Dokuz Eylül Üniversitesi Mühendislik Fakültesi Fen ve Mühendislik Dergisi
Dokuz Eylul University Faculty of Engineering Journal of Science and Engineering

Basılı/Printed ISSN: 1302-9304. Elektronik/Online ISSN: 2547-958X

Performance Comparison of CNN Based Hybrid Systems Using UC Merced Land-Use Dataset

UC Merced Land-Use Veri Kümesi Kullanılarak CNN Tabanlı Hibrit Sistemlerin Performans Karşılaştırması

Fatma Günseli Yaşar Çiklaçandir [1*], Semih Utku [2]

[1] İzmir Katip Çelebi University, Faculty of Engineering and Architecture, Department of Computer Engineering, İzmir, TÜRKİYE
[2] Dokuz Eylül University, Faculty of Engineering, Department of Computer Engineering, İzmir, TÜRKİYE
Corresponding Author *: fatmagunseli.yasar@ikcu.edu.tr

Geliş Tarihi / Received: 03.01.2023
Kabul Tarihi / Accepted: 08.03.2023
Araştırma Makalesi/Research Article
DOI:10.21205/deufmd.2023257516
Atıf şekli/ How to cite: YAŞAR ÇIKLAÇANDIR, F.G., UTKU, S. (2023). Performance Comparison of CNN Based Hybrid Systems Using UC Merced Land-Use Dataset. DEUFMD, 25(75), 725-737.

Abstract

Remote sensing is the technology of collecting and examining data about the earth with special sensors. The data obtained are used in many application areas. The classification success of remote sensing images is closely related to the accuracy and reliability of the information to be used. For this reason, especially in recent studies, it is seen that Convolutional Neural Network (CNN), which has become popular in many fields, is used and high successes have been achieved. However, it is also an important need to obtain this information quickly. Therefore, in this study, it is aimed to get results as successful as CNN and in a shorter time than CNN. Hybrid systems in which features are extracted with CNN and then classification is performed with machine learning algorithms have been tested. The successes of binary combinations of two different CNN architectures (ResNet18, GoogLeNet) and four different classifiers (Support Vector Machine, K Nearest Neighbor, Decision Tree, Discriminant Analysis) have been compared with various metrics. GoogLeNet & Support Vector Machine (93.33%) is the method with the highest accuracy rate, while ResNet18 & Decision Tree (50.95%) is the method with the lowest accuracy rate.

Keywords: *Remote Sensing, Convolutional Neural Network, Machine Learning, Hybrid Systems*

Öz

Uzaktan algılama, yeryüzü ile ilgili verilerin özel sensörler aracılığıyla toplanması ve incelenmesi teknolojisidir. Elde edilen veriler birçok uygulama alanında kullanılmaktadır. Uzaktan algılama görüntülerinin sınıflandırma başarısı, kullanılacak bilgilerin doğruluğu ve güvenilirliği ile yakından ilgilidir. Bu nedenle özellikle son yıllarda yapılan çalışmalarda birçok alanda popüler hale gelen Konvolüsyonel Sinir Ağlarının (CNN) kullanıldığı ve yüksek başarılar elde edildiği görülmektedir. Ancak bu bilgilerin hızlı bir şekilde elde edilmesi de önemli bir ihtiyaçtır. Dolayısıyla bu çalışmada CNN kadar başarılı ve CNN'den daha kısa sürede sonuç alınması amaçlanmaktadır. CNN ile özniteliklerin çıkarıldığı ve daha sonra makine öğrenmesi algoritmaları ile sınıflandırmanın yapıldığı hibrit sistemler test edilmiştir. İki farklı CNN mimarisi (ResNet18, GoogLeNet) ve dört farklı sınıflandırıcının (Support Vector Machine, K Nearest Neighbor, Decision Tree, Discriminant Analysis) ikili kombinasyonlarının başarıları çeşitli metriklerle karşılaştırılmıştır. GoogLeNet & SVM (%93,33) en yüksek doğruluk oranına sahip yöntem olurken, ResNet18 & DT (%50,95) en düşük doğruluk oranına sahip yöntemdir.

Anahtar Kelimeler: *Uzaktan Algılama, Konvolüsyonel Sinir Ağı, Makine Öğrenmesi, Hibrit Sistemler*

1. Introduction

Photogrammetry and remote sensing were not used until the 1960s, although studies on land use and land cover determination date back to the early 1800s [1]. With the activation of Landsat-1, the first earth observation satellite, in 1972, thematic maps such as land use types and land cover types began to be produced widely as a result of the evaluation of the images obtained using various classification algorithms. Since these years, many satellites have been in operation and solutions are produced for different purposes with different imaging options according to different user requests. With the production of thematic maps based on the classification of remotely sensed data, the accuracy of the maps obtained has been questioned and many studies have been carried out to determine the accuracy of the classified images from the 1970s to the present.

Satellite images, which are the basic products of remote sensing technologies, are an important source for obtaining important information about the Earth [2]. Thanks to the evaluation of the data obtained by remote sensing technologies, reliable information about the planet and its surroundings is obtained [3]. Remotely sensed image analysis can be used in different areas such as producing maps. The most commonly used method for producing these maps is the classification of satellite images. The accuracy rates of the maps obtained as a result of the classification have direct effects on the results. Therefore, many classification methods have been tested on remotely sensed images.

The success rates of Multi-Layer Perceptron (MLP), K Nearest Neighbor (KNN), J48, Naive Bayes, Bayes Net, KStar algorithms have been compared in the study using the Urban Land Cover data set [4]. The MLP algorithm is the algorithm that performs the classification with the highest success. IKONOS satellite images were used to do object-based crop pattern recognition in agricultural areas [5]. After the features extracted using three different deep learning architectures (AlexNet, VGG-16, and GoogleNet) have been reduced by neighborhood component analysis, these features have been classified by SVM in the study of Özyurt [6]. In the study of Bilgilioğlu et al. [7], in which the algorithms of KNN and random forest (RF) have been used for the thematic map production of tea gardens, it is obtained that RF give better results than KNN. Spectral Angle Mapper method is used for land use change detection in Denizli City between 1984 and 2018 [8]. In the study in which three different machine learning algorithms (RF, KNN, and Support Vector Machine) are tested in order to determine hazelnut gardens, the highest success has been achieved from SVM [9].

In order to make a comparison with our study, the performances of the studies using the UC Merced Land-Use dataset have been examined. Akram et al. have combined the features extracted by using different CNN-based methods (AlexNet, VGGNet-16, VGGNet-19, GoogLeNet, ResNet) [10]. All extracted features have been classified using Support Vector Machine (SVM) and KNN after feature selection and dimensionality reduction. Yuan et al. have performed their tests on two different data sets in the CNN-based system where LASC is used as the pooling mechanism [11]. A similar study is presented in the study of [12], in which seven different datasets (WHU-RS19, UC-Merced Land Use, SIRI-WHU, RSSCN7, AID, PatternNet, NWPU-RESISC45) have been tested. Features are extracted by using deep learning models (AlexNet, VGGNet, GoogLeNet, Inception-V3, ResNet101). Then, these features are classified using machine learning based methods (NB, DT, RF, KNN, SVM). ResNet101 and SVM are the best performing pair. Iorga and Neagoe have tested the performance of VGG16-based CNN architectures on same dataset [13]. They obtain 86.61% accuracy rate. Xu et al. have used four different datasets (UC Merced Land-Use, AID, NWPU-45, OPTIMAL-31) to test the performance of Multi-structure Joint Decision Convolutional Neural Network (MJDCNN) [14]. They combine three different pretrained CNN architecture (AlexNet, Inception v3, and ResNet18) in this method. It is observed that the performances of the classifiers after cascading increased in the study of [15].

Information can be accessed in a very short time with the help of remote sensing data. Monitoring major natural events with remote sensing methods is important in order not to repeat the damage caused by this event and to result in minimum damage to this event. With consideration for the research issues in previous studies, this study aims to advance the field of remote sensing. CNN models have disadvantages as well as advantages. One of them is running time. In this study, the advantage of being successful in feature extraction in CNN and the advantages of fastness of classifiers such as Support Vector Machine, K Nearest Neighbor, Decision Tree, Discriminant Analysis are combined.

The rest of this paper is as follows. Chapter 2 includes the information about the dataset used and the background for the methods used. The proposed system is included in Chapter 3. The findings obtained in the experiments and the discussion are given in Chapter 4. Chapter 5 contains the conclusions and future works.

2. Material and Methods

2.1. Dataset

The methods used in the study were compared using UC Merced land-use dataset [16]. It is derived from large optical images obtained from various locations of the United States by the US Geological Survey as seen in Figure 1.

Figure 1. Dataset Samples (a) agricultural (b) airplane (c) baseball diamond (d) beach (e) buildings (f) chaparral (g) dense residential (h) forest (i) freeway (j) golf course (k) harbor (l) intersection (m) medium residential (n) mobile home park (o) overpass (p) parking lot (q) river (r) runway (s) sparse residential (t) storage tanks (u) tennis court

In the data set consisting of 21 classes, there are 100 images in the sizes of 256×256 pixels in each class. The dataset contains the images belonging the areas of agricultural, airplane, baseball diamond, beach, buildings, chaparral, dense residential, forest, freeway, golf course, harbor, intersection, medium residential, mobile home park, overpass, parking lot, river, runway, sparse residential, storage tanks, and tennis court. As can be depicted from here, it contains samples of various types of land cover and various types of land use, both homogeneous and inhomogeneous.

2.2. CNN Architectures

When object detection/recognition is mentioned in images, deep learning comes to mind first, especially in recent years ([17]-[20]). Convolutional Neural Network (CNN) is one of the classes of deep learning algorithms that help us to distinguish objects in images, and allow us to analyze the images. CNNs are feed-forward neural networks consisting of convolutional layers, Re-Lu (Rectified Linear Units) correction layers, pooling layers, and fully-connected (FC) layers. Each layer has its own function.

Convolutional layer is the first layer that handles the image in CNN architectures. As it is known, images are matrices made up of pixels. In the convolution layer, certain features are tried to be captured with the help of a filter smaller than the dimensions of the original image. Re-Lu is a non-linear function that works as f(x) = max(0,x) whose main purpose is to get rid of negative values. The pooling layer, like the convolutional layer, lowers dimensionality. In this manner, both the needed processing power and the useless features captured are minimized, while more relevant aspects are concentrated on. In the CNN architecture, the fully-connected layer comes after the successive convolution, ReLu, and pooling layers. This layer is dependent on all fields of the previous layer. The image, which is in the form of a matrix, is converted into a vector in this layer. A CNN architecture with seven layers for six classes is as in Figure 2.

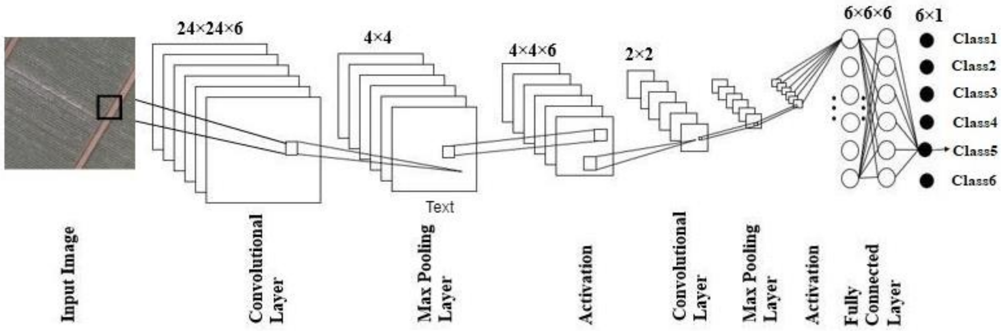

Figure 2 . A CNN Architecture

LeNet [21], AlexNet [22], ZF Net [23], GoogLeNet [24], VGGNet [25], and Microsoft ResNet [26] are among the frequently used CNN architectures. There are several forms of the ResNet architecture (ResNet18, ResNet34, ResNet50, ResNet101, ResNet110, ResNet152, ResNet164). Although they are similar to each other, they have a different number of layers. For example, ResNet18 has 18 layers, while ResNet50 has 50 layers. Increasing the number of layers usually leads to higher success, but also leads to increased completion time. For this reason, it should be preferred by making a time-performance comparison.

ResNet18 and GoogLeNet have been used in this study (Figure 3). These architectures will be briefly discussed in Section 2.2.1 and Section 2.2.2.

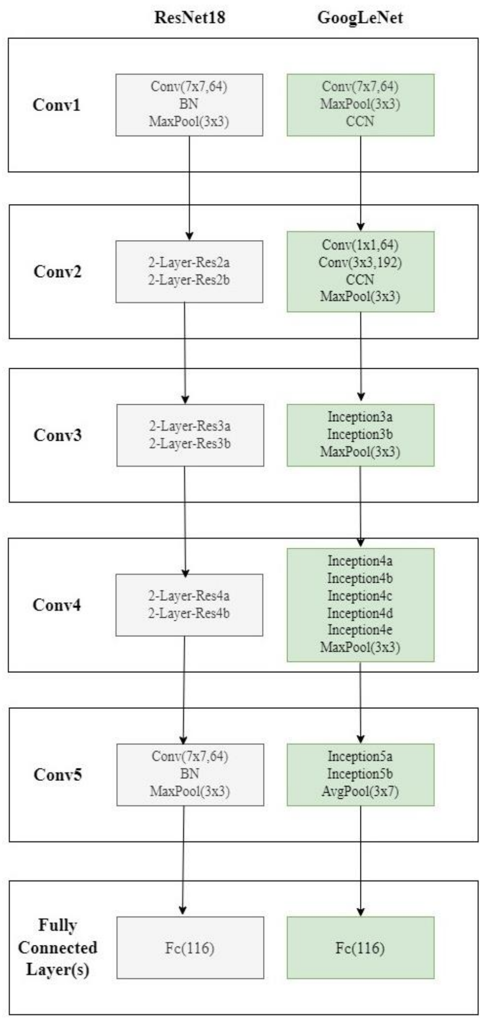

Figure 3. CNN models used for our study

2.2.1. ResNet

ResNet, which is obtained by adding some shortcuts to the classical network, consists of residual blocks. In the residual block, an $f(x)$ function is obtained by taking the x value as an input and passing it through the series of convolution-activation-convolution. Then, $h(x) = f(x) + x$ is produced by adding the original input x to the $f(x)$ function. In the classical convolution operation, while the $h(x)$ function is equal to the $f(x)$ function, in this network, the original data is added after the convolution operation is applied to the input.

2.2.2. GoogLeNet

GoogLeNet offers a deeper and broader structure without increasing the cost required for computation. This network is basically based on pooling operations applied to the same input and application of multiple convolution operations. It is a complex architecture containing inception modules. Images to be used as input in this network consisting of 22 layers must be 224×224 in size.

2.3. Machine Learning

Machine learning (ML) is a field of computer science that combines many different fields, examining methods for automating the solutions of complex problems. It is divided into three categories as supervised, unsupervised, and semi-supervised machine learning (Figure 4). Supervised machine learning algorithms can assign a label to an unlabeled element using existing labeled instances. In unsupervised machine learning algorithms, all elements are unlabeled and unclassified. No prior training process is carried out. Finally, a combination of both labeled and unlabeled data is used for training in semi-supervised learning.

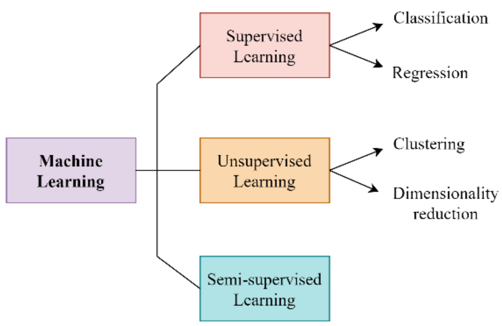

Figure 4. Machine Learning

In this study, in which the performances of four different classification algorithms (SVM, KNN, DT, and DA) are compared, these classifiers will be briefly mentioned before proceeding to the proposed system section.

2.3.1. Support Vector Machine (SVM)

Support Vector Machine is one of the supervised learning based algorithms [27-28]. The dataset is divided into two classes: training and test. Using a labelled training set, SVM finds the best hyperplane dividing the classes. There might be several planes dividing the two classes. An ideal hyperplane is the plane that is the furthest away from the classes' closest data points.

2.3.1. K Nearest Neighbor (KNN)

The K-NN method is a basic and commonly used classification technique [29-31]. The K value, which specifies the number of items to be looked at for classification, is the input value in this algorithm. Distances from other objects to the sample object are calculated. Different distance metrics such as Euclidean, Manhattan, and Minkowski can be used to determine the distances between the sample object and the other objects in the dataset. These distances are ordered and the k nearest neighbors are found. The class of this sample is selected according to these neighbors.

2.3.3. Decision Trees (DT)

Decision trees (DT) are a classification and pattern recognition technique that has gained popularity in recent years [31-32]. A decision tree's fundamental structure consists of three basic pieces known as nodes, branches, and leaves. The process of classifying the data and creating the tree structure begins at the root node, which is the initial node of the tree, and continues until nodes or leaves without branches are identified.

2.3.4. Discriminant Analysis (DA)

Discriminant analysis (DA) allows the variables in the dataset to be divided into two or more classes [33-34]. It examines the distribution of classes and uses the differences between their average values to distinguish between classes. By considering p attributes of observations, functions are created to assign these units to real groups.

3. Hybrid Systems

CNN-based architectures have been used in several papers to classify the remote sensing images. CNN has disadvantages as well as advantages:

- Completion time can be very long.
- Another disadvantage of CNN is that it can detect only one object in an image.
- When trying to develop a face recognition application, it loses the relationships between eyes, lips, mouth and face. Therefore, similar results with the original image can be obtained even when the position of the nose and mouth is changed.
- CNN cannot be used to find the position of an object.
- There is a requirement that the training data contain a large number of elements to use CNN. Using more data increases the achievement of the system.

Different studies have been carried out to overcome the disadvantages of CNN. Hybrid methods which is used in this paper is an alternative to deep learning approaches with long execution times. Feature extraction based on CNN and classification based on machine learning have been combined. It is aimed to take the advantages of using machine learning methods that respond in a shorter time instead of the long classification process of CNN. Since the target is not only time here, it is compared with CNN-based classification in terms of performance. Comparisons have been made between these methods by using more than one machine learning method like SVM, KNN, DT, DA. Various experiments have been carried out while trying to find the optimal combination that gives the best response in a short time.

4. Experiments

4.1. Preliminaries to experiments

The experimental results have been compared in terms of sensitivity, specificity, precision, accuracy, and F-score. Sensitivity is the percentage of samples assigned to the class “a” among the samples from the class “a”. The specificity is the percentage of samples that are not assigned to the “a” class among the samples that are not from the “a” class. Precision explores how many correct predictions out of all predictions. Accuracy seeks an answer to the question of "How many of all land samples have we labeled correctly?". The F-measure is calculated by taking the harmonic mean of accuracy and recall. The classification performance of the models was examined using 10-fold cross validation. 70% of the database is reserved for training and 30% of it is reserved for testing to make comparisons in this study.

This study was carried out using Matlab 2022a on a computer with Intel(R) Core (TM) i7-6700HQ CPU @ 2.60GHz 2.59 GHz.

4.2. Experimental results

In order to learn the results to be obtained when the classification is performed with the CNN-based approaches, classification has also performed using only ResNet18 and GoogLeNet algorithms. While ResNet18 classifies the images with 97.14% success, GoogLeNet performs the classification with 92.62% success. ResNet18 gives results in 24 minutes 34 seconds in this 21-class data set used, while GoogLeNet gives results in 11 minutes 11 seconds. The class recognized with the lowest accuracy in ResNet18 is 'medium residential'. The classes recognized with the lowest accuracy rates in GoogLeNet are 'dense residential' and 'buildings'.

Classification results of the combinations of CNN & machine learning are given in Table 1. The classification of the features drawn from different layers of the networks has also been examined. The most successful method is GoogLeNet & SVM (pool5-drop_7x7_s1) with an accuracy rate of 93.38% and the second successful method is ResNet18 & SVM (pool-5). The method with the lowest success is ResNet18 & DT (pool-5) with an accuracy rate of 50.95%. When the classification performances of the features drawn from the pool-5 and res3b-relu layers of ResNet18 are examined, it is seen that higher performance is obtained when the features are extracted from the pool-5 layer (except for DT). The same is true for GoogLeNet. When the classification performances of the features drawn from the pool5-drop_7x7_s1 and inception_3b-relu_1x1 layers of GoogLeNet are examined, it is seen that higher performance is obtained when the features are extracted from the pool5-drop_7x7_s1 layer (except DT). It can be concluded that the classification of features drawn from previous layers of CNN is generally less successful.

In Table 1, it has been determined that the latter layers in CNN show higher classification performance compared to the previous layers. Therefore, the next analyzes will be made using the pool-5 layer for ResNet18 and the pool5-drop_7x7_s1 layer for GoogleNet. In Table 2, the performance comparison of the methods in the classification process of 21 different classes is made. Here, the results obtained using the features extracted from the pool-5 layer for ResNet18 and from the pool5-drop_7x7_s1 layer for GoogLeNet are used. Images belonging to the 'chaparral' class have been the images classified with the highest success (96.81%). After the 'chaparral' class, the second and third classes with the high successes are 'agricultural' and 'harbor'. Additionally, classification successes of 'beach', 'forest', and 'parkinglot' classes are over 90%. On the other hand, 'denseresidential', 'mediumresidential', and 'buildings' are the classes with classification successes below 70% (61.90%, 67.61%, and 69.87%, respectively). It is seen that the performance of the methods decreases due to the classes that are more difficult to classify. It is obvious that the hard-to-recognize classes in CNN-based classification are the same classes as in machine learning-based classification.

Table 1. Classification results of different combinations

No	Method	Layer	Sensitivity	Specificity	Precision	Accuracy	F Score
1	ResNet18 & SVM	pool-5	93.45%	99.64%	93.18%	93.17%	93.08%
		res3b_relu	81.97%	98.32%	79.52%	79.52%	79.97%
2	ResNet18 & KNN	pool-5	89.18%	99.38%	88.73%	88.73%	88.58%
		res3b_relu	86.17%	99.11%	84.60%	84.60%	84.89%
3	ResNet18 & DT	pool-5	52.79%	95.46%	50.95%	50.95%	51.42%
		res3b_relu	61.25%	96.82%	0.60%	60.00%	60.14%
4	ResNet18 & DA	pool-5	93.16%	99.58%	92.22%	92.22%	92.30%
		res3b_relu	90.98%	99.45%	90.00%	90.00%	90.10%
5	GoogLeNet & SVM	pool5-drop_7x7_s1	93.65%	99.64%	93.33%	93.33%	93.38%
		inception_3b-relu_1x1	89.00%	99.36%	88.57%	88.57%	88.62%
6	GoogLeNet & KNN	pool5-drop_7x7_s1	88.65%	99.35%	88.25%	89.52%	88.10%
		inception_3b-relu_1x1	87.85%	99.29%	87.30%	87.30%	87.34%

7	GoogLeNet & DT	pool5-drop_7x7_s1	64.18%	97.08%	62.06%	62.06%	62.55%
		inception_3b-relu_1x1	64.59%	97.23%	63.33%	63.33%	63.38%
8	GoogLeNet & DA	pool5-drop_7x7_s1	88.32%	99.29%	87.30%	87.30%	87.61%
		inception_3b-relu_1x1	87.82%	99.27%	87.14%	87.14%	87.16%

Table 2. Distribution of f scores according to classes

	Method								
Class	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**Average**
(a)	98.31%	100%	88.14%	94.74%	92.86%	96.67%	95.08%	100%	95.73%
(b)	98.31%	98.31%	60.32%	98.31%	94.74%	95.08%	64.41%	94.74%	88.03%
(c)	98.31%	90.32%	49.06%	98.36%	100%	96.77%	60.32%	93.33%	85.81%
(d)	100%	98.36%	62.50%	96.67%	94.92%	96.77%	73.33%	98.31%	90.11%
(e)	83.08%	76.67%	31.58%	83.87%	93.33%	68.00%	50.00%	72.41%	69.87%
(f)	96.55%	100%	90.32%	100%	96.67%	100%	90.91%	100%	96.81%
(g)	66.67%	76.67%	28.57%	74.58%	82.76%	80.70%	28.57%	56.67%	61.90%
(h)	95.08%	95.24%	65.63%	100%	96.55%	89.23%	89.66%	96.77%	91.02%
(i)	96.77%	96.77%	48.28%	96.55%	98.31%	94.92%	54.55%	94.74%	85.11%
(j)	93.33%	87.10%	55.74%	93.10%	92.06%	93.33%	57.58%	91.23%	82.93%
(k)	100%	100%	77.78%	100%	98.31%	100%	81.36%	96.55%	94.25%
(l)	90.32%	80.00%	29.03%	93.33%	83.87%	89.23%	44.44%	68.66%	72.36%
(m)	81.82%	75.00%	36.92%	75.00%	81.25%	76.92%	48.28%	65.71%	67.61%
(n)	93.10%	88.14%	48.39%	80.00%	93.10%	83.58%	57.14%	75.41%	77.36%
(o)	94.74%	86.79%	50.00%	94.92%	98.31%	89.66%	49.28%	86.67%	81.30%
(p)	100%	96.67%	66.67%	98.36%	96.67%	96.67%	84.62%	98.31%	92.25%
(q)	92.06%	88.52%	36.67%	91.80%	92.06%	89.29%	66.67%	91.80%	81.11%
(r)	100%	96.77%	53.85%	95.24%	96.77%	88.89%	70.37%	100%	87.74%
(s)	95.24%	84.75%	23.73%	87.88%	96.67%	86.21%	46.15%	89.66%	76.29%
(t)	89.66%	70.83%	28.57%	88.89%	91.80%	92.86%	42.55%	83.64%	73.60%
(u)	91.53%	73.33%	48.00%	96.67%	90.00%	71.43%	58.18%	85.25%	76.80%

What will be the performance of the methods when the classes that are difficult to classify are removed from the dataset? Based on this question, a new dataset consisting of classes with high classification successes have been created. The results obtained when the methods used have been tested on the six-class dataset consisting of classes with performances above 90% are given in Table 3. These six classes are 'agricultural', 'beach', 'chaparral', 'forest', 'harbor', and 'parking lot'. While the highest success in the 21-class dataset is 93.38%, the highest success in the 6-class dataset is 100%. While GoogLeNet & SVM is the most successful method in the 21-class dataset, ResNet18 & SVM and ResNet18 & KNN have the best performances in the 6-class dataset (99.67%). In both, classification based on DT has the lowest success.

Table 3. Classification results of six-class dataset

No	Method	Sensitivity	Specificity	Precision	Accuracy	F Score
1	ResNet18 & SVM	99.67%	99.93%	99.67%	99.67%	99.67%
2	ResNet18 & KNN	99.67%	99.93%	99.67%	99.67%	99.67%
3	ResNet18 & DT	91.41%	97.97%	90.56%	90.56%	90.66%
4	ResNet18 & DA	98.92%	99.78%	98.89%	98.89%	98.89%
5	GoogLeNet & SVM	99.34%	99.87%	99.33%	99.33%	99.33%
6	GoogLeNet & KNN	98.92%	99.78%	98.89%	98.89%	98.89%
7	GoogLeNet & DT	94.63%	98.85%	94.47%	94.44%	94.47%
8	GoogLeNet & DA	99.46%	99.89%	99.45%	99.44%	99.45%

Table 4 shows the average success rates of the classifiers and the methods used in feature extraction. When the average accuracy rates of the methods are examined, it is seen that GoogLeNet is more successful than ResNet18 and SVM is more successful than other classifiers (KNN, DT, DA). While the average rate of DT in the 21-class dataset is remarkably low (56.51%), the rate of the same classifier in the 6-class dataset is 92.5%. Another remarkable situation is that the ratios of KNN and DA are close to each other in both data sets. In the 21-class dataset, average accuracy rates of KNN and DA are 89.13% and 89.76%, respectively, while their rates in the 6-class dataset are 99.28% and 99.17%. When the feature extraction successes of deep learning approaches are also examined in the study, it is seen that the successes of both ResNet18 and GoogLeNet are quite high in the 6-class dataset (97.20% and 98.04%, respectively). It is difficult to say the same for the ratios obtained in the 21-class dataset, since the results obtained with DT are very low, which lowers the averages of ResNet18 and GoogLeNet. The average accuracy rates of ResNet18 and GoogLeNet are 81.27% and 83.05%, respectively.

Table 4. Average accuracy rates

	Dataset	
Method	**21-class dataset**	**6-class dataset**
Classification based on SVM	93.25%	99.50%
Classification based on KNN	89.13%	99.28%
Classification based on DT	56.51%	92.50%
Classification based on DA	89.76%	99.17%
Feature extraction using ResNet18	81.27%	97.20%
Feature extraction using GoogLeNet	83.05%	98.04%

In this study, the classification of the dataset containing 2100 images was carried out in less than 1 minute (Table 5). When compared to the execution times of ResNet18 and GoogLeNet, there is a huge difference. It is obvious that one of the biggest advantages of hybrid models is time saving.

Table 5. Execution times

		Dataset	
No	**Method**	**21-class dataset**	**6-class dataset**
1	ResNet18 & SVM	51.51 sec	46 sec
2	ResNet18 & KNN	46.93 sec	46.33 sec
3	ResNet18 & DT	51.08 sec	26.87 sec
4	ResNet18 & DA	47.51 sec	37.81 sec
5	GoogLeNet & SVM	56.27 sec	40.68 sec
6	GoogLeNet & KNN	54.15 sec	29.01 sec
7	GoogLeNet & DT	52.72 sec	29.29 sec
8	GoogLeNet & DA	53.27 sec	29.29 sec

4.3. Discussion

Table 6 shows a comparison of our study with other studies. In some studies in this table, the methods used have been tested on more than one dataset. In the performance column of the table, only the results obtained for the dataset that we used are available. In this table, where the methods they used are also given, it is seen that the highest result has been obtained in the study conducted by Akram et al. (99.1%). The results of the hybrid models are given at the end of the table.

The results obtained by using the features extracted from different layers of CNN are given in the table. The features extracted from pool-5 layer of ResNet18 are classified with 81.35%

accuracy rate while the features extracted from res3b_relu layer of ResNet18 are classified with 78.78% accuracy rate on average. The features extracted from pool5-drop_7x7_s1 layer of GoogLeNet are classified with 82.91% accuracy rate while the features extracted from inception_3b-relu_1x1 layer of GoogLeNet are classified with 81.63% accuracy rate on average.

Table 6. Comparison with other studies

Study	Method		Performance
[35]	VLAT		94.3%
[36]	VGG16		93%
[10]	CNNs-Entropy controlled NCA		99.1%
[11]	LASC-CNN		97.14%
[13]	VGG16		86.61%
[37]	ResNet-50 GoogLeNet		96.42% 97.32%
[14]	MJDCNN		95.79%
[15]	CNN cascade feature + McODM		97.55%
Our Experiments	ResNet18 GoogLeNet		97.14% 92.62%
	ResNet18 & SVM ResNet18 & KNN ResNet18 & DT ResNet18 & DA	pool-5 layer	93.08% 88.58% 51.43 % 92.30%
	ResNet18 & SVM ResNet18 & KNN ResNet18 & DT ResNet18 & DA	res3b_relu layer	79.97% 84.89% 60.14% 90.10%
	GoogLeNet & SVM GoogLeNet & KNN GoogLeNet & DT GoogLeNet & DA	pool5-drop_7x7_s1 layer	93.38% 88.10% 62.55% 87.61%
	GoogLeNet & SVM GoogLeNet & KNN GoogLeNet & DT GoogLeNet & DA	inception_3b-relu_1x1 layer	88.62% 87.34% 63.38% 87.16%

5. Conclusion and Future Works

In this paper, hybrid systems using deep learning and machine learning have been tested on UC Merced land-use dataset which is the well-known and widely used. These systems, which were evaluated in terms of both time and performance, were also compared with CNN-based classification. In the experiments, it was observed that some classes reduced success. For this reason, a 6-class dataset was created by selecting the classes with high-successes and the same tests were performed on this dataset. It is seen that SVM (93.25%) performs the classification with the highest success among hybrid systems. DT (56.51%), which gave lowest accuracy rate in the 21-class dataset, also had the lowest accuracy rate in the 6-class dataset. The classification successes of KNN and DA are close to each other. Average classification success is higher when features are extracted using GoogLeNet (83.05%) than when features are extracted using ResNet18 (81.27%). On the other hand, the success achieved in classification with ResNet18 (97.14%) is higher than the successes of GoogLeNet and other hybrid systems. However, when the systems are evaluated not only in terms of performance but also in terms of time, it is seen that hybrid systems that give results close to ResNet18 in a short time are better alternatives. The two alternatives to CNN-based classification are GoogLeNet & SVM with an accuracy rate of 93.33% and ResNet18 & SVM with an accuracy rate of 93.17%.

The classification successes of the features drawn from different layers were also examined in the study. An average of 81.35% success was achieved when features drawn from the pool5 layer of ResNet18 were used, while an average of 78.78% success was achieved when the res3b_relu layer was used. An average of 82.91% success was achieved when features drawn from the pool5-drop_7x7_s1 layer of GoogLeNet were used, while an average of 81.63% success was achieved when the inception_3b-relu_1x1 layer was used. When the results of the pool-5 layer and res3b_relu layer in ResNet18 and the results of the pool5-drop_7x7_s1 layer and inception_3b-relu_1x1 layer in GoogLeNet are examined, it is seen that the features drawn from the later layers of the network are classified with higher success.

Despite the successes of this study, several issues warrant additional examination. GoogLeNet & SVM achieves 93.38% success in a short time. However, there are studies with higher success in the literature. Increasing the success rate even more is one of our goals from now on.

References

[1] Yener, H., Koç, A., Çoban, H.O. 2006. Assessment Methods for Classification Accuracy in Remote Sensing Data, Journal of the Faculty of Forestry Istanbul University, 56(2), 71-88.

[2] Çölkesen, İ., Kavzoğlu, T. 2013. Arazi Örtüsü Haritalarinin Üretilmesinde Yeryüzü Özelliklerinin Siniflandirma Doğruluğuna Etkilerinin İncelenmesi: Trabzon örneği, Türkiye Ulusal Fotogrametri ve Uzaktan Algılama Birliği VII. Teknik Sempozyumu.

[3] Döş, M.E., Uysal, M. 2019. Classification of Remote Sensing Data with Deep Learning Algorithms, Turkish Journal of Remote Sensing, 1(1), 28-34.

[4] Sabanci, K., Ünlerşen, M.F. Kayabaşi, A. 2016. Machine Learning Methods for Land Cover Classification from Multispectral Images, IMCOFE, 472-478.

[5] Tavus, B., Karataş, K., Türker, M. 2018. Object-Based Crop Pattern Detection from IKONOS Satellite Images in Agricultural Areas, Pamukkale University Journal of Engineering Sciences, 25(5), 603-614.

[6] Özyurt, F. 2019. Classification of Remote Sensing Images Based on Convolutional Neural Networks and Neighborhood Component Analysis Features, Afyon Kocatepe University Journal of Science and Engineering, 19(3), 669-675.

[7] Bilgilioğlu, B.B., Çömert, R., Yiğit, O., Bedir F. 2019. Extraction of Tea Gardens by Object-Based Classification Approach From High Spatial Resolution Satellite Images, Turkish Journal of Remote Sensing, 1(1), 21-27.

[8] Çan, T., Tekin, S., Traore, M., & Kumsar, H. 2020. Land Use Change Detection in Denizli City Center Using spectral angle mapper method and evaluations in terms of some earth science data, Pamukkale University Journal of Engineering Sciences, 26(8), 1360-1364.

[9] Apaydin, C., Abdikan, S., 2021. Determination of Hazelnut Gardens by Pixel Based Classification Methods Using Sentinel-2 Data, Geomatik, 6(2), 107-114.

[10] Akram, T., Laurent, B., Naqvi, S.R., Alex, M.M., Muhammad, N. 2018. A Deep Heterogeneous Feature Fusion Approach for Automatic Land-use Classification, Information Sciences, 467, 199-218.

[11] Yuan, B., Li, S., Li, N. 2018. Multiscale Deep Features Learning for Land-Use Scene Recognition, Journal of Applied Remote Sensing, 12(1).

[12] Shafaey, M.A., Salem, M.A.M., Ebeid, H.M., Al-Berry, M.N., Tolba M.F. 2018. Comparison of CNNs for Remote Sensing Scene Classification, 13th International Conference on Computer Engineering and Systems, 27-32.

[13] Iorga, C., Neagoe, V.E. 2019. A Deep CNN Approach with Transfer Learning for Image Recognition, 11th International Conference on Electronics, Computers and Artificial Intelligence, 1-6.

[14] Xu, L., Chen, Y., Pan, J., Gao, A. 2020. Multi-Structure Joint Decision-Making Approach for Land Use Classification of High-Resolution Remote Sensing Images Based on CNNs, IEEE Access, 8, 42848-42863.

[15] Zhang, Z., Cui, X., Zheng, Q., Cao, J. 2021. Land Use Classification of Remote Sensing Images Based on Convolution Neural Network, Arabian Journal of Geosciences, 14(4), 1-6.

[16] Yang, Y., Newsam, S. 2010. Bag-of-Visual-Words and Spatial Extensions for Land-Use Classification, Proceedings of the 18th SIGSPATIAL international conference on advances in geographic information systems, 270-279.

[17] Karahan, T., Nabiyev, V. 2021. Plant Identification with Convolutional Neural Networks and Transfer Learning, Pamukkale University Journal of Engineering Sciences, 27(5), 638-645.

[18] Özkan, K., Seke, E., Işık, Ş. 2021. Wheat Kernels Classification Using Visible-Near Infrared Camera Based on Deep Learning, Pamukkale University Journal of Engineering Sciences, 27(5), 618-626.

[19] Cevik, F., Kilimci, Z.H. 2021. The Evaluation of Parkinson's Disease with Sentiment Analysis using Deep Learning Methods and Word Embedding Models, Pamukkale University Journal of Engineering Sciences, 27(2), 151-161.

[20] Çelik, G., Talu, M.F. 2021. Generating the Image Viewed from EEG Signals, Pamukkale University Journal of Engineering Sciences, 27(2), 129-138.

[21] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P. 1998. Gradient-Based Learning Applied to Document Recognition, Proceedings of the IEEE, 86(11), 2278-2324.

[22] Krizhevsky, A., Sutskever, I., Hinton, G.E. 2012. Imagenet Classification with Deep Convolutional Neural Networks, Advances in neural information processing systems, 25, 1097-1105.

[23] Zeiler, M.D., Fergus, R. 2014. Visualizing and Understanding Convolutional Networks, In European conference on computer vision, 818-833.

[24] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Rabinovich, A. 2015. Going Deeper with Convolutions, In Proceedings of the IEEE conference on computer vision and pattern recognition, 1-9.

[25] Simonyan, K., Zisserman, A. 2014. Very Deep Convolutional Networks for Large-Scale Image Recognition, arXiv preprint arXiv:1409.1556.

[26] He, K, Zhan,g X, Ren, S, Sun, J. 2016. Deep Residual Learning for Image Recognition, Proceedings of the IEEE conference on computer vision and pattern recognition, 770-778.

[27] Cortes, C., & Vapnik, V. 1995. Support-vector networks. Machine learning, 20, 273-297.

[28] Wang, M, Zhang, X, Niu, X, Wang, F, Zhang, X. 2019. Scene classification of high-resolution remotely sensed image based on ResNet. Journal of Geovisualization and Spatial Analysis, 3, 1-9.

[29] Fix E., & Hodges, J. 1951. Discriminatory Analysis: Nonparametric Discrimination: Consistency Properties, 4.

[30] Cover, T., & Hart, P. 1967. Nearest neighbor pattern classification. IEEE transactions on information theory, 13(1), 21-27.

[31] Bouteldja, S, Kourgli, A. 2020. A comparative analysis of SVM, K-NN, and decision trees for high resolution satellite image scene classification. In Twelfth International Conference on Machine Vision (ICMV 2019), 11433, 410-416.

[32] Quinlan, J. R. 1987. Simplifying decision trees. International journal of man-machine studies, 27(3), 221-234.

[33] Fisher, R. A. 1936. The use of multiple measurements in taxonomic problems. Annals of eugenics, 7(2), 179-188.

[34] Huang, R, Taubenböck, H, Mou, L, Zhu, X X 2018. Classification of settlement types from Tweets using LDA and LSTM. In IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, 6408-6411.

[35] Negrel, R., Picard, D., Gosselin, P.H. 2014. Evaluation of Second-Order Visual Features for Land-Use Classification, 12th International Workshop on Content-Based Multimedia Indexin, 1-5.

[36] Alias, B., Karthika, R., Parameswaran, L. 2018. Classification of High Resolution Remote Sensing Images Using Deep Learning Techniques, International Conference on Advances in Computing, Communications and Informatics, 1196-1202.

[37] Helber, P., Bischke, B., Dengel, A., Borth, D. 2019. Eurosat: A Novel Dataset and Deep Learning Benchmark for Land Use and Land Cover Classification, IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 12(7), 2217-2226.