INVESTIGATION OF POLYPS IN ENDOSCOPY IMAGES BY USING DEEP LEARNING ALGORITHM

Emine CENGİZ[1] , Faik YAYLAK[2] , Eyyup GULBANDILAR[3*]

[1] Yalova University, Faculty of Engineering, Department of Computer Engineering, Merkez Campus, Cınarcık Street, Yalova, Turkey, ORCID No : http://orcid.org/0000-0002-6695-9500

[2] Kutahya Health Sciences University, Faculty of Medicine, Department of General Surgery, Evliya Celebi Campus, Kutahya, Turkey, ORCID No : http://orcid.org/0000-0002-1216-0429

[3] Eskisehir Osmangazi University, Faculty of Engineering and Architecture, Department of Computer Engineering, Meselik Campus, Odunpazari, Eskisehir, Turkey, ORCID No : http://orcid.org/0000-0003-0297-4864

Keywords

Deep learning, Activation function, Optimization method, Polyp, Endoscopy

Abstract

Recent advances in machine learning, particularly with regard to deep learning, help to recognize and classify objects in medical images. In this study, endoscopy images were examined and deep learning method was used to classify healthy and polyp cells. For the proposed system, a database was created from the archives of General Surgery Department Endoscopy Unit in Kutahya Evliya Celebi Training and Research Hospital. The database contains 93 polyps and 216 normal images from 54 archive records. For image multiplexing, a total of 1236 images were obtained by rotating each image 90 degrees around its axis. K-fold Cross Validation method was used to reduce the variability of performance results. In this study, 48 different models were created by using different activation and optimization functions to find the best classification model in deep learning. According to the experimental results, it was observed that accuracy of the models depends on the selected parameters; the best model with the accuracy rate of 91% was obtained with 64 neurons in the hidden layer, ReLU activation function and RmsProp optimization method whereas the worst model with the accuracy rate of 76% was obtained with 32 neurons in the hidden layer, Tanh activation and RMSprop optimization functions. Accordingly, classification performance of polyp images can be optimized by utilizing different activation and optimization methods during the design of deep learning models.

DERİN ÖĞRENME ALGORİTMASI KULLANILARAK ENDOSKOPİ GÖRÜNTÜLERİNDE POLİPLERİN ARAŞTIRILMASI

Anahtar Kelimeler

Derin öğrenme, Aktivasyon fonksiyonu, Optimizasyon metodu, Polip, Endoskopi

Öz

Makine öğrenimindeki, özellikle derin öğrenmeyle ilgili son gelişmeler, tıbbi görüntülerdeki nesneleri tanımaya ve sınıflandırmaya yardımcı olur. Bu çalışmada endoskopi görüntüleri incelenmiş, sağlıklı ve polipli hücrelerini sınıflandırılması için derin öğrenme yöntemi kullanılmıştır. Önerilen sistem için Kütahya Evliya Çelebi Eğitim ve Araştırma Hastanesi Genel Cerrahi Anabilim Dalı Endoskopi Ünitesi arşivlerinden bir veri tabanı oluşturulmuştur. Veri tabanı 54 arşiv kaydından; 93 polip ve 216 normal görüntü içermektedir. Görüntü çoğaltma için her görüntü kendi ekseni etrafında 90 derece döndürülerek toplam 1236 görüntü elde edilmiştir. Performans sonuçlarının değişkenliğini azaltmak için K-kat Çapraz Doğrulama yöntemi kullanıldı. Bu çalışmada, derin öğrenmede en iyi sınıflandırma modelini bulmak için farklı aktivasyon ve optimizasyon fonksiyonları kullanılarak 48 farklı model oluşturulmuştur. Deneysel sonuçlara göre, modellerin doğruluğunun seçilen parametrelere bağlı olduğu; %91 doğruluk oranı ile en iyi model gizli katmandaki 64 nöron, ReLU aktivasyon fonksiyonu

[*] Corresponding Author; e-mail : egulbandilar@ogu.edu.tr

ve RmsProp optimizasyon yöntemi ile elde edilirken, en kötü model %76 doğruluk oranı ile gizli katmandaki 32 nöron, Tanh aktivasyonu, RMSprop optimizasyon yöntemi ile elde edilmiştir. Buna göre, derin öğrenme modellerinin tasarımı sırasında farklı aktivasyon ve optimizasyon yöntemleri kullanılarak polip görüntülerinin sınıflandırma performansı optimize edilebilir.

Araştırma Makalesi		Research Article	
Başvuru Tarihi	: 28.05.2022	Submission Date	: 28.05.2022
Kabul Tarihi	: 13.10.2022	Accepted Date	: 13.10.2022

1. Introduction

The success of machine learning algorithms on image recognition in recent years coincides with the time when electronic medical records and diagnostic imaging have increased significantly. This highlights Convolutional Neural Networks (CNNs) that focus on machine learning algorithms applied to medical image analysis and emphasizes the clinical aspects of the field. Machine learning techniques cannot process unprocessed information without expert assistance and pre-processing. Unlike machine learning, in deep learning, the learning process takes place on raw data and the necessary information for this is obtained with the model created in different layers.

Medical image interpretation is mostly done by experts such as radiologists or doctors. Meanwhile, given the wide variety in pathology and the potential fatigue of experts, researchers and physicians have started to benefit from computer aids. Although the rate of progress in medical image analysis is not as fast as in medical imaging technology, this is improving thanks to machine learning techniques (Shen, Guoron and Heung-Il, 2017).

Endoscopy is considered a standard method for screening stomach polyps. Stomach polyps are rare lesions found incidentally during endoscopy procedures performed for different reasons. These lesions can be treated if detected early (Rustam et al., 2021). The use of endoscopy process with high resolution devices has increased the studies on the use of smart systems in this field.

The health sector is a high priority sector, completely different from other sectors. People expect the highest level of care and service regardless of cost. Therefore, it seems that deep learning in real world applications produces exciting, accurate solutions for medical imaging and is an important method for future applications in the healthcare industry (Yixuan and Meng, 2017).

This study aimed that reveals the design of the support system for physicians by classifying the polyp and normal images in endoscopy images using the CNN method, which is one of the deep learning algorithms that has been popular in the field of engineering in recent years.

2. Literature Review

The use of deep learning algorithms, a sub-branch of machine learning, is increasing rapidly and enhancing performance in various medical applications.

Deep learning is used in applications such as the detection of anatomical and cellular structures, tissue segmentation, computer-aided disease diagnosis.

Pannu, Ahuja, Dang, Soni and Malhi (2020) in their study described a CNN augmented supervised learning ensemble to detect bleeding symptoms in Wireless Capsule Endoscopy images. RGB pixel densities and data distribution are studied to illustrate the complexity of the classification problem. Color palette reduction using minimum variance quantization is used for back propagation in CNN. As a result of the study, an accuracy of 0.95 in the general endoscopy dataset and 0.93 in the real video dataset was obtained.

Ozawa et al. (2020) used DCNN architecture called Single Shot MultiBox Detector for the detection of colorectal polyps. This architecture has been used as both polyp detection and polyp classification method. In the study, 7077 colonoscopy images, including 1172 colorectal polyp images, were used. As a result of the study, a sensitivity of 92% and a positive predictive value (PPV) of 86% were obtained.

Byrne et al. (2019) studied a DCNN model for real-time polyp classification on endoscopic video images of colorectal polyps. The generated model was tested in a separate series of 125 videos with polyps. As a result of the study, 98% sensitivity, 83% specificity, 94% accuracy were obtained.

Tulum et al. (2019) aimed to develop the CNN classifier system for automatic detection of polyp structures. Polyp/false positive classification was performed with CNN using 1543 candidate polyp projection images obtained after colon segmentation, candidate region identification, and false positive reduction in images obtained from 30 different patients. The developed classification system performs with a sensitivity of 91.89% and a false positive rate of 0 per data set.

Shin and Balasingham (2017) took the shape and colour properties of the images and compared these two classification methods using Support Vector Machines

(SVM) and CNN. Three convolution layers and adamax optimization function were used for this study. In addition, the learning rate was 0.002 and the number of revolutions was 200. Since there were similar pictures in the same data set and so exaggerated results could be obtained, CVC-Clinic, ETIS-Larib and Asu-Maya data sets were used. 612 images in the CVC-Clinic dataset were used for training, while 196 images in ETIS-Larib and 170 images in Asu-Maya dataset were used for testing. At the end of the study, 84% accuracy was obtained with SVM and 91% accuracy with CNN.

Ribeiro, Uhl and Hafner (2016) studied on polyp classification using the CNN deep learning algorithm. For this, they obtained 25 healthy images from 18 patients and abnormal images from 56 patients. Since the increase in the number of data would improve the training, they obtained a total of 800 pictures by rotating the shapes and changing the size. In this study, different architectures were tested to evaluate the effect of filter size and number, and the number of output units in the fully connected layer in the classification.

Zou et al. (2015) examined the classification problem of digestive organs in capsule endoscopy images using the Deep Convolutional Neural Network (DCNN) model. One million images were used for this study. 60 thousand images for training and 15 thousand images for testing were randomly selected. There were 20 thousand pictures for the stomach, small intestine and colon in the training and five thousand pictures for each class's test. The learning rate was determined as 0.01. In Scholastic Gradient Descent (SGD), fully connected layers each had 64 neurons to train the system. An accuracy of 90.31% was obtained with the SVM method and 95.52% with CNN.

Sarraf and Tofighi (2016) used a deep learning method, CNN, to differentiate the Alzheimer's brain from a normal and healthy brain. In this study, images of 28 Alzheimer's patients and 15 healthy individuals, consisting of 24 women and 19 men, were selected from the Alzheimer's Disease Neuroimaging Initiative database. The data was divided into three parts. 60% for training, 20% for verification and 20% for testing. Using this method, the classification accuracy rate was found to be 96.85%.

Suzuki et al. (2016) used Deep Convolutional Neural Networks (DCNN), one of the deep learning methods, to identify masses in mammography. In the study, they formed eight weighted layers. Five of these layers were convolution layers and three were fully connected layers. The study was carried out on a total of 198 mammography images, 99 with a mass and 99 healthy images. Mass classification accuracy was 89.90%.

Yang et al. (2016) made a performance comparison of SVM and DCNN in their study. For this purpose, they used 300×300 pixel-images of 243 normal cells and 334 tumors in the kidney. To increase the number of data, 577 images were cropped or rotated, resulting in a total of 21,349 images. 14.233 samples of this data set were reserved for training, 3.558 samples for validation and 3.558 samples for testing. In the DCNN classification, the number of intermediate layers was tested in two ways as five and seven. The following accuracy rates were found: 85% for SVM, 97.37% for DCNN-5 and 97.91% for DCNN-7.

Ortac and Ozcan (2021) studied by multidimensional deep learning method such as Convolutional Neural Network (CNN) on hyperspectral images. In this study, they are evaluated one-dimensional, two-dimensional and three-dimensional convolution model approaches that can present efficient classification performance. As a result, their studies have presented to have achieved higher classification rates compared three-dimensional Convolutional Neural Networks with a state of art models.

3. Material and Method

3.1. Convolutional Neural Networks

Convolutional neural networks (CNN) are one of the deep learning algorithms to be used in vision and image processing, especially on computers, inspired by the brain's technique for processing visual information (Bengio, 2008). CNN is a specialized version of Artificial Neural Networks (ANN) in training pictures and videos, but how CNN works is slightly different from the standard ANN. In CNN, each neuron in a layer dense only on a small part of the image. While a neuron in one layer of the network is focused on some features of the image in the study, other neurons in the layer focus on the other features of the same image. CNNs use convolution in one or more of the layers instead of the general matrix multiplication.

A CNN model has two important advantages. The first is to reduce the weights of filters that are important to learn. This eliminates many problems that may occur due to the need for too many neurons and the large number of weights for these neurons to learn. The second advantage is weight sharing. The parameters learned with CNN can be shared as input to the next layer. Therefore, the same weights are used again in the layer and so there is no need to learn again. Thus, more layers learn more complex features and patterns.

CNN generally consists of three layers: convolution, pooling and fully connected layers. CNN design for endoscopy polyp classification Figure. 1 shows the general structure of CNN.

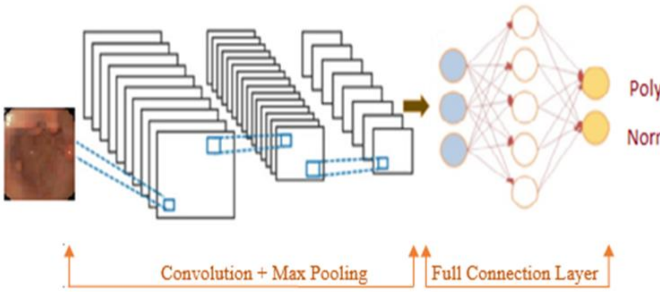


Figure .1. CNN for endoscopy polyp

3.1.1 Convolutional Layer

The convolution layer is the first hidden layer of CNN. In this layer, an image becomes a batch of filtered images. The convolution layer reduces the number of weights in the network, allowing the system to be trained faster. Filters are processed with the input image to generate a smaller data set and pass it to the next layer.

In the convolution process, 3x3, 5x5, 7x7, 9x9, 11x11 size filters are used on the image matrix. The specified filters circulate over the entire image matrix, highlighting the features in the image and a new image matrix in image size is obtained (Liu, Shen and Van den Hengel, 2015). However, as the size of the used filter increases, the size of the output image will decrease, so information will be lost. That is why the smallest filter size, 3x3, is usually used.

At first, the filter is placed in the upper left corner of the image. The image and filter indices are multiplied by each other and the results are summed, and the result is stored in the output matrix. Then the filter is moved one pixel to the right and the process is repeated. After the first line ends in the input image, the process is repeated in the second line. After all operations are completed, the output matrix is obtained.

3.1.2 Pooling Layer

Pooling layer is a layer that usually takes place after the convolution process (Castelluccio, Poggi, Sansone, and Verdoliva, 2015). It is the layer applied to gradually reduce the spatial size of the image to reduce the number of parameters and calculations in the network, and thus control overfitting. This layer operates independently in each depth slice of the input, so the depth dimension remains unchanged because the depth dimension represents the coloured part of the picture.

Filters are used in the pooling layer just as in the convolution layer. These filters can be realized by shifting the image according to a certain step-by-step value, taking the maximum value of the pixels in the image (max pooling), taking the smallest value (min pooling), and taking the average of the values (mean pooling).

3.1.3. Fully Connected Layer

It is the layer that follows the convolution and pooling layers and is called "dense". The classification of the model from the features obtained from the previous layers takes place in this layer. The fully connected layer takes the feature map from the previous layer and process it for the desired number of classes. Since the values to be obtained at the output of the network will be numbers, after processing with multi-dimensional data (matrices), the data obtained from the feature map is made unidimensional by reducing the dimension in this layer.

3.1.4. Dropout Layer

The dropout layer is a technique used to avoid the problem of memorizing data in artificial neural networks (Srivastava, Hinton, Krizhevsky, Sutskever and Salakhutdinov, 2014). The dropout technique is generally used later in fully connected layers. The bonds in the fully connected layer are broken using dropout. Although the dropout value varies according to the problem and the data set, it is generally taken as 0.5.

3.1.5. Classification Layer

After the fully connected layer, the classification layer is used. The number of objects to be classified and the output value of the classification layer must be the same. Different classifiers are used in this layer. Softmax classifier is mostly preferred because of its success. Softmax is a generalization of the sigmoid activation function for multiple classifications.

3.2. Data Enhancement

The success rate of the training is increased by applying various deformation and transformation processes to the data to be included in the training. It is possible to perform these operations in real time with the ImageDataGenerator class. Using ImageDataGenerator positively affects the educational success but extends the training period. Table 1 shows the data enhancement techniques used for this study.

Table 1

Data enhancement techniques

Data enhancement method	**Explanation**	**Value**
Rescale	It is the rescaling value. Its default value is set to 'None'. If the value is None or 0, no rescaling is applied. Otherwise, the input data is multiplied by the given value before all operations.	1./255
Shear_range	The shear angle, in degrees, applied counter clockwise.	0.2
Zoom_range	It is the random approximation range.	0.2
Horizontal_flip	flips inputs horizontal	True
Vertical_flip	flips inputs vertically	True

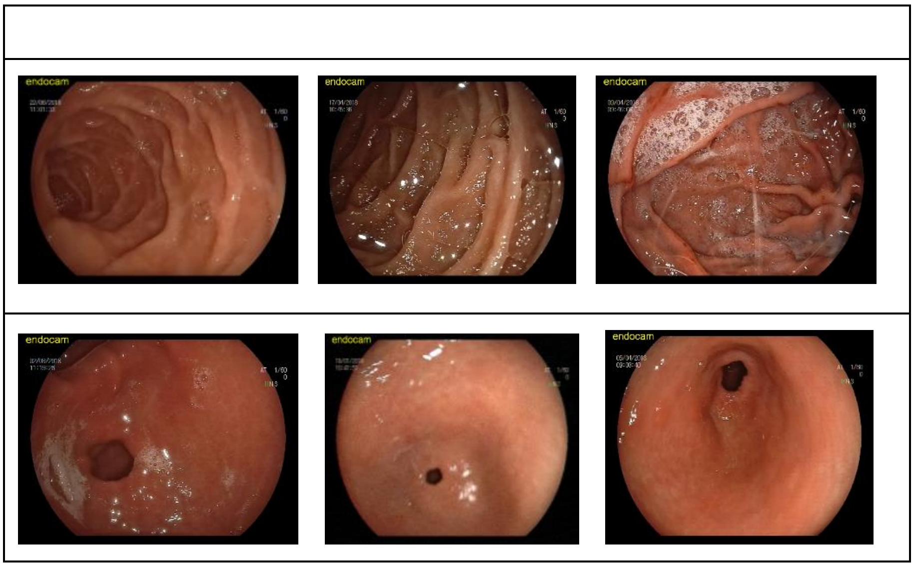


Figure 2. Normal images

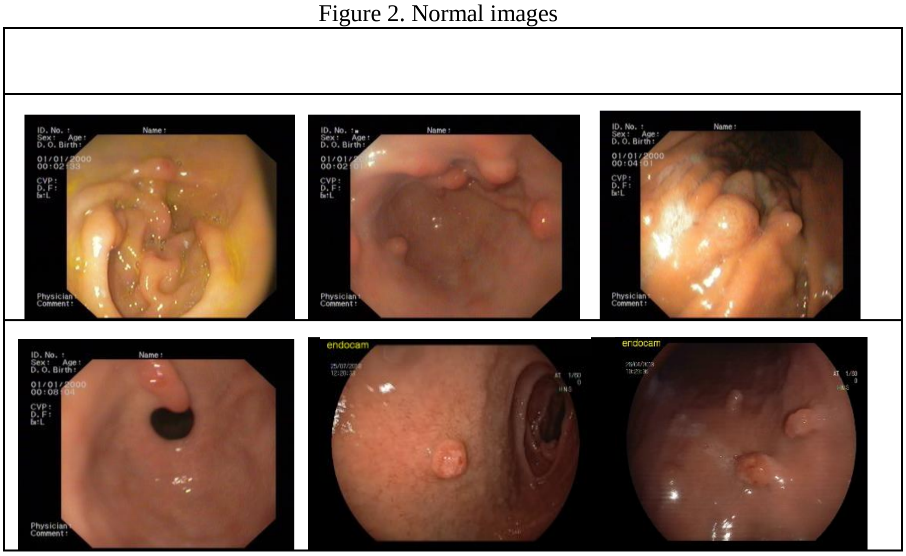


Figure 3. Polyp images

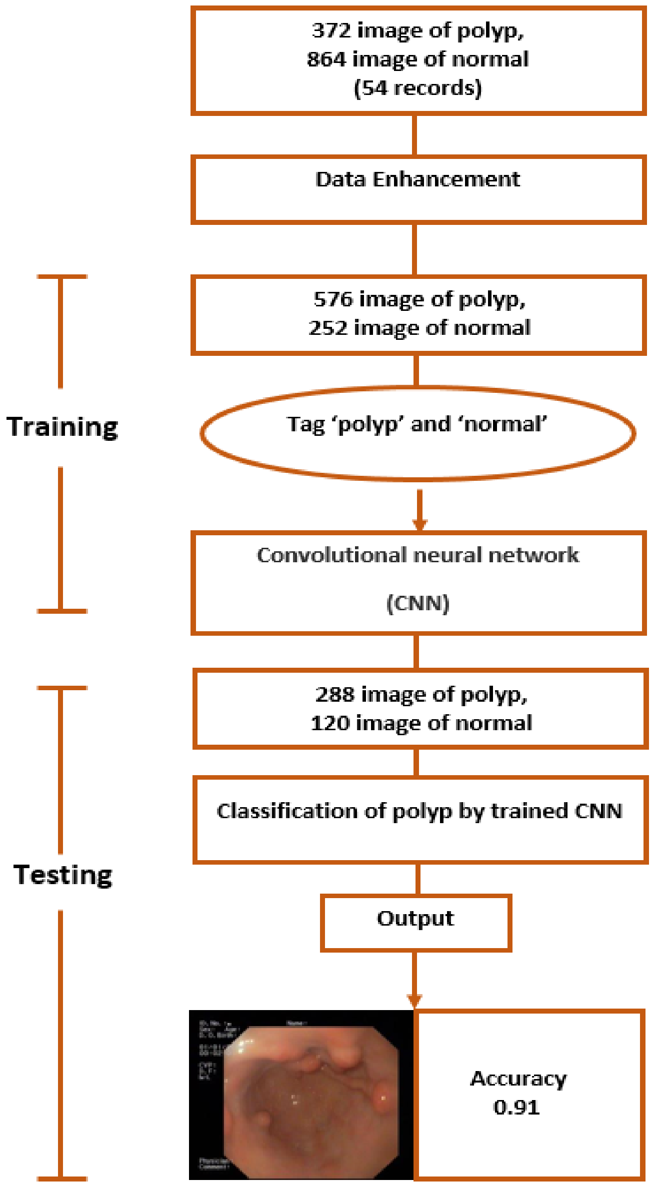


Figure 4. Flowchart of the study design

3.3. Dataset

In this study, a database was created by taking polyps and normal images from the 2018 archive records of Kütahya Health Sciences University General Surgery Department, Kutahya Evliya Celebi Training and Research Hospital Endoscopy Unit. This study was carried out in accordance with the Research and Publication Ethics. This study was approved by Kutahya Health Sciences University Ethic Committee on Non-Invasive Studies on date 17 July 2019 with approval number E.5269. In this database, there were 93 polyp and 216 normal images from 54 records. In order to increase the number of images in the data set, a total of 1236 images were obtained by rotating each image 90 degrees around its axis. Examples for normal images are given Figure 2 and polyp images are given in Figure 3.

For this study, a retrospective medical endoscopy image archive (of Kutahya Health Sciences University Surgical Endoscopy Unit) was used with permission of Evliya Celebi Research and Training Hospital. Data usage was permitted based on the frame of study protocol. In this frame all available data source was anonymized before image files were made available for the studies. This anonymization included the removal of all meta data, text reports, any ID numbers, names, and age or gender features which may be indicated in the whole endoscopy record. This anonymization procedure was performed in the endoscopy suite and then this framed data was ready to be transferred to the researchers. Thus, the image data used in this study was original with it's origin, but retrospective thus an informed consent case by case was not needed and available. After manager permission, ethic committee approval was available from Kutahya Health Sceinces University Ethic Committee on Non-Invasive Studies. After this permission and approval procedures image data for the study were transferred to the researchers. Figure 4 shows the flowchart of the study design (Cengiz, 2020).

3.4. Hyper Parameters

The performances of the models were evaluated according to the accuracy, precision, recall and f1-score of the test data.

Accuracy: The number of correct predictions over all predictions.

$$\text{Accuracy} = \frac{\text{TP+TN}}{\text{TP+FP+TN+FN}} \quad (1)$$

Precision: Precision is the ratio of correctly predicted positive observations to the total predicted positive observations.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \quad (2)$$

Recall: Recall is a measure of how many of the positive cases the classifier correctly predicted, over all the positive cases in the data.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (3)$$

F1 score: F1 Score is the weighted average of precision and recall.

$$\text{F}_1 = 2 * \frac{\text{precision}*\text{recall}}{\text{precision+recall}} \quad (4)$$

In the study, 48 different models were created using the neuron numbers, epoch numbers, activation functions and optimization methods given in Table 2.

Table 2

Hyper parameters belonging to the models created in the study.

Parameters used in the models	Parameter Value
Numbers of Neurons	32
	64
Epoch Numbers	5
	10
	15
Activation Functions	Relu
	Tanh
Optimization Methods	Sgd
	Adagrad
	Adam
	RMSprop

3.4.1 Activation methods

One of the activation functions is applied after the convolution layer. If the activation function is not implemented, the neural network becomes a linear function with limited learning abilities. In order to obtain results for nonlinear operations, the activation function is mostly a nonlinear function. Figure 5 shows x input, w weight and *f(x)* activation process sent to the output of the network. The result is either an output or an input to another layer. Sigmoid, Tanh and ReLU are generally used as activation functions.

3.4.2 Optimization methods

In ANN, the weight values must be updated at each step until optimum learning is achieved and this process takes place within certain methods. In weight updating, "Gradient Descent" is one of the widely used methods (Yazan and Talu, 2017). Different gradient-based optimization algorithms are used to update the weights in ANN. These algorithms have advantages and disadvantages compared to each other (Ruder, 2016). Optimization algorithms such as Adam, Stochastic Gradient Descent, Adagrad, Adamax are commonly used in ANN to update the weight coefficients.

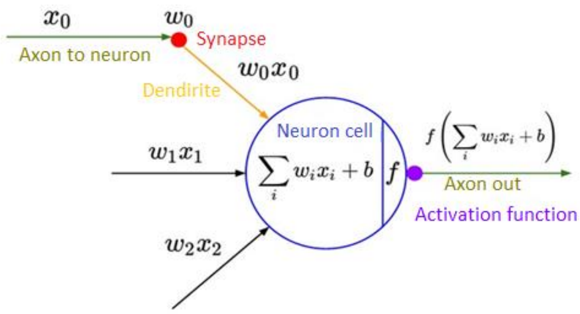


Figure 5. Activation function

4. Experimental Results

Network models prepared with hyper parameters given in Chapter 3.4 were applied to the training process with 5, 10 and 15 epoch numbers and sample data set with 50 iterations in each epoch. The accuracy, precision, recall and f1-score values of the models are shown in Table 3. K-Fold Cross Validation method was used to reduce the variability of performance results due to the random generation of training and test data. The K value was chosen as 10 (Cengiz, 2020).

According to Table 3, Model 24 is the most successful classification model with 64 neurons in its hidden layer, Relu activation function, 15 epoch number, and 91.1% accuracy using RmsProp optimization technique. Model 28 is classified with 76.3% accuracy and the lowest success rate among 48 models. The accuracy values of the test data show that the classification range of the models varies between 0.76 and 0.91. Table 3 also shows that the models with the highest and lowest success were created by Adam and Rmsprop optimization methods.

Figure. 6 shows the performance metrics of the top five models and Figure. 7 shows the performance metrics the least successful five models.

Table 3, 4 and 5 show that more successful results are obtained when RmsProp optimization technique is used with ReLu activation.

When the RmsProp optimization algorithm is used together with the Relu activation function, it has been observed that the success rate is increased when the epoch number is increased in the 64-neuron algorithm. (91.1% > 90.5 ; model 24, model 16).

Similarly, the model created with the RmsProp optimization algorithm with 32 neurons and Relu activation function was observed to be more successful than the models created with the Adam optimization algorithm. (90.6% > 90.5% ; model 12, model 3)

In addition, when the model 28 is compared with the model 40, and the model 36 with the model 48, it is observed that the success rate increases when the number of neurons is increased.

In this study, endoscopic images were examined, and the models created with different activation and optimization methods were compared for the classification of normal and polyp images. According to the data obtained, it was observed that the results of the models differed according to the parameters.

5. Conclusions

In order to find the best classification model in deep learning, different activation and optimization methods, different numbers of neurons and epochs were used. The results of the models differed according to the parameters.

Using optimization methods, activation functions, neuron and epoch numbers, 48 different deep neural network models were created. The most successful models were created with Rmsprop optimization methods. The results of the study are examined, it is seen that while the Relu activation function generally showed more successful results than the Tanh activation function.

It is known that, in studies with deep neural networks, the performance of the models varies according to the structure, size, optimization methods and activation functions of the data set used. In addition, the performance of optimization algorithms varies depending on the choice of parameters and how the neural network will be created.

Although the selection criteria of optimization algorithms in deep learning applications are not known exactly, the algorithms show different performances depending on the structure of the problem and the parameters. Therefore, if different combinations of optimization algorithms, activation functions, neuron and epoch numbers are created while modelling in deep learning applications, a more suitable architecture for the data set will be obtained.

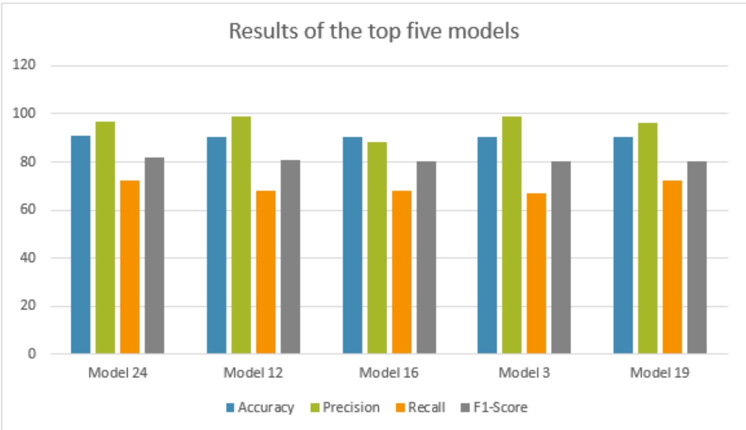


Figure 6. Performance metrics of the top five models.

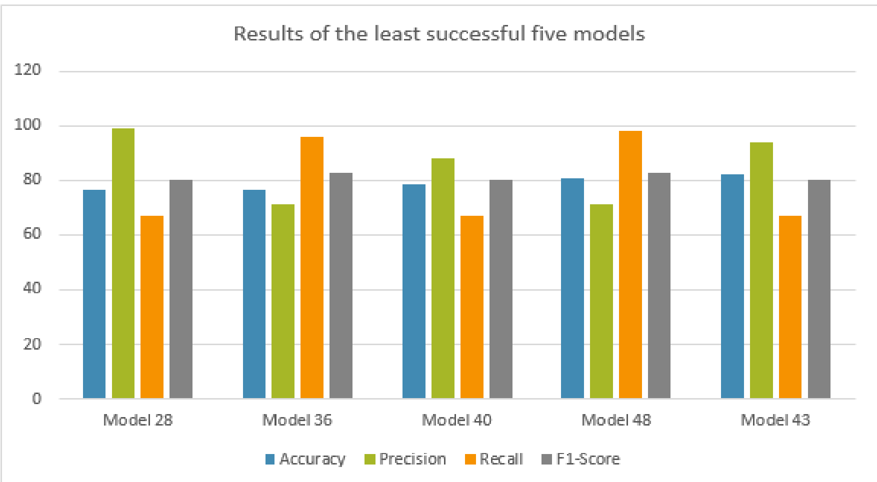


Figure 7. Performance metrics of the least successful five models.

Table 3

Hyper parameters belonging to the models created in the study

Model Number	Numbers of Neurons	Activation Functions	Epoch Number	Optimization Function	Accuracy (%)	Precision (%)	Recall (%)	f1-Score (%)
1	32	Relu	5	Sgd	86.27	71.0	70.0	83.0
2				Adagrad	87.08	98.0	61.0	76.0
3				Adam	90.54	99.0	67.0	80.0
4				RMSprop	90.32	99.0	67.0	80.0
5			10	Sgd	88.85	96.0	67.0	79.0
6				Adagrad	86.72	89.0	62.0	76.0
7				Adam	90.39	96.0	71.0	81.0
8				RMSprop	89.39	95.0	72.0	82.0
9			15	Sgd	89.26	96.0	68.0	79.0
10				Adagrad	89.07	79.0	70.0	74.0
11				Adam	89.51	86.0	72.0	79.0
12				RMSprop	90.64	99.0	68.0	81.0
13	64	Relu	5	Sgd	89.17	99.0	64.0	78.0
14				Adagrad	84.29	98.0	61.0	76.0
15				Adam	90.02	94.0	72.0	82.0
16				RMSprop	90.59	88.0	68.0	80.0
17			10	Sgd	88.82	92.0	70.0	80.0
18				Adagrad	89.41	99.0	66.0	79.0
19				Adam	90.47	96.0	72.0	82.0
20				RMSprop	90.27	92.0	73.0	81.0
21			15	Sgd	88.28	62.0	73.0	67.0
22				Adagrad	89.66	99.0	66.0	79.0
23				Adam	90.34	95.0	71.0	80.0
24				RMSprop	91.10	97.0	72.0	82.0
25			5	Sgd	88.77	98.0	63.0	78.0
26				Adagrad	89.36	98.0	64.0	78.0
27				Adam	86.42	99.0	68.0	81.0
28				RMSprop	76.35	99.0	67.0	80.0

Table 3

Hyper parameters belonging to the models created in the study (continue)

29				**Sgd**	**89.34**	**97.0**	**65.0**	**79.0**
30	**32**	**Tanh**	**10**	**Adagrad**	**89.73**	**94.0**	**63.0**	**78.0**
31				**Adam**	**82.79**	**96.0**	**67.0**	**80.0**
32				**RMSprop**	**82.28**	**71.0**	**95.0**	**83.0**
33				**Sgd**	**89.61**	**88.0**	**66.0**	**79.0**
34				**Adagrad**	**89.49**	**98.0**	**67.0**	**79.0**
35			**15**	**Adam**	**84.85**	**96.0**	**72.0**	**82.0**
36				**RMSprop**	**76.62**	**71.0**	**96.0**	**83.0**
37				**Sgd**	**89.90**	**95.0**	**67.0**	**80.0**
38			**5**	**Adagrad**	**89.39**	**99.0**	**67.0**	**80.0**
39				**Adam**	**86.42**	**93.0**	**73.0**	**82.0**
40				**RMSprop**	**78.50**	**88.0**	**67.0**	**80.0**
41				**Sgd**	**89.24**	**88.0**	**66.0**	**79.0**
42			**10**	**Adagrad**	**89.02**	**92.0**	**70.0**	**80.0**
43	**64**	**Tanh**		**Adam**	**82.11**	**94.0**	**67.0**	**80.0**
44				**RMSprop**	**86.10**	**98.0**	**74.0**	**84.0**
45				**Sgd**	**90.07**	**95.0**	**69.0**	**80.0**
46				**Adagrad**	**89.63**	**93.0**	**68.0**	**78.0**
47			**15**	**Adam**	**88.87**	**99.0**	**72.0**	**83.0**
48				**RMSprop**	**80.88**	**71.0**	**98.0**	**83.0**

Compliance with ethical standards

This study was approved by Kutahya Health Sciences University Ethic Committee on Non-Invasive Studies on date 17 July 2019 with approval number E.5269.

An original retrospective image data source was used with permission of Evliya Celebi Research and Training Hospital administration.

No informed consent was needed for this study.

Contribution of Researchers

In this study; Emine CENGİZ, preparation of codes, literature review; Faik YAYLAK, labeling of data, evaluation of results; Eyyup GULBANDILAR contributed to the writing of the codes, the evaluation of the results, and the preparation of the article.

Conflict of Interest

No conflict of interest was declared by the authors.

References

Bengio, Y. (2008) *Learning deep architectures for AI. Foundations and Trends in Machine Learning* 2(1): 1–127.

Byrne, M. F., Chapados, N., Soudan, F., Oertel, C., Pérez, M. L., Kelly, R., Iqbal, N., Chandelier, F. & Rex, D. K. (2019). Real-time differentiation of adenomatous and hyperplastic diminutive colorectal polyps during analysis of unaltered videos of standard colonoscopy using a deep learning model. *Gut*, *68*(1), 94-100. http://dx.doi.org/10.1136/gutjnl-2017-314547

Castelluccio, M., Poggi, G., Sansone, C. & Verdoliva, L. (2015) *Land use classification in remote sensing images by convolutional neural networks*, arXiv preprint arXiv:1508.00092. https://doi.org/10.48550/arXiv.1508.00092

Cengiz, E. (2020). (Master's Thesis). *Investigation of polyps in endoscopy images by using deep learning algorithm.* Eskisehir Osmangazi University Graduate School of Natural and Applied Sciences, Eskisehir, Turkey. (in Turkish)

Liu, L., Shen, C. & Van den Hengel, A. (2015) *The treasure beneath convolutional layers: Cross-convolutional-layer pooling for image classification*, In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, USA, pp. 4749-4757.

Ortac, G., & Ozcan, G. (2021). Comparative study of hyperspectral image classification by multidimensional Convolutional Neural Network approaches to improve accuracy. *Expert Systems with Applications*, 182, 115280. https://doi.org/10.1016/j.eswa.2021.115280

Ozawa, T., Ishihara, S., Fujishiro, M., Kumagai, Y., Shichijo, S., & Tada, T. (2020). Automated endoscopic detection and classification of colorectal polyps using convolutional neural networks. *Therapeutic advances in gastroenterology*, *13*,1756284820910659. https://doi.org/10.1177/1756284820910659

Pannu, H. S., Ahuja, S., Dang, N., Soni, S., & Malhi, A. K. (2020). Deep learning based image classification for intestinal hemorrhage. *Multimedia Tools and Applications*, *79*(29), 21941-21966. https://doi:10.1109/access.2021.3061592

Ribeiro, E., Uhl, A. & Hafner, M. (2016) *Colonic polyp classification with convolutional neural networks*, In: 2016 IEEE 29th International Symposium on Computer-Based Medical Systems, Belfast and Dublin, Ireland, pp.253-258.

Ruder, S. (2016) An overview of gradient descent optimization algorithms, arXiv preprint arXiv:1609.04747. https://doi.org/10.48550/arXiv.1609.04747

Rustam, F., Siddique, M. A., Siddiqui, H. U. R., Ullah, S., Mehmood, A., Ashraf, I. & Choi, G. S. (2021). Wireless capsule endoscopy bleeding images classification using CNN based model. *IEEE Access*, *9*, 33675-33688.

Sarraf, S. & Tofighi, G. (2016), 2016 IEEE Future Technologies Conference, pp. 816-820. https://doi:10.1109/FTC.2016.7821697

Shen, D., Guoron, W., Heung-Il, S. (2017) Deep learning in medical image analysis. *Annual review of Biomedical Engineering* 19, 221-248. https://doi.org/10.1146/annurev-bioeng-071516-044442

Shin, Y. & Balasingham, I. (2017) *Comparison of hand-craft feature based SVM and CNN based deep learning framework for automatic polyp classification*. In: 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Jeju Island, Korea, pp. 3277-3280. https://doi:10.1109/embc.2017.8037556

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. & Salakhutdinov, R. (2014) Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(1), 1929- 1958.

Suzuki, S., Zhang, X., Homma, N., Ichiji, K., Sugita, N., Kawasumi, Y., Ishibas, T. & Yoshizawa, M. (2016) *Mass detection using deep convolutional neural network for mammographic computer-aided diagnosis*, 55th Annual Conference of the Society of Instrument and Control Engineers of Japan, Tsukuba, Japan, pp. 1382-1386. doi:10.1109/sice.2016.7749265

Tulum, G., Osman, O., Bolat, B., Dandin, Ö., Ergin, T., & Cüce, F. (2019). Colonic Polyp Classification Using Projection Image and Convolutional Neural Network. In *2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT)* (pp. 1-4). IEEE. https://doi:10.1109/EBBT.2019.8741701 doi:10.1109/EBBT.2019.8741701

Yang, X., Chen, S., Ding, V. et al (2016) *A deep learning approach for tumor tissue image classification.* IASTED Biomedical Engineering – 2016, https://www.researchgate.net/publication/298929528_A_Deep_Learning_Approach_for_Tumor_Tissue_Image_Classification

Yazan, E. & Talu, M.F. (2017) *Comparison of the stochastic gradient descent based optimization techniques, 2017* International Artificial Intelligence and Data Processing Symposium (IDAP), Malatya, Turkey, pp. 1-5. https://doi:10.1109/IDAP.2017.8090299

Yixuan, Y. & Meng, M. (2017) Deep learning for polyp recognition in wireless endoscopy images. Medical *Physics*, 44(4), 1379-1389. https://doi.org/10.1002/mp.12147

Zou, Y., Li, L., Wang, Y. et al (2015) *Classifying digestive organs in wireless capsule endoscopy images based on deep convolutional neural network.* In: 2015 IEEE International Conference on Digital Signal Processing, Singapore, pp. 1274-1278. https://doi:10.1109/ICDSP.2015.7252086

Table 4

Results of the top five models

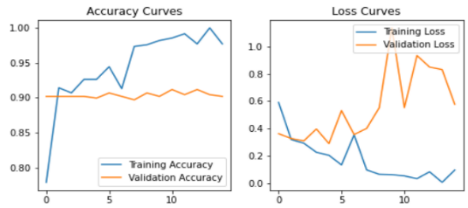
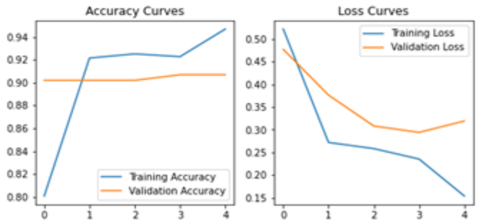
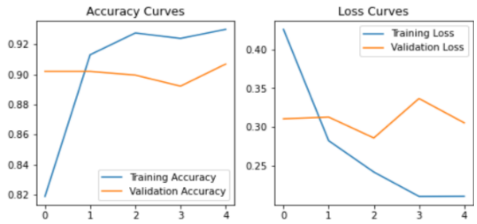
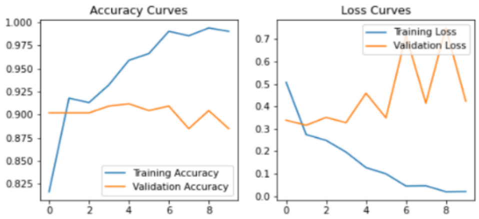
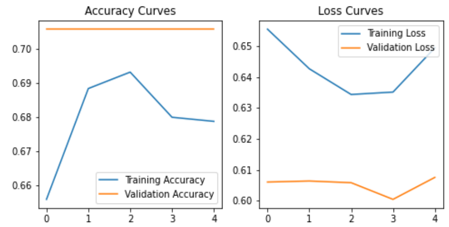
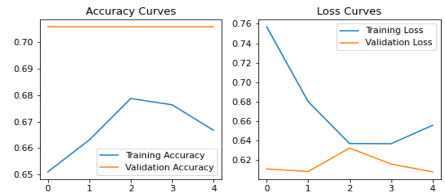
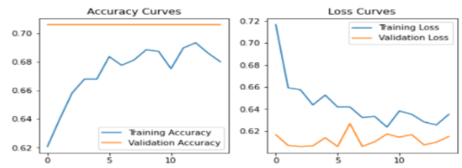
Model no	Numbers of Neurons	Activation Function	Optimization Method	Accuracy (%)	Accuracy and Loss Curves
24	64	Relu	RMSprop	91.10%	
12	32	Relu	RMSprop	90.64%	
16	64	Relu	RMSprop	90.59%	
3	32	Relu	Adam	90.54%	
19	64	Relu	Adam	90.47%	

Table 5

Results of the least successful five models.

Model no	Numbers of Neurons	Activation Function	Optimization Method	Accuracy (%)	Accuracy and Loss Curves
28	32	Tanh	RMSprop	76.35%	
36	32	Tanh	RMSprop	76.62%	
40	64	Tanh	RMSprop	78.50%	
48	64	Tanh	RMSprop	80.88%	
43	64	Tanh	Adam	82.11%