Research Article/Araştırma Makalesi

GAZİ
JOURNAL OF ENGINEERING SCIENCES

Performance Evaluation Of Basic Capsule Network Architecture In The Classification Of Biomedical Images

Sümeyra Büşra Şengül*[a], İlker Ali Özkan[b]

Submitted: 06.01.2023 Revised: 24.06.2023 Accepted: 15.07.2023 doi:10.30855/gmbd.0705067

ABSTRACT

Diagnose and treat diseases, a variety of imaging techniques are used, including X-ray, computed tomography (CT), mammography, ultrasound, magnetic resonance imaging (MRI) and diffused optical tomography (DOT). Correct analysis of these medical images is required for early disease detection and application of the appropriate treatment. In image analysis, the identification of the relevant area, as well as information such as its size, location, and direction, are critical in determining the best treatment methods. The convolutional neural network (CNN) architecture is one of the most widely used deep learning architectures in medical image analysis. However, it was stated that CNN was insufficient to measure the relationship between these features while extracting image features, and it could not hide features such as pose (position, direction, size), deformation, and texture. The Basic Capsule Network (CapsNet) Architecture was proposed to overcome CNN's disadvantage and increase success. In this study, MedMNIST dataset collection consisting of medical images was used. The RetinaMNIST, BreastMNIST, and OrganMNIST-A datasets included in MedMNIST were used to evaluate the classification performance of the CapsNet architecture. Capsnet succeeded in these with accuracy rates of 54%, 85.2%, and 89%, respectively. CapsNet has been shown to produce comparable results to advanced CNN models.

Keywords: CapsNet, Res-Net, Convolutional neural networks, Medical image analysis, Dynamic routing algorithm

[a*] Selcuk University, Technology Faculty, Dept. of Computer Engineering 42075 - Konya, Türkiye
Orcid: 0000-0003-1385-0920
e mail:sumeyra.sengul@selcuk.edu.tr

[b] Selcuk University, Technology Faculty, Dept. of Computer Engineering 42075 - Konya, Türkiye
Orcid: 0000-0002-5715-1040

*Corresponding author: sumeyra.sengul@selcuk.edu.tr

Biyomedikal Görüntülerde Temel Kapsül Ağ Mimarisinin Sınıflandırma Performansının İncelenmesi

ÖZ

Hastalıkların teşhis, tanı ve tedavisinde manyetik rezonans görüntüleme (MRI), difüz optik tomografi (DOT), pozitron emisyon tomografisi (PET), bilgisayarlı tomografi (CT), mamografi, ultrason, röntgen gibi çeşitli görüntüleme teknikleri kullanılmaktadır. Hastalığın erken evrelerde tespit edilebilmesi ve doğru tedavinin uygulanabilmesi bu medikal görüntülerin doğru bir şekilde analiz edilmesine bağlıdır. Görüntü analizinde ilgili alanın belirlenmesi ve o alana ait büyüklük, konum yön gibi bilgilerin tespiti doğru tedavi yöntemlerinin belirlenmesi açısından çok önemlidir. Son on yılda oldukça hızlı gelişim gösteren derin öğrenme yöntemleri, görüntüleri önişlem gerekmeden işleyebilmesi, hızlı ve doğru sonuçlar vermesi nedeniyle, görüntülerin sınıflandırması ve segmentasyonunda çokça kullanılmaktadır. Medikal görüntülerin analizinde en yaygın kullanılan derin öğrenme mimarilerinden birisi de evrişimli sinir ağları (ESA) mimarisidir. Ancak ESA'nın görüntü özelliklerini çıkarırken, bu özelliklerin birbirleri arasındaki ilişkiyi ölçmekte yetersiz kaldığı ve özelliklere ait poz (konum, yön, boyut), deformasyon, doku gibi özellikleri saklayamadığı belirtilmiştir. ESA'nın bu dezavantajını gidermek ve başarıyı artırmak için Temel Kapsül Ağ (CapsNet) Mimarisi önerilmiştir. Bu çalışmada medikal görüntülerden oluşan MedMNIST veri seti topluluğu kullanılmıştır. MedMNIST veri topluluğunda bulunan RetinaMNIST, BreastMNIST, OrganMNIST-A veri setleri ile CapsNet mimarisinin sınıflandırma performansı incelenmiştir. CapsNet RetinaMNIST veri setinde %54, BreastMNIST veri setinde %85.2, OrganMNIST-A veri setinde %89 başarı elde etmiştir. CapsNet'in gelişmiş ESA modelleri ile yakın sonuçlar verdiği görülmüştür.

Anahtar Kelimeler: CapsNet, Res-Net, Evrişimli sinir ağları, Medikal görüntü analizi, Dynamik yönlendirme algoritması

To cite this article: S.B. Sengul and İ. A. Ozkan, "Performance Evaluatıon Of Basıc Capsule Network Archıtecture In The Classıfıcatıon Of Bıomedıcal Images" *Gazi Journal of Engineering Sciences,* vol. 9, no. 2, pp. 238-247, 2023. doi:10.30855/gmbd.0705067

1. Introduction

Medical imaging techniques including magnetic resonance imaging (MRI), positron emission tomography (PET), computed tomography (CT), mammography, ultrasound, x-ray, diffused optical tomography (DOT) and digit pathology images are frequently used in the early diagnosis, diagnosis and treatment of various diseases[1, 26-30]. It is necessary to analyze the features of medical images quickly and accurately in order to detect the disease in the early stages, to determine the appropriate treatment, and predict the patient's response to treatment. Deep learning methods are structures that can learn these features from raw data without the need for preprocessing. With the advancement of deep learning techniques, convolutional neural networks (CNN) in medical image analysis have successfully classified data and have been widely used [2]. It also accelerates the treatment process of the patient by detecting small details that may not be noticed in the patient's symptoms [3]. However, CNN is insufficient to measure the relationship between features while extracting the features of the image. In the pooling layer, which is used to reduce the number of parameters, the location information of the features is lost due to pixel loss and the success of the method decreases. It needs more training data to improve the falling success. This leads to an increase in the number of parameters in the network and, accordingly, a prolongation of the training period [4].

In the diagnosis of the disease from medical images, it is very important to determine the relevant area and to detect the information (such as size, location, and direction) of that area. For example; the diagnosis and treatment of bleeding in intracranial hemorrhages depends on the location of the bleeding, one or more occurrences [5]. However, it is a disadvantage in the analysis of medical images that CNN cannot hide and ignore information such as pose (position, direction, size), deformation, and texture of the objects in the image.

In order to overcome these limitations of CNN and improve its success, CapsNet architecture was proposed in 2017 [6]. In CapsNet architecture, vector output capsules are used instead of CNN's scalar output feature maps, and a dynamic routing algorithm is used instead of pooling. CapsNet stores not only the objects in the image, but also the physical feature information of these objects, such as pose (position, direction, size), deformation, speed, color and texture, with capsules consisting of neuron groups [7]. Capsule network has been applied to medical images and successful results have been obtained due to these features and the ability to eliminate the deficiencies of CNN.

Radiologists must have a high level of expertise to examine and analyze brain images. However, some details may not be visible to the human eye. these cases, the image interpretation support of deep learning algorithms for correctly interpreting images can shorten the treatment process [31]. Furthermore, deep learning algorithms have been used successfully in image classification [32]. Afshar et al. [8] classified tumor types on brain MRI images. Due to CapsNet's ability to store information about each object on the image, they only segmented the tumor region to increase success. As a result of the training with the same data set, they obtained an accuracy rate of 72.13% with CNN, and an accuracy rate of 86.56% with CapsNet. Another study comparing CNN architectures and CapsNet is the detection of muscle tears in the shoulder. While they obtained 94.75% accuracy rate with CapsNet in the study, they obtained 93.21%, 88.45%, 87.63% and 85.20% accuracy rates from CNN, AlexNet and GoogLenet, respectively [9]. CapsNet also gave very good results on retinal OCT images. In the study conducted to detect glaucoma disease on OCT images, the images were gone through various preprocessing steps such as contrast enhancement, edge detection, and bilateral filtering. Images were given to CapsNet architecture with 3 convolution layers and 94% success was achieved. ResNet-50 architecture was trained with the same images and 84% success rate was obtained [10]. Kumar et al. [11] also presented a hybrid architecture by using CapsNet architecture with VGG16 architecture. In their study on the detection of diabetic retinopathy from retinal fundus images, an accuracy rate of 74% was obtained with CNN, while an accuracy rate of 96.24% was achieved with hybrid architecture. Madhu et al. [12] also increased the number of convolutional layers in the architecture they developed for the diagnosis of malaria and made changes in the dynamic routing function. Instead of the squash function used as the activation function in the classical dynamic routing function, the ReLU activation function was used. As a result of the study, they obtained an accuracy of around 93%-95% from various CNN architectures (ResNet-50, VGG19, VGG16, InceptionV3, DenseNet121, Xception, MobileNet) and 99.03% accuracy with the CapsNet-ReLU model they developed. Yadav et al.[33] classified pneumonia from radiograph images using capsule nets and achieved 93.80% accuracy. CapsNet has been applied for the diagnosis and classification of various types of cancer and very successful results have been obtained [22-25].

CapsNet has been used for feature extraction as well as classification. In the diagnosis of diseases such as breast cancer and skin cancer, this architecture was used to extract the features of the images after the relevant region was segmented [20,21].

This study was carried out to evaluate the classification performance of CapsNet architecture on benchmark datasets and to compare it with the classification performance of different deep learning models. The basic architecture is used without any innovation in the CapsNet architecture. RetinaMNIST, BreastMNIST and OrganMNIST-A benchmark datasets in the MedMNIST[13] data collection consisting of different medical images were used in the study.

In this study, the CapsNet architecture and dynamic routing algorithm in the 2nd section, the datasets used and the performance criteria examined in the 3rd section, the software, hardware and the way the study was implemented in the 4th section, and finally the results of the study were evaluated in the 5th section.

2. CapsNet Architecture

Sabour et al. have developed CapsNet architecture in 2017 [6]. The Capsule network is deep learning architecture that stores the objects in the image and information about the objects in the image, such as pose (position, direction, size), deformation, speed, color, texture, as well as physical feature information with capsules consisting of neuron groups [7]. It stores the objects in the image and the spatial relationships of the parts that make up the objects in the transformation matrix. Thus, the object can be accurately identified even when the angle and position of the object change [14].

Architecture starts with the convolution layer. In this layer, the basic features of the image such as edges and corners are extracted and switched to the primary capsule layer. The primary capsule layer consists of convolution, reshaping, and squash functions that the capsule network architecture uses as an activation function, unlike other deep learning architectures. With the squash function, images are converted into vectorial outputs with a length between zero and one. Each capsule represents an entity in the image. The length of the output vector of the capsule indicates the probability of that entity being in the image. The next layer is the digit capsule layer. The dynamic routing algorithm that provides matching between the capsule layers is implemented in this layer. And as a result, there are as many vectors as the number of classes.

2.1. Dynamic Routing Algorithm

Dynamic routing algorithm provides matching between capsule layers. The dynamic routing algorithm is shown in Algorithm 1.

Algorithm1. Dynamic Routing Algorithm [6]

procedure ROUTING ($\hat{u}_{j|i}$, *r*, *l*)

for all capsules *i* in layer *l* and capsule *j* in layer (*l*+1): $b_{ij} \leftarrow 0$.

for r iterations **do**

for all capsule *i* in layer *l*: $c_i \leftarrow$ **softmax** (b_i)

for all capsule *j* in layer (*l*+1): $s_j \leftarrow \sum_i c_{ij}\hat{u}_{j|i}$

for all capsule *j* in layer (*l*+1): $v_i \leftarrow$ **squash** (s_j)

for all capsule *i* in layer *l* and capsule *j* in layer (*l*+1): $b_{ij} \leftarrow b_{ij} + \hat{u}_{j|i}v_{ij}$

return v_j

In the algorithm, b_{ij} is the temporary similarity variable with an initial value of 0. Its value is updated in iteration. c_{ij} matching coefficients are calculated by passing b_{ij} through the softmax function

(Equation (1)).

$$c_{ij} = \frac{\exp(b_{ij})}{\sum_k \exp(b_{ik})} \quad (1)$$

$$\hat{u}_{j|i} = W_{ij} u_i \quad (2)$$

$$s_j = \sum_i c_{ij} \hat{u}_{j|i} \quad (3)$$

As shown in Equation (2), the input values of the capsule u_i are multiplied by the weight matrix W_{ij} to get the estimated output vector $\hat{u}_{j|i}$ of the capsule. The product of the prediction vector and the matching coefficient is summed up and the weighted sum s_j is found (Equation (3)). The value of s_j is passed through the squash function and the output value of the capsule, v_j , is calculated. The equation of the squash function is shown in Equation (4). The squash function brings short vectors closer to 0 and long vectors closer to 1.

$$v_j = \frac{\|s_j\|^2}{1+\|s_j\|^2} \cdot \frac{s_j}{\|s_j\|} \quad (4)$$

In the last step of the algorithm, b_{ij} is updated with the formula shown in Equation (5) and all steps are repeated from step 4 as many as *r*.

$$b_{ij} = b_{ij} + \hat{u}_{j|i} v_{ij} \quad (5)$$

2.2. Loss Function (Margin Loss)

In the last layer, the digit capsule layer, the margin loss is calculated for each capsule. With this calculation, it is determined whether the object is in the image or not.

$$L_k = T_k \max(0, m^+ - \|v_k\|)^2 + \lambda(1 - T_k)\max(0, \|v_k\| - m^-)^2 \quad (6)$$

The formula shown in Equation (6) is used to calculate the margin loss. If the prediction is correct, T_k=1 is taken, otherwise T_k=0. The lambda value is taken as 0.5. m^+ is 0.9 and m^- is 0.1. The total loss is found by adding the losses of the capsules.

2.3. Network Architecture

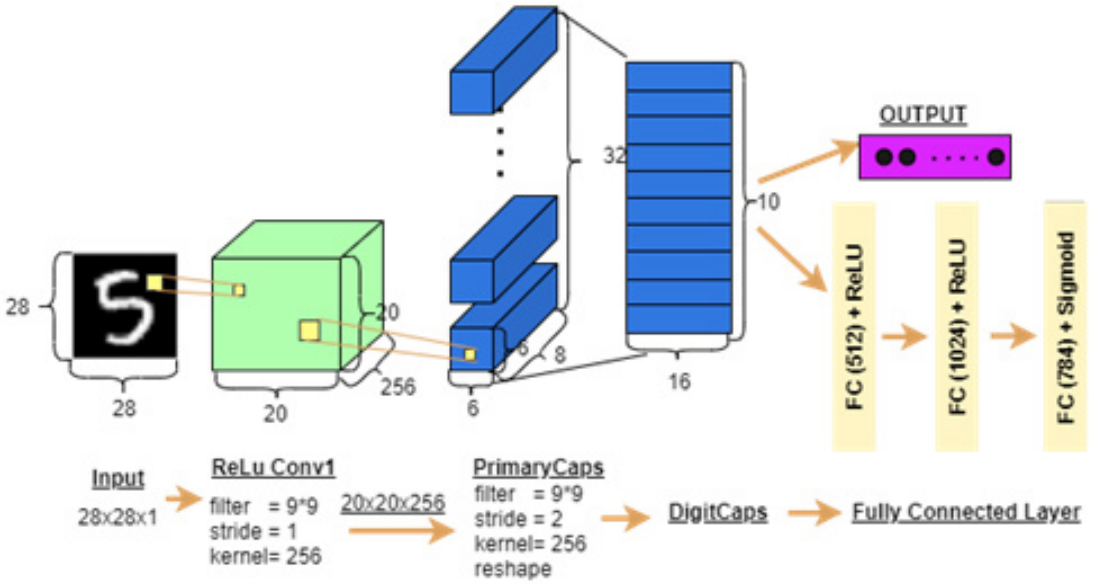

Figure 1. CapsNet Architecture

CapsNet Architecture consists of two parts, encoder and decoder. The decoder part reconstructs the image, while the encoder extracts the features and classifies the image.

Figure 1 shows the CapsNet architecture. In architecture, the first layer is the convolution layer, from which the basic features of the image such as edges and corners are extracted. It takes 28*28 single channel images as input. After convolution with 256 core size, 9*9 filter and 1 step number, 20*20*256 sized output tensor is obtained. The second layer is the primary capsule layer. The convolution operation is applied to the input tensor with the size of 20*20*256 as 256 kernel size, 9*9 filter and number of steps 2. From the obtained 6*6*256 dimensional tensor, 32 8-dimensional capsules (6*6*8*32) are encapsulated by the reshaping process. The squash function is used as the activation function at the layer output. The last layer is the digit capsule layer. This layer is the one where the dynamic routing function is used. As a result of the dynamic routing algorithm from the (6*6*8*32) input tensor, 10 capsules, each of which has 16 dimensions, are formed. The capsule with the highest probability value out of the capsules gives the output. After this layer, the loss calculation is made and the image is reconstructed with fully connected layer [6].

3. Experimental Setup

3.1. MedMNIST Dataset

MedMNIST is a 28*28 dimensional collection of medical open datasets [13]. It consists of 10 2-dimensional (28*28) data sets. All datasets are divided into train, validation and test data.

The MedMNIST data collection consists of various images such as X-ray, OCT, ultrasound, CT. All datasets included in MedMNIST are open-to-use datasets under a Creative Commons (CC) license.

In this study, 3 data sets in MedMNIST were used. The datasets used are RetinaMNIST, BreastMNIST, OrganMNIST-A. In Table 1, train, test and validation samples and rates of the data sets are given.

Tablo 2. The sample of train, test and validation data; Train, test and validation ratio

Datasets	Train Data	Test Data	Validation Data	Train Rate	Test Rate	Validation Rate
RetinaMNIST	34581	17778	6491	60%	30%	10%
BreastMNIST	1080	400	120	67.5%	25%	7.5%
OrganMNIST-A	546	156	78	70%	20%	10%

3.1.1. RetinaMNIST

RetinaMNIST consists of diabetic retinopathy fundus images from the DeepDR dataset. It consists of 5 classes as no apparent, mild non-proliferative, moderate non-proliferative, severe non-proliferative and proliferative diabetic retinopathy [15]. Samples from the data set are presented in Figure 2.

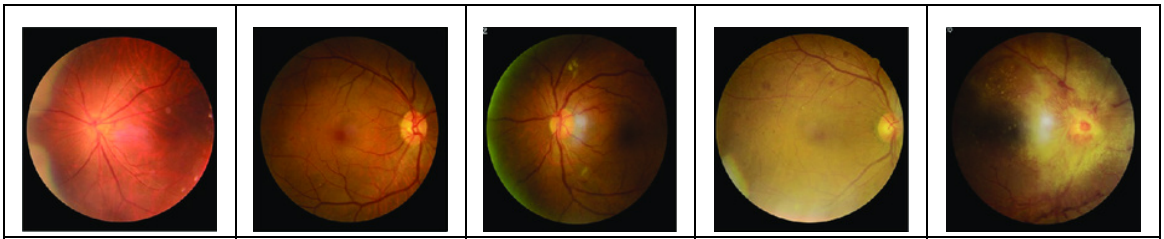

Figure 2. RetinaMNIST image samples [15]

3.1.1. BreastMNIST

BreastMNIST consists of 780 breast ultrasound images from the Bahaya Hospital Breast Ultrasound dataset. It is divided into 3 classes as normal, benign and malignant [16]. For classification, normal and benign images were classified as positive and malignant images as negative. Samples of normal, benign

and malignant cysts are shown in Figure 3.

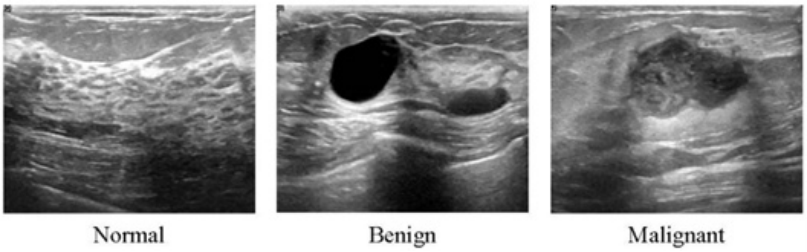

Figure 3. BreastMNIST image samples [16]

3.1.1. OrganMNIST-A

OrganMNIST- (Axial) consists of 131 abdominal CT images of the liver taken from the LiTS dataset. 11 organs are labeled in the axial plane [17]. An image sample in the axial plane is given in Figure 4.

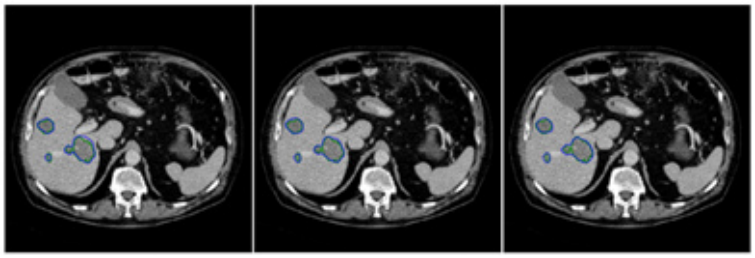

Figure 4. OrganMNIST-A image samples [17]

3.2. Performance Metrics

Accuracy criteria are used to evaluate the performance of the CapsNet architecture and compare it with known models.

Accuracy, precision, sensitivity and f1-score values were calculated to scale and evaluate the performance of the architecture on medical images. Since the datasets used are multi-class, macro-precision, macro-sensitivity and macro-f1 score formulas were taken into account while calculating the metric (2). The formulas are detailed in Table 2.

Tablo 2. Formulas of Performance Metrics [18]

Metric	Formula
Accuracy	$\frac{\Sigma_{i=1}^{l} = \frac{tp_i + tn_i}{tp_i + fn_i + fp_i + tn_i}}{l}$
Precision ($\boldsymbol{P_M}$)	$\frac{\Sigma_{i=1}^{l} = \frac{tp_i}{tp_i + fp_i}}{l}$
Recall ($\boldsymbol{R_M}$)	$\frac{\Sigma_{i=1}^{1} = \frac{tp_i}{tp_i + fn_i}}{l}$
f1-score ($\boldsymbol{F_M}$)	$\frac{(\beta^2 + 1)P_M R_M}{\beta^2 P_M + R_M}$

4. Results and Discussions

This study was carried out on a device with Python 3.8, Tensorflow2.3, CUDA10.1.243, cuDNN7.6.5 technologies, Intel Xeon Gold 6226R processor, Nvidia Grid RTX8000-12Q graphics card. In the creation of the CapsNet model, the study at the address "https://github.com/XifengGuo/CapsNet-Keras" was used [19].

In the study, BreastMNIST, RetinaMNIST and OrganMNIST-A datasets were used to evaluate the classification performance of the CapsNet architecture. The datasets in the MedMNIST data collection are given as 28*28 sized, and they are also divided into train, test and validation. Therefore, no sizing

or parsing was done in the datasets in the study. The datasets were used as is. Retina-MNIST and OrganMNIST-A datasets are multi-class, BreastMNIST dataset is a binary class.

In this study, different deep learning models were used to compare the performance of CapsNet Architecture. These models are Resnet-18, Resnet-50, Auto-sklearn, AutoKeras, Google AutoML Vision [13]. The accuracy values of these models and the accuracy value we obtained from the CapsNet architecture in our study are presented in Table 3. According to Table 3, CapsNet achieved an accuracy rate of 54% on retinal fundus images, 85.2% on breast ultrasound images, and 89% on Organ CT axial plane images, respectively. In addition, the parameter values (number of periods, batch size and learning rates) used in each data set to obtain the best results are given in Table 4. Table 5 shows the performance metrics of the CapsNet model on the RetinaMNIST, BreastMNIST and OrganMNIST-A datasets. The complexity matrices of the data sets used in Figure5, Figure6 and Figure7 are presented.

Tablo 3. Performance Metrics of CapsNet on Different Datasets

Methods	RetinaMNIST	BreastMNIST	OrganMNIST-A
Resnet-18	0.515	0.859	0.921
Resnet-50	0.490	0.853	0.916
Auto-sklearn	0.525	0.808	0.563
Auto-Keras	0.420	0.915	0.929
Google AutoML Vision	0.530	0.865	0.818
CapsNet	0.540	0.852	0.890

Tablo 4. Parameter Values

Datasets	The Numbers of Epoch	Batch Size	Learning Rate
RetinaMNIST	50	100	0.0008
BreastMNIST	50	8	0.0007
OrganMNIST-A	25	256	0.001

Tablo 5. Performance Metrics of CapsNet on Different Datasets

Methods	RetinaMNIST	BreastMNIST	OrganMNIST-A
Accuracy	0.540	0.852	0.890
Precision ($\boldsymbol{P_M}$)	0.340	0.895	0.901
Recall ($\boldsymbol{R_M}$)	0.376	0.735	0.882
f1-score ($\boldsymbol{F_M}$)	0.340	0.77	0.892

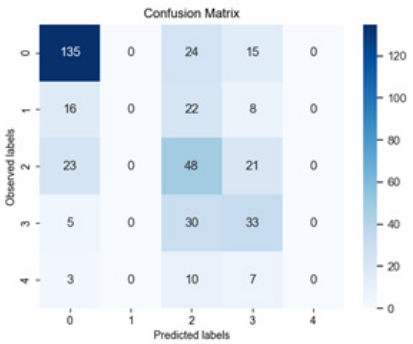

Figure 5. RetinaMNIST Confusion Matrix

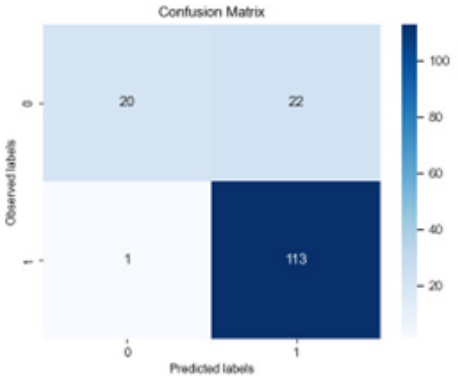

Figure 6. BreastMNIST Confusion Matrix

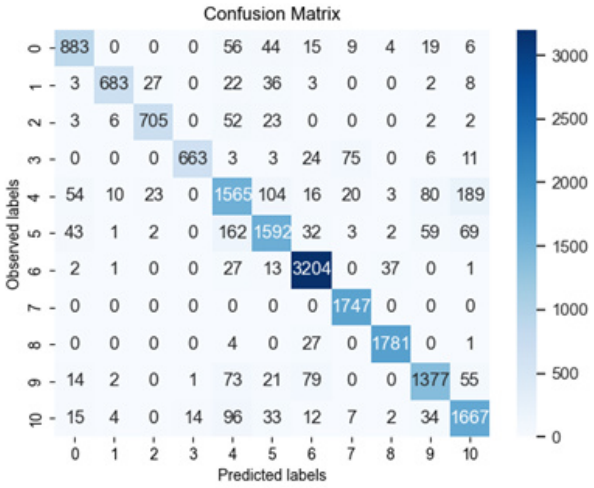

Figure 7. OrganMNIST-A Confusion Matrix

5. Conclusions

Although CNN models achieve very high success in image analysis, since they cannot store information such as the pose (position, direction, size) of the objects on the image, they cannot perform accurate analysis when they encounter images of the object at different angles. CapsNet is an architecture developed to overcome these shortcomings of CNN. In this study, the performance of CNN, which has achieved high success in the analysis of medical images, and CapsNet were compared.

In the literature, it has been seen that CapsNet architecture gets close results with CNN models in image classification. Examining the training results, the basic CapsNet architecture yielded very close results with 18-layer ResNet, 50-layer ResNet, and advanced multi-layered architectures such as Auto-sklearn, AutoKeras, Google AutoML Vision, which try to find the best model according to the data. While results close to these models were obtained in the BreastMNIST and OrganMNIST-A datasets, a better accuracy rate was obtained than these models in the RetinaMNIST dataset.

As a result of the literature review and training results, it is thought that higher success can be achieved with differences such as increasing the number of convolution layers, making innovations in the dynamic routing function, and using it with different architectures of the CapsNet architecture, which has fewer layers and a simpler structure than advanced architectures.

Acknowledgements

We would like to thank the Scientific Research Coordinatorship of Selcuk University for their support with the project titled “Data Intensive and Computer Vision Research Laboratory Infrastructure Project” numbered 20301027.

Conflict of Interest Statement

The authors declare that there is no conflict of interest.

References

[1] K. Raza and N. K. Singh, "A Tour of Unsupervised Deep Learning for Medical Image Analysis," *Current Medical Imaging Reviews*, vol. 17, no. 9, pp. 1059–1077, Sep. 2021. doi:10.2174/1573405617666210127154257

[2] S. M. Anwar, M. Majid, A. Qayyum, M. Awais, M. Alnowami, and M. K. Khan, "Medical Image Analysis using Convolutional Neural Networks: A Review," *Journal of Medical Systems*, vol. 42, no. 11, p. 226, Oct. 2018. doi:10.1007/s10916-018-1088-1

[3] M. M. Qanbar and S. Taşdemir, "Detection of Malaria Diseases with Residual Attention Network," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 7, no. 4, pp. 238–244, Dec. 2019. doi:10.18201/ijisae.2019457677

[4] M. Kwabena Patrick, A. Felix Adekoya, A. Abra Mighty, and B. Y. Edward, "Capsule Networks – A survey," *Journal of King Saud University - Computer and Information Sciences*, vol. 34, no. 1, pp. 1295–1310, Jan. 2022. doi: 10.1016/j.jksuci.2019.09.014

[5] D. Kirbaş, "İntraserebral Kanamanın Tıbbi Tedavisi," *Dusunen Adam The Journal of Psychiatry and Neurological Sciences*, vol. 7, no. 4, p. 54, 1994.

[6] S. Sabour, N. Frosst, and G. Hinton, "Dynamic Routing Between Capsules", *in 31st Conference on Neural Information Processing Systems (NIPS 2017), 4-9 Dec 2017*, [Online]. Available: https://proceedings.neurips.cc. [Accessed: 13 Agust 2023].

[7] X. Zhang and S.-G. Zhao, "Fluorescence microscopy image classification of 2D HeLa cells based on the CapsNet neural network," *Medical & Biological Engineering & Computing*, vol. 57, no. 6, pp. 1187–1198, Jun. 2019. doi:10.1007/s11517-018-01946-z

[8] P. Afshar, A. Mohammadi, and K. N. Plataniotis, "Brain Tumor Type Classification Via Capsule Networks," *in 2018 25th IEEE International Conference on Image Processing (ICIP), Oct. 2018, Athens, Greece.* [Online]. Available: https://ieeexplore.ieee.org/document/8451379. [Accessed: Aug. 13, 2023].

[9] A. Sezer and H.B. Sezer, "Capsule Network-Based Classification Of Rotator Cuff Pathologies From MRI," *Computers & Electrical Engineering,* vol. 80, pp. 106480, Dec. 2019. doi: 10.1016/j.compeleceng.2019.106480

[10] J. Gaddipati, A. Desai, J. Sivaswamy and K. A. Vermeer, "Glaucoma Assessment From OCT İmages Using Capsule Network", *in 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Jul. 2019, Berlin, Germany*. [Online]. Available: https://ieeexplore.ieee.org/document/8857493. [Accessed: Aug. 13, 2023].

[11] G. Kumar, S. Chatterjee, and C. Chattopadhyay, "DRISTI: A Hybrid Deep Neural Network For Diabetic Retinopathy Diagnosis," *Signal Image Video Process*, vol. 15, no. 8, pp. 1679–1686, Apr. 2021. doi: 10.1007/s11760-021-01904-7

[12] G. Madhu et al., "Imperative Dynamic Routing Between Capsules Network for Malaria Classification," *Computers, Materials & Continua*, vol. 68 no.1, pp. 903-919, Mar. 2021. doi:10.32604/cmc.2021.016114

[13] J. Yang, R. Shi and B. Ni, "MedMNIST Classification Decathlon: A Lightweight AutoML Benchmark for Medical Image Analysis," *in 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), Apr. 2021, Nice, France*. [Online]. Available: https://arxiv.org/abs/2010.14925. [Accessed: Aug. 13, 2023].

[14] G. E. Hinton, S. Sabour, and N. Frosst, "Matrix capsules with EM routing," *International Conference on Learning Representations*, 30 April- 3 May. 2018. [Online]. Available: https://openreview.net/ [Accessed: Aug. 13, 2023].

[15] R. Liu et al., "DeepDRiD: Diabetic Retinopathy—Grading and Image Quality Estimation Challenge," *Patterns*, vol. 3, no. 6, p. 100512, Jun. 2022. doi:10.1016/j.patter.2022.100512

[16] W. Al-Dhabyani, M. Gomaa, H. Khaled, and A. Fahmy, "Dataset Of Breast Ultrasound İmages," *Data in Brief,* vol. 28, pp. 104863, Feb. 2020. doi:10.1016/j.dib.2019.104863

[17] P. Bilic et al., "The Liver Tumor Segmentation Benchmark (Lits)," *Medical Image Analysis*, vol. 84, pp. 102680, Feb. 2023. doi:10.1016/j.media.2022.102680

[18] M. Sokolova and G. Lapalme, "A Systematic Analysis Of Performance Measures For Classification Tasks," *Information Processing and Management*, vol. 45, no. 4, pp. 427–437, July 2009. doi: 10.1016/j.ipm.2009.03.002

[19] X. Guo, "CapsNet-Keras." Aug. 10, 2023. Accessed: Aug. 13, 2023. [Online]. Available: https://github.com/XifengGuo/CapsNet-Keras

[20] D. Adla, G. V. R. Reddy, P. Nayak, G. Karuna, "Deep Learning-Based Computer Aided Diagnosis Model For Skin Cancer Detection And Classification," *Distrib and Parallel Databases*, vol. 40, no. 4, pp. 717–736, Dec. 2022. doi: 10.1007/s10619-021-07360-z

[21] T. Kavitha et. al., "Deep Learning Based Capsule Neural Network Model for Breast Cancer Diagnosis Using Mammogram

Images," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 14, no. 1, pp. 113–129, Mar. 2022. doi:10.1007/s12539-021-00467-y

[22] R. LaLonde, P. Kandel, C. Spampinato, M. B. Wallace and U. Bagci, "Diagnosing Colorectal Polyps in the Wild with Capsule Network", *in 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Apr. 2020, Iowa, USA*. [Online]. Available: https://arxiv.org/abs/2001.03305. [Accessed: Aug. 13, 2023].

[23] S. C. Satapathy, M. Cruz, A. Namburu, S. Chakkaravarthy, and M. Pittendreigh, "Skin Cancer classification using Convolutional Capsule Network (CapsNet)," *Journal of Scientific & Industrial Research*, vol. 79, no. 11, Art. no. 11, Apr. 2020. doi:10.56042/jsir.v79i11.35913

[24] M. Ali and R. Ali, "Multi-Input Dual-Stream Capsule Network for Improved Lung and Colon Cancer Classification," *Diagnostics,* vol. 11, no. 8, pp.1485, Aug. 2021. doi: 10.3390/diagnostics11081485

[25] C. Peng, Y. Zheng, and D.-S. Huang, "Capsule Network Based Modeling of Multi-omics Data for Discovery of Breast Cancer-Related Genes," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 17, no. 5, pp. 1605–1612, Sep. 2020. doi:10.1109/TCBB.2019.2909905

[26] F. Xue and J. Jiang, "Dynamic Enhanced Magnetic Resonance Imaging versus Ultrasonic Diffused Optical Tomography in Early Diagnosis of Breast Cancer," *Journal of Healthcare Engineering*, vol. 2022, p. e4834594, Apr. 2022, doi: 10.1155/2022/4834594

[27] Y. A. Üncü, G. Sevim, and M. Canpolat, "Approaches to preclinical studies with heterogeneous breast phantom using reconstruction and three-dimensional image processing algorithms for diffuse optical imaging," *International Journal of Imaging Systems and Technology*, vol. 32, no. 1, pp. 343–353, 2022. doi: 10.1002/ima.22648

[28] Y. A. Üncü, G. Sevim, T. Mercan, V. Vural, E. Durmaz, and M. Canpolat, "Differentiation of tumoral and non-tumoral breast lesions using back reflection diffuse optical tomography: A pilot clinical study," *International Journal of Imaging Systems and Technology*, vol. 31, no. 4, pp. 2023–2031, 2021. doi: 10.1002/ima.22578

[29] K. M. S. Uddin, M. Zhang, M. Anastasio, and Q. Zhu, "Optimal breast cancer diagnostic strategy using combined ultrasound and diffuse optical tomography," *Biomed. Opt. Express, BOE*, vol. 11, no. 5, pp. 2722–2737, May 2020. doi: 10.1364/BOE.389275

[30] E. Y. Chae et al., "Development of digital breast tomosynthesis and diffuse optical tomography fusion imaging for breast cancer detection," *Scientific Reports*, vol. 10, no. 1, Art. no. 1, Aug. 2020. doi:10.1038/s41598-020-70103-0

[31] M. Yeo et al., "Review of deep learning algorithms for the automatic detection of intracranial hemorrhages on computed tomography head imaging," *Journal of NeuroInterventional Surgery*, vol. 13, no. 4, pp. 369–378, Apr. 2021. doi: 10.1136/neurintsurg-2020-017099

[32] T. Danişman et al., "Predıctıng The Locatıon Of The Uterıne Cervıcal Os From 2d Images Wıth Cnn" *JESD*, vol. 8, no. 5, Art. no. 5, Dec. 2020. doi: 10.21923/jesd.828457

[33] S. S. Yadav and S. M. Jadhav, "Deep convolutional neural network based medical image classification for disease diagnosis," *J Big Data*, vol. 6, no. 1, p. 113, Dec. 2019, doi: 10.1186/s40537-019-0276-2.