Avrupa Bilim ve Teknoloji Dergisi
Özel Sayı, S. 309-318, Ekim 2019
© Telif hakkı EJOSAT'a aittir
Araştırma Makalesi

www.ejosat.com ISSN:2148-2683

European Journal of Science and Technology
Special Issue, pp. 309-318, October 2019
Copyright © 2019 EJOSAT
Research Article

Classification of Skin Lesions in Dermatoscopic Images with Deep Convolution Network

Emrah ÇEVİK[1*], Kenan ZENGİN[2]

[1]Tokat Gaziosmanpaşa Üniversitesi, Reşadiye Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Tokat, Türkiye (ORCID: 0000-0002-6199-5529)
[2]Tokat Gaziosmanpaşa Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Bilgisayar Mühendisliği Bölümü, Tokat, Türkiye (ORCID: 0000-0002-7940-6315)

(This publication has been presented orally at HORA congress.)

(First received 1 August 2019 and in final form 25 October 2019)

(**DOI:**10.31590/ejosat.638247)

ATIF/REFERENCE: Çevik, E. & Zengin, K. (2019). Classification of Skin Lesions in Dermatoscopic Images with Deep Convolution Network. *European Journal of Science and Technology*, (Special Issue), 309-318.

Abstract

In this study, a deep learning based solution with convolutional neural network is presented to solve the problem of classification of dermoscopic images containing skin lesions. Defining models with classical machine learning techniques takes a lot of time and with this model, it cannot make data meaningful without pretreatment. Thanks to deep learning, we have come a long way in problems that we think are difficult to solve for many years. Deep learning achieves results by processing the data at hand without any intervention by us.

Classification of dermoscopic images of skin lesions is a difficult task to distinguish between benign or malignant melanocytic tumors. Malignant melanoma is the deadliest form of skin cancer and is one of the fastest developing cancers in the world. If diagnosed early, it can be easily treated and ultimately, early diagnosis of melanoma is vital. Dermoscopy has become one of the most important tools in the diagnosis of melanoma and other pigmented skin lesions. Due to the inaccuracy, subjectivity, and poor reproducibility of human judgment, there has been a need to process the dermoscopy image with an automatic recognition algorithm.

In 2017, the support vector machine (SVM) classifier was used to differentiate 172 dermatoscopic images into two classes as "benign"and "malignant". Experiments on the dataset have 91% accuracy. However, the fact that we have thousands of images in our data set and that we will break them down into seven lesion classes required us to search for more effective methods.

Classroom inconsistency of melanomas is considered a challenging process due to the low contrast of skin lesions and artificial objects such as noise, presence of hair, air bubbles and similarity between non-melanoma cases in dermoscopy images. To solve these problems, we propose the VGGNET-16 architecture, which includes a powerful convolutional neural network model to classify seven different types of disease on dermoscopic images.

"HAM10000 "(Human Againist Machine) data set was used with VGGNET-16 architecture and the results were observed. The data set, which is published as an educational set for academic machine learning and made public through the ISIC archive, consists of 10015 dermatoscopic images. K-Fold Cross Validation technique was used to differentiate the data set consisting of seven lesion classes as training and test area. In the test phase of the educated model, the validation of the classes was obtained as 85.62%.

Keywords: Deep Learning, CNN, Convolution, Pooling, VGGNET, K-Fold

[1]Corresponding Author: Tokat Gaziosmanpaşa Üniversitesi, Reşadiye Meslek Yüksekokulu, Bilgisayar Teknolojileri Bölümü, Tokat, Türkiye (ORCID: 0000-0002-6199-5529)

Derin Konvolüsyon Ağı ile Dermatoskopik Görüntülerde Deri Lezyonlarının Sınıflandırılması

Öz

Bu çalışmada, cilt lezyonu içeren dermoskopik görüntülerin sınıflandırılması problemini çözmek için konvolüsyonel (evrişimsel) sinir ağı ile derin öğrenme temelli bir çözüm sunulmuştur. Klasik makine öğrenme teknikleri ile model tanımlamak hem çok zaman almakta hem de bu model ile veriyi ön işlem yapmadan anlamlı hale getirememektedir. Derin öğrenme sayesinde, uzun yıllar boyunca çözülmesinin zor olduğunu düşündüğümüz problemlerde büyük mesafe katedilmiştir. Derin öğrenme, elimizdeki veriyi bizim tarafımızdan herhangi bir müdahale olmadan kendisi işleyerek sonuca ulaşmaktadır.
Deri lezyonunun dermoskopik görüntülerin sınıflandırılması melanositik tümörlerin iyi huylu veya kötü huylu olarak ayırt edilmesi zor bir görevdir. Malign melanom cilt kanserinin en ölümcül şeklidir ve dünyadaki en hızlı gelişen kanserlerden biridir. Erken teşhis edilirse kolayca tedavi edilebilir ve sonuçta melanomun erken teşhisi hayati öneme sahiptir. Dermoskopi melanom ve diğer pigmentli cilt lezyonlarının teşhisinde en önemli araçlardan biri haline gelmiştir. İnsan kararının yanlışlığı, öznelliği ve kötü tekrarlanabilirliği nedeniyle, dermoskopi görüntüsünün otomatik bir tanıma algoritması ile işlenmesi bir ihtiyaç olmuştur.
2017 yılında Muhammed Shadid ve Salman Khan tarafından yapılan uygulama ile 172 adet dermatoskopik görüntünün "benign" ve "malignant" olarak iki sınıfa ayrıştırılmasında Support Vector Machine(SVM) sınıflandırıcısı kullanılmıştır. Veri kümesi üzerinde yapılan deneyler %91 doğruluk hassasiyetine sahiptir. Ancak bizim veri setimizde binlerce imge olması ve yedi adet lezyon sınıfına ayrıştıracak olmamız daha etkin yöntemler aramamızı gerektirmiştir.
Melanomların sınıf içi tutarsızlığı, cilt lezyonlarının düşük kontrastı ve dermoskopi görüntülerinde gürültü, saçın varlığı, hava kabarcıkları ve melanom olmayan vakalar arasındaki benzerlik gibi yapay nesneler nedeniyle zorlu bir süreç olarak kabul edilir. Bu problemleri çözmek için, dermoskopik görüntülerde farklı yedi hastalık tipini sınıflandırmak için güçlü bir evrişimsel sinir ağı modeli içeren VGGNET-16 mimarisini önermekteyiz.
"HAM10000" (Human Againist Machine) veri seti üzerinde VGGNET-16 mimarisi ile tasarlanan derin ağ modeli eğitilerek sonuçlar gözlemlenmiştir.Akademik makine öğrenmesi için bir eğitim seti olarak yayınlanan ve ISIC arşivi aracılığıyla kamuya açık olan veri seti 10015 adet dermatoskopik görüntüden oluşmaktadır. Toplam yedi adet lezyon sınıfının bulunduğu veri setinin eğitim ve test alanı olarak ayrıştırılmasında K-Fold Cross Validation tekniğinden faydalanılmıştır. Eğitilmiş modelin test aşamasında sınıfların onaylama doğruluğu%85,62olarak elde edilmiştir.

Anahtar Kelimeler: Derin Öğrenme, ESA, Konvolüsyon, Havuzlama, VGGNET, K-Fold

1. Introduction

In this study, a model was developed using deep learning architectures for early detection of skin cancer using dermoscopic images. Classification of dermoscopic images of skin lesions is a difficult task to distinguish between benign or malignant melanocytic tumors. Malignant melanoma is the deadliest form of skin cancer and is one of the fastest developing cancers in the world. If diagnosed early, it can be easily treated and ultimately, early diagnosis of melanoma is vital. Dermoscopy is an important tool in the early diagnosis of skin cancer. Due to the inaccuracy, subjectivity, and poor reproducibility of human judgment, there has been a need to process the dermoscopy image with an automatic recognition algorithm.

With the introduction of "Deep Learning" methods in the field of machine learning, with the help of GPU and other hardware developments, artificial intelligence based methods have become the preferred method. At the beginning of the 2000s, to produce solutions by increasing the number of layers and nodes in Artificial Neural Networks (ANNs) was not an efficient method as a result of insufficient hardware units. [1].

Through deep learning, it has provided an environment for development in many scientific fields such as voice recognition, visual object detection, drug prediction. [2]. The choice of the learning algorithm is an important factor in large data sets. [3]. Standard artificial neural network is a highly advanced method of classification. We can also talk about the existence of a spatial structure rather than the independence of inputs by the use of convoluted neural networks. [4]. By means of convolutional neural networks, the image input dataset is made more meaningful for the artificial neural network.

Although the incidence of skin cancer diseases has increased in recent years, there has been an increase in the studies conducted for early diagnosis. When we look at previous studies, we see that data sets are limited or class distributions are not balanced. In this sense, we have developed a more reliable model that provides comprehensive discrimination by stabilizing the data set and using k-fold validation.

2. Material and Method

The images in Ham10000 dataset are classified using convolution, centering and full connection layers in the model created with VGGNET architecture. Images in the size of 600x450 were reduced to 400x300 unit sizes and processed. In this application coded with Python programming language, Tensorflow, Keras, Sklearn machine learning libraries were used. PYCharm was preferred as the IDE environment.

2.1. Deep Learning and CNN

The images in Ham10000 dataset are classified using convolution, centering and full connection layers in the model created with VGGNET architecture. Images in the size of 600x450 were reduced to 400x300 unit sizes and processed. In this application coded with Python programming language, Tensorflow, Keras, Sklearn machine learning libraries were used. PYCharm was preferred as the IDE environment.

2.1.1. Definition

Many non-linear layers are utilized for feature extraction from the data set with deep learning. The output produced from one layer is transferred as input to another. [5]. Algorithms can be supervised (such as classification) or uncontrolled (such as pattern analysis). In this application, it is known which class of image of each lesion belongs to in the training stage of artificial neural network model. Therefore, supervised learning was conducted.

Supervised learning knows the input and results of the information in the data set. Performs a learning using the previously defined class labels. [6]. The deep learning approach is based on learning by extracting features from the whole data itself. With vectors and edge information, features are extracted on the image. The most ideal algorithms are used to extract these features through deep learning. [7]. In the application, “Relu” algorithm was used as activation function, “Binary Cross-Entropy” function was used ass loss function and “Adam” function was used as optimizer function in convolution and fully connected artificial neural network stages.

2.1.2. History

Supervised deep-feed multi-layer first learning algorithm was published in 1965 by Ivakhnenko and Lapa [8].

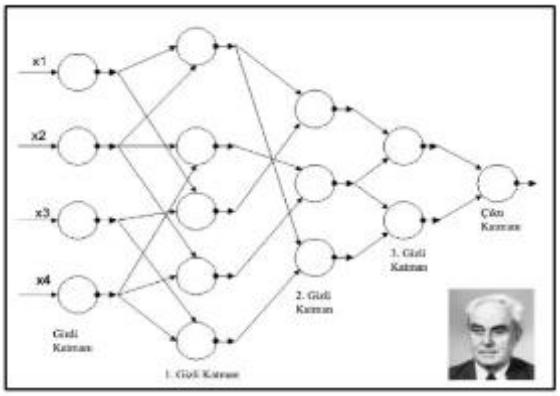


Figure 1. The first known deep network architecture trained by Ivankhnenko A. [9].

The first deep learning architecture “Neocognitron” was proposed by Fukushima in 1979. Fukushima's nets contain multiple bending and pond layers
[10]. Yann LeCun and friends have developed mailbox writings [11]. The network model gives good results, but it is seen as a disadvantage that the time required to train the network is high.[11] Although the network worked successfully, it was found to be unsuitable in practice since the training lasted approximately 3 days.

Yann LeCunn co-opted with curved webs to classify handwriting numbers (MNIST) through the Net LeNet network [12]. In 1995, Brendan Frey, Peter Dayan, and Geoffrey Hinton developed the wake-sleep algorithm, which showed that a fully connected network of hundreds of hidden layers could be trained, even if the training lasted two days. [13]. In 1997, there were some important developments such as long short term memory for recurrent neural networks proposed by Hochreiter and Schmidhuber [14].

For the first time in the context of ANN, the term learning deep learning deep was introduced in 2000 by Igor Aizenberg et al. [15]. In an article published by Geoffrey Hinton in 2006, he demonstrated how a multilayer feedforward neural network can effectively train a layer at each iteration (he trained each layer with an uncontrolled Boltzmann machine without supervision), and then fine-tuned it with a controlled back propagation method. [16].

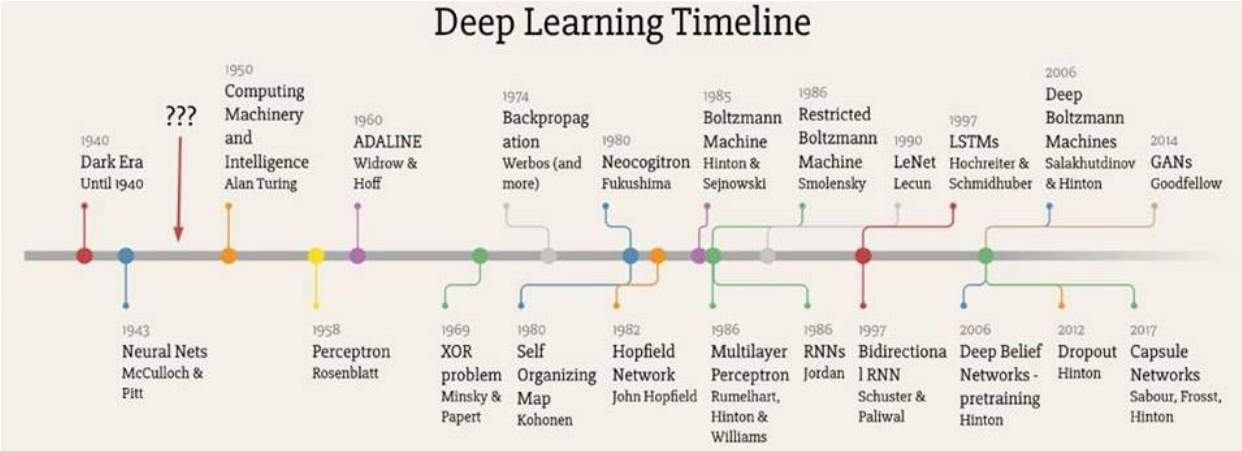


Figure2. Deep Learning History [20]

In 1994, Binder et al successfully used dermatoscopic images to successfully train an artificial neural network to differentiate melanomas from the melanocytic nevi, the most lethal skin cancer. [21]

In 2013, Mendonça et al obtained 200 dermatoscopic images as a PH2 dataset including 160 types and 40 melanomas. [22]

In 2018, the study by Philipp Tschandl, Cliff Rosendahl and Harald Kittler presented the HAM10000 data set as a large collection of multidimensional dermatoscopic images of diffuse pigmented skin lesions. [23]

Also in 2018, Nithin D Reddy trained an evolutionary neural network based on the ResNet50 architecture to accurately classify dermoscopy images of skin lesions into one of seven disease categories. The validation data set achieved a 91% balanced accuracy. [24]

In 2019, Muhammad Attique Khan, Muhammad Younus Javed, Muhammad Sharif, Tanzila Saba, Amjad Rehman proposed an automated system for skin lesion classification based on the extraction of deep neural network (DCNN) properties based on transfer learning and selection of optimal properties based on the kurtosis controlled principle component (KcPCA). The three data sets HAM10000 were used for the ISBI 2017 and ISBI 2016 test results and provided 89.8%, 95.60% and 90.20% accuracy, respectively. [25]

2.1.2. Convolutional Neural Networks (CNN)

A neural network consists of neurons that communicate with each other. Each neuron has weight values and the model will produce accurate results by updating these weights and training the network. As shown in figure 1, each layer processes the information from the previous layer to produce properties. These pattern recognition layers can be between 5 and 25 in a typical CNN. [18].

The first CNN network given in Figure 3 is LeNet architecture, which was introduced by Yann LeCun in 1988 and continued to be improved until 1998. [19]

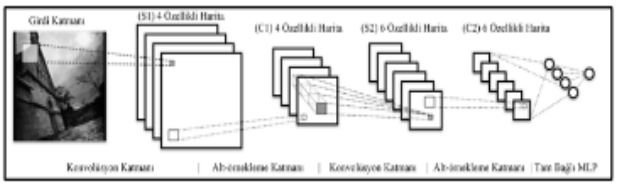


Figure 3. Architecture of the LeNet network.

CNN algorithms are used in many different fields such as image, sound processing, natural language processing (NLP) and biomedical.

A typical convolutional neural network is shown in Figure 4. In this study, the images were first introduced into the convolution and partnering processes consisting of five layers using VGGNET. Three artificial neural network layers were used in the full connection structure. In order not to memorize the model during the training phase, “dropout” method was applied and some nodes in the neural network were forgotten.

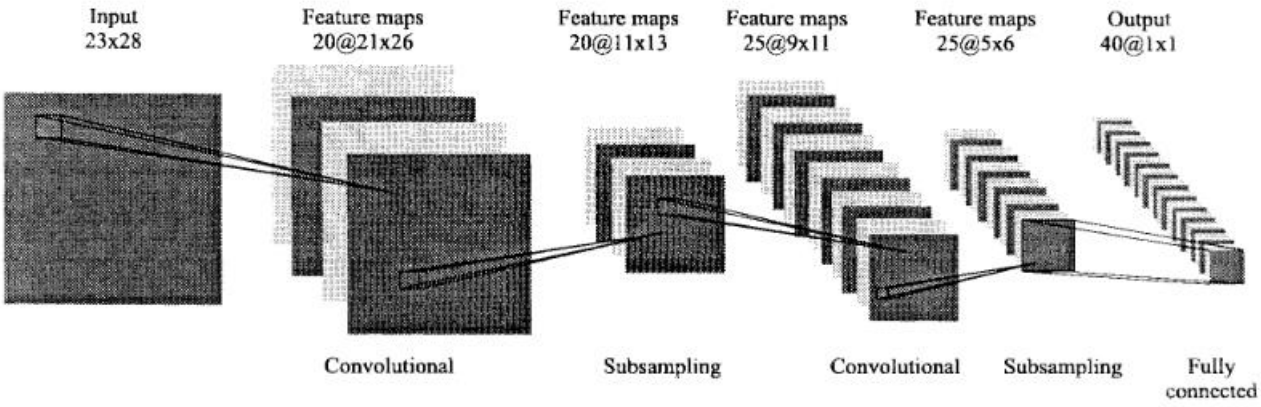


Figure 4. Typical Convolutional Neural Network [17]

2.2. Data Set and Model

The HAM 10000 dataset contains seven classes of skin lesions.

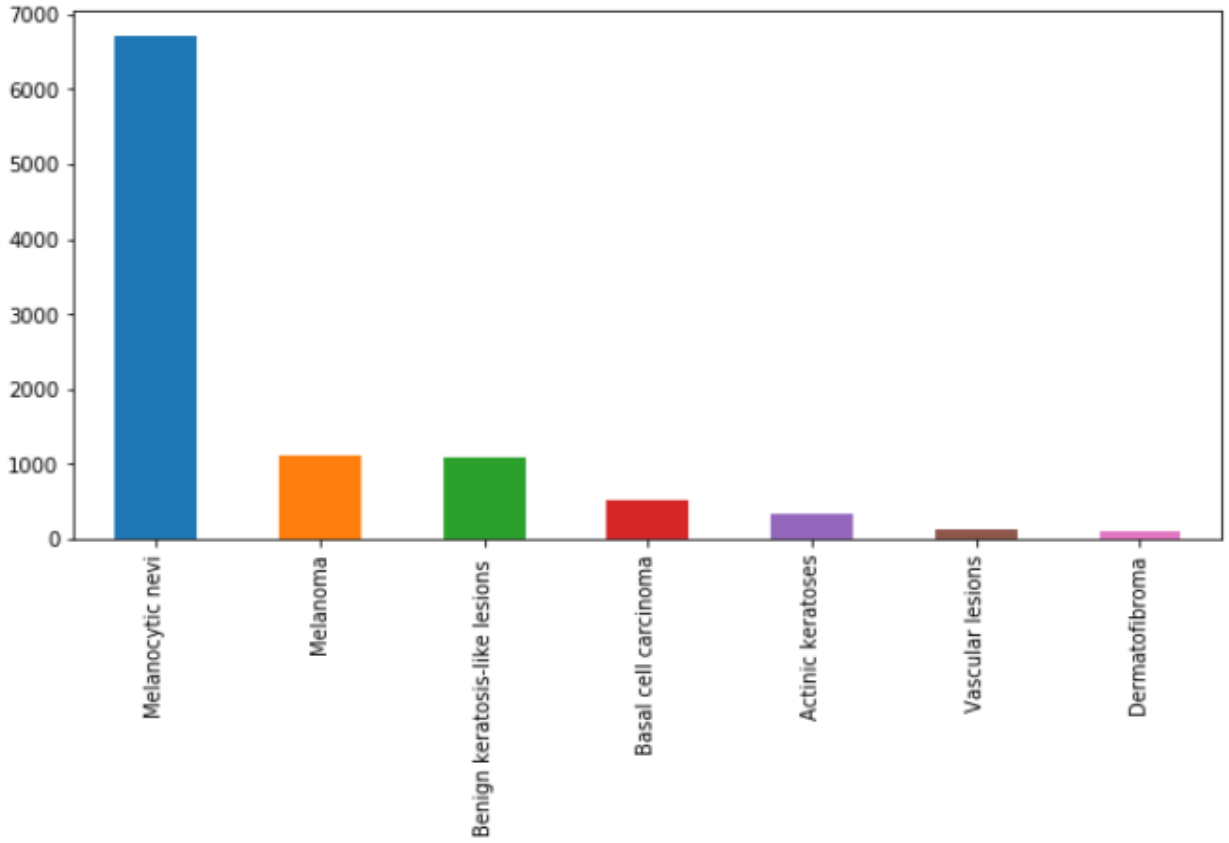


Figure 5. Classes in the HAM 10000 dataset

As can be seen, the excessive number of samples belonging to the "Melanocytic nevi" class in the distribution of classes on the data set disrupts the homogeneous structure of the data set. For this reason, the data of this class has been rearranged (some of it has been discarded) and the distribution of the classes for education has been made acceptable as in Figure 6.

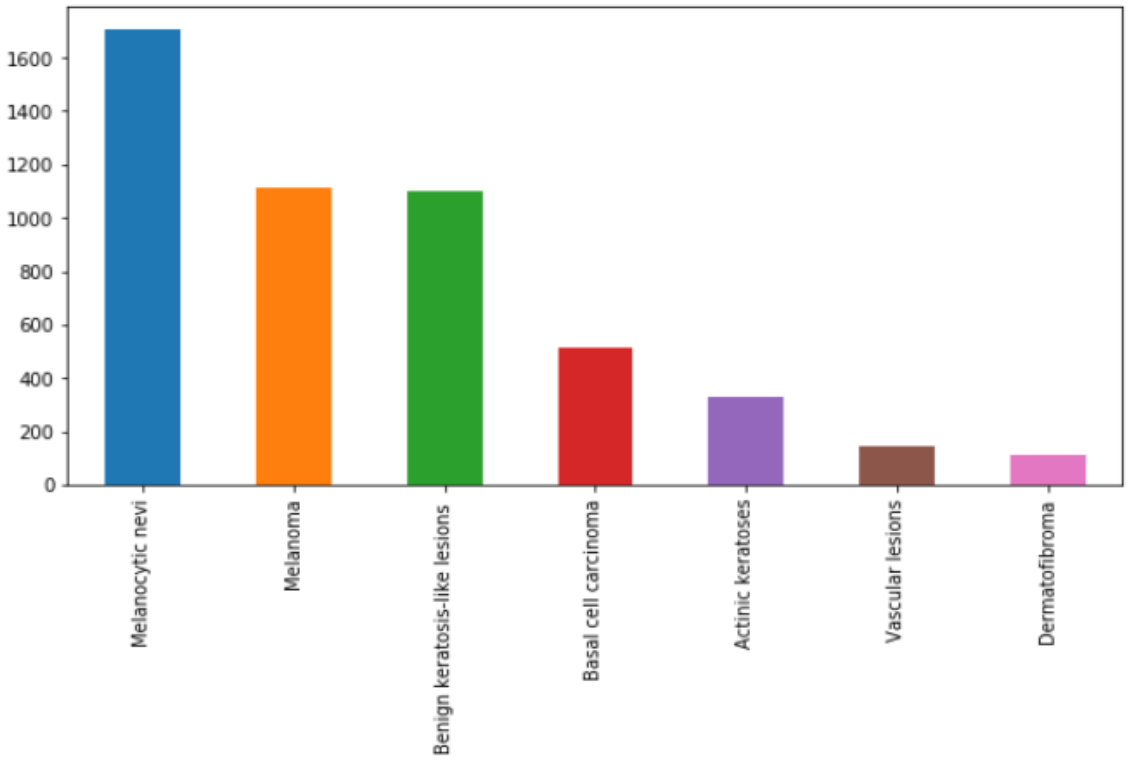


Figure 6. Regulated Raw 10000 Class Distributions

Sample images of these classes are given below.

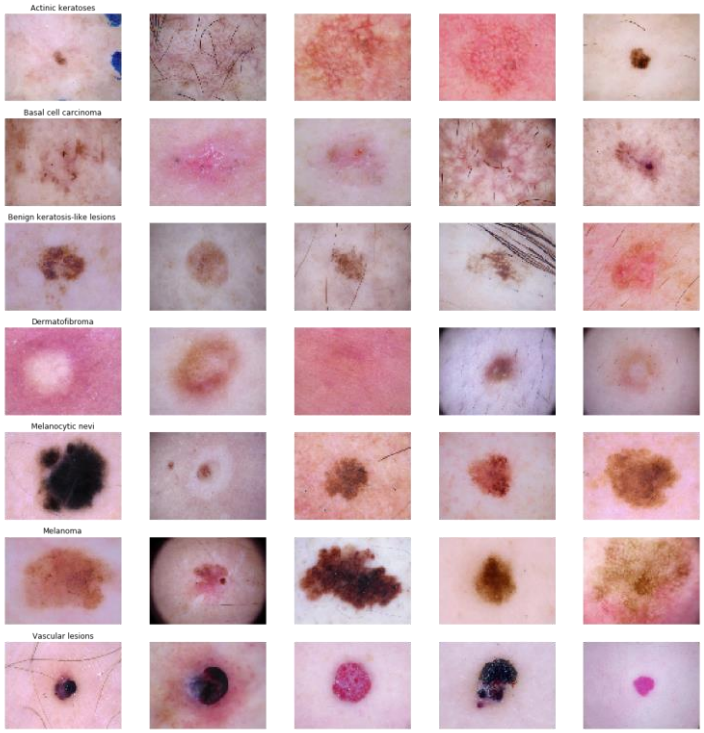


Figure 7. Sample images of classes

Statistical information of the parameters in the data set is shown in Table 1.Table 1 shows the statistical information about the data set. For example, let's examine the lesion_id column.

We see that there are 1015 lesion_id in total. We are aware that some lesion_ids are represented by more than one image, and therefore are 7470 kinds of lesion_ids. We also see that lesion_id of HAM_0001863 is the most repeated and repeated six times. So we have a total of six different images of this name lesion_id.

Table 1. General Statistics of HAM 10000 Data Set

	lesion_id	**image_id**	**dx**	**dx_type**	**sex**	**localization**	**cell_type**
count	10015	10015	10015	10015	10015	10015	10015
unique	7470	10015	7	4	3	15	7
top	HAM_0001863	ISIC_0034282	nv	histo	male	back	Melanocytic nevi
freq	6	1	6705	5340	5406	2192	6705

The VGGNET Architecture used in model building is as follows.

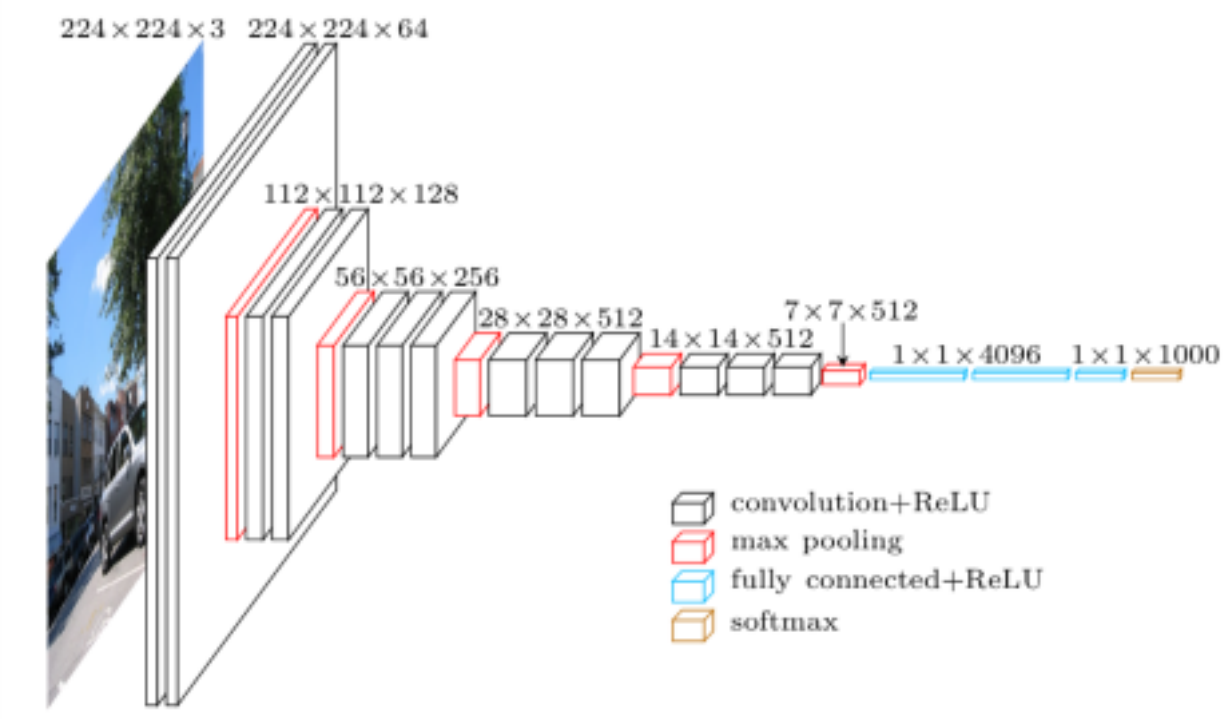


Figure 8. Methods and Neural Network Structure (VGGNET-16)

In practice, K-fold Cross Validation classification model was preferred for separation of data set as train and test. One of the main challenges we face when dealing with data is overfitting and underfitting.In this case, it is necessary to mention variance and bias.

Bias: Our guess is how far away our real value is.

Variance: It shows the distribution of the values in the data set according to the average.
ERROR = Bias^2 +Variance

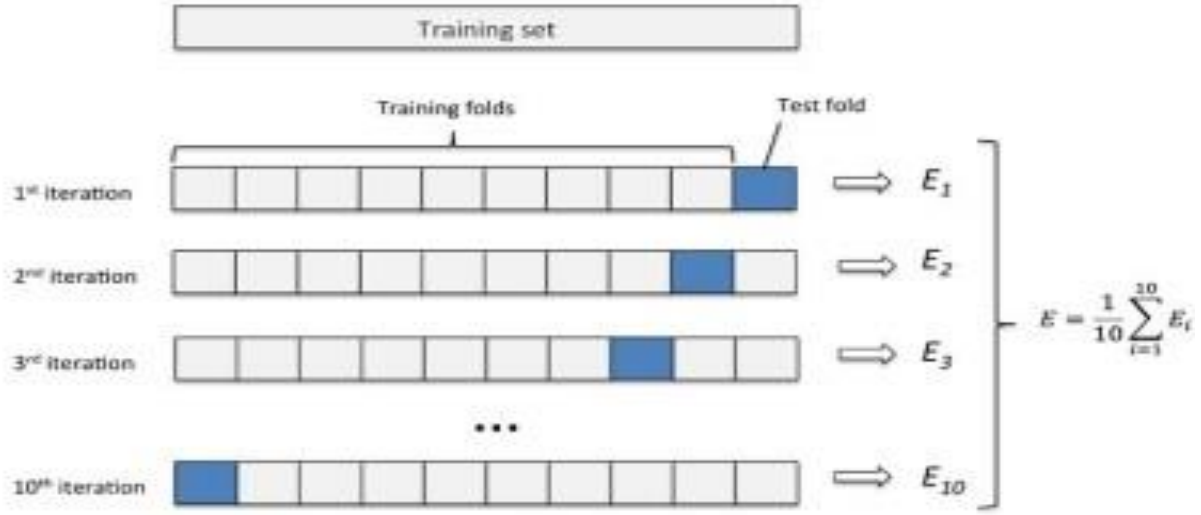


Figure 9. K-Fold Cross Validation

In the visual data set is divided into 10 parts. In each round, the blue painted area is allocated to the test set, while the other parts are reserved for training. When all rounds are over, E's arithmetic mean shows the performance of our model.

The purpose of separating the data set as a training and test set is to avoid possible overfitting and to understand how the model performs on a data set that it has not seen before. However, there may be some errors due to distribution during the training and testing phase of our model. K - Fold Cross Validation technique is used to minimize these errors.

3. Results and Discussion

3.1. Confusion Matrix

As a result, the graphs of the "loss" and "success" functions obtained during the training and verification stages of the model are as follows.

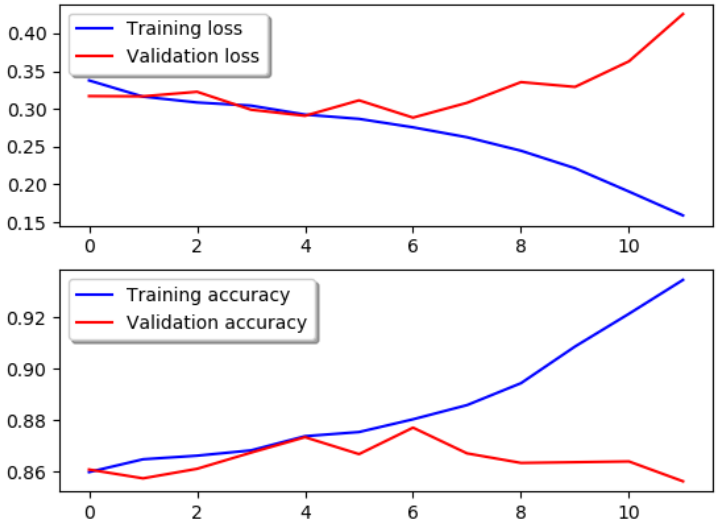


Figure 10. Loss Function and Accuracy Function

The confusion matrix obtained as a result of the application is as follows.

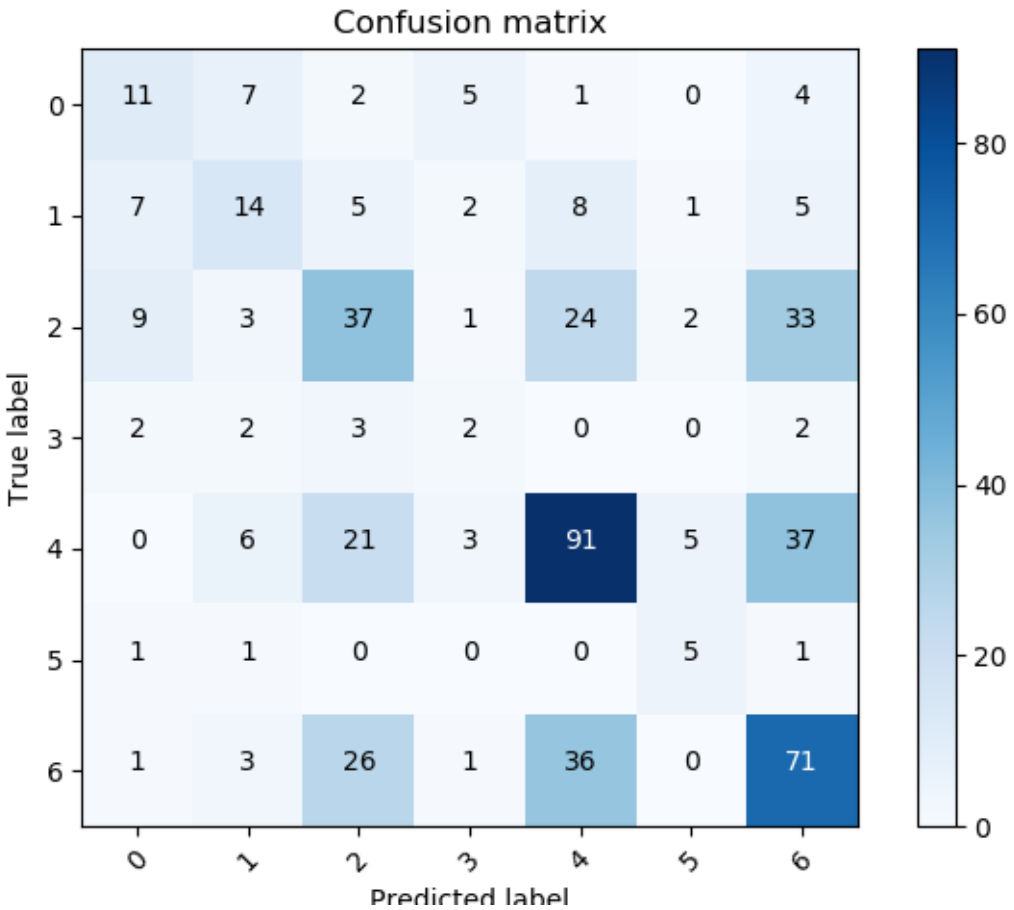


Figure 11. Confusion Matrix of

4. Conclusions and Recommendations

It is possible to increase the success of the application by changing convolution processes, activation functions, loss functions and using more hidden layers in the artificial neural network. Although the incidence of skin cancer diseases has increased in recent years, there has been an increase in the studies conducted for early diagnosis.When we look at previous studies, we see that data sets are limited or class distributions are not balanced. In this sense, we have developed a more reliable model that provides comprehensive discrimination by stabilizing the data set and using k-fold validation. In the test phase of the educated model, the validation of the classes was obtained as 85.62%.

4. Acknowledge

The authors wish to thank the Tokat Gaziosmanpasa University for its hardware support.

References

[1] Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, *61*, 85-117.
[2] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436.
[3] Bottou, L. (2010). Large-scale machine learning with stochastic gradient descent. In *Proceedings of COMPSTAT'2010*(pp. 177-186). Physica-Verlag HD.
[4] Simard, P. Y., Steinkraus, D., & Platt, J. C. (2003, August). Best practices for convolutional neural networks applied to visual document analysis. In *Icdar*(Vol. 3, No. 2003).
[5] Deng, L.,& Yu, D. (2014). Deep learning: methods and applications. *Foundations and Trends® in Signal Processing*, *7*(3–4), 197-387.
[6] Ozgur, A. (2004). Supervised and unsupervised machine learning techniques for text document categorization. *Unpublished Master's Thesis, İstanbul: Boğaziçi University*.
[7] Song, H. A.,& Lee, S. Y. (2013, November). Hierarchical representation using NMF. In *International conference on neural information processing*(pp. 466-473). Springer, Berlin, Heidelberg.
[8] Ivakhnenko, A. G.,& Lapa, V. G. (1966). *Cybernetic predicting devices* (No. TR-EE66-5). PURDUE UNIV LAFAYETTE IND SCHOOL OF ELECTRICAL ENGINEERING.
[9] Dettmers, T. (2015). Deep learning in a nutshell: Core concepts. *NVIDIA Devblogs*.
[10] Fukushima, K. (1980). Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological cybernetics*, *36*(4), 193-202.
[11] LeCun, Y., Boser, B., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W., & Jackel, L. D. (1989). Backpropagation applied to handwritten zip code recognition. *Neural computation*, *1*(4), 541-551.
[12] LeCun, Y., Boser, B. E., Denker, J. S., Henderson, D., Howard, R. E., Hubbard, W. E., & Jackel, L. D. (1990). Handwritten digit recognition with a back-propagation network. In *Advances in neural information processing systems*(pp. 396-404).
[13]Hinton, G. E., Dayan, P., Frey, B. J., & Neal, R. M. (1995). The" wake-sleep" algorithm for unsupervised neural networks. *Science*, *268*(5214), 1158-1161.
[14]Hochreiter, S.,& Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780.
[15]Aizenberg, I. N., Aizenberg, N. N., & Vandewalle, J. (2000). Multiple-Valued Threshold Logic and Multi-Valued Neurons. In *Multi-Valued and Universal Binary Neurons*(pp. 25-80). Springer, Boston, MA.
[16]Hinton, G. E. (2007). Learning multiple layers of representation. *Trends in cognitive sciences*, *11*(10), 428-434.
[17] Lawrence, S., Giles, C. L., Tsoi, A. C., & Back, A. D. (1997). Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks*, *8*(1), 98-113.
[18] Hijazi, S., Kumar, R., & Rowen, C. (2015). Using convolutional neural networks for image recognition. *Cadence Design Systems Inc.: San Jose, CA, USA*.
[19]Le Cun, Y., Jackel, L. D., Boser, B., Denker, J. S., Graf, H. P., Guyon, I., ... & Hubbard, W. (1989). Handwritten digit recognition: Applications of neural network chips and automatic learning. *IEEE Communications Magazine*, *27*(11), 41-46.
[20]F. Vazquez, " https://towardsdatascience.com/a-weird-introduction-to-deep-learning-7828803693b0"
[21] Binder, M., Steiner, A., Schwarz, M., Knollmayer, S., Wolff, K., & Pehamberger, H. (1994). Application of an artificial neural network in epiluminescence microscopy pattern analysis of pigmented skin lesions: a pilot study. *British Journal of Dermatology*, *130*(4), 460-465.
[22] Mendonça, T., Ferreira, P. M., Marques, J. S., Marcal, A. R., & Rozeira, J. (2013, July). PH 2-A dermoscopic image database for research and benchmarking. In *2013 35th annual international conference of the IEEE engineering in medicine and biology society (EMBC)*(pp. 5437-5440). IEEE.
[23] Tschandl, P., Rosendahl, C., & Kittler, H. (2018). The HAM10000 dataset, a large collection of multi-source dermatoscopic images of common pigmented skin lesions. *Scientific data*, *5*, 180161.
[24] Reddy, N. D. (2018). Classification of Dermoscopy Images using Deep Learning. *arXiv preprint arXiv:1808.01607*.
[25] Khan, M. A., Javed, M. Y., Sharif, M., Saba, T., & Rehman, A. (2019, April). Multi-Model Deep Neural Network based Features Extraction and Optimal Selection Approach for Skin Lesion Classification. In *2019 International Conference on Computer and Information Sciences (ICCIS)*(pp. 1-7). IEEE.