Turkish Journal of Science & Technology
19(2), 315-324, 2024

Research Paper
https://doi.org/10.55525/tjst.1386253

A YOLOV3-Based Method for Detecting Deepfake Manipulated Facial Images

Mehmet KARAKÖSE[1,] Hasan YETİŞ[2*], Mert ÇEÇEN[3]
[1,2,3] Department of Computer Engineering, Engineering Faculty, Fırat University, Elazığ, Türkiye
[1] mkarakose@firat.edu.tr, [*2] h.yetis@firat.edu.tr, [3] mert.cecen23@gmail.com

(Geliş/Received: 07/11/2023; Kabul/Accepted: 26/02/2024)

Abstract: With the advancement of technology and the development of applications that make it easier to transfer images, sounds and videos to the virtual environment, it has become much easier to access people's personal information, videos and images. Deepfake technology produces fakes of authentic images or sounds using deep learning and artificial intelligence techniques. Today, in addition to being used in the entertainment and film industries, it is also used in situations such as creating fake news and discrediting people. Different studies have been conducted in the literature to detect deepfake images and videos to prevent these situations. In this study, a comprehensive literature review was conducted. Real and fake images were collected and labelled from different datasets or videos, and a dataset was created by applying the necessary pre-processing steps. With the created dataset, training was carried out with YOLOv3 technology, which calculates class probabilities differently from traditional methods using Convolutional Neural Networks (CNN) and handles all operations in a single regression problem, which can make fast and high-accurate detection, and the modelling process is explained. With the tests performed in the study, the model that can detect fake images produced with deepfake technology with 95% accuracy was obtained.

Keywords: Convolutional neural networks, deepfake image detection, deep learning, YOLOv3.

Derin Sahte ile Manipüle Edilmiş Yüz Görüntülerin Tespiti için YOLOV3 Tabanlı Bir Yöntem

Öz: Teknolojinin ilerlemesi ve görüntü, ses ve videoların sanal ortama aktarılmasını kolaylaştıran uygulamaların gelişmesiyle birlikte insanların kişisel bilgi, video ve görsellerine ulaşmak çok daha kolay hale gelmiştir. Derin sahte teknolojisi, derin öğrenme ve yapay zekâ tekniklerini kullanarak gerçek görüntü veya seslerin sahtelerini üretmek için kullanılmaktadır. Günümüzde eğlence ve film endüstrilerinde kullanılmasının yanı sıra, sahte haber oluşturma ve insanları itibarsızlaştırma gibi durumlarda da kullanılmaktadır. Bu durumların önüne geçmek için literatürde derin sahte görsel ve videoların tespitine yönelik farklı çalışmalar yapılmıştır. Bu çalışmada kapsamlı bir literatür taraması yapılmış ve farklı veri setlerinden veya videolardan gerçek ve sahte görseller toplanmış, etiketlenmiş ve gerekli ön işleme adımları uygulanarak bir veri seti oluşturulmuştur. Oluşturulan veri seti ile Evrişimli Sinir Ağlarını kullanarak geleneksel yöntemlerden farklı bir şekilde sınıf olasılıklarını hesaplayan ve tüm işlemleri tek bir regresyon probleminde ele alan hızlı ve yüksek doğrulukla tespit yapabilen YOLOv3 teknolojisi ile eğitim gerçekleştirilmiş ve modelleme süreci anlatılmıştır. Çalışmada yapılan testlerle derin sahte teknolojisiyle üretilmiş sahte görüntüleri %95 doğrulukla tespit edebilen bir model elde edilmiştir.

Anahtar kelimeler: Evrişimli sinir ağları, deepfake görüntü algılama, derin öğrenme, YOLOv3.

1. Introduction

In today's society, most people use advanced lenses and cameras. In addition, with the applications developed for the digital environment, it has become very easy for users to share and upload images to the internet. It has become easier to access users'personal information, videos and images in these environments developed for people to share [1]. As a result, access to the personal images of government officials, business people, celebrities and many others has become much easier, and the opportunity to use these images has also become much easier. Deepfakes is a widely used technology to generate fake content from real images and sounds using deep learning techniques [2, 3]. The most frequently used deepfake production method uses face replacement with deep neural networks and automatic encoders [4]. In this method, the target video and several images of the face desired to be used in this video are generally used to create a deepfake [5]. Deepfakes are fake media content created using artificial intelligence to create fake news agendas, fake political agendas or personal attacks. When used for malicious purposes, deepfakes can harm individuals' reputations by sabotaging personal data security. Since there is no law prohibiting deepfakes today, detecting deepfakes is an important element in separating real images and fake image data and ensuring their security. The use of Generative Adversarial Networks (GANs) in the production of deepfake images is quite common [6]. Karras and colleagues proposed a controversial generative network model called StyleGAN to generate images of faces that have never existed before [7]. In another study, Zhu et al.

[*] Corresponding author: h.yetis@firat.edu.tr. ORCID Number of authors: [1] 0000-0002-3276-3788, [2] 0000-0001-7608-3293, [3] 0009-0008-3658-047X

introduced a face replacement method based on a generative adversarial network called CycleGAN [8]. Choi et al. introduced a technique called StarGAN, which can alter facial features like hair or skin colour, gender, age, and the presence of eyeglasses [9]. Thies et al. employed the Face2Face technique, which is based on a generative adversarial network, to manipulate the facial expressions of individuals in images [10]. As creating such fake images has become widespread and deepfake technology has developed dramatically, methods used to detect images created using this technology have also begun to be developed. Models were trained using data sets to detect deepfake images, and some signs and anomalies were tried to be detected to distinguish fake images from real images with the models created. Deepfake Detection technology detects fake images or videos [11-14]. Deepfake Detection is the process of detecting deepfake content. It was developed to detect fake or modified media content using deep learning techniques and artificial intelligence. There are different studies in the literature on detecting deepfake images. Deepfake detection studies, where the images used in these studies and the methods used to detect these images are explained in Table 1.

Table 1. Deepfake detection studies.

Reference	Images used in Deepfake Detection	Methods used in Deepfake Detection
[15]	Any video images that met the requirement of not exceeding a 50 Mb file size was used.	MesoInception4, FWA, VA-MLP, Xception-c23, ClassNSeg, Capsule, DSP-FWA, CNNDetection, Upconv, WM, Selim methods.
[16]	A new dataset containing high-quality Deepfake images with different models from DeepFakes was used.	The authors developed a method to identify high-quality Deepfakes by designing DMA-STA, a simple and effective Deepfakes model no-hold method based on spatial and temporal attention.
[17]	Images showing head posture and all features of the face were used using central areas from the images in the DARPA MediFor GAN Image/Video Challenge dataset and the images obtained from real and fake video frames in the UADFV dataset.	To determine whether the existing images in the datasets they used were fake or real, the authors performed training using Support Vector Machine (SVM), which is based on detecting differences between head poses estimated using facial landmarks and those in central facial regions.

An open-source online platform that can integrate Deepfake detection methods, called DeepFake-o-meter, was created by Yuezun Li et al. [15]. Users using this platform choose one of the methods that can upload any video that meets the condition of not exceeding 50 Mb file size and use it based on the advanced Inception modules, one of the deepfake detection methods offered by the system. These methods are MesoInception4, FWA, VA-MLP, Xception-c23, ClassNSeg, Capsule, DSP-FWA, CNNDetection, Upconv, WM, Selim methods [18]. After the uploaded video passes through the Docker containers, the faces in the video are extracted. After the captured images go through various pre-processing steps, it is determined whether the face is fake or not. Shan Jia et al. created a new data set by collecting different deepfake images. In this set, they used FaceSwap software and Autoencoder models to create five valid categories for Deepfake videos with examples in encoder, decoder, middleware, and input data. By designing a simple and effective Deepfakes model no-hold method, DMA-STA, based on spatial and temporal attention, achieved and evaluated over 70% accuracy in identifying high-quality Deepfakes on the DFDM dataset [16]. Yang and his colleagues compared head posture poses using facial features with central areas of the face in photographs and videos. To detect fake and real images or videos, they used the images contained in the DARPA MediFor GAN Image/Video Challenge dataset and the images in frames from the videos in the UADFV dataset. They used the differences in head poses as a feature vector to train the support vector machine (SVM). As a result of the studies, they revealed that the SVM classifier achieved 0.890 AUROC in the UADFV dataset and 0.843 AUROC in the DARPA MediFor GAN Image/Video Challenge dataset by using separate frames as the unit of analysis with the Area Under ROC (AUROC) as the performance measurement. [17].

Within the scope of these studies in the literature, different deep learning methods have been used to detect deepfake images and videos. Deep learning is a subset of machine learning that uses classification processes and learning methods to represent data in a specific format [19]. Deep learning is based on using the data set as an input and creating a model that can predict the outputs with the help of artificial intelligence. When classifying with deep learning, pre-processing steps significantly amplify the inputs and irrelevant variations are significantly suppressed [20]. Convolutional Neural Network (CNN), Recurrent Neural Network (RNN), Deep Belief Network, Deep Boltzmann Machine and Deep Autoencoder techniques are used in deep learning. Convolutional Neural

Networks have achieved great success in areas such as image processing, object detection, face recognition and video analysis [20, 21]. A convolutional Neural Network is a neural network that contains one or more convolutional layers, subsampling layers, and standard multiple layers. Although neural networks are not new, they are based on Alexnet [21] and Imagenet [22] technologies used to classify large-scale data. The working principle of Convolutional Neural Networks includes layers that can automatically extract and represent complex data features. YOLO technology takes the entire image simultaneously and estimates bounding box coordinates and class probabilities [23]. The training process of YOLO technology ensures that it has better generalisation ability overall as it is created using a large data set. In addition, better results can be achieved as it allows users to use various data augmentation techniques.

In this study, it is aimed to detect fake medical images created by deepfake methods. Chapter 2 provides information about YOLOv3 architecture and how to use it in practice. In Chapter 3, the application results are given experimentally and comparisons are made with current literature studies with similar purposes. In the Chapter 4, last section, the findings obtained as a result of the study are given and discussed.

1.1. Motivation

Deepfake is used to produce a different image by transferring the face of a source individual to the target person's body. Internet users first encountered the images produced by this technology in 2017 [24,25]. The first studies on producing deepfakes were carried out in 2014 by training Generative Adversarial Networks (GANs) with very large data sets. Deepfake, a type of artificial media that uses artificial intelligence to transform a person's image into a manipulated photo or video, can make people appear to say or do things they actually do not. Research reports state that images, sounds or videos produced with this technology may be used to facilitate crime in the coming years. In order to prevent these incidents from occurring, a study was carried out to detect the images produced by deepfake. In this study, a data set created with images taken from different data sets and videos and a model training was carried out using YOLOv3 technology to determine whether the images produced with deepfake are real or fake.

2. Proposed Method

In this study, it is aimed to detect an image if it is real or created by deepfake techniques. The proposed method contains of dataset collection, data regulation, detection (training-validation-testing the model) steps. In data collection step, the original images of the people and the fake images produced from these original images with deepfake technology were collected manually and turned into a data set. Orientation, image resizing, rotation, and brightness correction is applied on collected images. Because 416x416 YOLOv3 version is used, the images are resized according to the YOLO structure for faster training process. After applying the image pre-processing steps, the data set obtained. The obtained data set is divided into training, validation and testing data. 70% of all data is used for training, 10% of all data used for validation, and 20% of all data is used for test. Then training, validation, and test of the model actualized. After images are labelled as fake or real, the training process begins. Validation data was issued to measure the model's accuracy, while testing data was used to test the model after it was built. The general block diagram of the proposed method is given in Figure 1. In Figure 1, the last block refers to deep learning model. In the model, n is 1,2,4 and 8 respectively. When n is 2 then the 52x52 feature map is used for obtaining detailed (small) objects. When n is 4 medium-scale objects are detected from 26x26 feature map. And when n is 8, 13x13 feature map is used for less detailed (big) objects. The last block, model block, is shortening form of YOLOv3, given in Figure 2. At last, YOLO inferences from small, medium, and big objects, and determine the bounding box. Beside the bounding box, a class label (real or fake for our problem) is produced by YOLO.

Real and fake images produced with deepfake technology are collected to obtain the data set step. The bounding box labelling method was applied to the collected images. Automatic orientation was applied to the existing images to enrich the collected images before adding them to the data set. All image data was resized to 416 × 416 to be trained effectively and quickly in YOLOv3 technology. Then, a 10° rotation and a 25% brightness adjustment were made on the images for better recognition of each image. After the operations, the number of images in the dataset was obtained as 44,100. The clustered datasets are divided into 70% training, 10% validation, and 20% testing data.

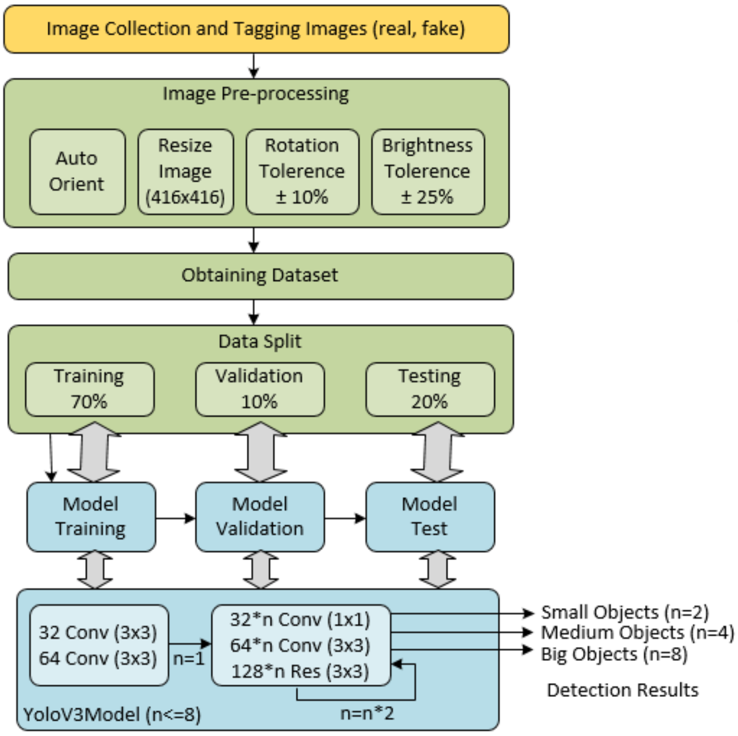

Figure 1. The block diagram of the proposed method.

In model stage, YOLOv3 technology using Convolutional Neural Networks was used. YOLOv3 technology was chosen among YOLO versions in order to train the model quickly and with high accuracy. Depending on different YOLO architectures, variable input sizes can be used: 320 × 320, 416 × 416 and 608 × 608. In this study, 416x416 sized inputs were used in order not to reduce image quality and increase processing speed. YOLOv3 uses Darknet-53 model for feature extraction. Figure 2 shows the main stages of the YOLOv3 algorithm with 416x416 image input. Convolution and residual layers are applied according to Darknet-53 model as shown in Figure 2. After reducing the image size to 52x52, 26x26 and 13x13, some other convolution layers are applied for detection of the interest area and the label. In YOLOv3, three different estimation scales are employed during the estimation process. The detection layer is employed to identify feature maps with three distinct dimensions, characterised by strides of 32, 16, and 8. The hyper parameters of the proposed method are given in Table 2.

Table 2. Hyper parameters of proposed method.

Parameter	Value
Input size	416x416
Classes	2
Threshold	0.3
Batch	64
Subdivisions	16
Channels	3
Momentum	0.9
Decay	0.0005
Learning rate	0.001
Max batches	4000

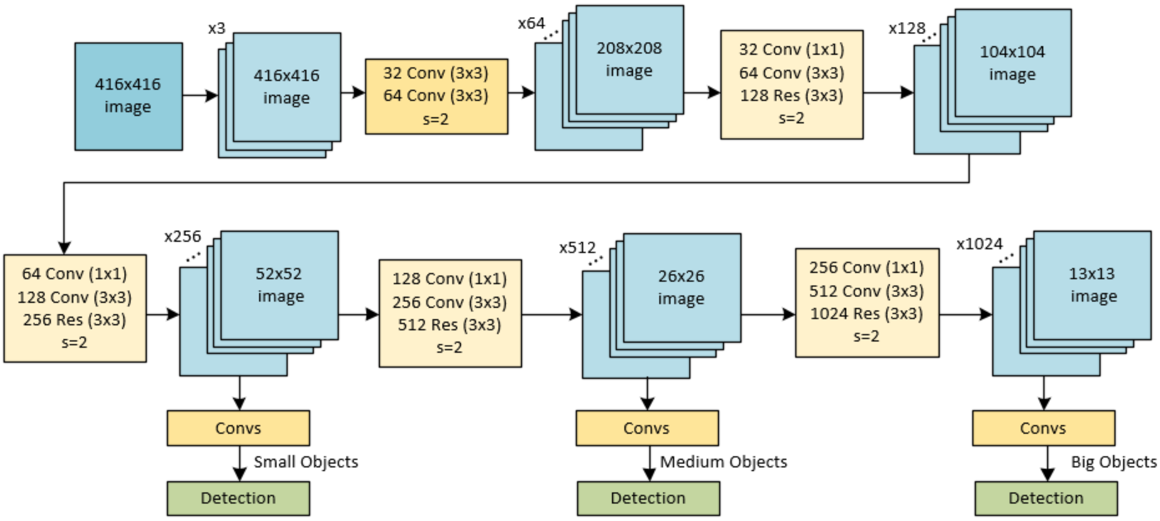

Figure 2. Architecture of YOLOv3 [26].

3. Experimental results

In order to achieve high success in the studies, the Graphics Processing Unit (GPU), which enables the creation of clear graphics, and the Google Collaboratory technology, which provides storage support by accessing the cloud environment, were used. Before the training was carried out, CUDA (Compute Unified Device Architecture) technology produced by NVIDIA was installed on the Jupyter Notebook created in the Colab environment. While creating the data set, labeling operations were performed with the bounding box method and Roboflow technology was used for these operations. Jupyter Notebook was created in the Colaboratory cloud environment created by Google and the necessary code blocks and parameter values were added. While making predictions with the created data set, YOLOv3 technology, which can detect objects by treating object detection as a single regression problem using Convolutional Neural Networks (CNN), was used. Thanks to the Darknet-53 network structure used by YOLOv3 in feature extraction, the evaluation is made more efficient and faster. The training process started by transferring the technologies and files to the cloud environment.

The training graphics of the created data set and the graphics showing the accuracy of the trained model as a result of the studies performed are explained in this section. The studies were carried out in the Google Colaboratory environment, which belongs to Google and provides users with GPU and cloud storage. After the images in the data set to be used for model training were collected, they were transferred to the Roboflow environment. Then, they were labeled with the bounding box method in this environment. Finally, the graphs obtained after applying the necessary pre-processing steps to the created data set are shown in Figure 3.

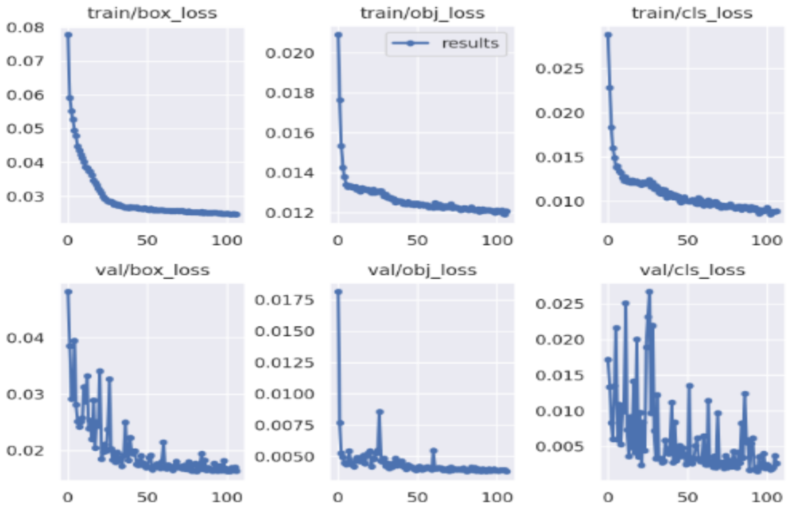

Figure 3. Training graphs loss values.

The train values in the graph in Figure 3 represent the measurements of the values of the training set, while the val values represent the measurements of the values of the validation set. The box_loss value focuses on the measure of the loss rate that will occur after the application of the bounding box technique used in labeling operations. Low values in the graph indicate that the model has improved to generalize and the data set is better labeled. The cls_los value shows the measure of the loss rate resulting from the classification. The decrease in the value in the graph indicates that a better classification is done.

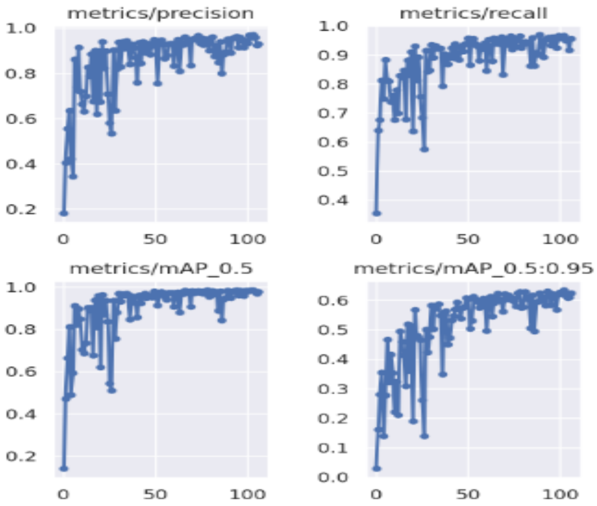

Figure 4. Training graphs metrics values.

The precision value in Figure 4 shows the precision values in the given prediction of the model, while the recall value shows the current performance of the system. The value of mAP_0.5 in the other graph indicates the average sensibility value, and the value of mAP_0.5:0.95 indicates the average precision. From the results in the graph, it is seen that the modeling process was successful and a good data set was created.

During the model training process, a Jupyter Notebook was created, and then the necessary codes were written in the code blocks in the created file to download the technologies and libraries to be used in deepfake detection. First of all, Darknet technology was installed on the cloud environment under study. Darknet technology is used to determine the detection rate of images entered into the model to be used in detecting deepfake images. After the darknet technology was cloned, the CUDA technology offered by NVIDIA, which we will use to carry out the training, was uploaded to the cloud environment by adding the necessary codes.

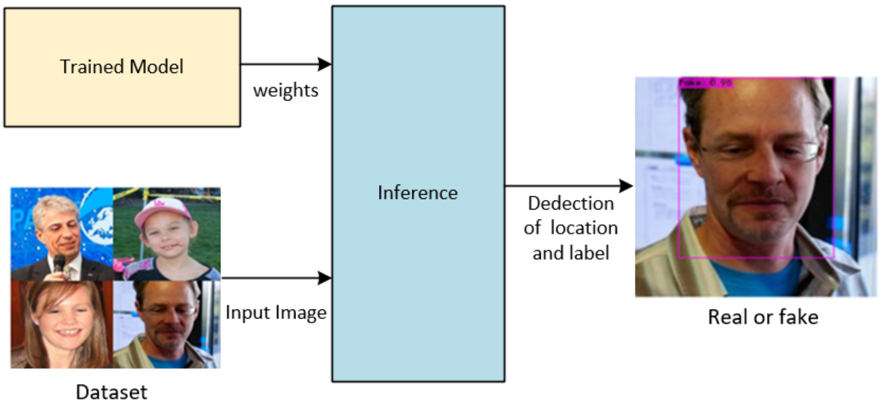

Figure 5. Model architecture.

As seen in Figure 5, first real and fake images were collected manually and turned into a data set. Data, names and cfg model files were created for the created dataset. Information on how many classes will be trained, the file paths of the test and training data, the file path containing the tag name of the class to be trained, and the file path where the backup file will be located when the training is completed have been entered into the data file. Then, the training process was started on the Google Colab platform by specifying the data, cfg network file and weight file that we stored under the Darknet main folder. The training process was terminated when the average loss value among the results obtained during training became very low as 0.1552. After the training was carried out with YOLOv3 technology, the weight values obtained as a result of the training were added to the Backup file of the Google Drive account. Different images were given to the created model, which was determined using the OpenCV library of the Python programming language. It was observed that the probability of the images being deepfake was over 95% due to the predictions.

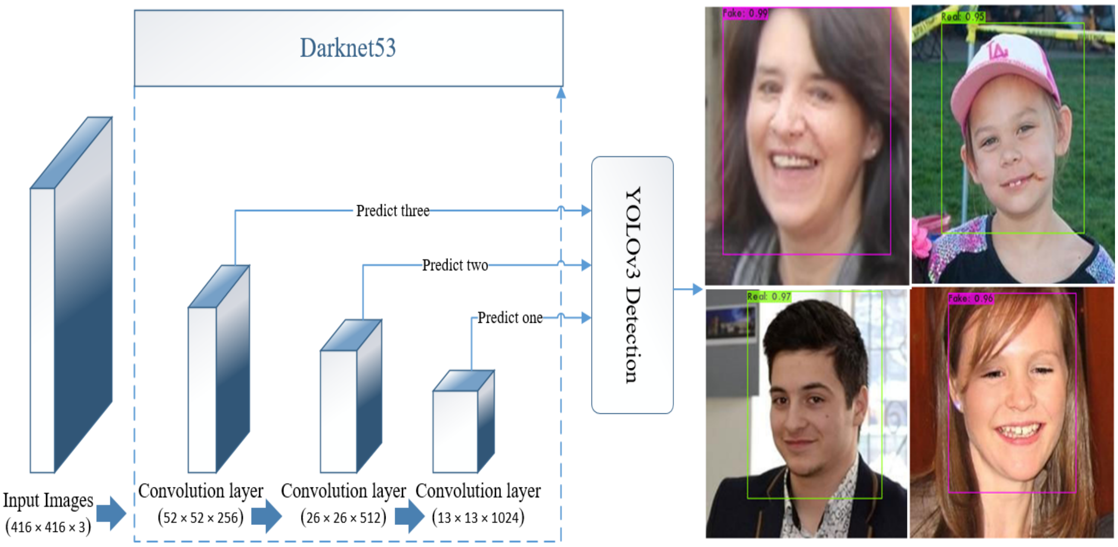

Figure 6. Rates of detecting whether some of the images are fake or real as a result of the YOLOv3 detection method.

Figure 6 shows the accuracy percentages obtained as a result of determining the fakeness and reality status of the visual by transferring the images into the model after creating and training the images in the data set, which includes fake images created with deepfake technology and real human images. The images entered into the model with a size of 416 × 416 were detected more efficiently by feature extraction processes of the Darknet-53 network used by YOLOv3, and then the probability of being a deepfake was estimated. After estimation was made with the obtained weight values, these ratios were printed using the OpenCV library of the Python programming language. The results obtained by testing the study show that the model can detect images with high accuracy.

Different studies have been conducted in the literature using different data sets and Convolutional Neural Networks to detect deepfake images and videos. In this article, a table has been created showing the accuracy rates obtained from studies using deep learning techniques and Convolutional Neural Networks, the data sets used by other studies in the literature, and the accuracy rates obtained from these data sets. Comparison results with studies in the literature are shown below.

Table 3. Comparison results with studies in the literature.

Reference	Method	Advantages	Disadvantages	Dataset	Accuracy Rates
[27]	The images were compressed by dividing into 8x8 blocks and training was carried out.	- High Accuracy - Works with Scaled and Compressed Images	They filtered the images preserving only those with the smaller side equal or greater than 30 pixels	OpenForensics	99.20%
[28]	Model training was performed using low-resolution video image data.	Works on low-resolution and short-time clips	-Need face detection before giving the model -It is done on video clips which has many variants of one pose	Kaggle Deepfake Detection Challenge (DFDC)	94.93%
				Face Forensics++	93.20%
[29]	Fake face detection was performed in images using two-stream face classification and patch trio.	- Ability to detect videos forged with traditional and deepfake methods.	- Weak against to low resolutions -It is done on video clips which has many variant of one pose	A dataset created by the authors using the Deepfake tools.	91.70%
This Study	Model training was carried out with real and fake image data by performing resizing and automatic orientation.	- High Accuracy -Detection and training on single images -No need extra step	Since it works on a single image basis, clip frame transition features cannot be used.	A mixed dataset created by the authors	95.00%

In the first study in Table 3, the authors aimed to design a deepfake detector that is robust to background and image size variability by accurately detecting resized and compressed images. While performing this process, they first took the images in the OpenForensics dataset as input data and divided them into 8x8 blocks. The resulting outputs were processed with Discrete Cosine Transform. When they tested the CNN architecture they designed, they achieved 99.2% accuracy. While performing different transformations in the study increases the processing load, filtering the data set causes the success rate to be high. In the second study in the table, the authors trained a CNN to detect deepfakes from low-resolution and short-duration videos in the Kaggle Deepfake Detection Challenge (DFDC) and Face Forensics++ datasets. As a result of their training, they reached an accuracy value of 94.93% for the DFDC dataset and 93.2% for the FaceForensics++ dataset in detecting fake videos. In the fourth study in the table, the authors proposed a two-stream network model for fake face detection, including two-stream face classification and patch triplets. They trained a CNN model by performing two different classifications of facial images as real and fake on a dataset collected for the first time using the FaceSwap and SwapMe tools. An accuracy of 92.70% was achieved in the data sets they created. Finally, in this study, a data set was created by combining real images of people taken from different videos and data sets on the internet and fake images obtained from these images using deepfake technology. Using the created data set and YOLOv3 technology working with CNN, a model that can detect whether the images are fake or real was created. As a result of the studies, it was seen that this model was 95% successful and when compared to other studies in the literature, an effective model was developed among the models trained using CNN.

4. Conclusion

The development of applications and devices that make it easier to share personal images on the internet and the increase in digitalization have greatly increased the opportunities for individuals to access their personal data and images. With the increase in images in the virtual environment, deepfake technology is used maliciously to manipulate photos or videos of famous people, to create and spread fake news content, and to defame or blackmail politicians or government officials. In this respect, it has the potential to cause serious social, political and economic problems. Studies show that the use of deepfake technology in illegal events will increase. For these reasons, deepfake images have become an increasing concern and pose a major threat to people's personal spaces. To prevent these situations, techniques developed for detecting and identifying deepfake images are gaining importance day by day. Although deepfake technology has potential risks, it may be possible to reduce them by developing detection and prevention methods.

In this study, detailed research was conducted on detecting deepfake videos and images, and then a method was developed to detect these fake images. First, a data set was created, and images created with different deepfake techniques were added to this data set. Then, using CNN, one of the deep learning methods, model training was carried out to detect deepfake images with YOLOv3 technology, which calculates class probabilities differently than traditional methods and handles all processes in a single regression. problem that can be detected quickly and with high accuracy. The model obtained from the training was tested to evaluate the possibility of deepfake in new images. As a result of the tests, model training, and evaluation, over 95% of the results were successful. This situation shows that an effective method has been developed for detecting deepfake images compared to literature studies. The continuous development of deepfake applications requires the model to be updated accordingly to detect the images produced by these applications.

Future studies aim to expand the data set and obtain better results by integrating new technologies into our model. In this context, different deepfake images will be collected manually and added to the data set to enlarge the data set. Then, as a result of training the expanded data set in different versions of YOLO technology, the rates obtained in these versions will be compared. As a result of the studies, the YOLO version that gives the best results will be selected to help researchers working in this field. Also potential improvements or adaptations of the YOLOV3 method for deepfake detection will be investigated.

Acknowledgements

This study was supported by the TUBİTAK (The Scientific and Technological Research Council of Turkey) under Grant No: 122E676.

References

[1] Çeçen M, Karaköse M. A Deepfake Image Detection Approach Based on YOLOv3. In: 2th International Conference on Advances and Innovations in Engineering; 21-23 September 2023. pp. 10-18.

[2] Franklin RJ, Mohona. Traffic Signal Violation Detection using Artificial Intelligence and Deep Learning. In: International Conference on Communication and Electronics Systems; 10-12 June 2020. pp. 839 - 844.

[3] İlhan İ., Balı E., Karaköse M. An Improved DeepFake Detection Approach with NASNetLarge CNN. In: IEEE International Conference on Data Analytics for Business and Industry; 25-26 October 2022. pp. 598-602.

[4] Seow JW, Lim MK, Phan R, Liu J. A comprehensive overview of Deepfake: Generation, detection, datasets, and opportunities. Elsevier Neurocomputing 2022; 513: 351–371.

[5] İlhan İ, Karaköse M. Derin Sahte Videoların Tespiti ve Uygulamaları için Bir Karşılaştırma Çalışması. Adıyaman Üniversitesi Mühendislik Bilimleri Dergisi 2021; 8(14): 47-60.

[6] John J, Sherif B. Comparative Analysis on Different DeepFakeDetection Methods and Semi Supervised GAN Architecture for DeepFake Detection. In: Proceedings of the Sixth International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud); 10-12 November 2022.

[7] Karras T, Laine S, Alia T. A Style-Based Generator Architecture for Generative Adversarial Networks. IEEE Transactions on Pattern Analysis and Machine Intelligence 2021; 43: 4217-4228.

[8] Zhu JY, Park T, Isola P, Efros AA. Unpaired Image-to-Image Translation Using Cycle-Consistent Adversarial Networks. Image Translation Using Cycle-Consistent Adversarial Networks. In: IEEE/CVF International Conference on Computer Vision; 22-29 October 2017; Venice, Italy.

[9] Choi Y, Choi M, Munyoung K, Ha JW, Kim S, Choo J. StarGAN: Unified Generative Adversarial Networks for Multi-Domain Imageto-Image Translation. In: IEEE/CVF Conference on Computer Vision and Pattern Recognition; 18-23 June 2018. pp. 8789-8797.

[10] Thies J, Zollhöfer M, Stamminger M, Theobalt C, Niebner M. Face2face: Real-Time Face Capture and Reenactment of RGB Videos. In: IEEE/CVF Conference on Compute Vision and Pattern Recognition; 27-30 June 2016. pp. 2387 – 2395.

[11] Khatri N, Borar V, Garg R. A Comparative Study: Deepfake Detection Using Deep-learning. In: 13th International Conference on Cloud Computing, Data Science & Engineering; 19-20 January 2023.

[12] Pipin SJ, Purba R, Pasha MF. Deepfake Video Detection Using Spatiotemporal Convolutional Network and Photo Response Non Uniformity. In: IEEE International Conference of Computer Science and Information Technology (ICOSNIKOM); 19-21 October 2022.

[13] Zhang J, Cheng K, Sovernigo G, Lin X. A Heterogeneous Feature Ensemble Learning based Deepfake Detection Method. In: IEEE International Conference on Communications; 16-20 May 2022. pp: 2084 - 2089

[14] Budhiraja R, Kumar M, Das MK, Bafila AS, Singh S, MeDiFakeD: Medical Deepfake Detection using Convolutional Reservoir Networks. In: IEEE Global Conference on Computing, Power and Communication Technologies (GlobConPT); 23-25 September 2022.

[15] Li Y, Zhang C, Sun P, Ke L, Ju Y, Qi H, Lyu S. DeepFake-o-meter: An Open Platform for DeepFake Detection. In: IEEE Security and Privacy Workshops (SPW); 27 May 2021; China. pp. 277-281.

[16] Jia S, Li X,Siwei L. Model Attribution of Face-Swap Deepfake Videos. In: IEEE International Conference on Image Processing (ICIP), 16-19 October 2022. pp: 2356 - 2360.

[17] Yang X, Li Y, Lyu S. Exposing deep fakes using inconsistent head poses. In: IEEE Int. Conf. Acoust., Speech and Signal Process. (ICASSP*)*, 25-30 March 2012. pp. 8261 –8265.

[18] Ataş S, İlhan İ, Karaköse M. An Efficient Deepfake Video Detection Approach with Combination of EfficientNet and Xception Models Using Deep Learning. In: 26th International Conference on Information Technology (IT); 13-15 December 2023.

[19] Bar NF, Yetis H, Karakose M. An efficient and scalable variational quantum circuits approach for deep reinforcement learning. Quantum Information Processing 2023; 22(8): 300.

[20] Srivastava N, Salakhutdinov RR. Multimodal Learning with Deep Boltzmann Machines. Advances in Neural Information Processing Systems 25 (NIPS 2012); 3-8 December 2012.

[21] Krizhevsky A, Sutskever I,Geoffrey EH. ImageNet Classification with Deep Convolutional Neural Networks. Adv.Neural Inf. Process. Syst. 2012; *25*: 1–9.

[22] Deng J, Dong W, Socher R, Li LJ, Li K, Fei LF. Imagenet: A large-scale hierarchical image database. In: IEEE Conference on Computer Vision and Pattern Recognition; 20-25 June 2009. pp. 248-255.

[23] Rajput, SK, Patni JC, Alshamrani SS, Chaudhari V, Dumka A, Singh R, Rashid, M, Gehlot A, AlGhamdi AS. Automatic Vehicle Identification and Classification Model Using the YOLOv3 Algorithm for a Toll Management System. Sustainability 2022; 14(15): 9163.

[24] Salih ZA, Thabit R, Zidan KA, Khoo Be. A new face image manipulation reveal scheme based on face detection and image watermarking. In: IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAIET); 13-15 September 2022.

[25] Chang X, Wu J, Yang T, Feng G. DeepFake Face Image Detection based on Improved VGG Convolutional Neural Network. In: 39th Chinese Control Conference; 27-29 July 2020.

[26] Belhi A, Gasmi H, Al-Ali AK, Bouras A, Foufou S, Yu X, Zhang H. Deep Learning and Cultural Heritage: The CEPROQHA Project Case Study. In: International Conference on Software, Knowledge Information, Industrial Management and Applications (SKIMA); 26-29 August 2019.

[27] Concas S, Perelli G, Marcialis GL, Puglisi G. Tensor-Based Deepfake Detection in Scaled and Compressed Images. In: IEEE International Conference on Image Processing (ICIP); 16-19 October 2022. pp. 3121 – 3125.

[28] Rahman A, Siddique N, Moon MJ, Tasnim T, Islam M, Shahiduzzaman Md, Ahmed S. Short and Low Resolution Deepfake Video Detection using CNN. In: IEEE Region 10 Humanitarian Technology Conference (R10-HTC) 16-19 September 2022. pp. 259 - 264.

[29] Afchar D, Nozick V, Yamagishi J, Echizen I. MesoNet: a Compact Facial Video Forgery Detection Network. In: IEEE International Workshop on Information Forensics and Security (WIFS), 11-13 December 2018.