accuracy	F1-Score	AUC
InceptionV3	0.65	0.49	%85.00
InceptionV3 + Additional Capacity	0.63	0.47	%84.00
InceptionV3 + Additional Capacity + Data Augmentation	0.68	0.50	%87.00

Table 2. Stages of the InceptionV3 model.

In figures 3, 4, and 5 below, confusion matrix and validation loss, validation AUC graphs of different periods can be observed, respectively, of the base model InceptionV3, InceptionV3 with additional capacity, and later InceptionV3 models with both additional capacity and data augmentation.

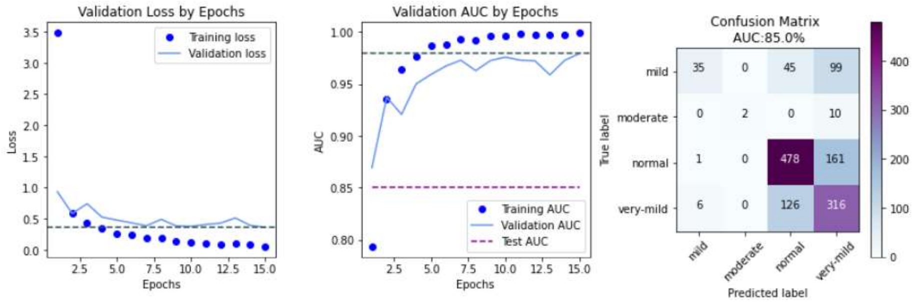

Figure 3. Base model InceptionV3.

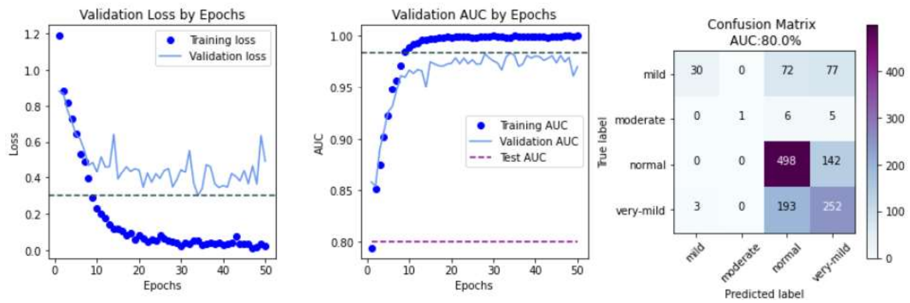

Figure 4. InceptionV3 + Additional Capacity.

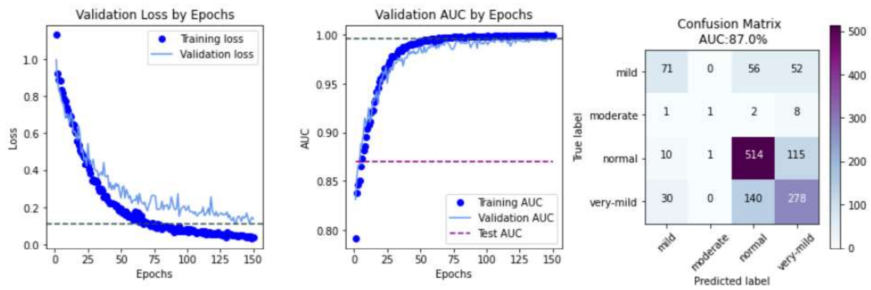

Figure 5. InceptionV3 + Additional Capacity + Data Augmentation.

CONCLUSION

In this study, the task of Alzheimer's disease classification is evaluated using the Alzheimer's Disease dataset consisting of 4 classes from Kaggle. Since InceptionV3 has the best AUC among ResNet50V2, Xception, and InceptionV3 models, InceptionV3 was used as the base transfer learning model, and then applied some tunings on the InceptionV3. Additional convolution layers were added to increase the capacity of the model and at this stage, the epoch value was set to 50 and the learning rate was taken as 0.001. After that process, the model was overfitting early in the epochs. Because of that problem, a data augmentation process was added to help generalize the model to new data. Some changes have been made to the rescale, width shift range, rotation range, height shift range, and zoom range properties of images for data augmentation. Finally, the model has trained again with the tuned parameters using data augmentation and additional convolution layers. During this process, the epoch value was set to 150 to give the model as much time to train as possible without overfitting.

Based on the data in the confusion matrix, it can be said that the model does best at identifying normal and very-mild MRIs and poorest at identifying moderate cases. Also, adding capacity to the model, searching for optimal hyperparameters, and adding data augmentation resulted in better performance on the test dataset. In future work, different datasets can be used for this problem. Different deep learning approaches and hybrid models may achieve a better result.

REFERENCES

Aderghal K, Khvostikov A, Krylov A, Benois PJ, Afdel K, Catheline G 2018. Classification of Alzheimer's disease on imaging modalities with deep CNNs using cross-modal transfer learning. in Proceedings of the IEEE 31st International Symposium on Computer-Based Medical Systems (CBMS), pp. 345-350.

Angkoso C, Purnama I, Purnomo M 2018. Analysis of Brain Tissue and Cerebrospinal Fluid Feature for Alzheimer’s disease Detection. IEEE International Conference on Computer Engineering Network and Intelligent Multimedia (CENIM), vol. 31 pp. 285-288.

Arendt T 2009. Synaptic degeneration in Alzheimer’s disease. Acta neuropathologica, vol 118 no 1 pp. 167-179.

Bhagyashree SR, Sheshadri HS 2018. Dementia: Advances and Treatment, Chapter 5 Alzheimer’s Disease.

Burkea S, Maramaldib P, Kukull W 2016. Neuropsychiatric symptoms and Apolipoprotein E: Associations with eventual Alzheimer’s disease development. Archives of Gerontology and Geriatrics, vol. 65 pp. 231-238.

Chen H, Ni D, Qin J, Li S, Yang X, Wang T, Heng PA 2015. Standard plane localization in fetal ultrasound via domain transferred deep neural networks. IEEE J Biomed Health Inform, 19(5):1627–1636.

Gupta A, Ayhan M, Maida A 2013. Natural image bases to represent neuroimaging data. in Proceedings of the International Conference on Machine Learning, pp. 987-994.

Jellinger KA 2003. Mild cognitive impairment. Aging to Alzheimer's disease. Eur. J. Neurol, 10 466.

Kazemi Y, Houghten SK 2018. A deep learning pipeline to classify different stages of Alzheimer's disease from fMRI data. in Proceedings of the IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 2018 pp. 1-8.

Litjens G 2017. A survey on deep learning in medical image analysis. Medical Image Analysis, vol. 42 pp. 60-88.

Liu M, Cheng D, Wang K, Wang Y 2018. Multi-modality cascaded convolutional neural networks for Alzheimer's disease diagnosis. Neuroinformatics, vol. 16 pp. 295- 308.

Lui YW, Chang PD, Zaharchuk G, Barboriak DP, Flanders AE, Wintermark M, Hess CP, Filippi CG 2020. Artificial Intelligence in Neuroradiology: Current Status and Future Directions. AJNR Am J Neuroradiol 41:E52–E59.

Mazo C, Bernal J, Trujillo M, Alegre E 2018. Transfer Learning for Classification of Cardiovascular Tissues in Histological Images. Computer Methods and Programs in Biomedicine, vol. 18 pp. 1-20.

McGeer PL 1986. Brain Imaging in Alzheimer’s Disease. British Med Bull, 42(1):24–28.

Mufidah R, Wasito I, Hanifah N, Faturrahman M 2017. Structural MRI classification for Alzheimer's disease detection using deep belief network. vol. 17 pp. 37-42.

Nair J, Mohan N 2017. Alzheimer's disease Diagnosis in MR Images using Statistical Methods. IEEE International Conference on Communication and Signal Processing, vol. 17 pp. 1232-1235.

Prince M, Wimo A, Guerchet M, Ali GC, Wu Y, Prina M 2015. World Alzheimer Report 2015. The Global Impact of Dementia: An analysis of Prevalence, Incidence, Cost, and Trends. Alzheimer's Disease International: London UK.

Rabeh AB, Benzarti F, Amiri H 2017. New Method of Classification to detect Alzheimer's Disease. IEEE 14th International Conference on Computer Graphics Imaging and Visualization, vol. 17 pp. 111-116.

Shimizu S, Hirose D, Hatanaka H, Takenoshita N, Kaneko Y, Ogawa Y, Sakurai H, Hanyu H 2018. Role of neuroimaging as a biomarker for neurodegenerative diseases. Frontiers in neurology, 9:265.

Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM 2016. Deep convolutional neural networks for computer-aided detection: Cnn architectures, dataset characteristics, and transfer learning. IEEE Trans Med Imaging, 35(5):1285–1298.

Song T, Chowdhury S, Yang F, Jacobs H, Fakhri G, Li Q, Johnson K, Dutta J 2019. GRAPH CONVOLUTIONAL NEURAL NETWORKS FOR ALZHEIMER'S DISEASE CLASSIFICATION. vol. 19 pp. 414-417.

Szegedy C, Liu W, Jia Y, Sermanet P, Reed S 2015. Going deeper with convolutions. CVPR.

Xiao T, Liu L, Li K, Qin W, Yu S 2018. Comparison of Transferred Deep Neural Networks in Ultrasonic Breast Masses Discrimination. Hindawi BioMed Research International Volume 2018, Article ID 4605191 9 pages.