ESKİŞEHİR TECHNICAL UNIVERSITY JOURNAL OF SCIENCE AND TECHNOLOGY
A- APPLIED SCIENCES AND ENGINEERING

16th Digital Design In Architecture Symposium
16th DDAS (MSTAS) - Special Issue 2022

2022, Vol.23, pp. 31-41, DOI:10.18038/estubtda.1165890

CLASSIFICATION OF SATELLITE IMAGES WITH DEEP CONVOLUTIONAL NEURAL NETWORKS AND ITS EFFECT ON ARCHITECTURE

Fadime DİKER [1, *], İlker ERKAN [2]

[1] Department of Architecture, Faculty of Architecture, Süleyman Demirel University, Isparta, Turkey
[2] Department of Architecture, Faculty of Architecture, Süleyman Demirel University, Isparta, Turkey

ABSTRACT

Unlike traditional machine learning methods, deep learning methods that can learn from image, video, audio, and text data, especially recently with the increase in hardware power, are also increasing in success. Considering the success and benefits of deep learning methods in many different fields with increasing data, similar effects are expected in architecture. In this study, we focused on textures by going down to specifics rather than general images. In this direction, a total of 4500 satellite images belonging to cloud, desert, green areas and water bodies were classified in the model developed using deep convolutional neural networks. In the developed model, 0.97 accuracy for cloud images, 0.98 accuracy for desert images, 0.96 accuracy for green areas images and 0.98 accuracy for water bodies images were obtained in the classification of previously unused test data (675 images). Although there are similarities in the images of cloud and desert, and images of green areas and water bodies, this success in textures shows that it can be successful in detecting, analyzing, and classifying architectural materials. Successful recognition, analysis and classification of architectural materials and elements with deep convolutional neural networks will be able to facilitate the acquisition of appropriate and useful data through shape recognition among many data, especially at the information collection phase in the architectural design process. Thus, it will help to take more accurate decisions by obtaining more comprehensive data that cannot be obtained from manual data analysis. Learning the distinctive features for classification of data in deep convolutional neural networks also explains architectural design differences and similarities. This situation reveals the hidden relationship in the designs and thus can offer architects the opportunity to make creative and original designs.

Keywords: Deep learning, Deep convolutional neural network, Image classification, Detection of material textures, Architecture

1. INTRODUCTION

The architectural design process is the process of solving design problems and making decisions using data and experience. Today, with the development of computer technologies and hardware, architectural design tools and processes are also developing and changing. Therefore, these tools and methods used by architects or designers also affect the architectural design process. With the continuous development and advancement of technology, data is also increasing rapidly. Machine learning algorithms, an artificial intelligence application belonging to the sub-branch of computer science, are used to recognize these increasing data, analyze them, learn from these data and make decisions in accordance with the information they have learned. Machine learning is an artificial intelligence method that enables the computer to make inferences and gain the ability to learn with various algorithms by using mathematical and statistical operations without using traditional programming methods. Machine learning, which is used in various fields, analyzes much more data than a human can analyze and produces better results, thus making more accurate predictions. It also contributes to automating ordinary tasks and helping in decision-making processes.

The concept of deep learning emerged for the first time with Hinton and Salakhutdinov [1] proposing algorithms that can learn the properties of data hierarchically with deep neural networks. Deep learning is an artificial intelligence method that uses multi-layer artificial neural networks and is a sub-branch of machine learning methods. The use of deep learning is increasing rapidly today with the increase in

*Corresponding Author: fadimediker@gmail.com
Received: 05.09.2022 Published: 23.12.2022

hardware power (graphic processor units - GPU) and processing power that can process increasing data. Unlike traditional machine learning methods, it can also learn from data belonging to images, videos, audio, and texts. In addition, unlike traditional machine learning, deep learning can perform attribute extraction itself with little or no computer intervention. With deep learning, which is a sub-branch of machine learning, it has become easier to define, classify and process the increased data. Considering the success and benefits of machine learning in various fields it is expected that similar success and benefits will be achieved in the discipline of architecture.

Various machine learning methods are used to classify architectural designs and elements visually with computer vision techniques. Learning the distinguishing features for classification also helps to explain the design differences and similarities. Yoshimura et al., [2] used a deep convolutional neural network (DCNN) model to classify the works of 34 different architects, and this model divided the images into classes depending on the visual similarities measured by the algorithm. Llamas et al., [3] used convolutional neural networks (CNN) for classification of architectural heritage images and stated that the application of these techniques can significantly contribute to the digital documentation of architectural heritage. Similarly, Obeso et al., [4] used convolutional neural network for classification of architectural styles of buildings in digital photographs of Mexican cultural heritage and stated that style identification with this technique can make a wide contribution in video description tasks, especially in automatic documentation of cultural heritage.

The classification process is used to automate some tasks at the architectural design stage. With the development of technology, the number of architectural elements increases as the building becomes more sophisticated in 3D models, and therefore, architects usually separate each geometry into semantically correct layers after the draft model is finished in order to model faster in the first stage of schematic drawings. The work of separating and labeling the geometries into individual layers is an ordinary task that does not require special knowledge. Yetis et al., [5] aimed to automate this work of architects and designers in order to reduce this workload and improve work performance. For this purpose, they applied and compared 5 machine learning models, logistic, k-nearest neighbors algorithm (KNN), support vector machines (SVM), naive nayes and decision tree, to label architectural elements in various parametric design environments (Rhinoceros, Grasshopper, Grasshopper Python and Grasshopper Python Remote).

The classification process can also be used in the architectural design process to correct situations that are impossible or difficult to change after the design process. Diker and Erkan [6], in their study, divided the window design efficiency of classrooms into 7 classes using fuzzy logic method, which is an artificial intelligence method. They stated that by using the developed model in the early design stage, it can enable the creation of window designs that provide sufficient visual comfort in classrooms by having pre-design knowledge. Some decisions made for the structural system during the architectural design process may need corrections in the future, and this causes losses in both time and cost. Bingöl et al., [7] created an Irregularity Control Assistant that can provide general information to architects about the compliance of structural system decisions with earthquake regulations by using deep learning and image processing methods to solve such problems at an early stage of architectural design. They stated that with the Irregularity Control Assistant they have created, it will enable correct decisions to be taken at the early stage of architectural design and reduce the corrections that may occur during the implementation project stage.

The development of hardware power that can process the increasing data enables the processing of visual data. This development will play an important role, especially in the processing of architectural data. Contrary to the previous studies, the images used in this study focus on the textures by going down more specifically, rather than the more general image (images based on function, design style and building type). Accordingly, in this study, the success of classifying satellite images with different textures of nature using deep neural networks was evaluated. It is expected that the use of image processing in the visuals of architectural structures will provide a different perspective to the detection and classification of materials, with the performance obtained because of the classification of satellite images with different properties of deep convolutional neural networks. Thus, various deep learning methods can be used to classify architectural designs, elements, and materials visually with computer vision techniques. In addition, a total of 4500 satellite images of cloudy (1125), desert (1125), green area (1125) and water (1125) were used to use a balanced data set for the accuracy of the model.

2. MACHINE LEARNING AND DEEP LEARNING

Machine learning is an application of artificial intelligence that provides computers the ability to learn and improve automatically from experience without explicitly programming it. In line with learning through experience, Mitchell [8] defines machine learning as follows: If the performance of a computer program on tasks at T (task), as measured by P (performance), increases with experience of E (experience), some task class learns from experience of E with respect to T and performance measure P. That's why machine learning focuses on the development of computer programs that can access data and use them to learn. Traditional programming logic focuses on obtaining output data from the input data, and therefore the output data to establish the appropriate program for the type of problem while using the software or by training machine learning machine with logic input and output data for the type of problem is obtained from the appropriate program or software [9]. In this direction, the primary purpose of machine learning is to enable computers to learn automatically without human intervention or assistance and to adjust actions accordingly. Machine learning allows them to make more accurate inferences based on their past experiences, now or in the future, and thus machines can help make the right decisions.

Deep learning is a subfield of machine learning concerned with algorithms inspired by the structure and function of the brain, called artificial neural networks. Multilayer deep neural networks are used in deep learning. Deep neural network algorithms are a multi-layered form of artificial neural networks developed by taking inspiration from the human brain. Deep neural networks are expressed as a neural network with an input layer, an output layer, and multiple hidden layers. In deep neural networks, the data learned at each layer constitutes the input data of the next layer, and thus a network structure is formed in which features are learned from the first layer to the last layer. Neurons in the structure of deep neural networks have weight values, bias and activation function. After the neurons are multiplied by the input value and the weights, an output value is obtained by adding the bias value. Various activation functions (tanh, sigmoid, relu) are used to control this output value obtained (to decide whether the neuron can be active or not).

3. METHODOLOGY

3.1. Deep Convolutional Neural Network (DCNN)

Deep convolutional neural networks contain 3 basic layers as convolution layer, pooling layer and fully connected layer, as well as some hyperparameters such as stride, pixel padding, filter (core), activation functions.

3.1.1. Convolution layer

It is the first layer used to extract different properties from an input image. The input image presented to this layer is treated as a matrix. In the layer, the properties of the image are obtained by applying a certain number of filters [10]. Filters are moved over the images and perform matrix multiplications. The values obtained from the multiplications are added and the resulting value is revealed.

3.1.2. Pooling layer

The image is transferred to the pooling layer after the convolution layer. The pooling layer uses the incoming data to create an output vector containing smaller and more meaningful information. A sizing matrix of the size specified in the pooling layer is applied. This sizing matrix is applied according to the step shift value on the image. Different pooling processes are available. In general, maximum pooling and average pooling are used. If the maximum values contained in the matrix are taken, the maximum pooling method is applied, or if the average of the values is taken, the average pooling method is applied. The main purpose of this layer is to reduce the number of parameters in the network [11]. As a result of this layer, there is a decrease in size, information loss occurs. The resulting loss of information is beneficial for the neural network, as it creates less computational load on the next network layers and prevents the system from memorizing [12, 13]. Reducing the parameters also provides incompatibility control in the network.

3.1.3. Fully connected layer

In convolutional neural network architectures, the flatten process is applied before this layer to use the matrices obtained from the convolution and pooling layers one after the other in the fully connected layer. As a result of the flattening process, the input data of the fully connected layer becomes ready. Operations are performed on coefficients from hidden layers. After the coefficient operations, the data are correlated with the selected density function and produce an output value [14]. The output layer of the network is the part where the result values are labeled. In this layer, neurons are connected to each subsequent neuron one-to-one, and therefore this layer is called the fully connected layer. The purpose of the fully connected layer is to use these high-level features to classify the input image into various classes based on the training dataset.

3.1.4. Classification layer

It is the last layer of the deep convolutional neural network model applied for classification problems. Since this layer is the layer where the classification process is performed, the number of output values is equal to the number of classes of the data used in the model. In the architecture of deep convolutional neural networks, the SoftMax classifier is generally used in this layer [15]. In this layer, the classifier predicts probability values between 0 and 1 for each class, and as a result, the class with the highest probability value becomes the class predicted by the model [16].

3.2. Performance Evaluation Criteria

It is necessary to evaluate the performance of the models to measure the success of the developed models or to decide whether the model is a good model. Different performance evaluation methods are used for classification in supervised learning. In this study, confusion matrix, accuracy, precision, recall (sensitivity) and F1-score ratios were used as performance evaluation criteria.

3.2.1. Confusion matrix

It is an analysis tool that shows the extent to which a classifier can classify different class labels. In other words, it is an n*n matrix that shows the number of correct and incorrect predictions made by comparing the results obtained by a classification model with the actual results. If the number of class labels in the data set is n, the matrix size is also formed in the form of n*n. The following four evaluations are used for classification estimates [17]:

I. True Positive (TP): They are positive class labels that have been correctly predicted by the model.
II. True Negative (TN): They are negative class labels that have been correctly predicted by the model.
III. False Positive (FP): They are positive class labels that have been incorrectly predicted by the model.
IV. False Negative (FN): They are negative class labels that have been incorrectly predicted by the model.

3.2.2. Accuracy

It is expressed as the ratio of correct predictions in the model to all predictions (Equation 1).

$$Accuracy = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative}} \tag{1}$$

3.2.3. Precision

It is a success criterion that shows how many of all samples predicted as positive in the model are actually classified correctly (Equation 2).

$$Precision = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \tag{2}$$

3.2.4. Recall (Sensitivity)

It is a criterion that shows how successfully positive situations are predicted (Equation 3).

$$Recall\ (Sensitivity) = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{3}$$

3.2.5. F1 score (F score)

Precision and recall (sensitivity) are the harmonic mean of the criteria (Equation 4).

$$F1\ score\ (F\ score) = 2 * \frac{\text{Precision} * \text{Recall (Sensitivity)}}{\text{Precision} + \text{Recall (Sensitivity)}} \tag{4}$$

3.3. Developed Model

In the model, the data set shared as "Satellite Image Classification Dataset-RSI-CB256" on the "Kaggle" platform was used. A total of 5631 satellite images of cloudy (1500), desert (1131), green area (1500) and water (1500) mixed from sensors and google map snapshot are available in this data set. However, 1125 images belonging to each class were used in order to use a balanced data set (Figure 1).

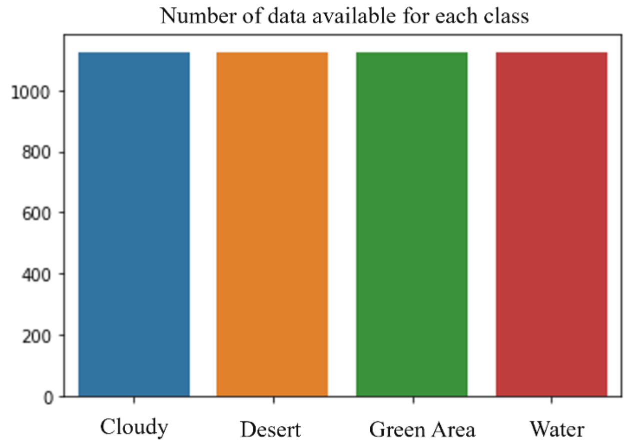

Figure 1. Number of data available for each class

Some sample images of the classes from the data set used are shown in Figure 2.

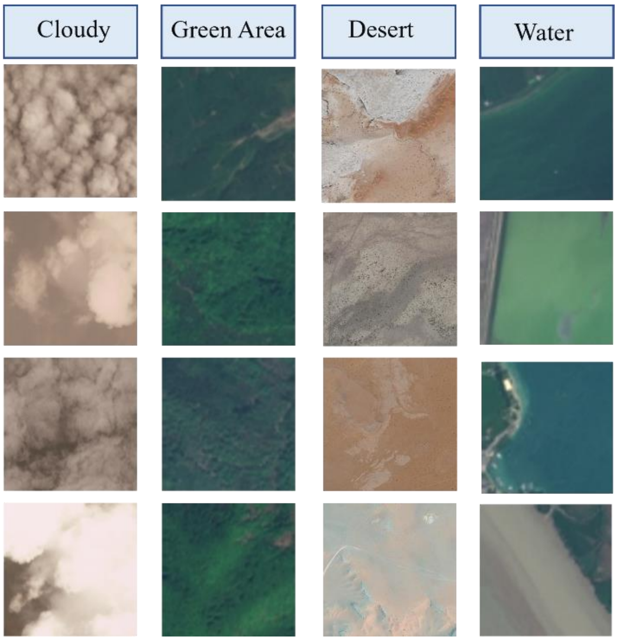

Figure 2. Examples of data sets of satellite images used

The design of the developed deep convolutional neural networks consists of 4 double convolution layers and a maximum pooling layer repeated 4 times, followed by a flattening layer and 4 densely connected layers. The core size is 3×3 for all convolutional layers, and the core size is 2×2 for all maximum pooling layers. ReLU (Rectified Linear Unit) activation function is used as an activation function in hidden layers. Since the output layer has 4 different classes, the SoftMax function is used as the activation function. For the developed deep convolutional neural network model, the Google Colaboratory program, which is offered free of charge by Google, was used with the Python programming language. The general architecture of the developed deep convolutional neural network model is shown in Figure 3.

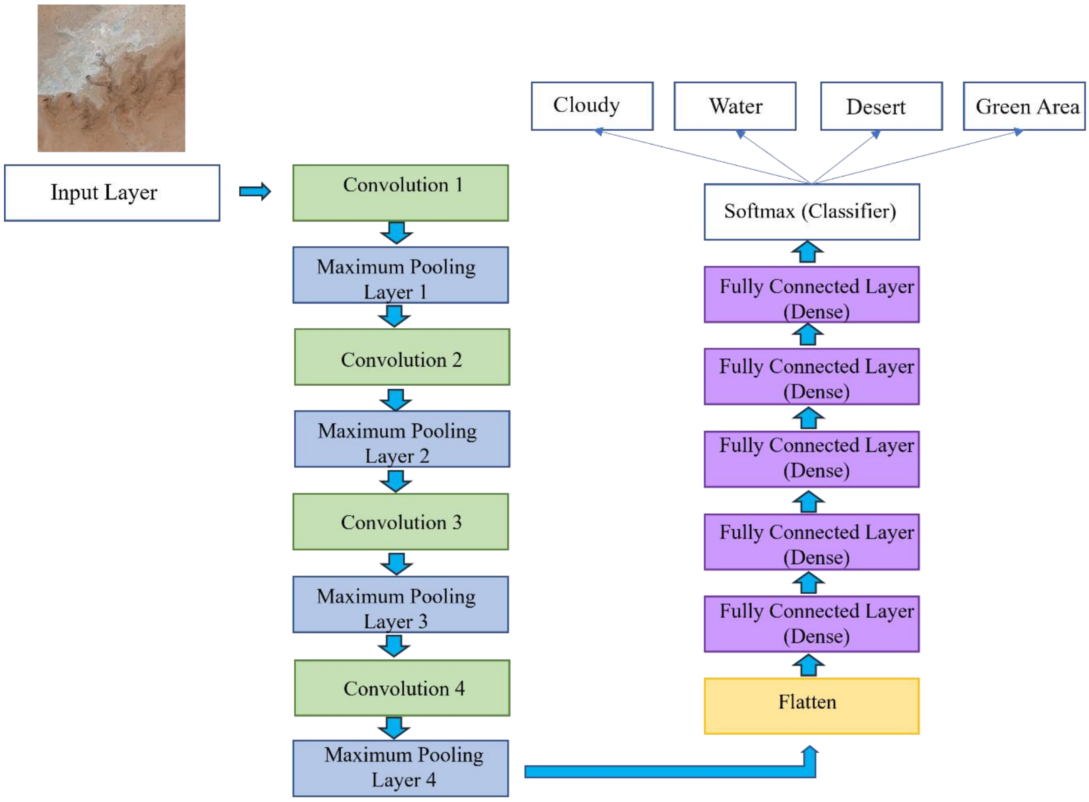

Figure 3. The developed Deep Convolutional Neural Network architecture

3.4. Results of the Model

75% of the dataset was used for training, 15% for testing, and the remaining 15% for validation. In the training of the developed model, 100 epochs, 0.0001 learning rate and Adam optimization algorithm were used. Accuracy and loss rates in training and validation during the epochs are shown in Figure 4-5.

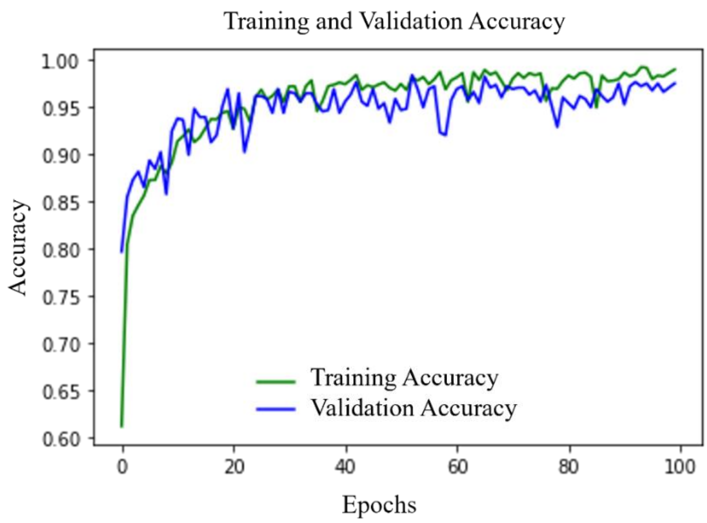

Figure 4. Training and validation accuracy graph obtained from the developed model

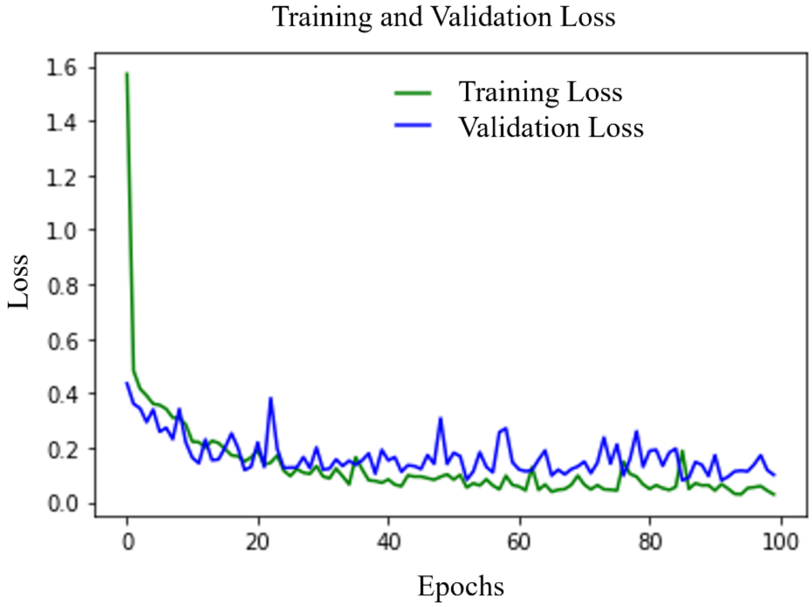

Figure 5. Training and validation loss graph obtained from the developed model

The performance of the developed deep convolutional neural network model was performed using different evaluation criteria. The estimation efficiency of the model was evaluated using four different performance evaluation criteria, namely accuracy, precision, recall (sensitivity) and F1-score. These values are calculated based on confusion matrices for each class. The confusion matrix obtained from the developed model is shown in Figure 6 and the results of the performance criteria evaluated according to this matrix are shown in Table 1.

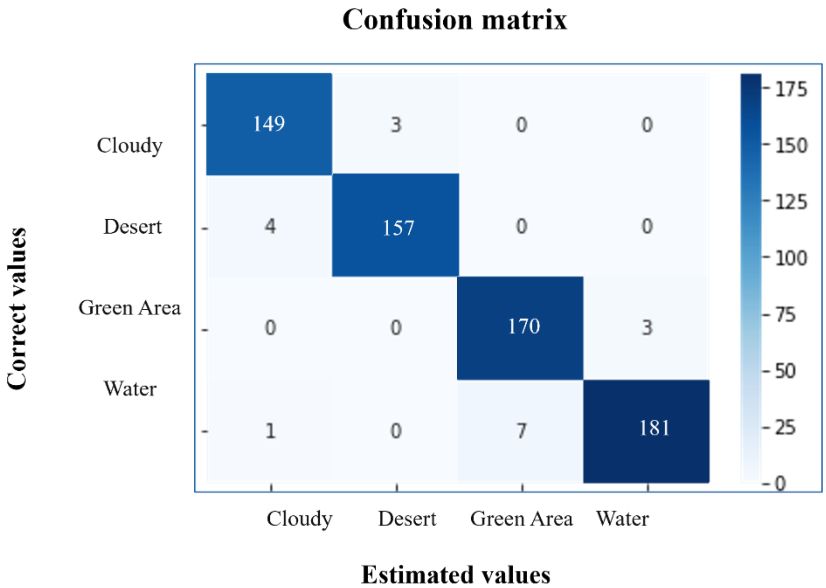

Figure 6. The confusion matrix obtained from the developed model

Table 1. Performance results of the developed model

	Cloudy	Desert	Green area	Water
Accuracy	0.97	0.98	0.96	0.98
Precision	0.97	0.98	0.96	0.98
Recall (Sensitivity)	0.98	0.98	0.98	0.96
F1 score (F score)	0.97	0.98	0.97	0.97

4. CONCLUSIONS AND RECOMMENDATIONS

With the continuous development and progress of technology, machine learning algorithms, which are an artificial intelligence application belonging to the sub-branch of computer science, are used to recognize, analyze, learn from these data, and make decisions based on what they have learned. One of the machine learning methods, especially with the increase in hardware power, the processing of visual data with deep learning can facilitate the acquisition of appropriate and useful data with shape recognition among many data, especially in the information collection stage in the architectural design process. In addition, it can help to make more accurate decisions by obtaining more comprehensive information (obtaining information that may be missed in manual searches). While inconsistent and incorrect results can be obtained with traditional methods, the error rate can be reduced using machine learning methods.

As a result of the model developed with deep convolutional neural networks, satellite images were successfully classified. Even if the desert and cloudy images are like each other, all but 8 images were classified correctly in the test data (675 images). Only one of these 8 images is classified as water. Similarly, although the images of the green area and the water are similar, all images were classified correctly, except for 10 images. This situation has shown that successful results can be obtained in the visual processing of architectural structures. In addition, unlike other studies, achieving this success in textures has shown that it can be successful in detecting, analyzing, and classifying architectural materials found in the visuals of architectural structures. Learning the distinctive features of data for classification in machine learning also explains design differences and similarities, and thus reveals hidden relationships in designs. This will allow architects and designers to make more creative and original designs. In addition to these, the ability to process visual data can facilitate the acquisition of appropriate and useful data with shape recognition among many data, especially during the information collection stage in the architectural design process, and thus, it will help to take more accurate decisions by obtaining more comprehensive information.

The classification process can be used to automate some work during the architectural design stage. For example, with the development of technology, the number of architectural elements increases as the building becomes sophisticated in 3D modeling, and therefore architects usually divide each geometry into semantically correct layers after the draft model is finished to model faster in the first stage of schematic drawings. By automating some monotonous work in the architectural design process with classification, it can reduce the workload of architects and designers and improve working performance. It can also be used for archiving increasing data (architectural artifacts, styles of buildings, etc.) with classification. This reduces the workload in digital documentation and contributes to its automation. With the model used in this study, it will be possible to classify, separate and even determine the distinguishing features of architectural elements. The increase in the number of architectural elements, the diversity of construction systems together with the developing technology reveals the importance of classification processes. In the model developed with the study, it is foreseen that façade applications, material classification and even data archiving can be done even by being inspired by the satellite images that are successfully classified. In addition, it is thought that additional design parameters may arise with the acquisition of data that cannot be obtained with manual data, which will provide designers with different perspectives.

CONFLICT OF INTEREST

The author(s) stated that there are no conflicts of interest regarding the publication of this article.

REFERENCES

[1] Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks, Science 2006; 313(5786): 504-507.

[2] Yoshimura Y, Cai B, Wang Z, Ratti C. Deep learning architect: classification for architectural design through the eye of artificial intelligence. In International Conference on Computers in Urban Planning and Urban Management, 2019; 249-265.

[3] Llamas JM. Lerones P, Medina R, Zalama E, Gómez-GarcíaBermejo J. Classification of architectural heritage images using deep learning techniques, Applied Sciences, 2017; 7(10): 992.

[4] Obeso AM, Vázquez MSG, Acosta AAR, Benois-Pineau, J. Connoisseur: classification of styles of Mexican architectural heritage with deep learning and visual attention prediction. In Proceedings of the 15th international workshop on content-based multimedia indexing, 2017; 1-7.

[5] Yetiş G, Yetkin O, Moon K, Kılıç Ö. A novel approach for classification of structural elements in a 3d model by supervised learning. In Proceedings of the 36th eCAADe Conference, 2018; 129-136.

[6] Diker F, Erkan İ. The fuzzy logic method in assessing window design for the visual comfort of classrooms at the early design stage, Journal of Architectural Engineering, 2022; 28(2), 04022013.

[7] Bingöl K, Aslı ER, Örmecioğlu HT, Arzu ER. Depreme dayanıklı mimari tasarımda yapay zeka uygulamaları: Derin öğrenme ve görüntü işleme yöntemi ile düzensiz taşıyıcı sistem tespiti, Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, 2022; 35(4): 2197-2210.

[8] Mitchell TM. Machine learning. McGraw-Hill,1997; 1(9): 414p.

[9] Kavuncu SK. Makine öğrenmesi ve derin öğrenme: Nesne tanıma uygulaması. Kırıkkale Üniversitesi, Fen Bilimleri Enstitüsü, Master's Thesis, 2018; 157p.

[10] Narın D, Onur TÖ. Investigation of the effect of edge detection algorithms in the detection of covid-19 patients with convolutional neural network-based features on chest x-ray images. In 2021 29th Signal Processing and Communications Applications Conference (SIU), 2021; 1-4.

[11] Önal MK, Avci E, Özyurt F, Orhan A. Classification of minerals using machine learning methods. In 2020 28th Signal Processing and Communications Applications Conference (SIU), 2020; 1-4.

[12] İnik Ö, Ülker E. Derin öğrenme ve görüntü analizinde kullanılan derin öğrenme modelleri, Gaziosmanpaşa Bilimsel Araştırma Dergisi, 2017; 6(3): 85-104.

[13] Ergün GB, Güney S, Ergün TG. Köpeklerdeki uzun kemiklerin evrişimsel sinir ağları kullanılarak sınıflandırılması, Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 2021; 33(1): 125-132.

[14] Yamashita R, Nishio M, Do RKG, Togashi K. Convolutional neural networks: an overview and application in radiology, Insights into imaging, 2018; 9(4): 611-629.

[15] Bozkurt F, Yağanoğlu M. Derin evrişimli sinir ağları kullanarak akciğer x-ray görüntülerinden covid-19 tespiti, Veri Bilimi, 2021; 4(2): 1-8.

[16] Mining WID. Data mining: Concepts and techniques. Morgan Kaufinann, 2006; 10, 559-569.

[17] Uğuz S. Makine öğrenmesi teorik yönleri ve python uygulamaları ile bir yapay zeka ekolü. Nobel Akademik Yayıncılık, 2019; 298p.