DOI: 10.37990/medr.1093272 Med Records 2022;4(3):297-303

MEDICAL RECORDS-International Medical Journal

Research Article

Investigation of Usability of Artificial Intelligence Semantic Video Processing Methods in Medicine

Yapay Zekâya Dayalı Anlamsal Video İşleme Yöntemlerinin Tıpta Kullanılabilirliğinin Araştırılması

Zeynep Kucukakcali, Hasan Ucuzal, Emek Guldogan

Inonu University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, Malatya, Turkey

Abstract

Aim: The goal of this study is to produce user-friendly software for healthcare professionals with various approaches such as detection, identification, classification, and tracking of polyps contained in endoscopic images utilizing appropriate video/image processing techniques and CNN architecture.

Material and Method: There were 345 photos in total in the study. These photographs are images depicting anatomical milestones, clinical findings, or gastrointestinal procedures in the digestive tract that have been documented and validated by medical specialists (skilled endoscopists). Each class has hundreds of images. The photos were downloaded from https://datasets.simula.no/kvasir, which is a free source for educational and research purposes. In the modeling phase, CNN and the Max-Margin object detection technique (MMOD), one of the deep neural network designs in the Dlib package, were employed. The data set was separated as 80% training and 20% test dataset using the simple cross-validation method (hold-out). Precision, recall, F1-score, average precision (AP), mean average precision (mAP), ideal localization recall precision (oLRP), mean optimal LRP (moLRP), and intersection over union (IoU) were used to evaluate model performance.

Results: When the previously described steps were performed on the open-access video image dataset of endoscopic polyps in the current study, all performance metrics examined in the training dataset received a value of 1, whereas, in the test dataset precision, sensitivity, F1-score, AP, mAP, oLRP, and moLRP were 98%, 90%, 94%, 89%, 89%, 48%, and 48% respectively.

Conclusion: The proposed approach was found to make accurate predictions in the diagnosis of gastrointestinal polyps based on the values of the calculated performance criteria.

Keywords: Object recognition, deep learning, decision support system, gastrointestinal polyps, convolutional neural networks

Öz

Giriş: Bu çalışmada endoskopik görüntülerde yer alan poliplerin tespiti, tanımlanması, sınıflandırılması ve takibi için uygun video/görüntü işleme teknikleri ve CNN mimarisi kullanılarak sağlık profesyonelleri için kullanıcı dostu bir yazılımın geliştirilerek sunulması amaçlanmıştır.

Material ve Methot: Çalışmada yer alan veri seti 345 görüntü içermekte olup görüntüler anatomik olarak bilinen dönüm noktaları, patolojik bulgular veya sindirim sistemindeki gastrointestinal prosedürler gibi her sınıf için yüzlerce görüntüden oluşmakta ve çeşitli tıp doktorları (deneyimli endoskopistler) tarafından açıklanmış ve doğrulanmıştır. Görseller araştırmalarda ve eğitimlerde kullanılmak amacıyla açık kaynak olan https://datasets.simula.no/kvasir adresinden alınmıştır. Modelleme esnasında Dlib kütüphanesinde yer alan derin sinir ağı mimarilerinden olan CNN ve Max-Margin nesne algılama yöntemi (MMOD) kullanılarak modellemeler yapılmıştır. Veri seti basit çapraz geçerlilik yöntemi (hold-out) kullanılarak %80'i eğitim, %20'si test veri seti olacak şekilde ayrıştırılmıştır. Model performansının değerlendirilmesinde ise kesinlik, duyarlılık, F1-skor, ortalama kesinlik (average precision, AP), ortalama kesinlik değerlerinin ortalaması (mean average precision, mAP), kesiştirilmiş bölgeler ölçütleri (intersection over union, IoU), en uygun konumlandırma kesinliği ve duyarlılığı (optimal localization recall precision, oLRP), ortalama en uygun LRP (Mean Optimal LRP, moLRP) kullanılmıştır.

Bulgular: Mevcut çalışmada endoskopik poliplerin açık erişimli video görüntü veri kümesi üzerinde daha önce açıklanan adımlar gerçekleştirildiğinde, eğitim veri kümesinde incelenen tüm performans metrikleri 1 değerini alırken, test veri kümesinde kesinlik, duyarlılık, F1-skoru , AP, mAP, oLRP ve moLRP sırasıyla %98, %90, %94, %89, %89, %48 ve %48 idi.

Sonuç: Çalışmada sonucunda elde edilen performans metriklerine ait değerler dikkate alındığında, önerilen sistemin gastrointestinal poliplerin tanısında başarılı tahmin sonuçları verdiği belirlenmiştir.

Anahtar Kelimeler: Nesne tanıma, derin öğrenme, karar destek sistemi, gastrointestinal polipler, evrişimsel sinir ağları

Received: 25.03.2022 **Accepted:** 07.05.2022

Corresponding Author: Zeynep Kucukakcali, Inonu University, Faculty of Medicine, Department of Biostatistics and Medical Informatics, Malatya, Turkey, **E-mail**: zeynep.tunc@inonu.edu.tr

INTRODUCTION

Clinical decision support systems provide support to physicians in diagnosing diseases and providing appropriate treatments with computer software, using the patient's health-related data such as text, audio, video, and images. Thus, while the data of the patient is evaluated clinically, the relevant decision support software helps the physician by trying to predict the best possible among the appropriate options. In this way, it affects the change in the result between death and life with early diagnosis and appropriate treatment (1-4).

Machine learning and deep learning networks that make up artificial intelligence are used in clinical decision support systems. Machine learning is a system that can make predictions on data and reveal inferential results by applying various algorithms (5-9). This system is used in the solution of many problems and has also enabled the use of deep learning networks with the development of new algorithmic approaches. Deep Learning is the technique of obtaining a learning representation from the data by using computational methods in each layer of the architecture consisting of successive layers and discovering the hidden features in data sets. With the advancement of this technique in recent years, deep learning algorithms have started to be used in areas such as video processing, image processing, voice recognition, and robotics. Among these areas, video/image processing is one of the most important research topics in the development of clinical decision support systems (10-12).

Polyps are lesions that can be detected as mucosal growths within the intestine. Polyps are flat, raised, or peduncled and can be distinguished from normal mucosa by color and surface patterns. Most bowel polyps are harmless, but some have the potential to turn into cancer. Therefore, detecting and removing polyps is important to prevent the development of colorectal cancer (13).

Convolutional Neural Networks (CNN) are one of the most important neural networks in deep learning. These networks can perform complex tasks such as images, sounds, text, videos, and are most commonly applied for analyzing visual images. CNN uses an adaptation of multi-layer sensors designed to require minimal preprocessing. In computer vision, CNNs are known to be powerful visual models that provide feature hierarchies that enable accurate segmentation. It is also reported that CNNs perform classification estimates relatively faster than other algorithms (11,14).

In this study, it is aimed to develop a user-friendly software for healthcare professionals with various methods such as detection, identification, classification and tracking of polyps in endoscopic video images using appropriate video/image processing techniques and CNN architecture. In addition, this study aims to help healthcare professionals make clinical decisions about the disease, as well as provide support for disease diagnosis, follow-up, and development.

MATERIAL AND METHOD

Data Set of the Study

The Kvasir data set was collected by the Bærum Hospital of the Vestre Viken Health Foundation in Norway and checked by the experts at the Norwegian Cancer Registry Center. The data set consists of a total of 4000 images and 8 classes, and there are 500 images in each class. These images are described and verified by medical practitioners (experienced endoscopists).

In our study, a total of 345 images in JPG image compression format belonging to the polyp class in the Kvasir data set were used. The data were obtained by decomposing the video images in the gastrointestinal tract into pictures. Images were obtained from https://datasets.simula.no/kvasir , which is open access for research and education purposes.

Deep Learning

Deep learning, which is a sub-field of machine learning (ML), has increased in popularity in recent years due to the fact that its computing power has greatly increased and large new data sets are increasing day by day. The field of deep learning has shown and demonstrated groundbreaking performances in a variety of complex tasks, including image classification, object detection, speech recognition, language translation, natural language processing, and gameplay. The ability of deep learning models to work on the graphics processing unit has enabled them to outperform many classical machine learning approaches in terms of modeling large data sets. Deep learning systems can accept multiple types of data as input, which is particularly important for the same type of health data (15). It uses many layers of nonlinear processing units for deep learning, feature extraction, and conversion. Each successive layer uses the output from the previous layer as input (16).

CNN deep learning architecture is frequently used in the processing of medical image and video data. Recently, deep learning algorithms such as CNN, detection of breast cancer on mammograms, segmentation of liver metastases with computed tomography (CT), brain tumor segmentation with resonance (MR) imaging, classification of high-resolution chest CT images of interstitial lung patients as a decision support system started to be used (17). The units in the layers in the CNN are locally connected, meaning each unit receives weighted inputs from a narrow range known as the receiving area in the previous layer. Stacks layers to create multi-resolution pyramids. The higher-level layers learn from the increasingly wider receptive fields. The main computational advantage of CNNs is that all receiver fields in a layer share weights, resulting in far fewer parameters than fully connected neural networks. Some of the best-known CNN architectures are AlexNet, VGGNet, ResNet, GoogLeNet, MobileNet, and DenseNet (18).

Image Processing

Image processing is the process of digitizing the image by applying different methods and techniques to extract useful information from images and obtain advanced images. In other words, image processing can be defined as computer studies aimed at changing the digital image data with the help of a computer or software in accordance with the targeted situation. Image analysis, on the other hand, can be summarized as the process of obtaining the numerical data needed for the targeted purpose from the available data (19). The most widely used open-source image processing libraries are OpenCV, Scikit-image, Dlib and BoofCV.

Modeling Stage

The methods used in this study can be examined in 5 stages. These;

1. Image pre-processing,
2. Drawing bounding boxes and labeling the images,
3. Establishing a deep learning model,
4. Training of the model,
5. Development of desktop video processing software,

Image Pre-Processing

Images in the data set contain green or black boxes that help determine the current position of the endoscope tip within the length of the bowel. Since these boxes do not make any sense for the whole image, they have been cropped out of the image. In addition, the process of converting images with different resolutions from 720x576 to 1920x1072 pixels into 350x350 pixels has been applied. Standardized images were obtained with this process.

It is important that the images in the data set have the same dimensions. Because the CNN input layer creates a pyramid for each image. This pyramid enables the detection of an object of any size, so the CNN model learns to detect larger or smaller objects in the image data set.

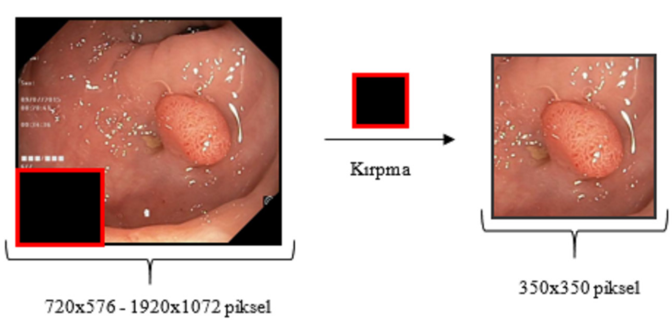

Figure 1. Illustration of polyp preprocessing

Drawing Bounding Boxes and Labeling the Images

An XML data set was created with the Imglab tool to draw bounding boxes on pre-processed images and start the labeling process. The Imglab graphical interface has been accessed to draw bounding boxes and label the images in the XML data set created. A window containing a list of images in the Imglab tool folder was opened and a bounding box was drawn and labeled for the boxes for each image.

Deep Learning Model Used in the Study

In our study, Convolutional Neural Networks (CNN) and Max-Margin object detection method (MMOD), which are among the deep neural networks architectures in Dlib Library, were used. In this thesis, Lenet architecture, which has a special multi-layered neural network, was used (20) .

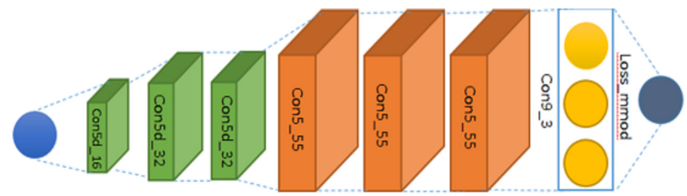

Figure 2. CNN architecture used in the study

MMOD is a method used to learn to detect objects in images. In this method, the detector takes every window in the image during the training phase and is scored with an appropriate objective function that aims to balance false perceptions and overlooked perceptions. It does not perform any subsampling with the MMOD method, instead, it optimizes all sub-windows.

Training of the Model

Images in the data set were taken and an image pyramid was created for each image. With the created pyramid image, it has been provided to find objects at the scale expected by the detectors. Thus, the image pyramid, which is several times larger than the normal image, has a rectangular structure, making it easy to work and process at high speeds on the graphics processing unit(GPU) using CNN.

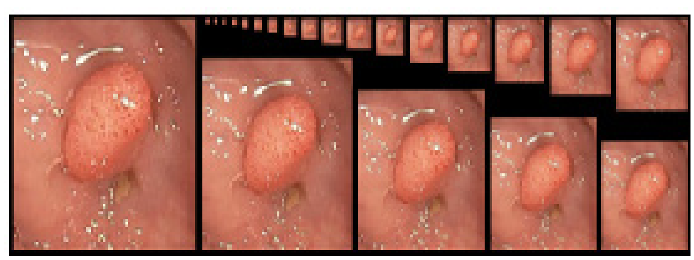

Figure 3. Image pyramid of the model

CNN took the image pyramid as its input and created a new image set using the convolution layers. A heat map was created in all parts of the image set that may contain polyp. The bright red areas in the heat map are the places that CNN thinks contain polyps, while the dark blue regions are the places that they think do not contain polyps. It was observed that the heat map detected the polyp correctly when placed on the image pyramid.

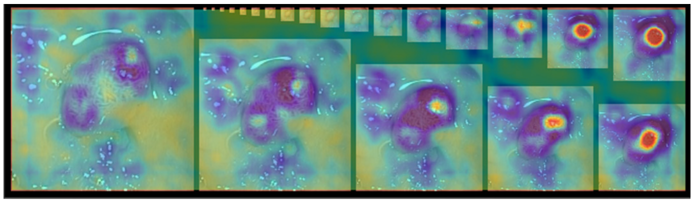

Figure 4. The heatmap image pyramid of the model

After finding all hot spots in the CNN outlet, non-maximum suppression method was applied and the sections corresponding to the determined hot spots were removed. Thus, a model that recognizes polyps was created using CNN+MMOD.

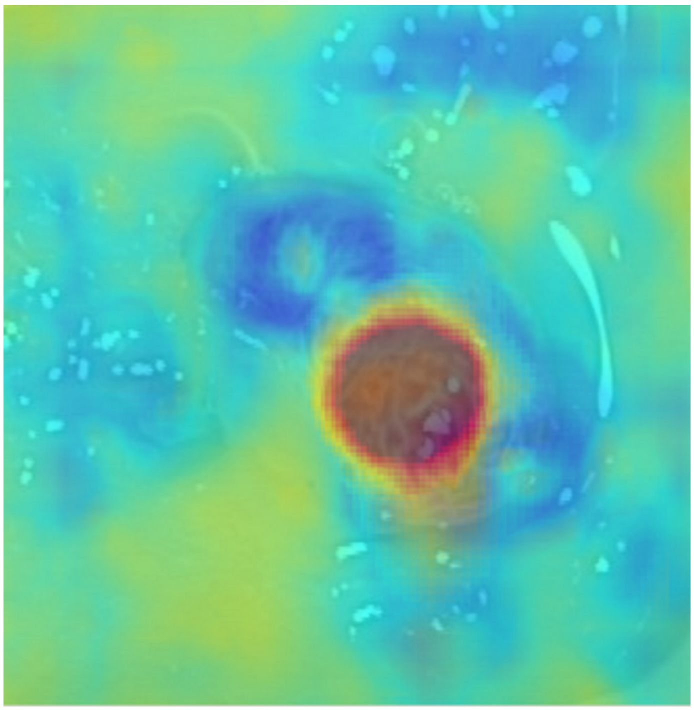

Figure 5. CNN + MMOD heat map

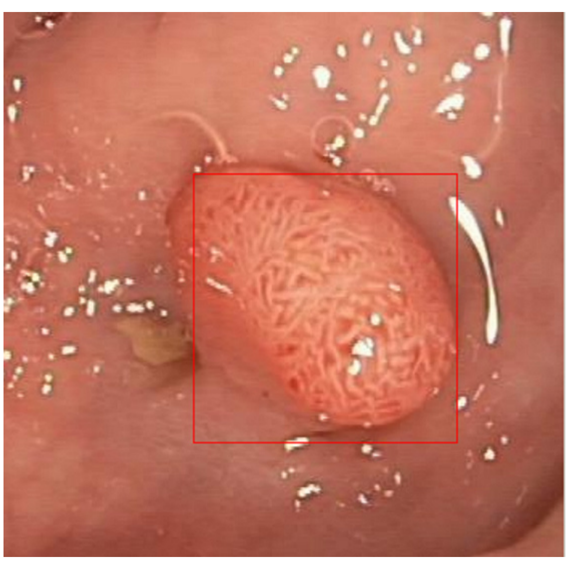

Figure 6. Determination of the polyp

Development of Desktop Video Processing Software

Video processing software was developed in Python language using Dlib and OpenCV libraries. In the software, the model trained by CNN+MMOD created in the C++ programming language was used.

Video processing software divided the images taken from the uploaded video or camera into frames and reduced in size. The images divided into frames were scanned by CNN detectors belonging to our model. Thus, the polyps in the video images were found and their positions were determined. Locations are enclosed in a red frame and visibility is provided. The working principle of the software is given in figure 7.

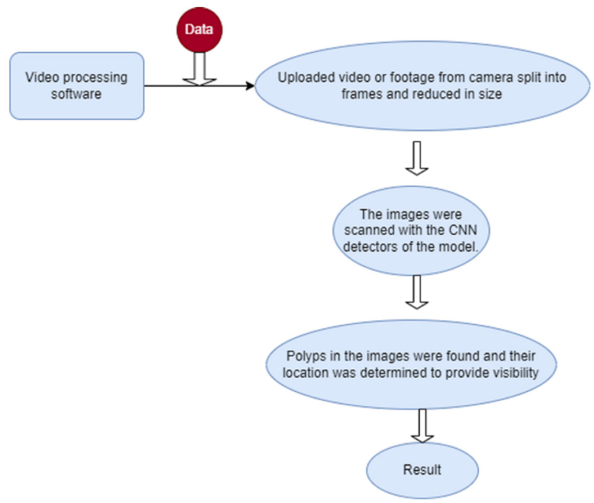

Figure 7. The working principle of the software

The procedures performed are given in Figure 8.

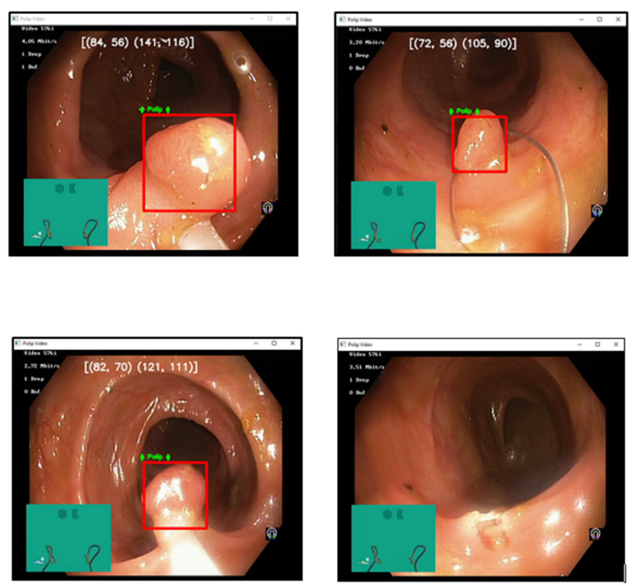

Figure 8. Finding polyps with video processing software

Video processing software also includes video recording feature. Thus, it is ensured that the processed videos are archived at the same time.

Performance Evaluation Metrics

In this study, Precision, Recall, F1-Score, Average Precision (AP), Mean Average Precision (mAP), IoU (Intersection over Union), LRP (Localization Recall) Precision - Positioning Precision and Precision) performance metrics were used. Of these metrics, mAP is a metric that combines recall and sensitivity for sequential retrieval results. The average precision criterion is used to evaluate detection algorithms. mAP metric is the sensitive and recall product of detected bounding boxes. The mAP value ranges from 0 to 1. The higher the better. It is an evaluation metric used to measure the accuracy of the object detector in the IoU dataset. The IoU is calculated by dividing the area where the predicted bounding box and the real bounding box intersect the area where these two limiters meet. It can be said to be a good estimate if the IoU has a value above 0.5. LRP moves along the recall sensitivity (RP) curve by specifying a performance score for each point consisting of positioning, precision, and sensitivity errors, and finally finds the best configuration the detector can achieve.

RESULTS

The images used in our study were divided into 276 for training and 69 for testing. Separated images were reduced to a resolution of 350x350 pixels, standardization was achieved and made suitable for the CNN model. Initial parameter values of the CNN model are given in Table 1.

Table 1. Beginning parameter values of the CNN model

Parameter	Value
Net Layer	21
Net size	0.955304 MB
Overlaps_nms	(0.1, 0.1)
Overlaps_ignore	(0.5, 0.95)
İterations without progress threshold	
	50000
Test iterations without progress threshold	
	1000
Optimization algorithm	SGD (Stochastic gradient descent)
Learning rate	0.0001
Min learning rate	0.0001
Learning rate shrink factor	
	0.1
Weight decay	0.0001
Momentum	0.9
Truth match IOU thresh	0.5
Loss per miss	1
Loss per FA	1

In the testing of the model created with CNN+MMOD, 69 polyp images and 71 bounding boxes were created as test data. Training and testing of the model were carried out on an Intel® Xeon® E5-1630 v3 (8 Core, 3.70 GHz) processor, 32GB memory, and NVIDIA GeForce GTX1080 Ti 11GB video card. The training lasted approximately 21 hours and was completed in 140581 iterations.

The CNN network is trained by convolutions of 21 layers and 5x5 filter sizes. The learning rate shrink factor was chosen as 0.1 as the initial value and continued until 127915 iterations. The training is completed when the minimum learning speed reaches 0.0001.

When the training of the model is completed, the test loss value is 0.46462 and the training loss value is 0.01058.

An example of the specified reference border boxes and the limit boxes estimated by the model are given in Figures 8 and 9, and the results of 69 polyp images belonging to the test data and the estimation results of 71 border boxes are given in Table 2.

Table 2. Estimated values

Prediction	Border Box
True Positive	64
False Positive	1
False Negative	7

While calculating the estimation values, the images of the test data were checked. In Figures 8, red boxes show reference boundary boxes, and blue boxes show the boxes predicted by the model. Box number 1 is false positive, box number 2 is true positive, and box number 3 is false negative.

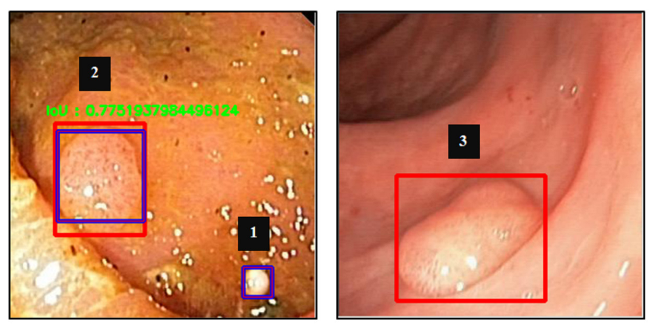

Figure 9. Diagnostic Values

The diagnostic performance of the model was calculated according to the specified evaluation criteria. Localization Recall Precision (LRP), a new performance criterion, was used for object detection in the calculation. The IoU value obtained from the training and test data was used to calculate the LRP. Performance Evaluation Metrics values of the training and test image data set are given in Table 3.

Table 3. Performance evaluation metric values of training and test image dataset

Metric	Training	Test
Precision	1	0.9846
Recall	1	0.9014
F1-Score	1	0.9412
AP	1	0.8942
mAP	1	0.8942
oLRP (τ=0.5)	0.4064	0.4862
moLRP (τ=0.5)	0.4064	0.4862

Since it is a single object class, the AP value gave the same results as the mAP value and the oLRP value gave the same results as the moLRP value at the end of the calculation.

DISCUSSION

Medical image processing is one of the most important components of artificial intelligence applications in medicine, and the development of deep learning has made great contributions to this field. There are three main types of objectives in medical image processing, namely classification, perception and segmentation, and they are closely related to each other. Object detection, in particular, forms the basis of many medical image processing tasks. Developments in deep learning in recent years have contributed greatly to the object detection performance of models. (21). In this study, a user-friendly software was developed for healthcare professionals by providing the detection, definition, classification, and tracking of polyps using open-access video image data set video/image processing techniques and deep learning architectures for polyps contained in endoscopic video images. The aim of this study is to help healthcare professionals make clinical decisions about the disease, as well as provide support for disease diagnosis, follow-up, and development.

In this study, all performance metrics examined in the training data set took the value 1. In the test data set, precision, recall, F1 score, AP, mAP, oLRP, and moLRP obtained from the model were 98%, 90%, 94%, 89%, 89%, 48% and 48% respectively.

In a study, a computer-based system based on color wavelet and convolutional neural network properties of endoscopic video frames has been proposed to support gastrointestinal polyp detection. In this context, it was determined that the proposed system detects and classifies gastrointestinal polyps from endoscopic video outperforms existing methods in datasets from different open databases (22).

In a study, a computer-based system based on color wavelet and convolutional neural network properties of endoscopic video frames was proposed to support gastrointestinal polyp detection. In this context, it has been determined that the proposed system for detecting and classifying gastrointestinal polyps from endoscopic video performs better than existing methods in data sets from different open databases (23).

CONCLUSION

As a result, very successful results were obtained in the prediction of endoscopic polyps with the model used according to the performance evaluation metrics. The proposed computer-aided system will be able to provide clinical support to clinicians during the diagnosis, treatment, and follow-up process.

Financial disclosures: *The authors declared that this study hasn't received no financial support.*

Conflict of Interest: *The authors declare that they have no competing interest.*

Ethical approval: *Ethics committee approval is not required.*

REFERENCES

1. Yagin FH, Cicek IB, Kucukakcali Z. Classification of stroke with gradient boosting tree using smote-based oversampling method. Med. 2021;10:1510-5.
2. Yagın FH, Kucukakcali Z, Cicek IB, Bag HG. The effects of variable selection and dimension reduction methods on the classification model in the small round blue cell tumor dataset. Middle Black Sea J Health Science. 2021;7:390-6.
3. Cicek İB, İlhami S, Yagın FH, Colak C. Development of a Python-Based Classification Web Interface for Independent Datasets. Balkan J Electrical Computer Engineering. 2022;10:91-6.
4. Akbulut S, Cicek IB, Colak C. Classification of Breast Cancer on the Strength of Potential Risk Factors with Boosting Models: A Public Health Informatics Application. Medical Bulletin of Haseki/Haseki Tip Bulteni. 2022;60:196-203.
5. Cicek İB, Kucukakcali Z, Yagın FH. Detection of risk factors of PCOS patients with Local Interpretable Model-agnostic Explanations (LIME) Method that an explainable artificial intelligence model. J Cognitive Systems. 2021;6:59-63.
6. Cicek IB, Kucukakcali Z. Classification of hypothyroid disease with extreme learning machine model. The Journal of Cognitive Systems. 2020;5:64-8.
7. Cicek İB, Kaya MO, Colak C. Assessment of COVID-19-Related Genes Through Associative Classification Techniques. Konuralp Medical Journal. 2022;14:1-8.
8. Percin İ, Yagın FH, Arslan AK, Colak C. An interactive web tool for classification problems based on machine learning algorithms using java programming language: data classification software. 2019 3rd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT); 2019: IEEE.
9. Cicek İB, Kucukakcali Z, Colak C. Associative classification approach can predict prostate cancer based on the extracted association rules. J Cognitive Systems. 2020;5:51-4.
10. Goodfellow I, Bengio Y, Courville A, Bengio Y. Deep learning: MIT press Cambridge; 2016
11. Karaman U, Cicek IB. Determination Of Cryptosporidium Spp. Risk Factors Using Multilayer Perceptron Neural Network And Radial Based Functional Artificial Neural Network Method. J Cognitive Systems.2020;5:83-7.
12. Kucukakcali Z, Cicek İB, Guldogan E. Performance evaluation of the deep learning models in the classification of heart attack and determination of related factors. J Cognitive Systems. 2020;5:99-103.
13. Hu W, Li C, Li X, Rahaman MM, Ma J, Zhang Y, et al. GasHisSDB: A new gastric histopathology image dataset for computer aided diagnosis of gastric cancer. Computers in biology and medicine. 2022;142:105207;1-9.
14. Yagın FH, Guldogan E, Ucuzal H, Colak C. A computer-assisted diagnosis tool for classifying COVID-19 based on chest X-Ray images. Konuralp Med J. 2021;13:438-45.

15. Yu J, Liu G. Knowledge-based deep belief network for machining roughness prediction and knowledge discovery. Comput Ind. 2020;121:103262.

16. Deng L, Yu D. Deep learning: methods and applications. Foundations Trends Signal Processing. 2014;7:197-387.

17. Chartrand G, Cheng PM, Vorontsov E, et al. Deep learning: a primer for radiologists. Radiographics. 2017;37:2113-31.

18. Işın A, Direkoğlu C, Şah M. Review of MRI-based brain tumor image segmentation using deep learning methods. Procedia Computer Science. 2016;102:317-24.

19. İnce Ö, Senel I, Yılmaz F. Image processing and analysis in health: advantages, challenges, threats and examples. Arc Health Sci Res. 2020;7:66-74.

20. LeCun Y. LeNet-5, convolutional neural networks 2015 [Available from: http://yann.lecun.com/exdb/lenet/.

21. Li Z, Dong M, Wen S,et al. CLU-CNNs: Object detection for medical images. Neurocomputing. 2019;350:53-9.

22. Zuluaga-Gomez J, Zerhouni N, Al Masry Z, et al. A survey of breast cancer screening techniques: thermography and electrical impedance tomography. J Med Eng Technol. 2019;43:305-22.

23. Billah M, Waheed S. Minimum redundancy maximum relevance (mRMR) based feature selection from endoscopic images for automatic gastrointestinal polyp detection. Multimedia Tools Applications. 2020;79:23633-43.