New Trends in Mathematical Sciences

Face and emotion recognition using deep learning based on computer vision methods

Celal Akcelik[1,2]*, Ali Okatan*[3] *and Ali Cetinkaya*[*,4,5]

[1]Istanbul Gelisim University, Faculty of Engineering and Architecture, Computer Engineering, Istanbul, 34310, Turkey
[2]GTech (G Teknoloji Bilişim San ve Tic AŞ.), Big Data & Analytics, Istanbul, 34485, Turkey
[3]Istanbul Aydin University, Faculty of Engineering, Software Engineering, Istanbul, 34295, Turkey
[4,*]Istanbul Gelisim University, Technology Transfer Office Application and Research Centre, Istanbul, 34310, Turkey
[5]Kocaeli University, Faculty of Engineering, Electronics and Communication Engineering, Kocaeli, 41380, Turkey

Abstract: Deep learning studies are among the discipline that are rapidly increasing and developing today. Especially after the concept of big data enters our lives, deep learning methods have used to process the data. In the study, it has aimed to detect the face on the picture determined by the user and to conduct emotional analysis and gender determination with the deep learning methods of the detected face. Viola-Jones algorithm has used for face recognition. "Mini_Xception" model of Convolutional Neural Networks (CNN) has used for emotion analysis and gender detection. Estimation rates have measured with 18 different experiments performed. The most successful emotion recognition has measured as 93,11% and the most successful gender recognition has measured as 90,75%. Experiments within the study have supported by visual studies.

1 Introduction

Computer vision is widely used today in the fields of face recognition systems and object classification. Face recognition is the automatic identification or verification of the people in the data obtained from images or videos. There are four basic stages of face identification processes. These operations are face detection, normalization, feature extraction and classification. No matter how successful the normalization and classification algorithms are in face detection, if the feature extraction stage is not successful, that system cannot achieve the desired success [1]. Commonly used combinations for both model training and evaluation are tested in recent studies [2]. Using real-time convolutional neural networks architecture, facial recognition, gender classification, and emotion classification are performed simultaneously in a single step [3], [4], [5]. Applications that make age estimation from face images over pictures are developed and data sets are created from face pictures for these applications [6], [7], [8]. Processing of human faces; It consists of face recognition, face tracking and face creation stages. Parallel applications are performed using CUDA and OpenCV on the GPU (Graphics Processing Unit) with Viola-Jones face detection algorithm. In this way, the calculation speed of the algorithm that can be obtained is increased, and comparative analyzes are made between serial and parallel applications in order to obtain better calculation results [9], [10]. Facial recognition is the first step in various applications in computer vision, such as human-computer interaction. In a study, a powerful GPU-based face detection application based on Viola Jones face detection algorithm has been developed by starting the computation in the graphics processing unit (GPU) using NVIDIA CUDA (Compute Unified Device Architecture) in terms of computing speed of face detection algorithms [11], [12].

* Corresponding author e-mail: alcetinkaya@gelisim.edu.tr

Material and metod

1.1 Viola-Jones algorithm

Viola-Jones Algorithm is a real-time detection algorithm with computer vision. For practical applications, at least two frames per second should be worked on. Being able to distinguish faces from non-faces from images is the first step of the detection and recognition process. The steps of Viola-Jones algorithm are determination of haar-like features, creation of integral image, classifier training (adaboost) and cascading classifier [13], [14].

1.2 Convolutional neural networks

The input layer on the CNN structure is the layer that is converted to the desired structure by entering the data of the model to be trained. Convolution layer is a customized linear process. Basically, these are networks that do convolution rather than matrix multiplication in at least one layer.

$$S(i,j) = \sum_m \sum_n I(i,j) K(i-m, j-n) \tag{1}$$

The terms "i" and "j" on Equation (1) indicate the position of the new matrix to be obtained after the convolution process. "S" is the output value, "I" is the input value and "K" is the filter value. Convolutional input matrix and filter matrix may not be the same size. In this case, the size of the output matrix is smaller than the input matrix. If you want the output matrix and the input matrix to be of the same size, pixel addition is done [15]. The values of the results produced in the output layer are transferred to the system and the obtained results become ready for use [3], [15].

1.3 Model and Datasets

Fig. 1. Real face samples from the UTKFace dataset

One hundred examples from the UTKFace data set are shown on Fig. 1. For sentiment analysis, the data set named "fer2013" published in 2013 is used. In this data set, there are pictures for detecting seven emotions. These emotions are angry (4593 pictures), disgust (547 pictures), fear (5121 pictures), happy (8989 pictures), sad (6077 pictures), surprise (4002 pictures), neutral (6198 pictures) Stop. The structure of the pictures is arranged in 48x48 size and shades of gray [16], [17]. Face samples are taken randomly from the UTKFace data set. This dataset is used to test the operation of the system [7], [18]. Two gender types (male and female) are defined for gender determination in the UTKFace data set. People in the data set are between 0 and 116 years old. There are 20000 face images. In order for the structure of the data set to be the same as the structure of the study, the pictures are processed in 48x48 size and gray tones. The images in the UTKFace data set show differences such as facial expression, resolution and illumination. The flow chart of the system is as shown in Fig. 2.
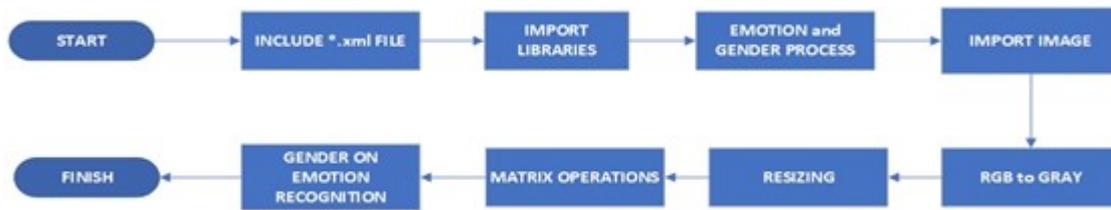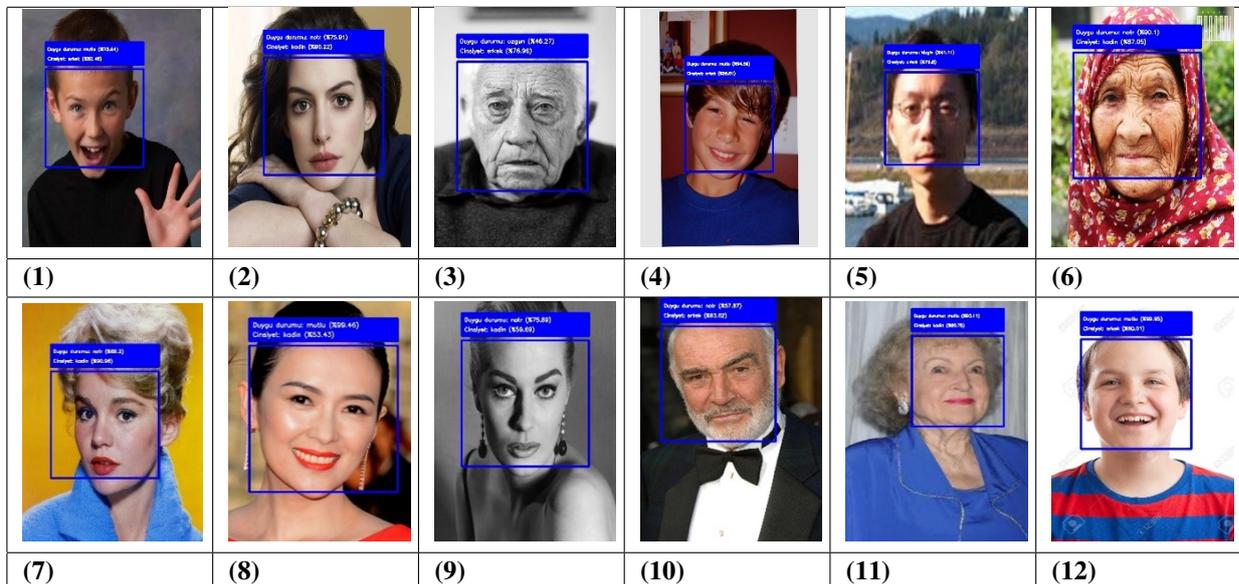
Fig. 2. Flowchart of the study

2 Experimental results

Table 1. Accepted results.

The system is tested with images in the UTKFace dataset. Valid results from experiments made from data sets are in Table 2, invalid results are in Table 4 and are given.

Table 2. Details of accepted results.

Experiments	Emotion	Emotion Rate %	Gender	Gender Rate %		Experiments	Emotion	Emotion Rate %	Gender	Gender Rate %
1	happy	75,94	male	82,48		**7**	neutral	68,2	female	90,98
2	neutral	75,91	female	82,48		**8**	happy	99,46	female	53,43
3	sad	46,27	male	76,96		**9**	neutral	75,89	female	59,69
4	happy	94,59	male	58,61		**10**	neutral	57,87	male	93,62
5	angry	41,11	male	79,8		**11**	happy	93,11	female	90,75
6	neutral	90,1	female	87,05		**12**	happy	99,95	male	60,01

The results of the pictures in the Table 1 given as input data in the application are explained in detail in Table 2. As an example, when the number one picture is given as input data on Table 1, the emotional state of the system is measured as "happy" and its rate as "75.94%". In the gender analysis, the ratio of "male" and gender prediction of the result is measured as "82.48".

Table 3. Incorrect results

| (1) | (2) | (3) | (4) | (5) | (6) |
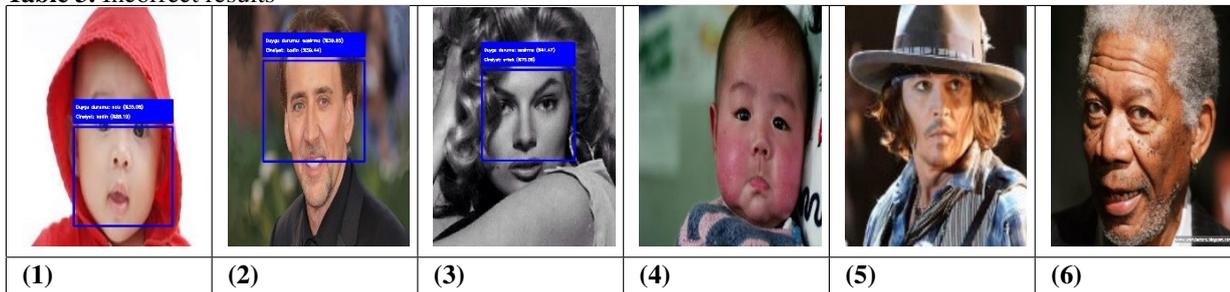
Tablo 4. Details of incorrect results.

Experiments	Emotion	Emotion Rate %	Gender	Gender Rate %		Experiments	Results
1	neutral	35,08	female	88,19		4	No results
2	surprise	39,65	female	59,44		5	No results
3	surprise	41,47	male	70,08		6	No results

When the picture number two is given as the input data on Table 3, the ratio of the emotional state of the system is measured as "39.65%". In the gender analysis, the rate of "female" and gender estimation of the result is measured as "59.44%". The pictures 4, 5 and 6 given in Table 3 do not have a printout on the system.

3 Result and Discussion

With this study, finding the faces on the picture determined by the user, finding the faces in the images, emotion analysis and gender determination with deep learning methods of the face found. Viola-Jones algorithm for face recognition, emotion analysis and gender determination are performed using CNN's "Mini_Xception" model. In order to define incorrectly defined pictures, similar pictures should be added during the teaching of the system. Therefore, more distinctive pictures should be added to the data set and taught to the model. The previously taught xml extension file used for face recognition can be trained with more data sets and better results can be obtained. The structure of CNN's "Mini_Xception" model used for sentiment analysis and gender determination can be changed, a different model can be used or the number of steps can be increased.

Competing interests

The authors declare that they have no competing interests.

Authors' contributions

All authors have contributed to all parts of the article. All authors read and approved the final manuscript.

References

[1] Singh, N., Sabrol, H. Convolutional Neural Networks-An Extensive arena of Deep Learning. A Comprehensive Study. Arch Computat Methods Eng (2021). https://doi.org/10.1007/s11831-021-09551-4

[2] I. Masi, Y. Wu, T. Hassner and P. Natarajan, Deep Face Recognition: A Survey, 31st SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI), 2018, pp. 471-478, 2018, doi: 10.1109/SIBGRAPI.2018.00067.

[3] Arriaga, O., Valdenegro-Toro, M., & Plöger, P. "Real-time convolutional neural networks for emotion and gender classification". arXiv preprint arXiv:1710.07557, 2017.

[4] Omkar M. Parkhi, Andrea Vedaldi, Andrew Zisserman, Deep face recognition, British Machine Vision Association, pp.1-12, 2015.

[5] Samira Ebrahimi Kahou, et al., Emonets: Multimodal deep learning approaches for emotion recognition in video, Journal on Multimodal User Interfaces 10(2), pp. 99-111, 2016.

[6] Zhang, Z., Song, Y., Qi, H., Age progression/regression by conditional adversarial autoencoder. arXiv preprint arXiv:1702.08423v2, 2017.

[7] Zhang, X., Zhao, R., Qiao, Y., Wang, X., Li, H., AdaCos: Adaptively scaling cosine logits for effectively learning deep face representations. arXiv preprint arXiv:1905.00292, 2019.

[8] Yi, D., Lei, Z., Liao, S., Li, S. Z., Learning face representation from scratch. arXiv preprint arXiv:1411.7923, 2014.

[9] Bhatia, A. R., Patel, N. M., Chauhan, N. C., Parallel implementation of face detection algorithm on GPU. 2016 2nd International Conference on Next Generation Computing Technologies (NGCT), 2016, Doi: 10.1109/NGCT.2016.7877497.

[10] Taner Uçkan, Deniz Dal, Opengl Tabanli Animasyonlarda Görüntü Kalitesinin Cuda Mimarisi İle İyileştirilmesi, Uludağ University Journal of The Faculty of Engineering, 21(1), pp. 79-96, 2016.

[11] Jain, V., Patel, D., A GPU based implementation of robust face detection system. Procedia Computer Science Volume 87, Pages 156-163, 2016, Doi: 10.1016/j.procs.2016.05.142.

[12] Minaee, S., Minaei, M., & Abdolrashidi, A., Deep-emotion: Facial expression recognition using attentional convolutional network. Sensors, 21(**??**), 3046, 2021.

[13] Viola, P., Jones, M. J., Robust real-time face detection. International Journal of Computer Vision, Volume 57, Issue 2, pp 137–154, 2004, Doi: 10.1023/B:VISI.0000013087.49260.fb

[14] Wang, Y.-Q., An analysis of the viola-jones face detection algorithm. Image Processing Online, 4 (2014), pp. 128–148, 2014, Doi: 10.5201/ipol.2014.104

[15] Kizrak, Merve Ayyüce, and Bülent Bolat. Derin öğrenme ile kalabalık analizi üzerine detaylı bir araştırma. Bilişim Teknolojileri Dergisi 11.3, 263-286, 2018.

[16] Kaggle, (2013). Challenges in representation learning: facial expression recognition challenge, https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/overview, Access date: 19 January 2021.

[17] Fatima, S. A., Kumar, A., & Raoof, S. S. Real Time Emotion Detection of Humans Using Mini-Xception Algorithm, In IOP Conference Series: Materials Science and Engineering (Vol. 1042, No. 1, p. 012027). IOP Publishing, 2021.

[18] Zhang, Z., Song, Y., Qi, H., (2017). UTKFace - large scale face dataset, https://susanqq.github.io/UTKFace, Access date: 10 March 2021.