International Journal of Computational and Experimental Science and ENgineering
(IJCESEN)

Vol. 9-No.1 (2023) pp. 1-5
http://dergipark.org.tr/en/pub/ijcesen

ISSN: 2149-9144

Research Article

A Novel Shape Descriptor for Object Recognition

Elif Ebru ÇAKI[1]*, Celal Onur GÖKÇE[2]

[1]Afyon Kocatepe Üniversity, Mechatronics Engineering Department, 03204, Afyonkarahisar-Turkiye
* **Corresponding Author** : **Email:** elif_ebru1999@hotmail.com - **ORCID:** 0000-0002-2225-5675

[2]Afyon Kocatepe Üniversity, Software Engineering Department, 03204, Afyonkarahisar-Turkiye
Email: cogokce@aku.edu.tr - **ORCID:** 0000-0003-3120-7808

Article Info:

DOI: 10.22399/ijcesen.1202300
Received : 20 November 2022
Accepted : 17 January 2023

Keywords:

Shape Descriptor
Object Recognition
MNIST

Abstract:

In this study a novel shape descriptor for object recognition is proposed. As a preprocessing stage, Canny edge detection is applied to input images. Output of Canny edge detector, namely edge image, is sampled and various number of points are selected. Chosen points are input to the new shape descriptor. Proposed shape descriptor is composed of deviations from average range and average angle. Shape descriptor is used as a feature extractor output of which is fed to linear classifier. Linear classifier is trained using pseudo-inverse and gradient descent techniques. Full MNIST dataset is used to test the system and results are reported.

1. Introduction

Object recognition is an important subfield of image processing. Several algorithms and methods are developed in this field. One of the most successful fast recognizers is that of Viola Jones [1].
The aim of this study is fast and accurate recognition of objects. Recognition of object means detecting the class of object [2]. In order to classify the image different features can be utilized. Traffic signs, medical images, handwritten character recognition, face recognition and fingerprint recognition are some variants of object recognition. Phases of object recognition process is shown in Figure 1. Image is converted to matrix data structure. In order to be processed by classifier a preprocessing is applied. After feature extraction classification is done.

2. Material and Methods

2.1 MNIST Dataset

One of the mostly used datasets in literature is MNIST (Modified National Institue of Standarts) dataset. It is composed of 60.000 training samples and 10.000 test samples. Each sample is a picture of handwritten digit between 0 and 9. Each digit picture has 28x28 dimension, with 784 pixels of gray levels. Some examples from the MNIST dataset are given in Figure 2 [3].

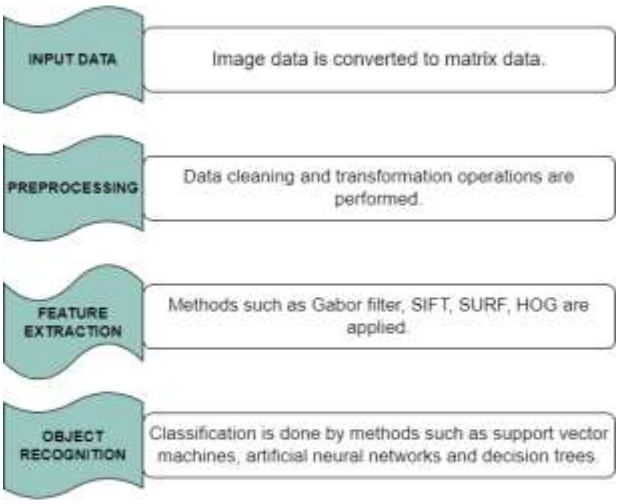


Figure 1. *Object recognition and classification phases*

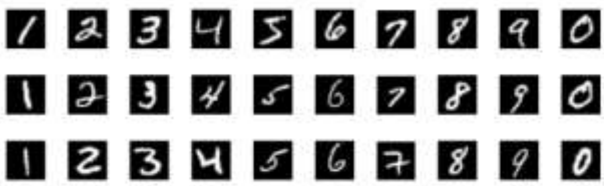


Figure 2. *MNIST dataset examples [3].*

2.2 Proposed Algorithm

Output of Canny edge detector, namely edge image, is sampled using a sub-optimal algorithm [4]. In sub-optimal sampling algorithm, all points are sorted with respect to x coordinates first and y coordinates after. Nearly equally spaced points are sampled within acceptable units of computation time. After sub-optimal algorithm, optimal algorithm is applied to representative points. In optimal algorithm, initially four edge points are selected as east, west, north and south. After these four points, subsequent points are selected optimally that are furthest from every point. Euclidean distance is used between two points $P = (x_1, x_2, \dots, x_n)$ ve $Q = (y_1, y_2, \dots, y_n)$ as shown in Equation 1 [5]. Optimal algorithm is illustrated in Figure 3.

$$\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} \quad (1)$$

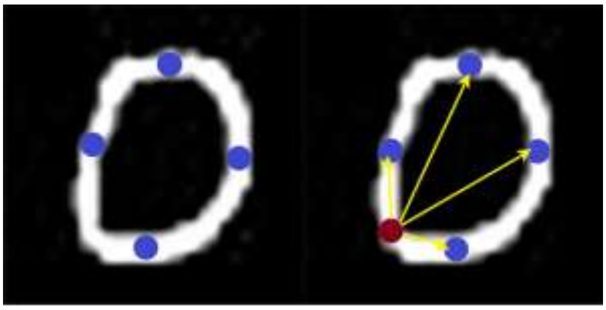


Figure 3. *Optimal algorithm illustrated*

After finding representative subset of edge points, center of mass is found. Distance of each point from center of mass is calculated using Equation 2. Angle of each point to center of mass is calculated using Equation 3. Average distance to center of mass and average angle from center of mass is found and represented as r_m and Θ_m, respectively. For the final descriptor, deviation from average distance and average angle is calculated and normalized as shown in Equation 4 and Equation 5. Distance and angle of representative points are illustrated in Figure 4.

$$r = \sqrt{(x - x_m)^2 + (y - y_m)^2} \quad (2)$$

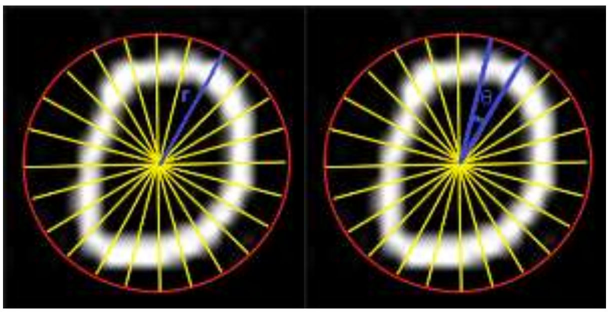


Figure 4. *Distance and angle of representative points*

$$\theta = arctan\left(\frac{y - y_m}{x - x_m}\right) \quad (3)$$

$$\%\Delta r = \frac{r - r_m}{r_m} x100 \quad (4)$$

$$\%\Delta\theta = \frac{\theta - \theta_m}{\theta_m} x100 \quad (5)$$

Since the performance of simple algorithm is low, a histogram based method is added. Histogram based data is entered as input to neural network classifier. As future work, optimal thresholds for histograms will be investigated.

2.3 Histogram Based Method

In image processing, a histogram is a bar graph representation of the color distribution of an image. In the graph, the x-axis gives the gray pixel values and the y-axis gives the number of pixels in that value. There is a Cumulative histogram by obtaining the histogram of the image. The Cumulative histogram contains the values obtained from the sum of each value of the histogram with the previous value and itself. It is normalized by dividing the obtained value by the total number of pixels. Thus, the image is improved by scattering the distribution in pixel values on average.

Let the variable r be the pixel value of the image, that is, r is in the range of $[0,\ L - 1]$. $r = 0$ is black and $r = L - 1$ is white. Since our images are 8 bits in our study, the maximum pixel value is 255. In this case, $T(r)$ gives the total number of transform function and s pixel values.

$$s = T(r) \quad 0 \leq r \leq L - 1 \quad (6)$$

The original data taken from the dataset we used between Figure 5 and Figure 14, the data with the specified 20 points, the angle and distance data, and the histograms of the angle-distance data are given.

Table 1. *Results for various number of neurons*

Number of neurons	Test accuracy	F1 score	Total Loss
128	67.39	0.674	0.968
256	67.9	0.679	0.966
512	69.41	0.694	0.934
1024	69.91	0.699	0.942
2048	68.61	0.686	0.988

Table 2. *Results for various number of epochs*

Number of Epochs	Test accuracy	F1 score	Total Loss
25	69.91	0.699	0.942
50	68.98	0.69	1.08
75	67.75	0.678	1.318
100	66.98	0.67	1.488

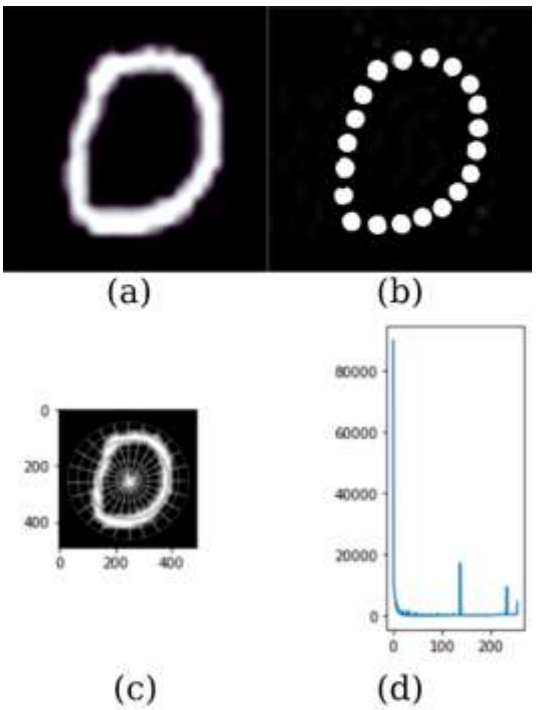


Figure 5. *MNIST dataset- number 0, (a) original data, (b) the data in which it determines the 20 point, (c) data determined by angle and distance, (d) histogram of data c*

MNIST veri seti 0 sayısı

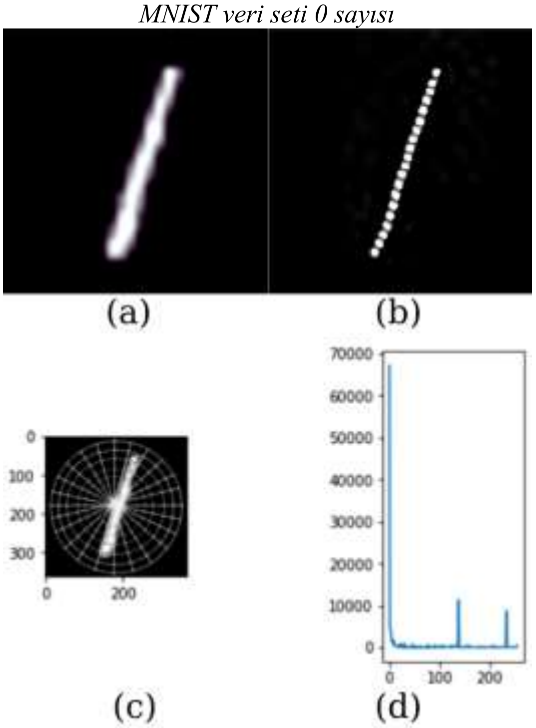
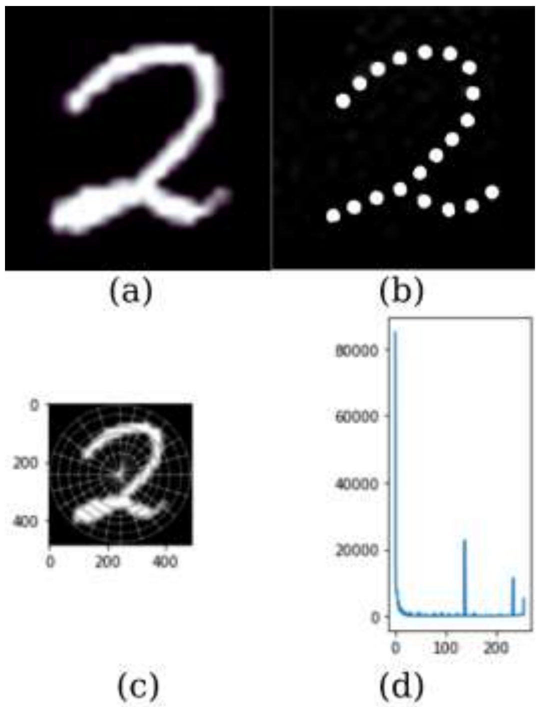


Figure 6. *MNIST dataset- number 1, (a) original data, (b) the data in which it determines the 20 point, (c) data determined by angle and distance, (d) histogram of data*



c

Figure 7. *MNIST dataset- number 2, (a) original data, (b) the data in which it determines the 20 point, (c) data determined by angle and distance, (d) histogram of data c*

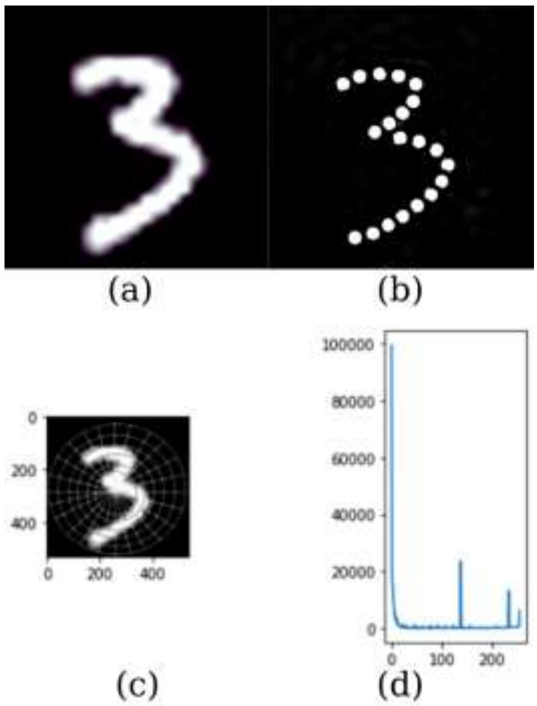


Figure 8. *MNIST dataset- number 3, (a) original data, (b) the data in which it determines the 20 point, (c) data determined by angle and distance, (d) histogram of data c*

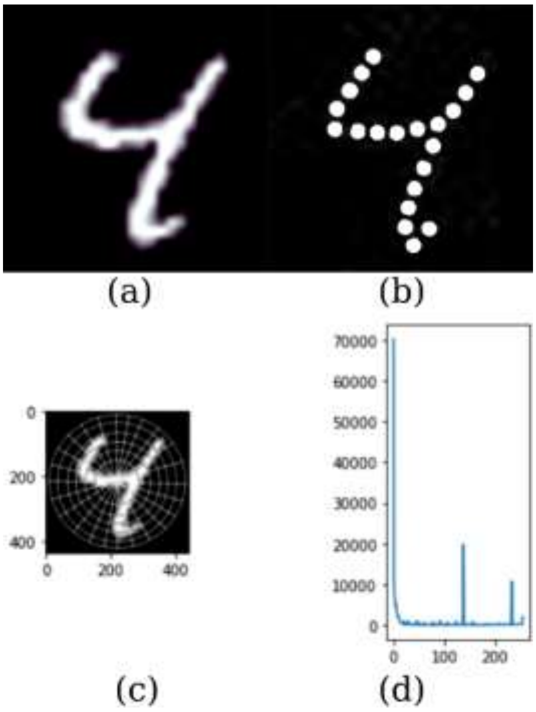


Figure 9. *MNIST dataset- number 4, (a) original data, (b) the data in which it determines the 20 point, (c) data determined by angle and distance, (d) histogram of data c*

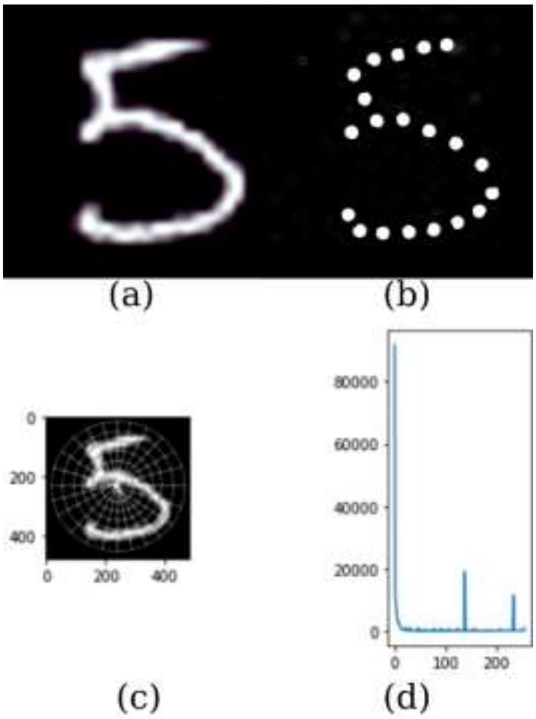


Figure 10. *MNIST dataset- number 5, (a) original data, (b) the data in which it determines the 20 point, (c) data determined by angle and distance, (d) histogram of data c*

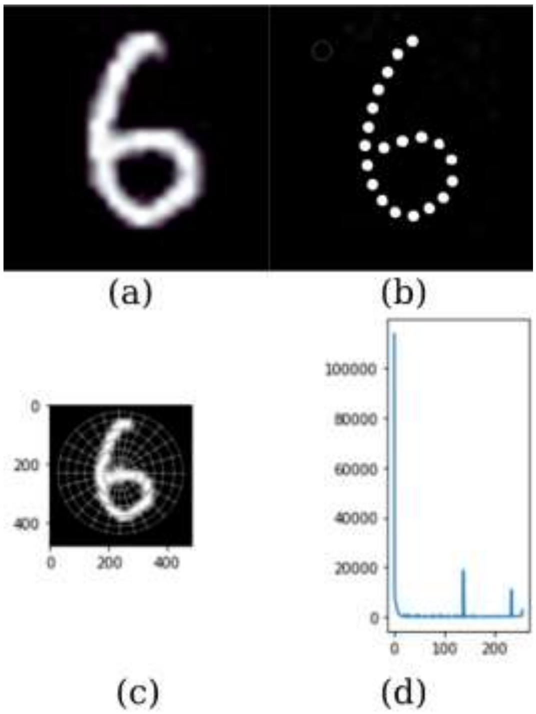


Figure 11. *MNIST dataset- number 6, (a) original data, (b) the data in which it determines the 20 point, (c) data determined by angle and distance, (d) histogram of data c*

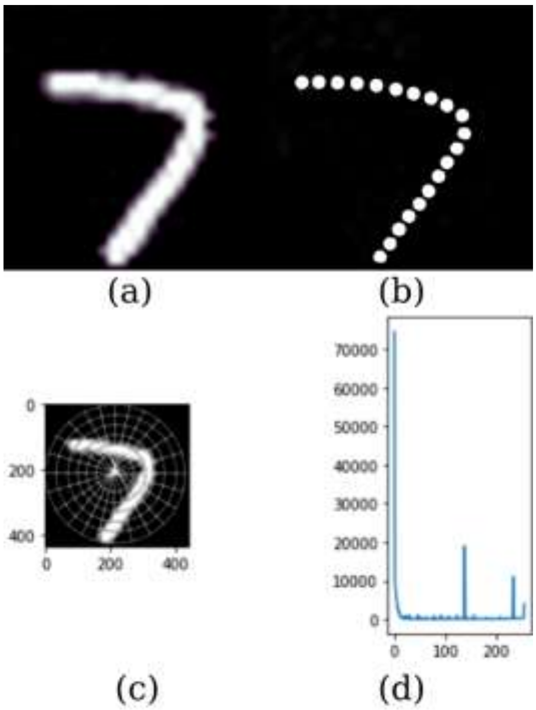


Figure 12. *MNIST dataset- number 7, (a) original data, (b) the data in which it determines the 20 point, (c) data determined by angle and distance, (d) histogram of data c*

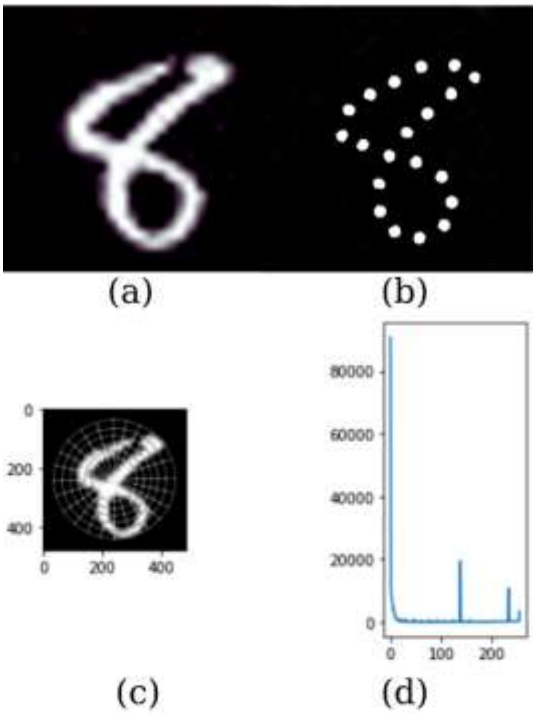


Figure 13. *MNIST dataset- number 8, (a) original data, (b) the data in which it determines the 20 point, (c) data determined by angle and distance, (d) histogram of data c*

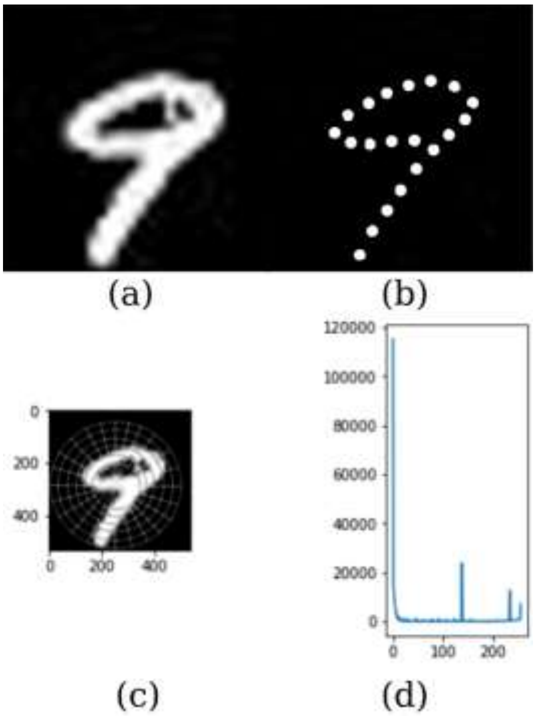


Figure 14. *MNIST dataset- number 9, (a) original data, (b) the data in which it determines the 20 point, (c) data determined by angle and distance, (d) histogram of data*

3. Results and Discussions

A fast shape descriptor is proposed in this study. MNIST dataset is used. A success of 69,91% is obtained using ReLU activation function with 1024 neurons and 25 epochs. Results for various numbers of neurons is given in Table 1. Results for various number of epochs is given in Table 2.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.
- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1]Viola, P. ve Jones, M. (2001, Aralık). Basit özelliklerin artırılmış bir dizisini kullanarak hızlı nesne algılama. Bilgisayarla görü ve örüntü tanıma üzerine 2001 *IEEE bilgisayar topluluğu konferansının Bildirilerinde. CVPR 2001* (Cilt 1, s. II). IEEE.

[2]V. V. Nabiyev, Artificial Intelligence. *Sözkesen Matbaacılık*, Ankara, 2005

[3]Wu, M. ve Zhang, Z. (2010). Mnist veri setini kullanarak el yazısı rakam sınıflandırması. *Ders projesi CSE802: Örüntü Sınıflandırma ve Analizi* .

[4]Sonugur, G., Çaki, E. E., Akan, S. A., & Gökçe, C. O. Gerçek Zamanlı İnsan Davranışı Anlamaya Doğru: Optimal-Altı Bir Şekil Tanımlayıcı. *Afyon Kocatepe Üniversitesi Fen ve Mühendislik Bilimleri Dergisi*, *22*(4); 769-777.

[5]Taşcı, E., & Onan, A. (2016). K-en yakın komşu algoritması parametrelerinin sınıflandırma performansı üzerine etkisinin incelenmesi. *Akademik Bilişim*, 1(1); 4-18.