

MARBLE CLASSIFICATION USING DEEP NEURAL NETWORKS

Murat CANAYAZ^{1*}, *Fatih ULUDAĞ*²


Deep learning, which has been described as the processing and interpretation of data, is now widely used. In this study, deep neural networks are used for the classification of marbles which can be used in the industry. For this purpose most used marbles images were obtained from companies in Turkey and 28-class dataset was created. Then VGG16, ResNet and LeNet models were trained on this dataset. Data augmentation was performed to have class balance. To evaluate the models performance accuracy metric is used. In the VGG16 model, fine tuning was applied and %97 accuracy was achieved. In experimental studies, models were trained with different parameter settings. The performances of the models are given comparatively. The fact that both new dataset and deep neural networks are used for the first time in marble classification are among the positive aspects of this study. It is planned to integrate the models produced in the future studies into mobile based expert systems.

Key words: Marble Classification, Deep Neural Networks, VGG16, ResNet, LeNet

1. Introduction

The increase in the use of natural stones in architecture and decoration, the simplicity in the processing of them as a result of technological developments and the obtaining of them in a more economical way has caused more marble production worldwide. Turkey is located in the Alpine-Himalayan belt the richest marble deposits in the world. Turkey's total marble reserves are estimated to be about 5.1 billion cubic meters, which corresponds to approximately 33 percent of the world reserves [1]. Marble industry has an important role in Turkey's economy. Overall marbles can be divided into eight major groups in Turkey. These marble groups are Marmara types, beige, blacks, whites, cherry colors, onyxs, pinks and greens [2]. In 2016, approximately 860 million dollars of exports were made [3]. Turkey, in 2016, took the first place by achieving %45,7 of world stone exports as marble, travertine etc.

Marble can be found in different colors and textures in nature due to the minerals in it and differences in its formation. In Turkey, more than 80 in different structures and more than 120 in different colors and patterns of marble reserves are determined. Although there is no general standard

¹ Computer Engineering Department, Engineering Faculty, Van Yuzuncu Yil University, 65000, Van, Turkey, (mcanayaz@yyu.edu.tr)  <https://orcid.org/0000-0001-8120-5101>

² Econometrics Department, Faculty Of Economics And Administrative Sciences, Van Yuzuncu Yil University, 65000, Van, Turkey, (fatihuludag@yyu.edu.tr)

for naming marble types, the color in general or the region where they are extracted or both are considered together Elazig Cherry, Mugla White, Burdur Beige etc.

Even marble types have with the same name, there may be color and texture differences. Even field experts have difficulty identifying the type of marble they encounter. This situation may cause problems between suppliers and customers[4].

The classification of the marble species and the determination of the product quality are generally made by experts in the field. Based on the variability of marble types from country to country and from region to region and the range of the color scale, companies need automatic systems.

In the literature, many studies have been done to classify marble species. Hernandez et al. compared the classification success of artificial neural networks with deterministic statistical classification algorithms for control of marble slabs[5]. Machine learning based computer vision applications have recently been used in the classification of marble types and determination of quality levels[6-9].

Doğan and Akay proposed a method based on the AdaBoost algorithm and histogram differences and totals for the classification of marble slabs and achieved a high success for four different marble classes[10]. Lopez et al. used functional support vector machine and functional neural network using data obtained from spectrometer for classification of ornamental Stones[11]. Topalova and Tzokev used multilayered artificial neural networks to classify marble textures in real time[12]. Selver et al. classify the marble slabs using hierarchical clustering algorithm to increase the quality control standards[9].

In the field of computer vision, Deep Learning models have been quite successful [13]. Since the artificial neural networks are insufficient to extract information from the pictures, the Convolutional Neural Networks have been developed[14]. The architecture of convolutional neural networks is inspired by the biological vision system. Compared to standard artificial neural networks with similar layer size, CNNs have less connection, lower number of parameters to be trained [13]. The success of these networks in vision applications is due to two important features: spatially shared weights and spatial pooling[15]. CNN is the more complex form of neural networks. In this multi-layered network, for feature extraction there are 3 main layers, such as convolution layer, pool layer, fully connected layer. Normally in the neural networks, each neuron in the input layer is bound to the next layer by the output neuron. This is called fully connected. In the CNN structure, the FC structure is not used until the last layer [20,21]. In the literature, the most popular area in which CNN is used is image classification.

In this study, LeNet, Visual Geometry Group Convolutional Neural Networks (VGG16) and Residual Neural Network(ResNet) architectures which are frequently used in the literature are used for the classification of marble slabs and their success is compared. The main aim of the study is to investigate models for the automatic identification systems needed in the sector and to integrate effective models with mobile platforms.

The contributions of this paper are as follows:

- It is the first multi-class study in marble classification using Deep Neural Networks.
- It provides new dataset for marble classification
- Models obtained at the end of the study can be used as decision support mechanism
- The models can be easily integrated into mobile applications

2. Related Work

When image classification problems for marble samples are examined, we see that geometric and morphological transformation, color and textures of the image are obtained by using image processing applications before CNN networks. Most of the works are designed to be operated on the production line based on marble industry. From these studies;

Alajarin et al. has designed a system to classify marbles in the production line at the marble industry. In the prototype he used, he analyzed the color texture of the marble surfaces and classified them according to the corresponding group. RGB, XYZ, YIQ, and K-L for image analysis. Color spaces are used, and then feature extraction is provided by basic component analysis. The extracted features and the artificial neural network were used to classify the marbles[8].

Selver et al. followed a two-stage method for marble classification. First, features were extracted using textures and color spaces. These features have been tested with many artificial neural networks. Then classification was made by using hierarchical radial basis function network (HRBFN)[9].

Benavente and Pina presented a mathematical method based on morphology, dividing and classifying polished marble samples. In this method, separate textural regions were determined by watershed method. Morphological procedures were applied to find 3 elements on image (vein, background and transition) and classification was made according to these properties[16].

Akkoyun performed a program to be used in marble classification. This program makes marble quality classification by using image processing applications on marble images obtained from production band[17].

In their study, Doğan and Akay suggested a new hierarchical classification method based on the use of AdaBoost classification algorithms of various types for automatic classification of marble slab images. Features obtained by total and difference histograms method are classified by using Adaboost algorithms. This classification was compared with neural networks and SVM[10].

Selver et al. have developed an automated industrial conveyor belt system using image processing and hierarchical clustering for classifying marble slabs in their study[18].

We could not find a marble classification multi-class study using CNN networks in the literature. However, Ferreira and colleagues used CNN to classify granite tiles[19] and Torun et al. [22] have trained 600 images in 3 classes by Alexnet and Local Binary Pattern method in their study.

3. Experiments

3.1. Experimental Setup

CNN networks such as LeNet, VGGNet16, ResNet were used for the application. Two methods were used in the use of these networks. In the first method, the existing network structures for LeNet and ResNet were preserved. The training was provided by the transfer learning method for the VGG16 network. A Geforce Gtx 1070 GPU graphic card was used for applications on the Ubuntu 16.04 operating system. This reduces the calculation time for each iteration. The application is coded in python language using the Keras library. The used network architectures and dataset will briefly be introduced.

3.1.1 LeNet

LeNet[23] is a CNN structure which is designed to recognize handwriting and machine-printed character recognition by Yann LeCun. The structure of LeNet is simple and small. LeNet as consisting of two parts: (i) a block of convolutional layers; and (ii) a block of fully-connected layers. The LeNet architecture consists of two sets of convolutional, activation, and pooling layers, followed by a fully-connected layer, activation, another fully-connected layer, and finally a softmax classifier.

3.1.2 VGG16

VGGNET [24] developed by Simonyan and Zisserman was first used in ILSVRC 2014. VGG16, consisting of 16 layers deep (13 convolutional layer), is the most preferred network structure for the extraction of attributes from images due to its homogeneous structure. It includes $3 * 3$ filters. The VGGNET structure is publicly available. It requires high calculation costs with 138 million parameters.

3.1.3 ResNet

This model[25], created by the Microsoft team, was developed to eliminate the problem of the convergence of gradient values to 0 in multi-layered deep networks. Thanks to this structure, which is known as residual now, it is possible to get successful results in deep layer networks.

3.1.4 Transfer Learning

Transfer learning is a method used for feature extraction from data set by using VGGNet, Alexnet[13], GoogleNet[26] etc. In this method, the attributes of the images are subtracted without using the upper layers of the network. In addition, in a method called Fine-tuning, upper layers are removed and replaced by other layers in this method. In transfer learning method, firstly, the existing layers must be frozen until the appropriate layers are found. After the network is trained in certain iterations, the existing layers are released and the network is continued with the new layers. In this study, this method is applied to VGGNet. Initially, new layers were added instead of the upper layers, then the existing layers were frozen up to 25 epoch. In the VGG16 model, Flatten, Dense (256), Dropout (256), Dense (classes = 28) layers are added instead of the upper layers that are extracted in the fine tuning process. With this training, the network reaches a certain accuracy rate. A new training process is continued as much as the desired epoch. Fig. 1 shows the transfer learning process.

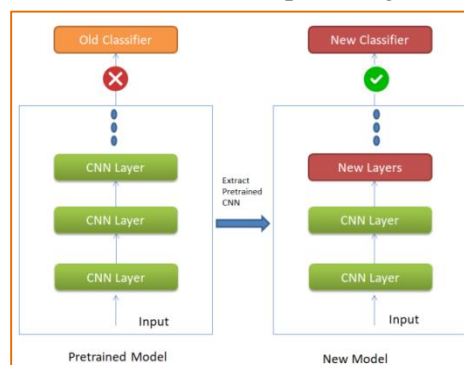


Figure 1. Transfer learning process

3.1.5 Dataset

In Turkey there are more than 300 varieties of marble. In terms of industry, there are 28 marble types which are the most widely used among these types of marble. A data set has been created from marble images obtained from the firms operating in the sector. These images were resized to 224 * 224 by using OpenCV[27] library based on the size of the image used in the CNN network. The number of images is increased via the data agumentation method. The data set consists of nearly 5600 marble images. The number of images in each class is set to be equal to 200. 150 of them were used for training and testing. The remaining 50 is used in the validation process for the first time by the network. Our dataset, which contains a sample from each class, is shown in Fig. 2. The original dataset we have created can be reached below link publicly available [Google Drive Link](#).

In the data agumentation process which we use to increase our data set, class ImageDataGenerator in the Keras library is used. The methods and ratios used are as follows; rotation_range=40,width_shift_range=0.2, height_shift_range=0.2, shear_range=0.2, zoom_range=0.2,horizontal_flip=True, fill_mode='nearest'.

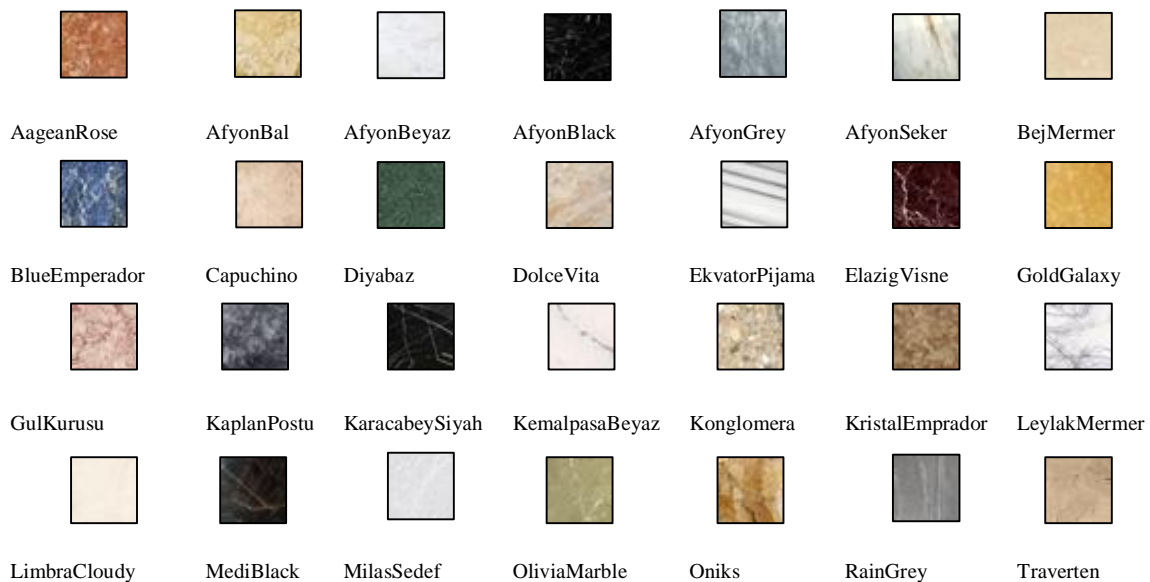


Figure 2. Marble images of 28 class

3.2. Evaluation Metric

After completing the model trainings, accuracy, precision, recall and f1 score metrics are used in the model estimation process. A brief description of these is given below. These metrics are easily obtained with the scikit-learn and keras library in python programming language.

Accuracy(acc) compares the index of the maximum true value to the index of the predicted maximum value, indicating how often the estimates are at the same point as the true values.

Precision (P) is defined as the number of true positives(Tp) over the number of true positives plus the number of false positives(Fp)

$$P = \frac{T_p}{T_p + F_p} \quad (14)$$

Recall (R) is defined as the number of true positives(Tp) over the number of true positives plus the number of false negatives(Fn)

$$R = \frac{T_p}{T_p + F_n} \quad (15)$$

The quantities are also related to the (F1) score, which is defined as the harmonic mean of precision and recall

$$F1 = 2 \frac{P \times R}{P + R} \quad (16)$$

A macro-average calculates the metric independently for each class, and then takes the average, whereas a micro average collects contributions to measure the average measurement of all classes[28].

3.3. Experimental results under different parameter settings

This section shows the results from the models. The models were run until they reached 100 and 1000 epochs. The parameter values used in these models are given in Tab. 1.

Working steps in summary;

- In the study, firstly marble images were collected to create dataset.
- These images were resized to 224 * 224 through OpenCV
- These images are enhanced by data augmentation
- Data set is divided into train, test.
- Marbles are classified by using CNN networks.

In the experiments, 2 pathways were observed in the use of CNN networks. Firstly, the data set that we have created has been trained in CNN networks and the appropriate weight values and high accuracy model have been saved. The second way was the classification of marble by using VGG16 model with the method called fine tuning. The results in the tables were obtained from the test data.

Table 1. Parameter values

Models	Optimizer	Loss Function	Metrics	Batch_Size	Test_Split_Size
LeNet	SGD learning rate:0.01 Adam learning rate:0.001	Categorical_Cross_entropy	Accuracy	128	0.25
VGGNet	RMSprop learning rate:0.001* SGD learning rate:0.001	Categorical_Cross_entropy	Accuracy	32	0.25
ResNet	SGD learning rate:0.1** Momentum:0.9 Adam learning rate:0.001	Categorical_Cross_entropy	Accuracy	128	0.25

*This optimizer is used for the first 25 epoch of training where the layers are frozen. SGD optimizer was used for further training. **In the ResNet model, overfitting was prevented by using L2 norm regularization. For this purpose, the coefficient of regularization for weight reduction is 0.0005

Table 2. Models with SGD optimizer(100 epochs)

Models		precision	recall	f1-score
LeNet	Micro Avg	0.73	0.73	0.73
	Macro Avg	0.75	0.75	0.72
	WeightedAvg	0.75	0.73	0.71
VGG16	Micro Avg	0.96	0.96	0.96
	Macro Avg	0.97	0.96	0.96
	WeightedAvg	0.96	0.96	0.96
ResNet	Micro Avg	0.82	0.82	0.82
	Macro Avg	0.85	0.83	0.82
	WeightedAvg	0.86	0.82	0.81

Table 3. Models with SGD optimizer(1000 epochs)

Models		precision	recall	f1-score
LeNet	Micro Avg	0.93	0.93	0.93
	Macro Avg	0.94	0.94	0.93
	Weighted Avg	0.93	0.93	0.93
VGG16	Micro Avg	0.97	0.97	0.97
	Macro Avg	0.98	0.98	0.98
	Weighted Avg	0.98	0.97	0.97
ResNet	Micro Avg	0.96	0.96	0.96
	Macro Avg	0.97	0.97	0.97
	Weighted Avg	0.97	0.96	0.96

When we look at the results in Tab. 2 of 100 epoch using SGD optimizer, it is seen that VGG16 achieves a better result with approximately 96% precision performance. LeNet model needs more iteration due to its simple structure.

It is possible to see a better performance of VGG16 in Tab. 3 where 1000 epoch is used. ResNet achieved a result close to VGG16 with a performance of 96%. Thanks to its high epoch value, the LeNet model provided a significant performance improvement over the results of 100 epochs.

Tab. 4 shows only the results of LeNet and ResNet models obtained with Adam optimizer. The reason why the VGG16 model is not in this table is that the data set we use with Adam optimizer is worse than the LeNet model in the experimental observations. Looking at the table, it seems that the ResNet model performs well from the Lenet model, but this is not a very satisfactory result for a model.

When we examine Tab. 5, it is possible to say that the Lenet model, which performs 92%, will be suitable for the data set that we use with Adam optimizer. Fig. 3 shows the confusion matrix obtained with validation data on the VGG16 model which provides the highest performance.

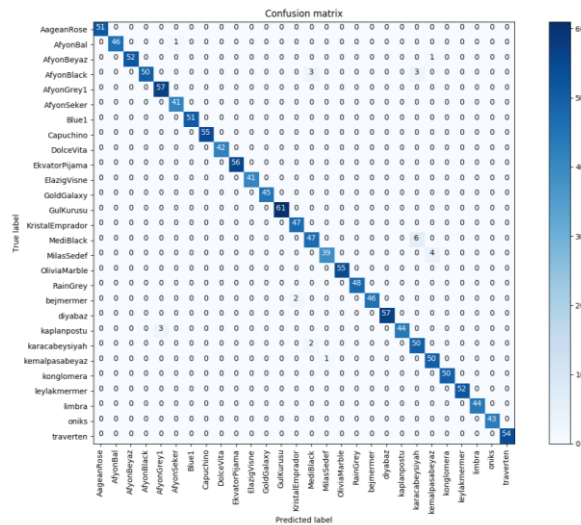


Figure 3. Confusion matrix of validation data on VGG16

Table 4. Models with Adam optimizer(100 epochs)

Models		precision	recall	f1-score
LeNet	Micro Avg	0.77	0.77	0.77
	Macro Avg	0.78	0.77	0.77
	WeightedAvg	0.78	0.77	0.77
ResNet	Micro Avg	0.88	0.88	0.88
	Macro Avg	0.90	0.89	0.87
	WeightedAvg	0.89	0.88	0.86

Table 5. Models with Adam optimizer(1000 epochs)

Models		precision	recall	f1-score
LeNet	Micro Avg	0.92	0.92	0.92
	Macro Avg	0.92	0.92	0.92
	WeightedAvg	0.92	0.92	0.92
ResNet	Micro Avg	0.89	0.89	0.89
	Macro Avg	0.90	0.90	0.90
	WeightedAvg	0.90	0.89	0.89

When the performance of the models are evaluated, it is possible to say that the VGG16 model is better than the other models when it is used with SGD optimizer on the marble data set. In the second performance criterion in which Adam Optimizer is used, LeNet performed better than ResNet in the high epoch number. It was observed that the VGG16 gave even worse results than LeNet when it is used with Adam optimizer. Fig. 4 shows the visualization of activation maps for the first 5 blocks in VGG16. Fig. 5 shows experimental studies on the model obtained by performing transfer learning on VGG16. The red label indicates the estimated label and the green label indicates the actual label value.

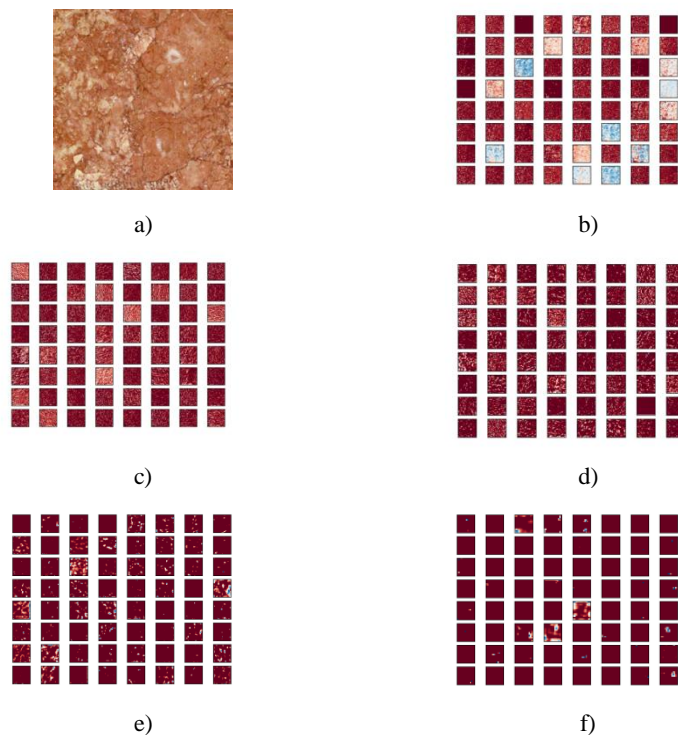


Figure 4. a) Original Image b) Visualization of the feature maps extracted from block 1 c) Visualization of the feature maps extracted from block 2 d) Visualization of the feature maps extracted from block 3 e) Visualization of the feature maps extracted from block 4 f) Visualization of the feature maps extracted from block 5

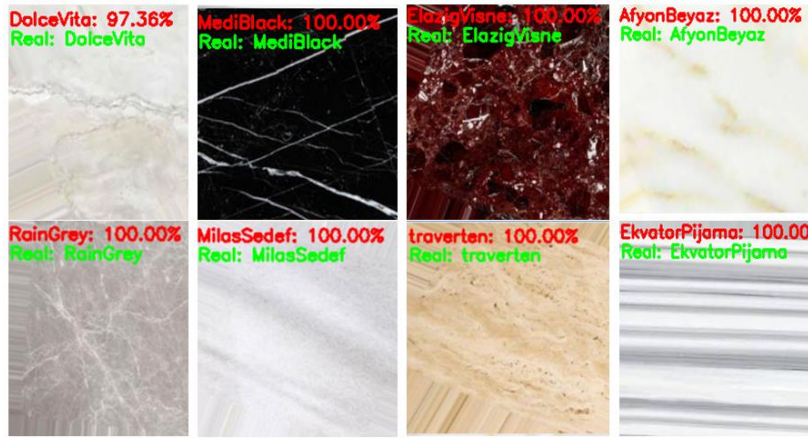


Figure 5. Experimental results of validation data on VGG16

4. Conclusion

The aim of this study, which we used to classify marble images with CNN networks, is preparing dataset for expert systems which can be used in marble industry and presenting our trained model with these datasets to the service of researchers. Within the framework of the research, 28 most familiar images of marble were obtained and these images were extended by data augmentation method and data set was created. The dataset is trained with LeNet, ResNet, VGGNet16, which is one of the most known CNN network architectures. Looking at the results of the training, it is observed that high accuracy rate is achieved. Because some networks used require computational costs, the GPU requirement was created and this requirement was met with the Geforce GTX 1070 GPU. VGGNet16 and ResNet models were completed on an epoch with an average of 18 seconds with this GPU. In LeNet model, this rate was 2 seconds. The evaluation metrics are not only satisfied with accuracy, but also the results are given in metrics such as f score, precision, recall in model estimation process. Since the use of CNN in the multi-class marble classification field was first tried in this study, it is suggested that researchers will give an idea for future studies. The dataset were published as public. The mobile expert system is designed to be developed for future applications.

Acknowledgment

This study is supported Scientific Research Projects Department Project No. FBA-2018-6915 by Van Yuzuncu Yıl University.

References

- [1] Yalçın, S., Uyanık, T. (2001). Dünya mermer ticaretinde Türkiye'nin yeri. *Türkiye III. Mermer Sempozyumu*, Afyon, Türkiye, 397-416.
- [2] Gökay, M. K., Gundogdu, I. B. (2008). Color identification of some Turkish marbles. *Construction and Building Materials*, 22(7), 1342-1349.
- [3] TURKSTAT, Turkey Marble Export, Turkey, 2016
- [4] Bianconi, F., González, E., Fernández, A., Saetta, S. A. (2012). Automatic classification of granite tiles through colour and texture features. *Expert Systems with Applications*, 39(12), 11212-11218.

- [5] Hernández, V. G., Perez, P. C., Pérez, L. G. G., Balibrea, L. M. T., Pina, H. P. (1995). Traditional and neural networks algorithms: applications to the inspection of marble slab. *In Systems, Man and Cybernetics, Intelligent Systems for the 21st Century., IEEE International Conference on*, 5, 3960-3965.
- [6] Martinez-Alajarin, J. (2004). Supervised classification of marble textures using support vector machines. *Electronics Letters*, 40(11), 664-666.
- [7] Juan, M.C.V.A., Balibrea, T., Manuel, L.(1999). Marble Slabs Quality Classification System using Texture Recognition and Neural Networks Methodology. *European Symposium on Artificial Neural Networks Bruges,Belgium*, 75-80.
- [8] Martinez-Alajarin, J., Luis-Delgado J. D., Tomas-Balibrea, L. M. (2005). Automatic system for quality-based classification of marble textures. *In IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 35(4), 488-497, doi: 10.1109/TSMCC.2004.843236
- [9] Selver, M. A., Akay, O., Ardali, E., Bahad, A., Onal, O., Ozden, G. (2009). Cascaded and hierarchical neural networks for classifying surface images of marble slabs. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, 39(4), 426-439.
- [10] Doğan, H., Akay, O. (2010). Using AdaBoost classifiers in a hierarchical framework for classifying surface images of marble slabs. *Expert Systems with Applications*, 37(12), 8814-8821.
- [11] López, M., Martínez, J., Matías, J. M., Taboada, J., Vilán, J. A. (2010). Functional classification of ornamental stone using machine learning techniques. *Journal of Computational and Applied Mathematics*, 234(4), 1338-1345.
- [12] Topalova, C., Tzokev, A. *Adaptive marble plate classification system based on neural network and Plc Implementation*, Annals of DAAAM & Proceedings, 2011.
- [13] Krizhevsky, A., Sutskever, I., Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *In Proceedings of the 25th International Conference on Neural Information Processing Systems*, USA, 1097-1105.
- [14] LeCun, Y. Generalization and network design strategies. Technical Report CRG-TR-89-4, Department of Computer Science, University of Toronto, 1989.
- [15] Kim, Y. (2014). Convolutional neural networks for sentence classification. *In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, Doha, Qatar, 1746–1751.
- [16] Benavente, N., Pina, P. (2009). Morphological segmentation and classification of marble textures at macroscopical scale. *Computers & Geosciences*, 35(6), 1194-1204.
- [17] Akkoyun, O. (2010). An evaluation of image processing methods applied to marble quality classification. *2nd International Conference on Computer Technology and Development*, Cairo, 158-162. doi: 10.1109/ICCTD.2010.5646128
- [18] Selver, M. A., Akay, O., Alim, F., Bardakçı, S., Ölmez, M. (2011). An automated industrial conveyor belt system using image processing and hierarchical clustering for classifying marble slabs. *Robotics and Computer-Integrated Manufacturing*, 27(1), 164-176.
- [19] Ferreira, A., Giraldi, G. (2017). Convolutional Neural Network approaches to granite tiles classification. *Expert Systems with Applications*, 84, 1-11. <https://doi.org/10.1016/j.eswa.2017.04.053>.
- [20] Rosebrock, A. *Deep Learning for Computer Vision*, PyImageSearch, 2017.
- [21] Goodfellow, I., Bengio, Y., Courville, A. *Deep Learning*, MIT Press, 2016.
- [22] Torun, Y., Akbaş, M. R., Çelik M. A., Kaynar, O. (2019). Development a Machine Vision System For Marble Classification. *27th Signal Processing and Communications Applications Conference (SIU)*, Sivas, Turkey, 1-4. doi: 10.1109/SIU.2019.8806419

- [23] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. (1998). Gradient-based learning applied to document recognition. *In Proceedings of the IEEE*, 86(11), 2278-2324.
- [24] Simonyan K., Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *In: CoRR abs/1409.1556*
- [25] He, K., Zhang, X., Ren, S., Sun, J. (2015). Deep residual learning for image recognition. *CoRR, abs/1512.03385*.
- [26] Szegedy C. et al. (2015). Going deeper with convolutions. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, 1-9.
- [27] OpenCV, (2019). <https://opencv.org/>
- [28] Scikit-Learn Documentation, (2019).