HORA 2019
www.horacongress.com

International Congress on Human-Computer Interaction, Optimization and Robotic Applications

July 5-7, 2019, Urgup, Nevşehir, Turkey

SETSCI
Conference Proceedings
www.set-science.com
8: 125-131, 2019

Performance Comparison of Carotid Artery Intima Media Thickness Classification by Deep Learning Methods

Serkan Savaş[1*], Nurettin Topaloğlu [2], Ömer Kazcı [3] and Pınar Nercis Koşar[3]

[1]*Faculty of Technology, Gazi University, Ph.D, Turkey*
[2]*Faculty of Technology, Gazi University, Turkey*
[3]*Radiology Department, Ankara Training and Research Hospital, Turkey*
[*]*(serkan_savas@hotmail.com) Email of the corresponding author*

Abstract – Deep learning is a machine learning sub-field that uses deep neural networks. Instead of customized algorithms for each study in this field, it is aimed to cover the wider data set of solutions based on learning the data. Deep learning is a promising approach to solving artificial intelligence problems in machine learning. Nowadays, deep learning algorithms have begun to show themselves in many applications also being studied in biomedical fields. In this medical image processing study, Carotid Artery Intima Media Thickness Ultrasound images were used. Carotid Artery is a type of cardiovascular disease that can result in stroke. If stroke is not diagnosed early, it is in the first place among the disabling diseases. On the other hand, it is the third most common cause of death after cancer and heart disease. For an early diagnose, biomedical image classification performances of VGGNet architecture, which had successful results in the Imagenet competition and an original convolutional neural network model were compared in this study. 501 ultrasound images from 153 patients were used to test the models' classification performances. It is seen that VGG16, VGG19 and CNNcc models achieved rates of 93%, 90% and 89.1% respectively. These results showed that deep architectures can provide proper classification on biomedical images and this can help clinics to diagnose the disease.

Keywords – Deep learning, carotid artery, intima media thickness, vggnet, convolutional neural network, machine learning, artificial intelligence.

I. INTRODUCTION

Deep learning is a machine learning sub-field that uses deep neural networks. Deep neural networks are multilayer neural networks containing two or more hidden layers [1]. In addition to be a sub-field of machine learning, it is the application field of deep neural networks, which is becoming more common day by day. Instead of customized algorithms for each study in this field, it is aimed to cover the wider data set of solutions based on learning the data. Deep learning is a promising approach to solving artificial intelligence problems in machine learning. In recent years, the techniques developed in deep learning research have influenced a wide range of information processing, both in traditional and new forms, in expanded contexts, including the most effective and important aspects of machine learning and artificial intelligence.

In deep learning, there is a structure based on learning multiple levels or representations of data. Top-level properties are derived from lower-level properties to form a hierarchical representation. This representation learns multiple levels of representation that correspond to different levels of abstraction [2].

It is difficult for artificial intelligence methods to solve problems such as picture and/or sound identification, which can easily be done by humans. These heuristic problems can be solved by the computer learning of the ability to understand and experience the world in a hierarchical way of the simplest concepts that can be defined in relation to each other. With the knowledge gained from experience, there is no need for formulas and computations specific to each problem used by computers. When the hierarchical structure is considered as a graph, a deep multi-layered structure is formed, each of which is installed on top of the other. Therefore, artificial intelligence methods based on hierarchical structure emerge as deep learning [3].

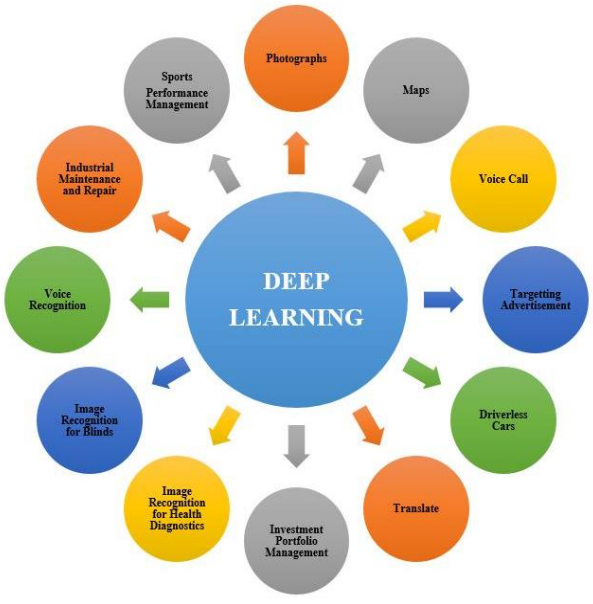


Fig. 1 Deep learning circle

DOI: 10.36287/setsci.4.5.025

The research areas of deep neural networks are at the intersection between artificial intelligence, graphic modeling, optimization, pattern recognition and signal processing [4]. Nowadays, these algorithms have begun to show themselves in many applications such as driverless vehicles, health services, film suggestions, translation services, chatbot, page suggestions, and advertising services (Fig. 1).

The elements that make deep learning architectures such an attractive field of study are as follows:

- Introducing worldwide text, image and audio data sets for research.
- Start to produce high-performance graphics cards (GPUs).
- Introduction of deep architectures such as AlexNet, ZFNet, ResNet, GoogLeNet, VGG16-19, Inception and so on.
- Starting to use deep learning platforms and libraries like Keras, Tensorflow, Theano, Caffe, Pytorch, MatConvNet etc.
- Activation functions, data training and data enhancement methods, and the development and use of effective optimizers by researchers.

When creating deep learning architectures, the algorithm to be used has great importance. These algorithms may vary according to the type, size, volume and structure of the data and the parameters to be used. Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Restricted Boltzmann Machines (RBM) and Deep Belief Nets (DBN) are some of them.

There are many libraries used to run deep learning algorithms, carry out studies and solve problems. Today, existing libraries are constantly updated and new libraries are also available. However, there is no fully dominant one among these libraries. Users are working in different libraries based on their experience. Some of the libraries used to carry out deep learning activities are:

- TensorFlow
- Theano
- Caffe
- Caffe2
- Keras
- PyTorch
- MXNet
- CNTK
- KNet

In addition, The VELES, DIGITS, Chainer, PaddlePaddle, Covnetjs, Deeplearning4j, PyLearn2, Deep Learn Toolbox-Matlab, Sci-Kit Learn, Accord.NET, Apache Spark, Accord.MachineLearning and more machine learning libraries and tools are available. They are offered to users in different areas of use.

Deep learning algorithms are also being studied in biomedical fields. Microscopic image, biomedical image or automatic abdominal multi-organ segmentation [5]-[7], detection of metastatic breast cancer and detection of mitosis in histology of this cancer [8], [9], diagnosis in diabetic retinal fundus photographs [10], retinopathy with false positive reduction in detection of pulmonary nodules [11] and automatically seizure detection/diagnosis in encephalogram signals [12] are some of these studies. Deep learning algorithms have produced particularly important results in image processing studies in recent years. Although it has recently started to be used in medicine, it is not yet sufficient. In order to prevent human error in medical imaging techniques, studies in this area are important.

In this medical image processing study, Carotid Artery (CA) Intima Media Thickness (IMT) Ultrasound (US) images were used. The CA (jugular vein) is the first vessel to be separated from the large vein that emerges from the heart and as it shown in Fig. 2, they carries clean blood to the brain.

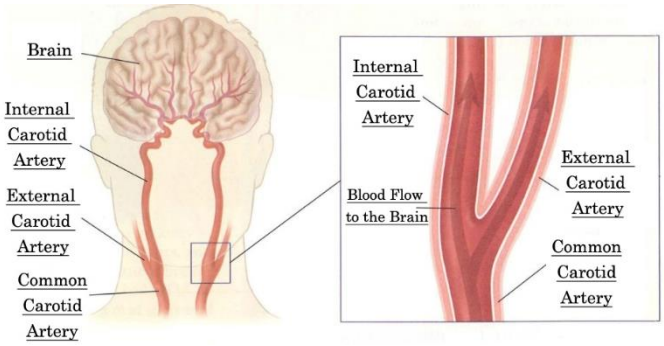


Fig. 2 Internal structures of the CA

CA is a type of cardiovascular disease that can result in stroke. Atherosclerosis is the most important cause [13]. A number of factors accelerate the natural process of aging in all vessels of the body, causing some arteries to contract, blockage or plaque in the vessels at an early age. Considering that the blood flow in the vessel during the development of the disease, the breakage in the plaque and the clot formed after the clogging of the vessels in the brain start to show signs, it is the most important study to determine the patients at the time when the blood flow decreases and the symptoms do not occur completely [14].

If stroke is not diagnosed early, it is in the first place among the disabling diseases [15]. On the other hand, it is the third most common cause of death after cancer and heart disease [16]. Approximately 16 million people have a stroke each year in the world [17]. Cerebrovascular Diseases (CVD) caused by stroke is calculated 64780 at the beginning of 2000s in Turkey [18].

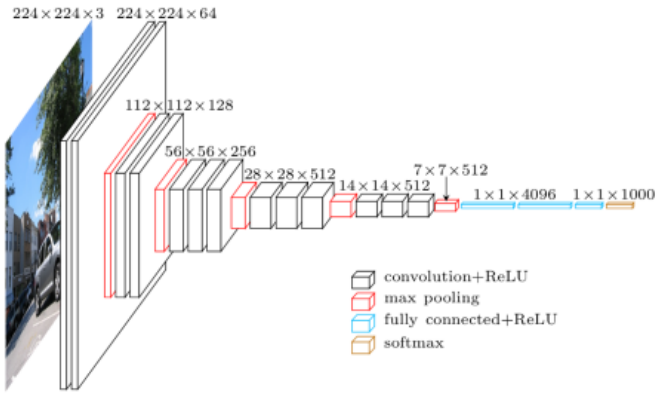
II. Materials and Method

In this study, biomedical image classification performances of VGGNet architecture which had successful results in the Imagenet competition and an original CNN model were compared. Since different imaging techniques and different types of images are used in the studies carried out in the field of medicine, it has not yet come to the stage of producing solutions with common models and algorithms. However, the architectures created in the Imagenet competition can classify from thousands of different image types. The deep aspects of deep architectures that are open to development in this area will be an important area for researchers in the near future. After the congress, an article comparing the performance of all the models that succeeded in the Imagenet competition on biomedical images will be prepared with the results and feedback from the congress. This paper is a preliminary study to investigate the performance of deep architectures on biomedical images.

VGGNet architecture has two different types, including 16 and 19 layers; VGG16, VGG19. The number of layers is determined by the number of weight layers. VGG16 architecture is an architecture that consists of 13 convolution 3 fully connected layers used for better results in ImageNet

2014 competition [19]. There are a total of 41 layers with Maxpool, Fullconnectedlayer, Relulayer, Dropoutlayer and Softmaxlayer layers. The image to be placed in the input layer is 224*x*224*x*3. The last layer is the classification layer [20].

The VGGNet architecture uses a 3*x*3 filter on all layers and uses the convolution-ReLU layers one above the other before the docking layer. As in other deep architectures, VGG architecture decreases the height and width dimensions of the matrices from the input layer to the exit, while the depth value increases. In 2014, it achieved a top-5 error rate of 7.3%. VGGNet architecture and structures of VGGNet Models Fig. 3.a. and Fig. 3.b.



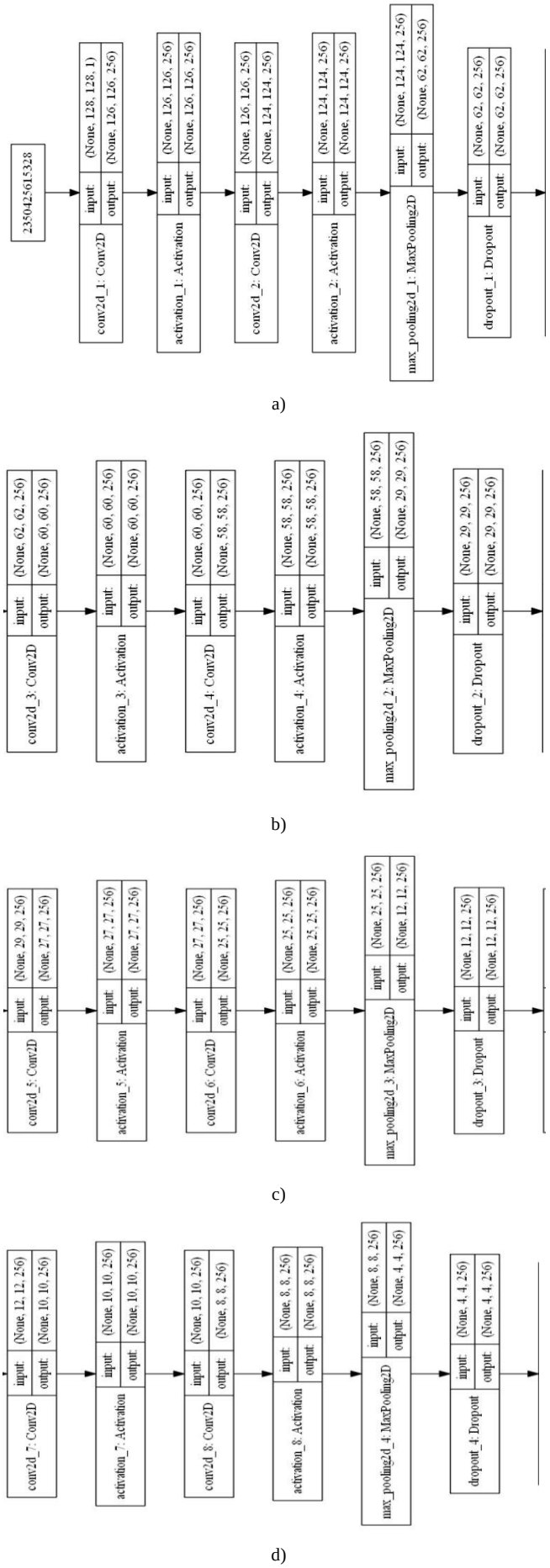
a)

ConvNet Configuration					
A	A-LRN	B	C	D	E
11 weight layers	11 weight layers	13 weight layers	16 weight layers	16 weight layers	19 weight layers
input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

b)

Figure 3. a) VGGNet architecture b) VGG16 (column D) and VGG19 (column E) models

Together with the VGG16 and VGG19 models, a unique CNN model for processing images on CA IMT US images was also used. The diagram of the original model prepared is shown in Fig. 4.

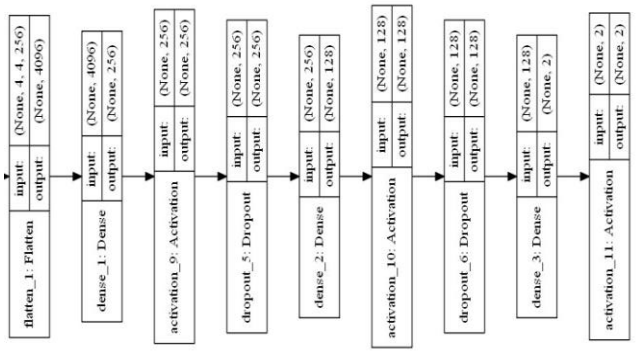


a)

b)

c)

d)



e)

Fig. 4 Original CNN model diagram a) 1st and 2nd convolutional layers b) 3rd and 4th convolutional layers c) 5th and 6th convolutional Layers d) 7th and 8th convolutional layers e) Flattening and fully connected layers

In this study, 501 ultrasound images of 153 patients who were treated in the Radiology Clinic of Ankara Training and Research Hospital were obtained with the Ethics Approval Certificate of Gazi University Ethics Commission dated 08/05/2018 and numbered 2018-217. These images were classified as "IMT: 1" and "IMT: 0" by two specialist doctors. 203 images were classified as "IMT: 1" while the remaining 298 were classified as "IMT: 0".

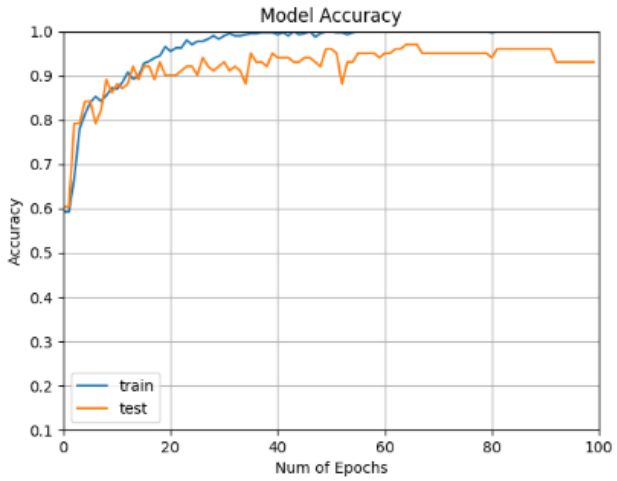
III. Results and Discussion

Several studies have tried various methods for early diagnosis and treatment of CA. These studies were mostly performed for segmentation operations on varying numbers of patient images using different machine learning algorithms [21]-[28]. These studies can be examined in two different sections:

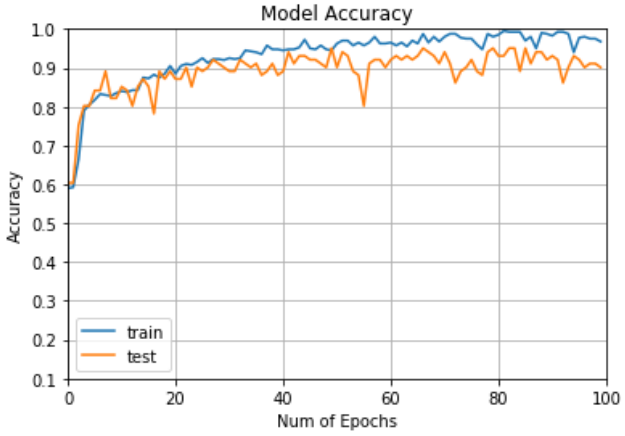
- CA Intima Media Segmentation Studies
- CA Intima Media Classification Studies

IMT classification studies of CA were mostly done with machine learning methods of Support Vector Machines (SVM) and Neural Networks (NN), and these studies produce different results. Classification studies with NN were gave accuracy between 71% and 73% when working with more than 200 images [29], [30] that was 99.1% in a study with 54 images [31]. It is understood from these results that the accuracy rate decreases as the number of images increases when working with NN. The use of SVM is slightly different. In different studies ranging from 270 to 350 images, different rates of performance rates were obtained between 73% and 83% [32], [33]. SVM methods have achieved better results than NN. After the NN and SVM algorithms used in the literature, the performances of the deep architectures are explained in this study.

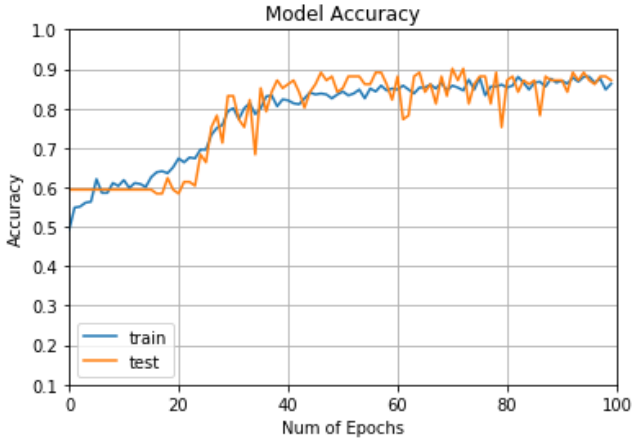
The performances of VGG16 and VGG19 models in classifying CA IMT US images were compared with a deep learning model created by the authors (CNNcc - CNN Carotid Classifier). In the post-congress period, with the feedback received from the congress, in addition to these models, new analyzes will be realized with the models that have succeeded in Imagenet competition and the study will be turned into an article. The accuracy graphs of the deep learning models run on the available images are shown in Fig. 5 and the results are shown in Table 1.



a)



b)



c)

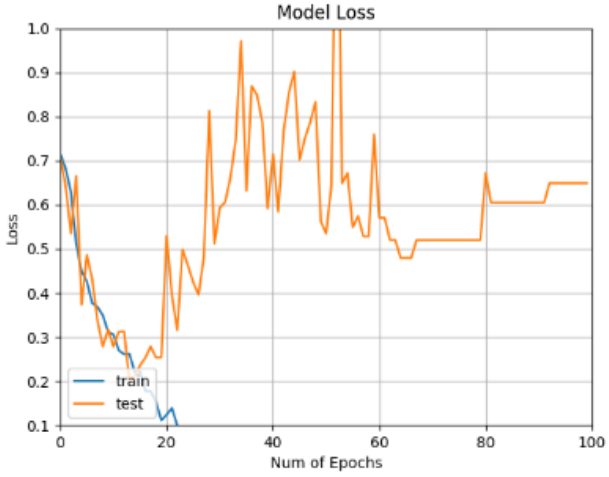
Fig. 5 Accuracy graphs a) VGG16 b) VGG19 c) CNNcc

Table 1. Accuracy rates and Loss results

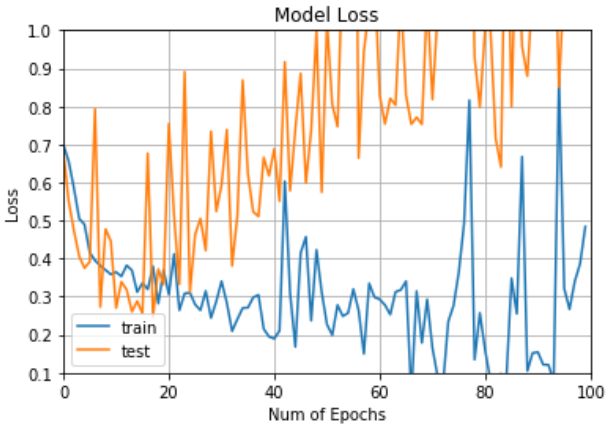
Model	Test Accuracy	Loss	Train Accuracy
VGG16	%93	0.65	%100
VGG19	%90	1.46	%94,5
CNNcc	%89,1	0.29	%89,4

When the graphics and results of the models used in the classification process on CA IMT US images are examined, it

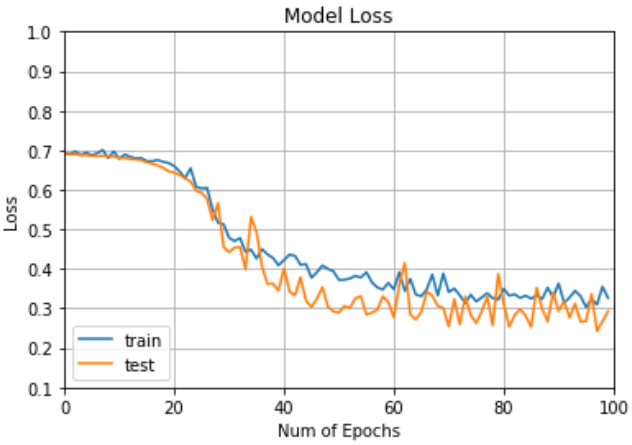
is seen that VGG16, VGG19 and CNNcc models reached rates of 93%, 90% and 89.1% respectively. While considering the accuracy of the model, the accuracy rate and loss parameters during the training are also taken into consideration. The graph of the loss functions of the models is shown in Fig. 6. The loss function is an important indicator because it is used to measure the inconsistency between the predicted value and the actual label. It is stated that the robustness of the model increases with decreasing value of loss function [34]. In this case, the VGG16 model produced the best result from the test accuracy, while the consistency in the loss parameter lagged behind the CNNcc model. Another decisive factor is the 100% accuracy achieved during training in the VGG16 model. From this result and the graph, it is easily seen that the VGG16 model has over-fitted while training on images. Thus, the result obtained is not a reliable result. On the other hand, the VGG19 is different from the VGG16. The VGG19 model has a 90% accuracy rate close to the CNNcc model. The accuracy rate obtained during the training was acceptable at 94.5% and no overfitting was performed on the data. On the other hand, the loss parameter, which is the other determinant parameter, has a value of 1.46 which is well above the CNNcc model. When all of these results were examined, the CNNcc model did not fall into the overfitting state, produced a result close to the VGG19 model at the accuracy rate, and consistency with the lowest result in the loss parameter provided the most feasible model on biomedical images.



a)



b)



c)

Fig. 6 Loss graphs a) VGG16 b) VGG19 c) CNNcc

In order to perform a more in-depth analysis of the results of the models on the images, the results of the confusion matrix were calculated separately for each model. The confusion matrix rows and columns are a square matrix (G x G) representing experimental and predicted classes, respectively. The confusion matrix contains all the information about the distribution of samples in the class and the classification performance [35]. The confusion matrix results for VGG16, VGG19 and Original CNNcc models are shown in Fig. 7.a, Fig. 7.b and Fig. 7.c, respectively.

Number = 101		Predicted		
		Negative	Positive	
Actual	Negative	TN = 56	FP = 5	61
	Positive	FN = 2	TP = 38	40
		58	43	

a)

Number = 101		Predicted		
		Negative	Positive	
Actual	Negative	TN = 59	FP = 2	61
	Positive	FN = 8	TP = 32	40
		67	34	

b)

Number = 101		Predicted		
		Negative	Positive	
Actual	Negative	TN = 57	FP = 3	60
	Positive	FN = 8	TP = 33	41
		65	36	

c)

Fig. 7 Confusion matrix a) VGG16 b) VGG19 c) CNNcc

Calculations such as accuracy, true/false positive rate, true negative rate, error rate, precision, prevalence and f1 score can be done with confusion matrix. Confusion matrix calculations for each model are shown in Table 2.

Table 2. Confusion matrix results

Value	Calculation	VGG16	VGG19	CNNcc
Accuracy (Acc)	$\frac{TP+TN}{Total}$	94/101 = 0,93	91/101 = 0,90	90/101 = 0,891
Error Rate (ER)	$\frac{FP+FN}{Total}$	7/101 = 0,07	10/101 = 0,10	11/101 = 0,108
Recall	$\frac{TP}{TP+FN}$	38/40 = 0,95	32/40 = 0,80	33/41 = 0,804
False Positive Rate (FPR)	$\frac{FP}{FP+TN}$	5/61 = 0,08	2/61 = 0,03	3/60 = 0,05
True Negative Rate (TNR)	$\frac{TN}{FP+TN}$	56/61 = 0,918	59/61 = 0,967	57/60 = 0,95
Precision	$\frac{TP}{TP+FP}$	38/43 = 0,88	32/34 = 0,94	33/36 = 0,916
Prevalence	$\frac{TP+FN}{Total}$	40/101 = 0,40	40/101 = 0,40	41/101 = 0,405
F1 Score	$\frac{2TP}{2TP+FP+FN}$	76/83 = 0,916	64/74 = 0,865	66/77 = 0,857

In confusion matrix, accuracy calculation gives the results of overall, how often is the classifier correct. Error rate is a measure of how often the classifier has incorrectly predicted. True positive rate indicates when it is actually positive and how often does it predict positive. False positive rate gives the results of negative actual value's positive prediction. True negative rate indicates when it is actually negative and how often does it predict negative. Precision is a measure of how accurately all classes are predicted. Prevalence is estimation of how often "positive" value is found at the end of the prediction and F1-score is the harmonic mean of precision and recall. The predicted performance results of the models on the labeled data for each class are shown in Table 3.

Table 3. Performance measures of models

	Precision			Recall		
	VGG16	VGG19	CNN	VGG16	VGG19	CNN
IMT:0	0.97	0.88	0.88	0.92	0.97	0.95
IMT:1	0.88	0.94	0.92	0.95	0.80	0.80
avg/total	0.93	0.90	0.89	0.93	0.90	0.89
	F1-score			Support		
	VGG16	VGG19	CNN	VGG16	VGG19	CNN
IMT:0	0.94	0.92	0.91	61	61	60
IMT:1	0.92	0.86	0.86	40	40	41
avg/total	0.93	0.90	0.89	101	101	101

The estimated results of the models on the labeled images according to the performance criteria shown in Table 3 are shown in Table 4.

Table 4. Test and predict results

VGG16								
Test values	0	0	1	1	1	0	0	0
Predicted Values	[1. 0.]	[1. 0.]	[0. 1.]	[0. 1.]	[0. 1.]	[1. 0.]	[1. 0.]	[1. 0.]
Result	√	√	√	√	√	√	√	√
VGG19								
Test values	0	0	1	1	0	0	0	0
Predicted Values	[1. 0.]	[1. 0.]	[0. 1.]	[0. 1.]	[0. 1.]	[1. 0.]	[1. 0.]	[1. 0.]
Result	√	√	√	√	X	√	√	√
CNNcc								
Test values	0	1	0	0	1	0	0	0
Predicted Values	[0. 1.]	[0. 1.]	[1. 0.]	[1. 0.]	[0. 1.]	[1. 0.]	[1. 0.]	[1. 0.]
Result	X	√	√	√	√	√	√	√

The results of the study showed that the performance of deep architectures in the biomedical field is promising. The classification results of the deep learning models used in the study were more pronounced than the machine learning algorithms. This study is especially important in terms of revealing certain deep learning models that can be applied to different imaging techniques in biomedical field. Thus, researchers and clinicians will be able to use a common model instead of using separate solutions and algorithms for each problem.

IV. CONCLUSION

In this study, the classification performance of deep learning models on CA IMT US images was compared. The proposed classification methods are important for early diagnosis and treatment of CVD. VGG16 and VGG19 models, which were successful in the Imagenet competition, and a CNNcc model, prepared by the authors, were tested and their classification performances were compared. In addition, the results obtained were compared with the results of the machine learning algorithms, which were previously performed classification studies.

501 US images from 153 patients in Ankara Training and Research Hospital were used to test the models. While the VGG16 model achieved an accuracy of 93%, the reliability of the result was reduced due to overfitting to the data during training. Changes and additions to the hyper-parameters

should be made in this model and adapted to the data. The VGG19 model achieved 90% performance and produced positive results. The loss parameter of the VGG19 model is 1.46. This value means inconsistency in the detection of labeled data for the model and some layer arrangements are required in the model by drop out method. Finally, the accuracy rate of the CNNcc model was measured as 89.1% and it was determined to be the most suitable model for classification with 0.29 loss parameter ratio.

In this study, it is important to show that the performance of deep architectures to make classification using medical imaging methods can be more efficient than the machine learning algorithms used before. It has been proven that deep architectures can be used in different types of images in various competitions. From this point of view, the formation of certain common deep learning models on biomedical images will be an important development in terms of artificial intelligence discipline. This study will be expanded with the feedbacks from the Congress and new models to be added, and the results will be published.

ACKNOWLEDGMENT

The authors would like to thank the Radiology Department of Ankara Training and Research Hospital for their kindly cooperation and providing all the ultrasound images used.

REFERENCES

[1] K. Phil, MATLAB Deep Learning: With Machine Learning, Neural Networks and Artificial Intelligence. Seoul, Soul-t'ukpyolsi, Korea: Apress. 2017.

[2] Y. Bengio, Learning Deep Architectures for AI. *Foundations and Trends in Machine Learning*, vol. 2, pp. 1-127. 2009. doi:http://dx.doi.org/10.1561/2200000006.

[3] I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning*. MIT Press. 2016. [Online]. Available http://www.deeplearningbook.org

[4] L. Deng, & D. Yu, Deep Learning Methods and Applications. *Foundations and Trends in Signal Processing*, vol. 7, pp. 197-387. 2013. doi: http://dx.doi.org/10.1561/2000000039

[5] O. Z. Kraus, J. L. Ba, & J. Brendan, Classifying and segmenting microscopy images with deep multiple instance learning. Bioinformatics, vol. 32, pp. 52–59. 2016.

[6] O. Ronneberger, P. Fischer, & T. Brox, U-Net: Convolutional Networks for Biomedical Image Segmentation. 2015. arXiv: https://arxiv.org/pdf/1505.04597.pdf

[7] P. Hu, F. Wu, J. Peng, Y. Bao, F. Chen, & D. Kong, Automatic abdominal multi-organ segmentation using deep convolutional neural network and time-implicit level sets. International Journal of Computer Assisted Radiology and Surgery, vol. 12, pp. 399–411, 2017.

[8] D. Wang, A. Khosla, R. Gargeya, H. Irshad, & A. H. Beck, Deep Learning for Identifying Metastatic Breast Cancer. 2018. arXiv: https://arxiv.org/pdf/1606.05718.pdf

[9] D. C. Cireşan, A. Giusti, L. M. Gambardella, & J. Schmidhuber, Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013. Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks, pp. 411-418. Berlin: Springer, 2013.

[10] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, R. D. Webster, Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA, vol. 316(22), pp. 2402-2410, 2016.

[11] Q. Dou, H. Chen, L. Yu, J. Qin, & P. A. Heng, Multilevel Contextual 3-D CNNs for False Positive Reduction in Pulmonary Nodule Detection. IEEE Transactions on Biomedical Engineering, vol. 64(7), pp. 1558-1567, 2017.

[12] U. R. Acharya, S. L. Oh, Y. Hagiwara, J. H. Tan, & H. Adeli, Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. Computers in Biology and Medicine, vol. 100, pp. 270–278, 2018.

[13] A. Civelek, (2014). *Karotis Arter Hastalığı*. [Online]. Available: http://www.alicivelek.com/karotis-arter-hastaligi/

[14] Ö. Kocamaz, (2016). *Şah Damar Tıkanıklığı "Karotis Arter Hastalığı"*. [Online]. Available: http://www.drkocamaz.com/karotis-arter-hastaligi/

[15] M. G. Bousser, Stroke prevention: an update. *Frontiers of Medicine, vol. 6*(1), pp. 22-34, 2012.

[16] K. Strong, C. Mathers, & R. Bonita, Preventing stroke: saving lives around the world. *Lancet Neurol, vol. 6*, pp. 182-187, 2007.

[17] A. Demirci Şahin, Y. Üstü, & D. Işık, Serebrovasküler Hastalıklarda Önlenebilen Risk Faktörlerinin Yönetimi. *Ankara Medical Journal, vol. 15*(2), pp. 106-113, 2015.

[18] N. Ünüvar, S. Mollahaliloğlu, N. Yardım, B. Bora Başara, V. Dirimeşe, E. Özkan, & Ö. Varol, *Türkiye Hastalık Yükü Çalışması.* T.C. Sağlık Bakanlığı. Refik Saydam Hıfzıssıhha Merkezi Başkanlığı Hıfzıssıhha Mektebi Müdürlüğü, 2004.

[19] K. Simonyan, & A. Zisserman, *Very Deep Convolutional Networks for Large-Scale Image Recognition.* (2014). [Online]. Available: arXiv:1409.1556

[20] F. Doğan, & İ. Türkoğlu, Derin Öğrenme Algoritmalarının Yaprak Sınıflandırma Başarımlarının Karşılaştırılması. *Sakarya Universıty Journal Of Computer And Informatıon Scıences, vol. 1*, pp. 10-21, 2018.

[21] R. M. Menchón-Lara, J. L. Sancho-Gómez, & A. Bueno-Crespo, Early-stage atherosclerosis detection using deep learning over carotidultrasound images. *Applied Soft Computing*, pp. 616-628, 2016.

[22] R. Rocha, A. Campilho, J. Silva, E. Azevedo, & R. Santos, Segmentation of the carotid intima-media region in B-mode ultrasound images. *Image and Vision Computing, vol. 28*, pp. 614-625, 2010.

[23] F. Molinari, G. Zeng, & J. S. Suri, Inter-Greedy Technique for Fusion of Different Segmentation Strategies Leading to High-Performance Carotid IMT Measurement in Ultrasound Images. *Journal of Medical Systems, vol. 35*, pp. 905-919, 2011.

[24] M. C. Bastida-Jumilla, R. M. Menchón-Lara, J. Morales-Sánchez, R. Verdú-Monedero, J. Larrey-Ruiz, & J. Sancho-Gómez, Frequency-domain active contours solution to evaluate intima–mediathickness of the common carotid artery. *Biomedical Signal Processing and Control*, pp. 68-79, 2015.

[25] R. M. Menchón-Lara, & J. L. Sancho-Gómez, Fully automatic segmentation of ultrasound common carotid artery images based on machine learning. *Neurocomputing*, pp. 161–167, 2015.

[26] U. Kutbay, F. Hardalaç, M. Akbulut, & Ü. Akaslan, A Computer Aided Diagnosis System for Measuring Carotid Artery Intima-Media Thickness (IMT) Using Quaternion Vectors. *Journal of Medical Systems, vol. 40*(149), 2016.

[27] F. Milletari, S. A. Ahmadi, C. Kroll, A. Plate, V. Rozanski, J. Maiostre, N. Navab, Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound. *Computer Vision and Image Understanding*, pp. 1–11, 2017.

[28] N. Ikeda, N. Dey, A. Sharma, A. Gupta, S. Bose, S. Acharjee, J. S. Suri, Automated segmental-IMT measurement in thin/thick plaque with bulb presence in carotid ultrasound from multiple scanners: Stroke risk assessment. *Computer Methods and Programs in Biomedicine, vol. 141*, pp. 73-81, 2017.

[29] E. C. Kyriacou, M. S. Pattichis, C. I. Christodoulou, C. S. Pattichis, S. K. Kakkos, M. Griffin, & A. Nicolaides, Ultrasound imaging in the analysis of carotid plaque morphology for the assessment of stroke. *Studies in health technology and informatics*, pp. 241-275, 2005.

[30] C. I. Christodoulou, C. S. Pattichis, M. Pantzaris, & A. Nicolaides, Texture-based classification of atherosclerotic carotid plaques. *IEEE Trans Med Imaging*, pp. 902-912, 2003.

[31] S. Mougiakakou, S. Golemati, I. Gousias, A. Nicolaides, & K. Nikita, Computer-aided diagnosis of carotid atherosclerosis based on ultrasound image statistics, laws' texture and neural networks. *Ultrasound in Medicine & Biology*, pp. 26-36, 2007.

[32] E. C. Kyriacou, M. S. Pattichis, C. S. Pattichis, A. Mavrommatis, C. I. Christodoulou, S. K. Kakkos, & A. Nicolaides, Classification of atherosclerotic carotid plaques using morphological analysis on ultrasound images. *Applied Intelligence*, pp. 3-23, 2009.

[33] U. R. Acharya, O. Faust, A. Alvin, V. S. Sree, F. Molinari, L. Saba, J. S. Suri, Symptomatic vs. Asymptomatic Plaque Classification in Carotid Ultrasound. *Journal of Medical Systems, vol. 36*, pp. 1861–1871, 2012.

[34] Z. Hao, (2017). Loss Functions in Neural Networks. [Online]. Available: https://isaacchanghau.github.io/post/loss_functions/

[35] D. Ballabio, F. Grisoni, & R. Todeschini, Multivariate comparison of classification performance measures. Chemometrics and Intelligent Laboratory Systems, vol. 174, pp. 33-44, 2018. doi:https://doi.org/10.1016/j.chemolab.2017.12.004