Classification of Carotid Artery Intima Media Thickness Ultrasound Images with Deep Learning

Serkan Savaş [1] · Nurettin Topaloğlu [2] · Ömer Kazcı [3] · Pınar Nercis Koşar [3]

Abstract

Cerebrovascular accident due to carotid artery disease is the most common cause of death in developed countries following heart disease and cancer. For a reliable early detection of atherosclerosis, Intima Media Thickness (IMT) measurement and classification are important. A new method for decision support purpose for the classification of IMT was proposed in this study. Ultrasound images are used for IMT measurements. Images are classified and evaluated by experts. This is a manual procedure, so it causes subjectivity and variability in the IMT classification. Instead, this article proposes a methodology based on artificial intelligence methods for IMT classification. For this purpose, a deep learning strategy with multiple hidden layers has been developed. In order to create the proposed model, convolutional neural network algorithm, which is frequently used in image classification problems, is used. 501 ultrasound images from 153 patients were used to test the model. The images are classified by two specialists, then the model is trained and tested on the images, and the results are explained. The deep learning model in the study achieved an accuracy of 89.1% in the IMT classification with 89% sensitivity and 88% specificity. Thus, the assessments in this paper have shown that this methodology performs reasonable results for IMT classification.

Introduction

Cerebrovascular disease (CVD) is one of the leading causes of death and permanent damage in the world [1]. CVD due to Carotid Artery (CA) disease is the most common cause of death in developed countries following heart disease and cancer [2]. According to North American Symptomatic Carotid Endarterectomy Trial data, the risk of ischemic CVD in the first year is 11% in CA stenosis of 70–79%, and this risk reaches 35%

when there is 90% or more stenosis [3, 4]. Approximately 75% of all patients with CA stenosis are asymptomatic. Several studies have shown that patients with internal CA stenosis benefit from endarterectomy compared to medical treatment with 50–99%, 60–99%, and 70–99%, depending on whether they are symptomatic [5–7]. Internal CA stenosis is one of the most important causes of ischemic cerebral palsy [8]. 80% of the ischemic processes occur due to atherosclerosis.

CVD is the primary pathological damage to one or more blood vessels of a brain region that are permanently or temporarily affected by ischemia or bleeding [9]. These diseases are caused by the obstruction or bleeding of the vessels feeding the brain and give symptoms related to the damaged brain area. CA passes through the two sides of the neck and takes clean blood to the brain. The first branch separated from the main veins is divided into two vessels. One to the right arm and the other is the right carotid vein, which carries blood to the right side of the brain. The left CA leading to the left brain is separated from the main vessels as a single vessel as shown in Fig. 1a [10].

CA disease is the atherosclerotic disease of our jugular veins, which occur on both sides of our neck, which may result in stroke. A sudden occlusion of these vessels results

---

This article is part of the Topical Collection on *Image & Signal Processing*
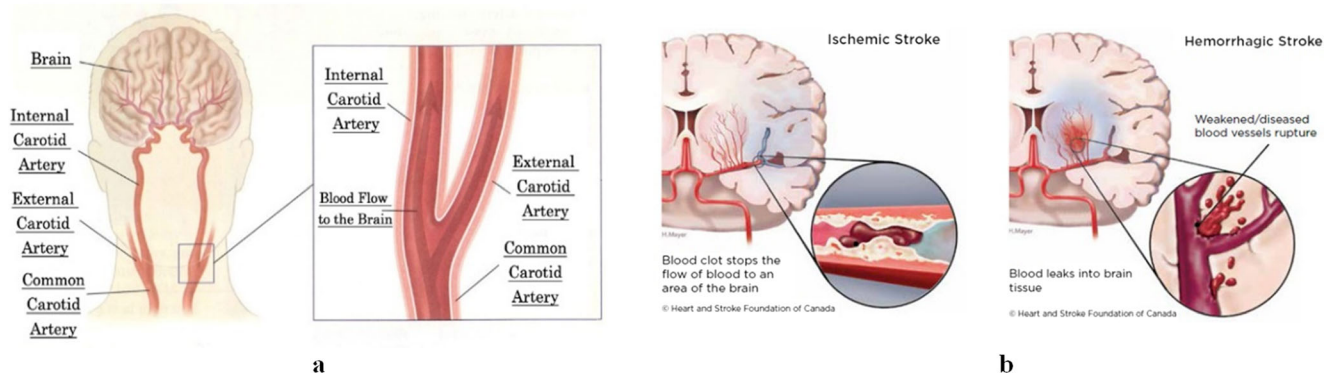
✉ Serkan Savaş
serkan_savas@hotmail.com

[1] Faculty of Technology, Computer Engineering Department Ph.D, Gazi University, Ankara, Turkey

[2] Faculty of Technology, Computer Engineering Department, Gazi University, Ankara, Turkey

[3] Department of Radiology, Ankara Training and Research Hospital, Ankara, Turkey

Fig. 1 **a** Carotid artery interior structures **b** Ischemic stroke and hemorrhagic stroke [10, 11]

in stroke [12]. CA disease is a form of CVD, which results most importantly from atherosclerosis. Oil particles, cholesterol, calcium and some other substances and cells accumulate in the artery wall causing set to be called plaque. These plaques increase their volume over time or cause clotting on the irregular surface and cause complete occlusion of the vessel. Fully occluded jugular veins or small clots that are detached from the atherosclerotic plaque may obstruct the small veins in the brain and can cause stroke [12].

If measures are not taken and not intervened early, stroke becomes the first cause of disability [13]. Every year approximately 16 million people in the world are undergoing a stroke [9]. The number of deaths because of CVD is 64,780 in Turkey at the beginning of the 2000s. First five diseases distribution causing death in Turkey is shown in Table 1.

80% of all strokes are occlusive strokes and 20% are brain hemorrhages [15]. Occlusive stroke and brain hemorrhage shapes [11] are shown in Fig. 1b.

Ultrasonography (USG) is the most sensitive and reliable method for the morphological evaluation of CA. In addition, colored and spectral Doppler USG can show flow changes caused by vascular lesions in real time [1]. CA USG is a common imaging study for the diagnosis of CA disease [16]. Doppler USG is a method of examining vascular structures with sound waves, which provides hemodynamic information about carotid and vertebral arteries. Although its sensitivity is 92.6% and specificity is 97%, angiography is accepted as the gold standard [17]. Doppler equation is defined as:

$$f_d = f_t - f_r = 2 f_t v \cos\theta / c \tag{1}$$

where f_t and f_r are the transmitted and received ultrasound frequencies, v is the speed of the target, c is the velocity of the sound in the environment, and θ is the angle between the ultrasound beam and the direction of movement of the target [18].

Various methods of early diagnosis and treatment of CA disease have been tried in different studies. These studies have mostly been used to perform segmentation operations on different numbers of patient images using different machine learning algorithms [19–27]. Segmentation process is a problem of image analysis and it is the process of preparing the image to display and diagnostic stages of image processing [28]. In addition to these studies, Intima Media Thickness (IMT) measurements were also carried out to predict automatic artery recognition and lumen diameter changes in some studies. [29–31]. Moreover, a cloud computing based platform creation study was conducted for risk assessment and stroke follow-up of CA patients [32]. Risk assessment is the most important stage of the disease. For this stage, classification studies were carried out using various methods such as Neural Networks (NN) [33–35], Support Vector Machine (SVM) [36, 37], and Enhanced Activity Index [38].

Different studies have begun in the field of medicine by using deep learning (DL) algorithms. Automatic abdominal multiple organ segmentation [39], classification and division of microscope images [40], segmentation of biomedical images [41], segmentation of brain tumor [42], identification of metastatic breast cancer [43], mitosis detection in breast cancer histology images [44], detection of diabetic retinopathy in retinal fundus photographs [45], diagnosing melanoma skin lesions [46] false positive reduction in the detection of pulmonary nodules [47] and automated detection/diagnosis of seizure using encephalogram signals [48] are some of these studies. Although it is known that the performance of deep architectures in the mentioned areas is impressive, the use of DL in the medical field is not yet sufficient. DL is widely used in areas such as computer identification, image processing, natural language processing and speech recognition. Such new approaches are necessary in automated medical decision-making systems because non-automated processes are more

Table 1 First five diseases causing death number and percentage distribution (Turkey, 2004) [14]

N.	Causes Of Death	Death Number	Percentage in Total
1	Ischemic Heart Disease	93,260	21,7
2	Cerebrovascular Diseases	64,780	15
3	Chronic Obstructive Pulmonary Disease	25,104	5,8
4	Perinatal Causes	24,756	5,8
5	Lower Respiratory Tract Infections	18,225	4,2

expensive, demand intensive labor, and therefore subject to human-induced errors.

The study of IMT classification of the CA has not been done with DL methods yet. DL studies have been carried out so far only at the image segmentation level. In this study, a proposed model using Convolutional Neural Networks (CNN) from DL algorithms determined the increased IMT and arterial narrowing which is one of the causes of CA disease.

Deep learning

Deep Learning is a new and promising field for machine learning to solve artificial intelligence problems. It is a subspace of machine learning and a field of application of Deep Neural Networks (DNN). In this area, instead of the customized algorithms for each study, it is aimed that the solutions are based on learning data sets and also cover larger data sets. Artificial Neural Networks (ANN) were inspired by the human brain. They are information processing structures consisting of process elements which are connected to each other by changing weights and each of which has their own memory. ANNs, having superiorities such as

learning, generalization, non-linearity, fault tolerance, harmony, parallelism, are used in many different application areas such as medical applications like image and signal processing, disease prediction, engineering, production, finance, optimization and classification [49]. In DNN, there are two or more secret neural network layers and more extensive relationships are established from simple to complex data. Each layer tries to establish a relationship between itself and the previous layer. Thus, a more detailed examination of the inputs is made and decisions that are more accurate are made. As shown in Fig. 2a ANN produces an output with applying weights and activation function to the given inputs. Figure 2b shows a DNN structure with three secret layers.

Different activation functions can be used when forming the structure of DNN. These functions may vary according to the type, structure, size and model of the data. The activation function determines the output that the cell will produce in response to the input itself. A non-linear function is usually selected. The main activation functions used are given as follows:

$$\text{Sigmoid activation function}: f(x) = \frac{1}{1 + e^{-x}} \quad (2)$$
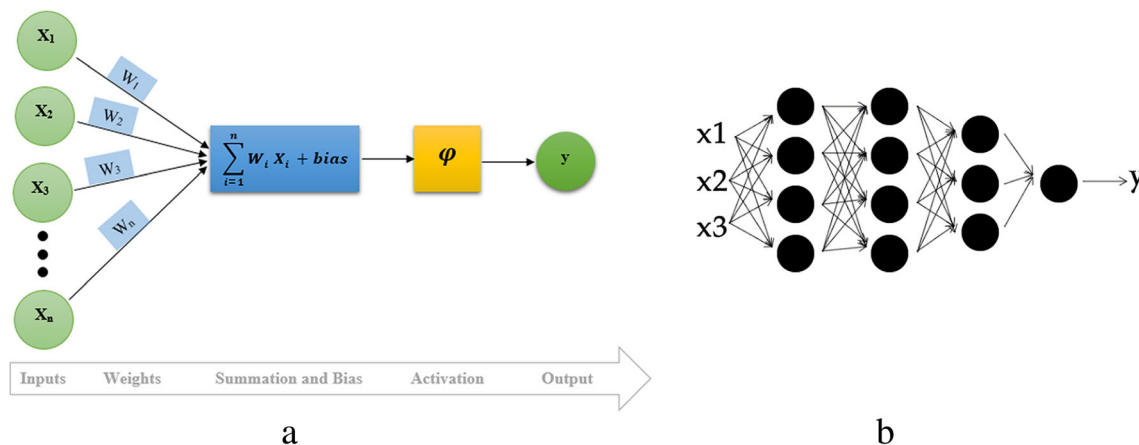
Fig. 2 a ANN structure b DNN structure
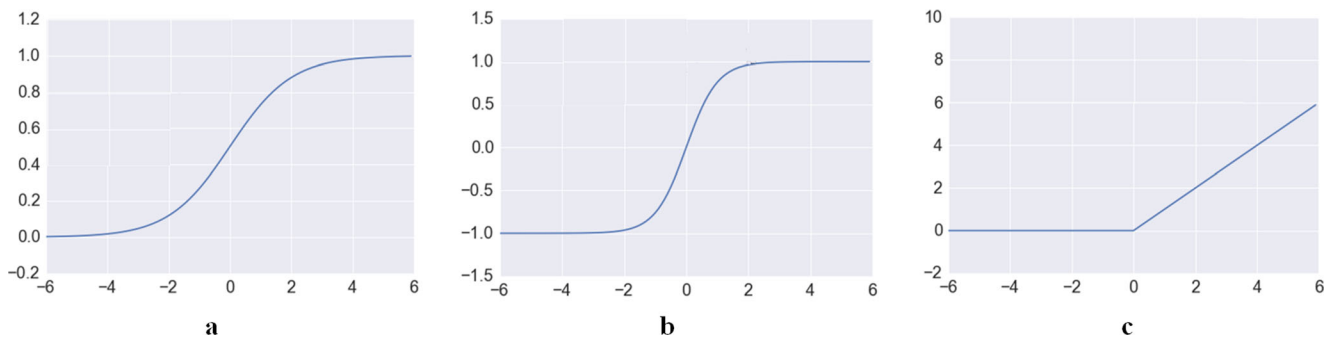
Fig. 3 **a** Sigmoid activation function **b** TanH activation function **c** ReLU activation function

TanH activation function : $\tanh(x) = \dfrac{2}{1 + e^{-2x}} - 1$ (3)

ReLU activation function : $f(x) = \begin{cases} 0 \; for \; x < 0 \\ x \; for \; x \geq 0 \end{cases}$ (4)

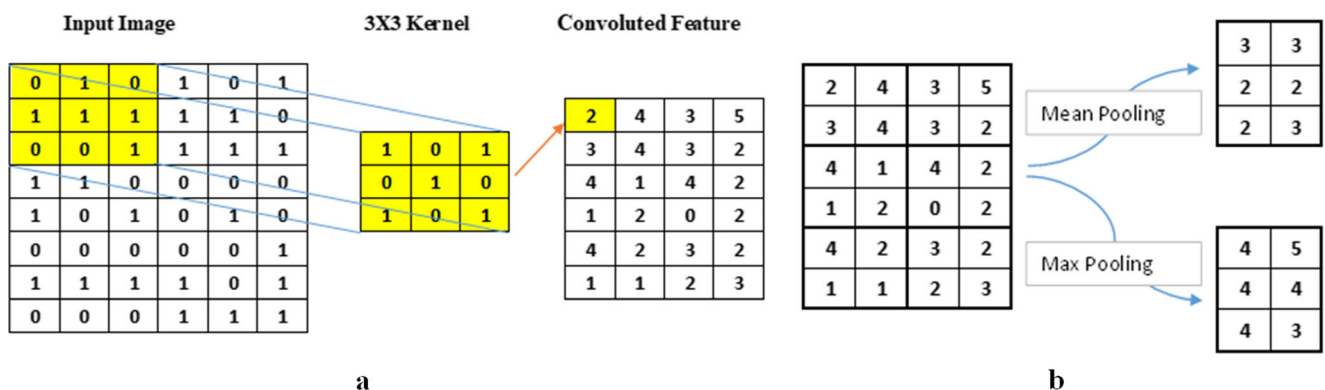The related curves of Sigmoid, TanH and ReLU are shown in Figs. 3a, b. and c respectively.

The most well-known of DL algorithms is CNN and often used in image classification problems. Kernels mostly used with dimensions of 3×3, 5×5, 7×7 on each layer in CNN as shown in Fig. 4a. Then the pooling process is done on the outputs of these kernels, which is shown in Fig. 4b. The data in the kernel is filtered by pooling. The max pooling method is the most commonly used pooling method, and the largest value is taken in the matrix with this method.

The convolution of two functions $(f * g)$ in the finite range $[0, t]$ is defined as follows:

$$[f*g](t) \equiv \int_0^t f(\tau)g(t-\tau)d\tau$$ (5)

where $[f * g](t)$ means the convolution of the functions f and g [50]. Alternatively, convolution is calculated in an infinite range mostly as:

$$[f*g] \equiv \int_{-\infty}^{\infty} f(\tau)g(t-\tau)d\tau$$ (6)

$$= \int_{-\infty}^{\infty} g(\tau)f(t-\tau)d\tau$$ (7)

Proposed strategy

A DL model has been created for the classification of images. In order to improve the performance of the model, hyper-parameters in the DL model were optimized with repeated analysis/test studies. The following steps were followed when creating a DL model.

In the first stage, which is called definition, the model parameters required for the DL model, such as numpy, os, matplotlib and sklearn, were included in the program. The image resolution is fixed at 128×128 and the image channel is determined as one since the grayscale image format will be used. The path, from which the images were taken, was identified and a second path was defined, in which the new images were saved after image preprocessing. Parameters such as batch size, number of classes, number of epochs, number of filters, pool size and convolution filter size, to be used in the model are defined in this part.

At the image pre-processing stage, the images taken from the image folder were resized to the 128×128 resolution, then converted to the grayscale image format and saved in the folder in the second path. The images were saved sequentially

Fig. 4 **a** Convolutional layer (3×3) **b** Pooling (2×2)

with the names they were previously classified. Then all images were flattened as a float and an array of images was created which was stored in a matrix. These stored images were first labeled "1" up to 203rd image and the next images were labeled "0". After the labeling process was completed, the images in the memory were mixed randomly and the sequential image list was changed to prevent the model from memorizing the data and to increase accuracy.

The image sequence obtained from the last step of the image pre-process stage was divided into two as train/test. While 80% of the images were used for training, the remaining 20% of the images were used for the test. During the definition of the training images, random selection was performed by preserving the proportions of "1" and "0" tagged images in the total image. There are 400 IMT US images in train dataset. 162 of the 400 IMT US images in the train dataset consist of images labeled as "1," while the remaining 238 images are labeled as "0". There are 101 IMT US images in test dataset, which is also selected randomly. 41 of the 101 IMT US images in the test dataset consist of images labeled as "1," while the remaining 60 images are labeled as "0". By this way, it is aimed to provide the learning sensitivity of the model on the images in the training set and to increase the accuracy of the model. In addition, the possibility of memorizing the data or over-fitting the data is prevented.

The model in the study was formed by determining the optimum parameters after repeated tests. CNN, which is the most known and frequently used in image classification problems, is used in the model. The model is designed sequentially. First, 256 filters were added to the model and the images were passed through a "3 × 3" convolution filter. There are eight convolution layers in total in the model. After each convolution layer, the "ReLU" activation function was added and the activation result was determined as the input of the new layer. After every two convolution layers, a maxpooling layer was added to prevent over-fitting and a pooling operation of "2 × 2" dimensions was applied. After the pooling layer, a drop out layer with "0.5" rate was added to the model. After these processes were completed, the model was flattened with the fully connected layers through 256, 128, and 2 outputs respectively. ReLU activation and drop out layers were added between each fully connected layer. The softmax activation function was used in the last layer of the model. When compiling the model, the loss parameter was selected as binary cross-entropy which is computed as follows:

$$L = -\frac{1}{n}\sum_{i=1}^{n}\left[y^{(i)}\log\left(\hat{y}^{(i)}\right) + \left(1-y^{(i)}\right)\log\left(1-\hat{y}^{(i)}\right) \right] \quad (8)$$

where n is number of samples and y is output of the related neuron [51]. The optimizer parameter of the model was selected RMSprop (learning rate, lr = 0.00001) and the metric is defined as accuracy. The summarize of the model is shown in Table 2.

Results and Discussion

In the study, to test the proposed model, from June 2018 to January 2019, 501 images of 153 patients were obtained from the patients who were treated at the Radiology Department of Ankara Training and Research Hospital. The images were taken with The Ethics Approval Certificate of Gazi University Ethics Commission dated 08/05/2018 and numbered 2018–217. Toshiba Aplio 400 Ultrasound device was used for ultrasound imaging. The images are classified as "IMT: 1" and "IMT: 0" by two doctors who are experts in the Department of Radiology at Ankara Training and Research Hospital. The summary and features of database are shown in Table 3.

In order to use the DL model that was created in the study on images, a system with the following features: i5–3.50GHz, 16 GB RAM and Nvidia graphical card (GP102, TITAN Xp) was used. The DL model was created using the Keras DL library with Tensorflow in the Python programming language on the Ubuntu operating system. Studies have been carried out using The Scientific Python Development Environment (Spyder) interface.

After operation of the proposed model, graphics and weights were recorded and the accuracy and loss parameters of the model were visualized. In order to evaluate the performance of the model, Receiver Operating Characteristic (ROC) Curve and Confusion Matrix were also created. Weights were recorded after training of the model for use in estimation model.

The accuracy of the model is 89.1% while loss of the model is 0.292. Loss function is an important indicator for CNN model, because it is used to measure the inconsistency between predicted value and actual label. It is a non-negative value, where the robustness of the model increases along with the decrease of the value of loss function [51]. The accuracy and loss graph of the model are shown in Fig. 5a and b respectively.

Our model starts to learn from the training data after 25 epochs. The accuracy of the model increases and the loss parameter decreases after this epoch. The classification results can also be represented in the so-called confusion matrix, also known as contingency table [52–54]. It is a square matrix (G x G), whose rows and columns represent experimental and predicted classes, respectively. The confusion matrix contains all the information related to the distribution of samples within the classes and to the classification performance [55]. Calculations such as accuracy, error rate (misclassification rate), true positive rate (also known as sensitivity or recall), false positive rate, true negative rate (also known as specificity), precision, prevalence and f1 score can be done with confusion matrix.

The confusion matrix shows where the classification model is confused when it makes predictions. It gives insight not only into the errors being made by the classifier model but more importantly the types of errors that are being made. There are two

Table 2 Summarized model parameters

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 126, 126, 256)	2560
activation_1 (Activation)	(None, 126, 126, 256)	0
conv2d_2 (Conv2D)	(None, 124, 124, 256)	590,080
activation_2 (Activation)	(None, 124, 124, 256)	0
max_pooling2d_1 (MaxPooling2D)	(None, 62, 62, 256)	0
dropout_1 (Dropout)	(None, 62, 62, 256)	0
conv2d_3 (Conv2D)	(None, 60, 60, 256)	590,080
activation_3 (Activation)	(None, 60, 60, 256)	0
conv2d_4 (Conv2D)	(None, 58, 58, 256)	590,080
activation_4 (Activation)	(None, 58, 58, 256)	0
max_pooling2d_2 (MaxPooling2D)	(None, 29, 29, 256)	0
dropout_2 (Dropout)	(None, 29, 29, 256)	0
conv2d_5 (Conv2D)	(None, 27, 27, 256)	590,080
activation_5 (Activation)	(None, 27, 27, 256)	0
conv2d_6 (Conv2D)	(None, 25, 25, 256)	590,080
activation_6 (Activation)	(None, 25, 25, 256)	0
max_pooling2d_3 (MaxPooling2D)	(None, 12, 12, 256)	0
dropout_3 (Dropout)	(None, 12, 12, 256)	0
conv2d_7 (Conv2D)	(None, 10, 10, 256)	590,080
activation_7 (Activation)	(None, 10, 10, 256)	0
conv2d_8 (Conv2D)	(None, 8, 8, 256)	590,080
activation_8 (Activation)	(None, 8, 8, 256)	0
max_pooling2d_4 (MaxPooling2D)	(None, 4, 4, 256)	0
dropout_4 (Dropout)	(None, 4, 4, 256)	0
flatten_1 (Flatten)	(None, 4096)	0
dense_1 (Dense)	(None, 256)	1,048,832
activation_9 (Activation)	(None, 256)	0
dropout_5 (Dropout)	(None, 256)	0
dense_2 (Dense)	(None, 128)	32,896
activation_10 (Activation)	(None, 128)	0
dropout_6 (Dropout)	(None, 128)	0
dense_3 (Dense)	(None, 2)	258
activation_11 (Activation)	(None, 2)	0

Total params: 5,215,106

Trainable params: 5,215,106

Non-trainable params: 0

possible predicted classes: "positive" and "negative". For this carotid classifier, "positive" would mean IMT and "negative" would mean no IMT. In this CA IMT case, confusion matrix results are shown in Fig. 6.

Depending on the results shown in Fig. 6, out of those 101 cases, the classifier predicted "positive" 36 times, and "negative" 65 times. In reality, 41 patients in the sample have the IMT, and 60 patients do not have the IMT. In confusion matrix, TP means observation is positive, and is predicted to be positive. FN means observation is positive, but is predicted negative. TN means

observation is negative, and is predicted to be negative. FP means observation is negative, but is predicted positive.

Accuracy calculation gives the results of overall, how often is the classifier correct. For this carotid classifier (cc), accuracy is calculated as follows:

$$
\begin{aligned}
Accuracy\ (Acc) &= \frac{TP + TN}{TP + TN + FP + FN} \\
Acc_{cc} &= \frac{90}{101} = 0,891
\end{aligned}
\tag{9}
$$

Error rate in general, is a measure of how often the classifier has incorrectly predicted. In addition, it is known as misclassification rate. Error rate is equivalent to one minus accuracy, and is also calculated as follows:

$$
\begin{aligned}
Error\ Rate\ (ER) &= \frac{FP + FN}{TP + TN + FP + FN} \\
ER_{cc} &= \frac{11}{101} = 0,108
\end{aligned}
\tag{10}
$$

True positive rate indicates when it is actually positive and how often does it predict positive. It is also known as "Sensitivity" or "Recall". Recall can be defined as the ratio of the total number of correctly classified positive examples divide to the total number of positive examples. High recall (small number of FN) indicates the class is correctly recognized. Recall is usually used when the goal is to limit the number of FN and is calculated as follows:

$$
\begin{aligned}
Recall &= \frac{TP}{TP + FN} \\
Recall_{cc} &= \frac{33}{41} = 0,804
\end{aligned}
\tag{11}
$$

False positive rate gives the results of negative actual value's positive prediction. It is calculated as follows:

$$
\begin{aligned}
False\ Positive\ Rate\ (FPR) &= \frac{FP}{FP + TN} \\
FPR_{cc} &= \frac{3}{60} = 0,05
\end{aligned}
\tag{12}
$$

True negative rate indicates when it is actually negative and how often does it predict negative. It is also known as "Specificity" and is equivalent to one minus FPR. Specificity is calculated as follows:

$$
\begin{aligned}
True\ Negative\ Rate\ (TNR) &= \frac{TN}{FP + TN} \\
TNR_{cc} &= \frac{57}{60} = 0,95
\end{aligned}
\tag{13}
$$

Table 3 Image database

PATIENT NUMBER	IMAGE NUMBER	CLASSES		US MACHINE	ETHICS
		IMT:1	IMT:0		
153	501	203	298	Toshiba Aplio 400	Gazi University Ethics Commission Date: 08/05/2018 Number: 2018–217
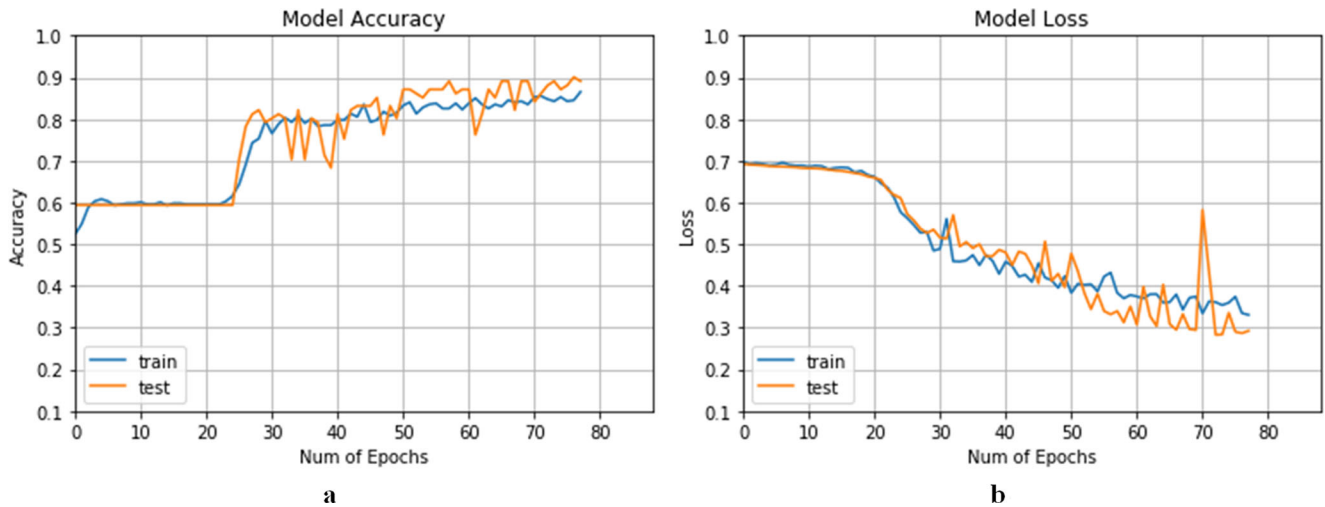
Fig. 5 **a** Accuracy graph of the model **b** Loss graph of the model

Precision is a measure of how accurately all classes are predicted. It is also known as positive predictive value. In order to get the value of precision, the total number of correctly classified positive examples is divided to the total number of predicted positive examples. High Precision indicates an example labeled as positive is indeed positive (small number of FP). Precision is calculated as follows:

$$Precision = \frac{TP}{TP + FP}$$
$$Precision_{cc} = \frac{33}{36} = 0,916 \tag{14}$$

Prevalence is estimation of how often "positive" value is found at the end of the prediction. It is calculated as follows:

$$Prevalence = \frac{TP + FN}{TP + TN + FP + FN}$$
$$Prevalence_{cc} = \frac{41}{101} = 0,405 \tag{15}$$

F1-score is the harmonic mean of precision and recall. It is a measure of how well the classifier performs and is often used to compare classifiers. If it is only tried to optimize recall, the algorithm will predict most examples to belong to the positive

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

TN: True Negative, TP: True Positive
FN: False Negative, FP: False Positive

Fig. 6 Confusion matrix of the model

Table 4 Performance measures

	precision	recall	f1-score	support
IMT:0	0.88	0.95	0.91	60
IMT:1	0.92	0.80	0.86	41
micro avg	0.89	0.89	0.89	101
macro avg	0.90	0.88	0.88	101
weighted avg	0.89	0.89	0.89	101

Table 5 Test and predict results

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.]
	X	√	√	√	√	√	√	√

class, but that will result in many false positives and, hence, low precision. On the other hand, if it is tried to optimize precision, the model will predict very few examples as positive results, but recall will be very low. Therefore, F1-score is useful when it is needed to take both precision and recall into account. F1-Score is calculated as follows:

$$F1\ Score = \frac{2TP}{2TP + FP + FN}$$
$$F1_{cc} = \frac{66}{77} = 0,857$$

(16)

After calculation of all parameters of confusion matrix, overall average performance measurements of the model for both of the classes "IMT:0" and "IMT:1" are shown in the Table 4.

It is seen in confusion matrix that the sensitivity of the model is 89% and specificity is 88%. There were 101 images to test the model. While testing after the training of the model, the number of images in both classes was determined by maintaining the ratio in the total image. The model is correctly predicted eight of nine test image classes as seen in the Table 5.

ROC analysis contributes to the process of clinical decision-making when the diagnosis process will take a long time, the cost will be high, special method-equipment and qualified human resources will be needed by determining appropriate cut-off values for indicators that will be determined in short-time, low-cost, and easily obtainable [56]. Sensitivity and specificity curves provide the comparison of the success of different tests in correct clinical diagnosis. In this CA IMT case, ROC curve is shown in Fig. 7.
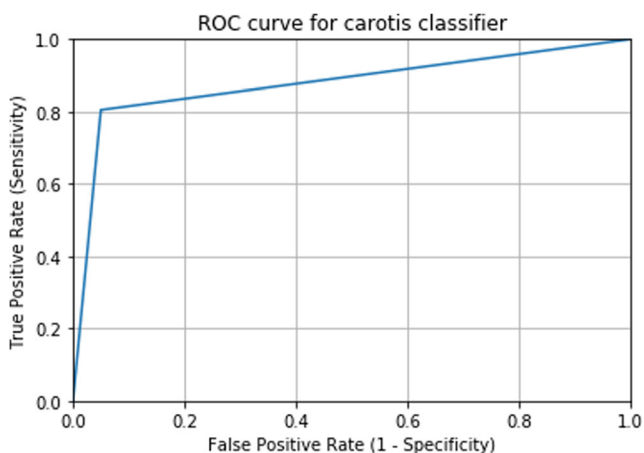
Fig. 7 ROC curve for model

The outputs of layers, while CNN model working on the IMT USG images, were also saved to show sample feature extraction. The original input image and feature extraction steps for 1st, 5th, 10th, 15th, and 20th layers activation outputs are shown respectively in Fig. 8.

Various studies have been carried out on CA images in the literature as shown in Table 6. These studies can be examined in two different sections:

• CA Intima Media Segmentation Studies
• CA Intima Media Classification Studies

According to Table 6, although the study of IMT classification of the CA has not been done with DL methods, DL studies have been carried out so far only at the image segmentation level. Classification studies were mostly carried out through machine learning methods with SVM and NN. These studies produced different results from each other. The classification studies carried out through NN were in the range of 71–73% when working with more than 200 images [33, 34], while this rate was 99.1% in a study with 54 images [35]. It is understood that by working with the NN, the accuracy rate decreases as the number of images increases. Instead, while working with DL methods, more input data means better accuracy because model learns features itself from input data. The CNN model proposed in this study achieved better accuracy rate of 89.1% from more image data. The use of SVM is slightly different from this situation. In different studies ranging image number from 270 to 350, performance rates of different ratios ranging from 73% to 83% were obtained with SVM [36, 37]. Although SVM methods produced better results than NN, the CNN method proposed in this study achieved better results than SVM methods. These results are an indication of a better achievement than previous studies when compared to the results given in the Table 6. Although the parameters such as number of patients, number of images, image quality etc. in all studies given in the Table 6 are different, it can be seen that our model is positive compared to the methods in other studies.

Conclusion

In this study, a new method for the classification of CA IMT on ultrasound images was proposed. The proposed method was to classify IMT for early diagnosis and treatment of
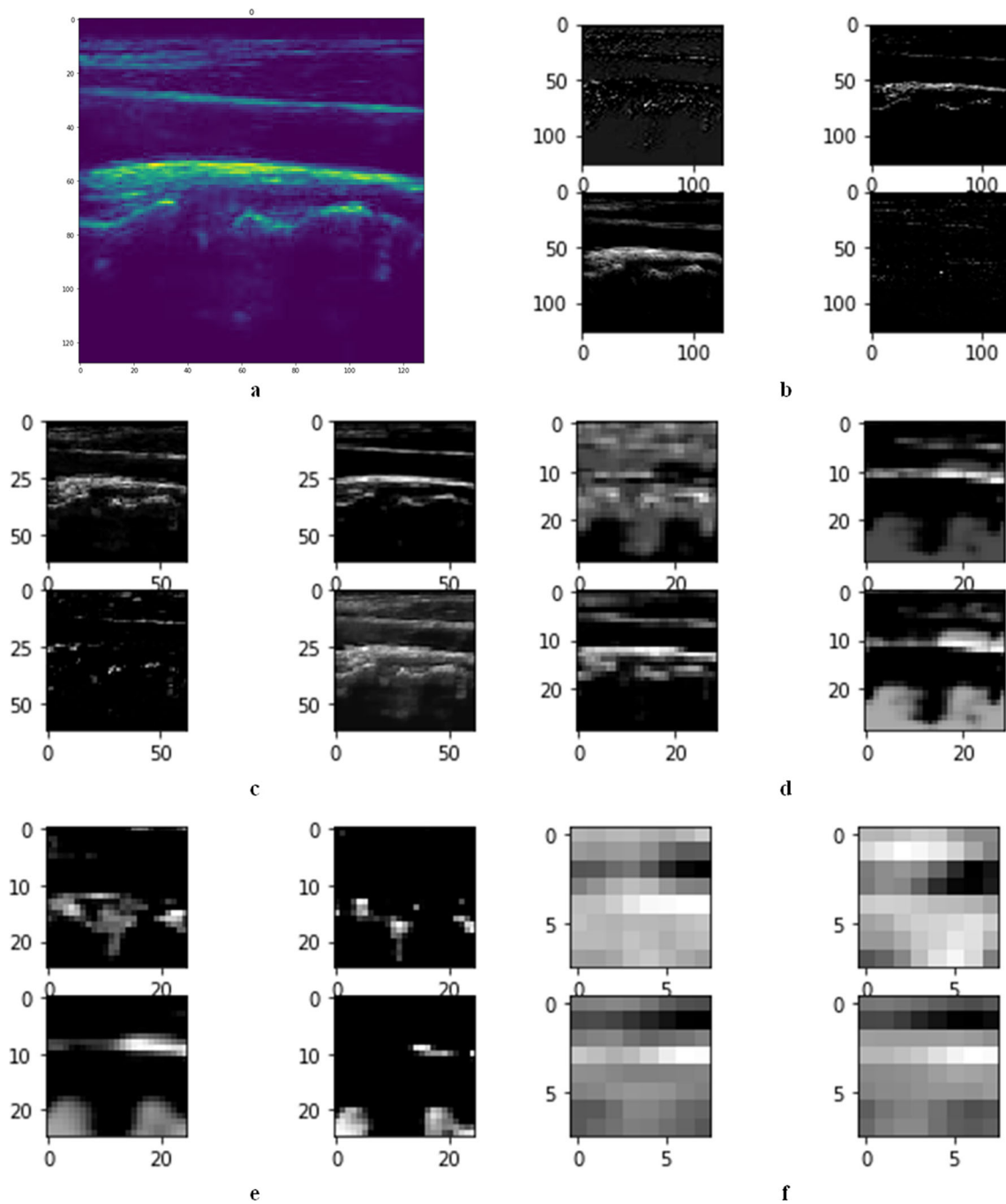
Fig. 8 Feature extraction **a** Input Image **b** 1st Layer **c** 5th Layer **d** 10th Layer **e** 15th Layer **f** 20th Layer

CVD. The proposal is based on a model in the field of DL, a new subfield of machine learning.

The model is tested on a database of 501 images from 153 patients treated in Ankara Training and Research Hospital during 8-months period. CNN algorithm, which is frequently used in image classification problems, is used in the model. The accuracy of the model was compared with the classification of the doctors. The results showed that the created DL model achieved a classification performance of 89.1%. The proposed model has 89% sensitivity and 88% specificity for IMT classification. The performance of our CNN model, based on the DL method, has been remarkably significant in the classification of the CA IMT. To summarize, the main contributions and developments of the proposed method are that this is an ongoing study of previous segmentation studies; it shows high performance in high number of images in classification

273   Page 10 of 12

J Med Syst (2019) 43: 273

Table 6   Carotid artery image processing studies

Author	Patient/Image number	Method	Results
Segmentation Studies			
Menchón-Lara et al. (2016) [19]	67	DL (ELM)	5.79 ± 34.42 μm (mean ± standard deviation)
Rui Rocha et al. (2010) [21]	47	A new approach to segmentation	Mean between: 0.07 mm and 0.11 mm Standard deviation between: 0.11 mm and 0.12 mm
M.C. Bastida-Jumilla et al. (2015) [23]	46	A new segmentation algorithm	13.8 ± 31.9 μm (mean ± standard deviation)
Kutbay et al. (2016) [25]	25	Quaternion Based Saliency Mapping	Mean square error: Artery diameter: 0.0014 mm2 Mass diameter: 0.000191 mm2
Nobutaka Ikeda et al. (2017) [27]	649	Full automatic bulb detection system	92.67% sensitivity
Alessandro C. Rossi et al. (2008) [29]	128	A new CCA detection algorithm	Success: 98.8%
Da-Chuan Cheng et al. (2010) [30]	78, 111, 86 images from 3 Patient Sequences	A new method	Mean error: 0.55 pixel (0.0583 mm) for the far wall and 0.63 pixel (0.0668 mm) for the near wall
Classification Studies			
Christodoulou et al. (2003) [33]	230	Modular NN	73,1%
Kyriacou et al. (2005) [34]	274	NN - SVM	71.2%
Mongiakakou et al. (2007) [35]	54	Hybrid NN	99.1%
Kyriacou et al. (2009) [36]	274	SVM	73.7%
Acharya et al. (2012) [37]	346	SVM/AdaBoost	82.4% SVM 81.7% AdaBoost
Pedro et al. (2013) [38]	146	Enhanced Activity Index	77%

studies, high performance in image quality differences and significant reliability and precision in the classification of IMT.

It is also important to emphasize that these positive aspects are very important for a method designed to help prevent CVD. The study showed that DL methods can produce effective results in medical research.

Acknowledgements 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.

Compliance with Ethical Standards

Research Involving Human Participants and/or Animals   All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee (The Ethics Approval Certificate of Gazi University Ethics Commission dated 08/05/2018 and numbered 2018–217) and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed Consent   Informed consent was obtained from all individual participants included in the study.

Conflict of Interest   There is no conflict of interest in this work.

References

1. Seçil, M., Carotid and Vertebral Doppler. Basic Ultrasonography and Doppler (pp. 479–498). Akademisyen Bookstore, 2013.
2. Centers for Disease Control and Prevention, Prevalence of disabilities and associated health conditions among adults. United States, 1999.MMWR. Morbidity and mortality weekly report, 50(7), 120.
3. Barnett, H., Taylor, D., Haynes, R., Sacket, D., Peerless, S., Ferguson, G., and Eliasziw, M., Beneficial effect of carotid endarterectomy in symptomatic patients with high-grade carotid stenosis. N. Engl. J. Med. 325(7):445–453, 1991.
4. Henry, J., Barnett, M., Taylor, D., Eliasziwq, M., Fox, A., Gary, G., and Meldrum, H., Benefit of Carotid Endarterectomy in Patients with Symptomatic Moderate or Severe Stenosis. N. Engl. J. Med. 339(20):1415–1425, 1998.
5. Benjamin, M., and Dean, R., Current Diagnosis & Treatment in Vascular Surgery. R. H. Dean, J. S. Yao, & D. C. Brewster içinde, Current Diagnosis & Treatment in Vascular Surgery (1st Edition b., pp. 1–5). Appleton & Lange, 1995.
6. Koçak, A., Comparison of Color Doppler Ultrasonography, Magnetic Resonance Angiography, Multislice Computed Tomography Angiography and Digital Subtraction Angiography Findings in Carotid Artery and Peripheral Artery Lesions.

🕿 Springer

İstanbul: T. C. Ministry of Health Dr. Siyami Ersek Thoracic and Cardiovascular Surgery Training and Research Hospital, 2009.

7. Burns, P., Gough, S., and Bradbury, A. W., Management of peripheral arterial disease in primary care. BMJ 326:584–588, 2003.

8. Phatouros, C. C., Higashida, R. T., Malek, A. M., Meyers, P. M., Lempert, T. E., Dowd, C. F., and Halbach, V. V., Carotid Artery Stent Placement for Atherosclerotic Disease: Rationale, Technique, and Current Status. Radiology:26–41, 2000.

9. Demirci Şahin, A., Üstü, Y., and Işık, D., Management of Preventable Risk Factors of Cerebrovascular Disease. Ankara Medical Journal 15(2):106–113, 2015.

10. Kocamaz, Ö., Jugular Veil Congestion "Carotid Artery Disease", 2016. Accessed: 05 01, 2018 Kalp ve Damar Cerrahisi Uzmanı Dr. Kocamaz: http://www.drkocamaz.com/karotis-arter-hastaligi

11. HSFC, What is stroke?, 2018. Accessed: 14 04, 2019 https://www.heartandstroke.ca/stroke/what-is-stroke

12. Civelek, A., Carotid Artery Disease, 2014. Accessed: 04 24, 2018, Prof. Dr. Ali Civelek: http://www.alicivelek.com/karotis-arter-hastaligi/

13. Bousser, M.-G., Stroke prevention: an update. Frontiers of Medicine 6(1):22–34, 2012.

14. Ünüvar, N., Mollahaliloğlu, S., Yardım, N., Bora Başara, B., Dirimeşe, V., Özkan, E., and Varol, Ö., Turkey Burden of Disease Study. T.C. Ministry of Health. Refik Saydam Hıfzıssıhha Center, 2004.

15. Caplan, L. R., Basic pathology, anatomy, and pathophysiology of stroke. In: Caplan's Stroke: A Clinical (4th ed. b.). Philadelphia: Saunders Elsevier, 2009.

16. Tahmasebpour, H. R., Buckley, A. R., Cooperberg, P. L., and Fix, C. H., Sonographic Examination of the Carotid Arteries. RadioGraphics 25:1561–1575, 2005.

17. Yurdakul, S., and Aytekin, S., Doppler ultrasonography of the carotid and vertebral arteries. Turkish Society of Cardiology Archive: 508–517, 2011.

18. Öztürk, A., Arslan, A., and Hardalaç, F., Comparison of neuro-fuzzy systems for classification of transcranial Doppler signals with their chaotic invariant measures. Expert Syst. Appl. 34:1044–1055, 2008.

19. Menchón-Lara, R.-M., Sancho-Gómez, J.-L., and Bueno-Crespo, A., Early-stage atherosclerosis detection using deep learning over carotid ultrasound images. Appl. Soft Comput.:616–628, 2016.

20. Santos, A. M., Santos, R. M., Castro, P. M., Azevedo, E., Sousa, L., and Tavares, J. M., A novel automatic algorithm for the segmentation of the lümen of the carotid artery in ultrasound B-mode images. Expert Syst. Appl. 40:6570–6579, 2013.

21. Rocha, R., Campilho, A., Silva, J., Azevedo, E., and Santos, R., Segmentation of the carotid intima-media region in B-mode ultrasound images. Image Vis. Comput. 28:614–625, 2010.

22. Molinari, F., Zeng, G., and Suri, J. S., Inter-Greedy Technique for Fusion of Different Segmentation Strategies Leading to High-Performance Carotid IMT Measurement in Ultrasound Images. J. Med. Syst. 35:905–919, 2011.

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

24. Menchón-Lara, R.-M., and Sancho-Gómez, J.-L., Fully automatic segmentation of ultrasound common carotid artery images based on machine learning. Neurocomputing:161–167, 2015.

25. Kutbay, U., Hardalaç, F., Akbulut, M., and Akaslan, Ü., A Computer Aided Diagnosis System for Measuring Carotid Artery Intima-Media Thickness (IMT) Using Quaternion Vectors. J. Med. Syst. 40(149), 2016.

26. Milletari, F., Ahmadi, S.-A., Kroll, C., Plate, A., Rozanski, V., Maiostre, J., and Navab, N., Hough-CNN: Deep learning for segmentation of deep brain regions in MRI and ultrasound. Comput. Vis. Image Underst.:1–11, 2017.

27. Ikeda, N., Dey, N., Sharma, A., Gupta, A., Bose, S., Acharjee, S., and Suri, J. S., Automated segmental-IMT measurement in thin/thick plaque with bulb presence in carotid ultrasound from multiple scanners: Stroke risk assessment. Comput. Methods Prog. Biomed. 141:73–81, 2017.

28. Kızılkaya, A., Image Segmentation. Denizli: Pamukkale University, 2008. Accessed: 20.08.2018 http://akizilkaya.pamukkale.edu.tr/B%C3%B6l%C3%BCm4_goruntu_isleme.pdf

29. Rossi, A. C., Brands, P. J., and Hoeks, A. P., Automatic recognition of the common carotid artery in longitudinal ultrasound B-mode scans. Med. Image Anal. 12:653–665, 2008.

30. Cheng, D.-C., Schmidt-Trucksäss, A., Liu, C.-H., and Liu, S.-H., Automated Detection of the Arterial Inner Walls of the Common Carotid Artery Based on Dynamic B-Mode Signals. Sensors 10: 10601–10619, 2010.

31. Loizou, C. P., Kasparis, T., Lazarou, T., Pattichis, C. S., and Pantziaris, M., Manual and automated intima-media thickness and diameter measurements of the common carotid artery in patients with renal failure disease. Comput. Biol. Med. 53:220–229, 2014.

32. Melillo, P., Orrico, A., Scala, P., Crispino, F., and Pecchia, L., Cloud-Based Smart Health Monitoring System for Automatic Cardiovascular and Fall Risk Assessment in Hypertensive Patients. J. Med. Syst. 39(109), 2015.

33. Christodoulou, C. I., Pattichis, C. S., Pantzaris, M., and Nicolaides, A., Texture-based classification of atherosclerotic carotid plaques. IEEE Trans. Med. Imaging:902–912, 2003.

34. Kyriacou, E. C., Pattichis, M. S., Christodoulou, C. I., Pattichis, C. S., Kakkos, S. K., Griffin, M., and Nicolaides, A., Ultrasound imaging in the analysis of carotid plaque morphology for the assessment of stroke. Studies in Health Technology and Informatics:241–275, 2005.

35. Mougiakakou, S., Golemati, S., Gousias, I., Nicolaides, A., and Nikita, K., Computer-aided diagnosis of carotid atherosclerosis based on ultrasound image statistics, laws' texture and neural networks. Ultrasound Med. Biol.:26–36, 2007.

36. Kyriacou, E. C., Pattichis, M. S., Pattichis, C. S., Mavrommatis, A., Christodoulou, C. I., Kakkos, S. K., and Nicolaides, A., Classification of atherosclerotic carotid plaques using morphological analysis on ultrasound images. Appl. Intell.:3–23, 2009.

37. Acharya, R. U., Faust, O., Alvin, A., Sree, V. S., Molinari, F., Saba, L., and Suri, J. S., Symptomatic vs. Asymptomatic Plaque Classification in Carotid Ultrasound. J. Med. Syst. 36:1861–1871, 2012.

38. Pedro, L. M., Sanches, J. M., Seabra, J., Suri, J. S., and Fernandes, J. F., Asymptomatic Carotid Disease—A New Tool for Assessing Neurological Risk. Echocardiography:353–361, 2013.

39. Hu, P., Wu, F., Peng, J., Bao, Y., Chen, F., and Kong, D., Automatic abdominal multi-organ segmentation using deep convolutional neural network and time-implicit level sets. Int. J. Comput. Assist. Radiol. Surg. 12:399–411, 2017.

40. Kraus, O. Z., Ba, J. L., and Brendan, J., Classifying and segmenting microscopy images with deep multiple instance learning. Bioinformatics 32:52–59, 2016.

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

42. Thillaikkarasi, R., and Saravanan, S., An Enhancement of Deep Learning Algorithm for Brain Tumor Segmentation Using Kernel Based CNN with M-SVM. J. Med. Syst. 43:84, 2019. https://doi.org/10.1007/s10916-019-1223-7.

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

44. Cireşan, D. C., Giusti, A., Gambardella, L. M., and Schmidhuber, J., 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.

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

46. Premaladha, J., and Ravichandran, K. S., Novel Approaches for Diagnosing Melanoma Skin Lesions Through Supervised and Deep Learning Algorithms. J. Med. Syst. 40:96, 2016. https://doi.org/10.1007/s10916-016-0460-2.

47. Dou, Q., Chen, H., Yu, L., Qin, J., and Heng, P.-A., Multilevel Contextual 3-D CNNs for False Positive Reduction in Pulmonary Nodule Detection. IEEE Trans. Biomed. Eng. 64(7):1558–1567, 2017.

48. Acharya, U. R., Oh, S. L., Hagiwara, Y., Tan, J. H., and Adeli, H., Deep convolutional neural network for the automated detection and diagnosis of seizure using EEG signals. Comput. Biol. Med. 100: 270–278, 2018.

49. Arı, A., and Berberler, M., Yapay Sinir Ağları ile Tahmin ve Sınıflandırma Problemlerinin Çözümü İçin Arayüz Tasarımı. Acta Infologica 1(2):55–73, 2017.

50. Weisstein, E.W., Convolution. Accessed: 08 27, 2018. MathWorld-A Wolfram: http://mathworld.wolfram.com/Convolution.html

51. Hao, Z., Loss Functions in Neural Networks, 2017. Isaac Changhau: https://isaacchanghau.github.io/post/loss_functions/

52. Ferri, C., Flach, P. A., & Hernández-Orallo, J., European Conference on Machine Learning. Improving the AUC of Probabilistic Estimation Trees (s. 121–132). Berlin, Heidelberg: Springer, 2003. 10.1007/978-3-540-39857-8_13

53. Provost, F., and Domingos, P., Tree Induction for Probability-Based Ranking. Mach. Learn. 52(3):199–215, 2003. https://doi.org/10.1023/A:1024099825458.

54. Rosset, S., ICML '04 Proceedings of the twenty-first international conference on Machine learning. Model selection via the AUC (s. 89). Banff: ACM New York, 2004. 10.1145/1015330.1015400