

Hiperspektral Görüntülerin Sınıflandırma Performanslarına Parametrelerin Etkisi: Oto-Kodlayıcılar ile Deneysel Bir Çalışma

Mert YİĞİT^{1,*}, Ümit Haluk ATASEVER¹

¹Erciyes Üniversitesi, Mühendislik Fakültesi, Harita Mühendisliği Bölümü, 38039 Kayseri

Özet

Son yıllarda, birçok farklı disiplinden araştırmacılar uzaktan algılama veri ve yöntemlerini çalışmalarında yoğun olarak kullanılmaktadırlar. Uzaktan algılama; haritacılık disiplininin yanında bölgesel planlama, tarım, güvenlik vb. birçok alanda kullanılmaktadır. Uzaktan algılama çalışmalarında kullanılan materyallerden olan hiperspektral görüntüleri günden güne daha fazla tercih edilmektedir. Yorumlanması zor olan görüntülerin, sınıflandırılması, kullanıcıların daha anlaşılabilir sonuçlar elde etmesine olanak sağlamaktadır. Bu çalışmada farklı makine öğrenmesi tekniklerinin farklı parametrelerine göre hiperspektral görüntülerin sınıflandırma performansına olan etkisi incelenmiştir. Destek Vektör Makineleri, Karar Ağaçları, Torbalama Karar Ağaçları, K-En Yakın Komşuluk, Alt Uzay En Yakın Komşuluk, Diskriminant Analizi yöntemleri çalışmada kullanılan geleneksel sınıflandırma yöntemleridir. Bu tekniklere ek olarak piksel temelli sınıflandırmada kullanılan bir derin öğrenme modeli olan Oto-Kodlayıcı tekniğinin parametrelerinin sınıflandırma performansına etkisi incelenmiştir. Yığınlanmış Oto-Kodlayıcı parametrelerin (transfer fonksiyonları, saklı katman sayıları vb.) farklı kombinasyonlarından sınıflandırma doğruluğunu görece yüksek yapan değişkenlerin belirlenmesi çalışmanın öncelikli amacını oluşturmaktadır. Uygulamalarda AVIRIS sensörü ile elde edilen 16 farklı sınıf etiketine sahip Indian Pines ile 6 farklı sınıf etiketine sahip Salinas-A veri setleri kullanılmıştır. Her bir yöntemin parametrelerinin farklı kombinasyonlar ile gerçekleştirilen deneysel çalışmalarda görece yüksek test doğruluğunu veren parametre kombinasyonları belirlenmeye çalışılmıştır. Bu deneysel çalışmalar sonucunda sınıflandırma doğruluğu açısından daha başarılı olan tekniğin Yığınlanmış Oto-Kodlayıcı tekniği olduğu görülmüştür.

Anahtar Sözcükler: Hiperspektral Görüntüler, Derin Öğrenme, Yığınlanmış Oto-Kodlayıcı, Makine Öğrenmesi

The Effect of Parameters on the Classification Performance of Hyperspectral Images: An Experimental Study with Auto-Encoders

Mert YİĞİT^{1,*}, Ümit Haluk ATASEVER¹

¹Erciyes University, Faculty of Engineering, Department of Geomatics Engineering, 38039 Kayseri

Abstract

In recent years, researchers from many different disciplines have been using remote sensing data and methods extensively in their studies. Remote sensing, in addition to the geomatics discipline, is used in many areas such as regional planning, agriculture, security, etc. Hyperspectral images, which are among the materials used in remote sensing studies, are used more and more day by day. Classification of images that are difficult to interpret allows users to obtain more understandable results. The classification stage is one of the basic applications in the process of producing meaningful information from images. In this study, the effect of different machine learning techniques on the classification performance of hyperspectral images was investigated. Support Vector Machines, Decision Trees, Bagging Decision Trees, Nearest Neighborhood, Subspace K-Nearest Neighborhood, Discriminant Analysis methods are the traditional classification methods used in the study. In addition to these techniques, the effect of parameters of Auto-Encoders technique, which is a deep learning model used in pixel-based classification, on classification performance was examined. The primary purpose of the study is to determine the variables that make the classification accuracy relatively high from different combinations of Stacked Auto-Encoder parameters (transfer functions, hidden layer numbers, etc.). Indian Pines with 16 different class and Salinas-A data sets with 6 different class were used in the applications. In experimental studies performed with different combinations of parameters of each method, it was tried to determine the parameter combinations that gave relatively high test accuracy. As a result of these experimental studies, it was seen that the technique that was more successful in terms of classification accuracy was the Stacked Auto-Encoder technique.

Keywords: Hyperspectral Images, Deep Learning, Stacked Auto-Encoder, Machine Learning

* Sorumlu Yazar: Tel: (0537)4642749

E-posta: mertyigit@erciyes.edu.tr (Yiğit M.), uhatasever@erciyes.edu.tr (Atasever Ü. Haluk)

1. Giriş

Geçmişten bugüne uzaktan algılama teknolojileri kullanılarak elde edilen uydu görüntüleri yeryüzü üzerindeki zamana bağlı değişimlerin analizi, çevrenin korunumu, tarım, madencilik, askeri gibi birçok dalda kullanılmaktadır (Shippert 2004). Birçok disiplin tarafından gerçekleştirilen çalışmalarda veri kaynağı olarak uydu görüntülerine ulaşabilmek amacıyla çok sayıda sensörlerle donatılmış hiperspektral uzaktan algılama uyduları vardır.

Elektromanyetik spektrumda görünür bölge ile kızılötesi bölge arasındaki alanın belirlenmesinde kullanılan hiperspektral görüntüleme sistemleri, klasik algılama sistemlerine nazaran çok daha fazla bant sayısı içermektedir. Hiperspektral görüntüleme sistemleri ile farklı özellikteki objelerin saptanmasında işlem yükü fazla olmasına rağmen sınıflandırma doğruluğu bakımından etkili sonuçlar ürettiği görülmektedir (Atasever vd. 2018).

Uydu görüntülerinden tematik bilgi elde edebilmek için genellikle en sık tercih edilen metodoloji sınıflandırmadır. Sınıflandırma kontrollü veya kontrolsüz olarak yapılabilir. Kontrollü sınıflandırmada eğitim veri seti kullanılmasından dolayı sınıflandırma doğruluğu kontrolsüz sınıflandırmaya göre genelde yüksek olabilmektedir. Sınıflandırma işleminde kullanılan yöntemler ve bu yöntemler içerisindeki değişken olan parametreler sınıflandırma performansını doğrudan etkilemektedir.

Bu çalışmanın amacı, makine öğrenmesi tekniklerinin parametrelerinin hiperspektral görüntülerin sınıflandırma performanslarına olan etkilerinin incelenmesidir. Çalışmada yer gerçeği görüntüsü mevcut olması sebebiyle Salinas-A ve Indian Pines hiperspektral veri setleri kullanılmıştır. Veri setleri bir derin öğrenme modeli olan Oto-Kodlayıcı, Destek Vektör Makineleri, Karar Ağaçları, K-En Yakın Komşuluk, Lineer Diskriminant, Torbalama Karar Ağaçları, Alt Uzaklık K-En Yakın Komşuluk teknikleriyle sınıflandırılmıştır.

2.Yöntemler

2.1. Destek Vektör Makineleri

İlk zamanlarında iki sınıflı lineer ayrılabilir veriler üzerinde sınıflandırma gerçekleştirmek için kullanılan Destek Vektör Makineleri günümüzde sınıf sayısı ikiden fazla olan ve lineer ayrılamayan verilerin sınıflandırılmasında da kullanılmaktadır (Huang vd. 2002). DVM'nin temelinde sınıflar arasında optimum ayırım yapabilen bir hiper düzlemin belirlenmesi yer almaktadır.

2.2. Karar Ağaçları

Karar Ağaçları, karmaşık ve büyük boyutlu verilerin sınıflandırılmasını hızlı bir şekilde yapmasından dolayı sınıflandırma teknikleri içinde sıklıkla tercih edilen bir yöntemdir (Safavian ve Landgrebe 1991). Düğüm, dal, yaprak olarak isimlendirilen üç kısımdan oluşan Karar Ağaçlarında bir düğüm her bir niteliği temsil eder. Kök düğüm kısmından başlanarak özniteliklere (band) göre şartla tanımlanır ve böylelikle ağaç yapısı oluşturulur.

2.3. Torbalama Karar Ağaçları

Breiman tarafından ortaya atılan makine öğrenmesi temelli bir grup algoritması olan Torbalama Karar Ağaçları yöntemiyle varyans en düşük değerlere düşürülerek aşırı öğrenme probleminin önüne geçilmiş olur (Breiman 1996). Bu yöntemin temelinde mevcut bir eğitim verisi üzerinden çok fazla miktarda alt eğitim örneğinin oluşturulması yatmaktadır.

2.4. Lineer Diskriminant

İstatistiksel bir yöntem olan Lineer Diskriminant yönteminin temelinde mevcut veri setindeki farklı sınıflara ait kısımların doğrusal şekilde ayrılabilceği duruma getirilmesi yer almaktadır (Teke ve Sakarya 2013).

2.5. K-En Yakın Komşuluk

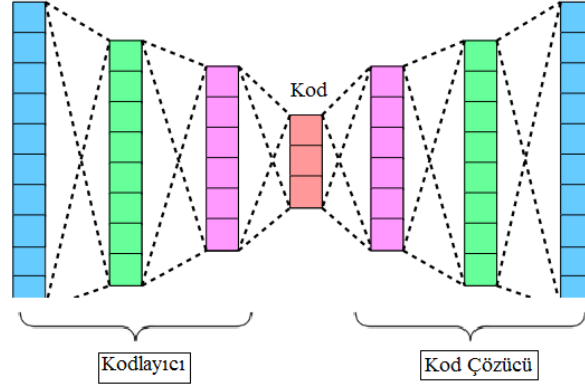
Sıklıkla kullanılan sınıflandırma tekniklerinden olan K-En Yakın Komşuluk tekniği, sınıflandırılacak verinin eğitim setindeki en yakın k adet veriye olan mesafesine göre sınıflandırma işlemini gerçekleştiren bir tekniktir. Bu teknik parametrik olmamasından dolayı yorumlanması basittir (Cunningham ve Delany 2007).

2.6. Alt Uzaklık En Yakın Komşuluk

Bir grup öğrenme yöntemi olan Alt-uzaklık KNN, ilk olarak eğitim veri setinden farklı alt uzaylar oluşturulur. Daha sonra bu alt uzaylar farklı KNN sınıflandırıcılar kullanılarak öğrenme işlemi yapılır ve bir tür çoğunluk uygulaması ile final kestirim belirlenir (Ho 1998).

2.7. Oto-Kodlayıcılar

Giriş katmanından gelen verileri yeniden kendi öğrenen bir model olan Oto-Kodlayıcılar kodlayıcı ve kod çözücü olarak iki kısımdan oluşurlar. Kodlayıcı kısmında bir transfer fonksiyonu ile verinin orijinal boyutundan da farklı sayıda veri üretilebilir (Günen vd. 2018). Kod çözücü kısmında yeni üretilen verinin kendisiyle aynı boyutta başka bir veriye dönüşüm işlemi gerçekleştirilir. Kod çözücü kısmının giriş vektörü, kodlayıcı kısmından çıkan çıkış vektörüdür. Şekil 1'de belirtilen Oto-Kodlayıcının kodlayıcı ve kod çözücü yapısıdır.



Şekil 1. Kodlayıcı ve Kod Çözücü Yapısı

Oto-Kodlayıcıdaki giriş-çıkış ilişkisi Eşitlik 1'de verilmiştir.

$$d = f(y; w, a) \quad (1)$$

Burada, w ve a kodlayıcının çıkış vektörlerine ait parametreleri, y giriş vektörünü ifade etmektedir. Kod çözücü kısmında kodlayıcı bölümünde olduğu gibi gizli katman seçerek çıkışa Eşitlik 2'de verilmiş olan bir fonksiyon aracılığıyla aktarılır.

$$\hat{y} = f(d; \hat{w}, \hat{a}) \quad (2)$$

Burada, \hat{w} ve \hat{a} kod çözücü kısmının çıkış vektörlerini etkileyen parametrelerdir. Oto-Kodlayıcıda giriş ve çıkış arasındaki etkileşim Eşitlik 3'te verilmiştir.

$$\hat{y} = h(w, b, \hat{w}, \hat{a}; y) \quad (3)$$

Oto-Kodlayıcıların birden fazla katmanının birleşiminden meydana gelen Yığılanmış Oto-Kodlayıcılarda ise her bir katmanda oluşan çıktılar bir sonraki katman girişine bağlanır. Yığılanmış Oto-Kodlayıcılar karmaşık ve çok sayıda öznelik (band) içeren verilerin sınıflandırılmasında oldukça etkindirler. Bu nedenle bu çalışmada Yığılanmış Oto-Kodlayıcılar tercih edilmiştir.

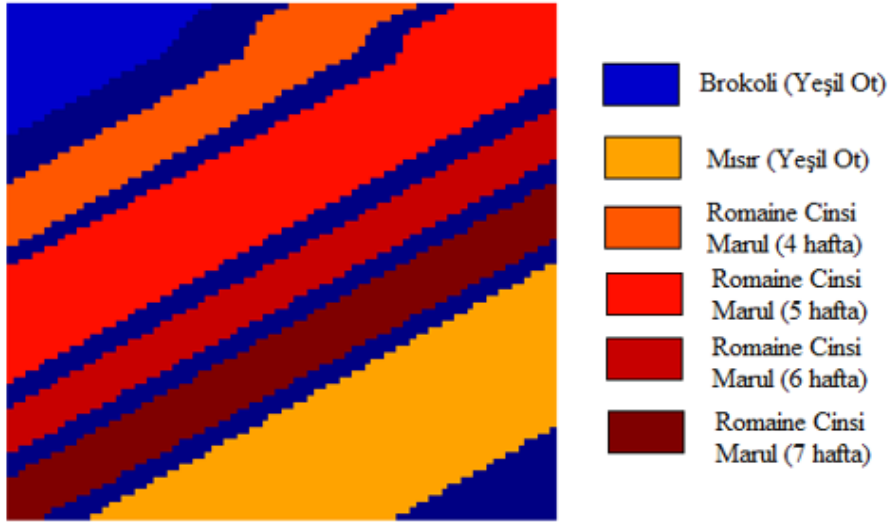
2.7.1. Ağın Tamamının Optimizasyonu

Oto-Kodlayıcılar ile oluşturulan Derin Öğrenme Ağlarında 1.gizli katman aracılığıyla çıktı olarak giriş verisi ile eğitim gerçekleştirilir. Ardından burada oluşturulan verinin diğer bir Oto-Kodlayıcı ile yeniden eğitimi gerçekleştirilir. Bu çıkış verisinin sınıflandırılması Softmax katmanında yapılır. Son olarak optimize edilmeden önce bütün yapılar birbiriyle bağlantılı hale gelir.

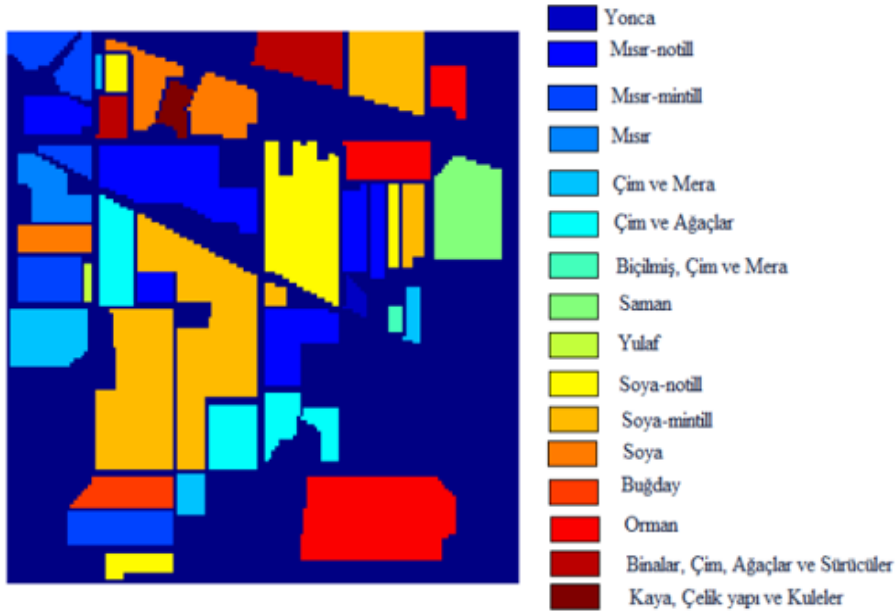
3. Veri Setleri ve Deneysel Uygulamalar

Bu çalışmada, önceki birçok karşılaştırma amaçlı uygulamada veri kaynağı olarak tercih edilmesi nedeniyle Salinas-A ve Indian Pines veri setleri kullanılmıştır. Salinas veri setinin küçük bir kısmı olan ve Şekil 2'de gösterilen Salinas-A veri seti 224 banda sahip ve 86x83 piksel çözünürlüğündedir. Salinas-A görüntüsünde 6 sınıf mevcuttur. Şekil 3'te verilen

200 banda sahip Indian Pines veri setinin ise 145x145 piksel çözünürlüğü mevcuttur. Görüntüde 16 farklı sınıf bulunmaktadır.



Şekil 2. Salinas-A Hiperspektral Uydu Görüntüsüne Ait Sınıflar



Şekil 3. Indian Pines Hiperspektral Uydu Görüntüsüne Ait Sınıflar

3.1. Deneysel Uygulamalar

Sınıflandırma performansını en fazla etkileyen unsurlardan biri, tercih edilen sınıflandırma tekniğinin kontrol parametreleridir. Özellikle çok fazla sayıda band içermesi nedeniyle hiperspektral görüntülerin sınıflandırılmasında kontrol parametreleri çok daha önemli bir hale gelmektedir. Band sayısı arttıkça yeryüzü objelerinin yansıtım karakterini tanımlamak daha efektif hale gelse de sınıflandırma tekniklerinin karar fonksiyonlarının tanımlanması güçleşmektedir. Örneğin Oto-Kodlayıcılar gibi yapay sinir ağı temelli yöntemlerde band sayısı arttıkça optimize edilmesi gereken ağırlık ve bias değerlerinin adedi dramatik olarak artmakta ve bu da tercih edilen optimizasyon algoritmasının performansını etkilemektedir. Her iki veri seti üzerinde karşılaştırma yapabilmek için kullanılan sınıflandırma teknikleri ve her tekniğin kendi parametreleri Tablo 1’de görüldüğü gibi belirtilmiştir. Parametrelerin belirlenmesinin ardından her iki veri seti üzerinde sınıflandırma işlemleri yapılmıştır.

Tablo 1. Sınıflandırma teknikleri ve değişken olan parametreleri

Oto-Kodlayıcı	Saklı Katman Sayısı: 30-50, Epok: 2000
	Encoder Transfer Fonksiyonu: Logsig (Log-Sigmoid) - Purelin (Linear) - Satlin (Saturating-Linear)
	Decoder Transfer Fonksiyonu: Logsig (Log-Sigmoid) - Purelin (Linear) - Satlin (Saturating-Linear)
	L2 Ağırlık Düzenlemesi Değeri: 1 – 10 - 20, Sparsity Oranı: 0.05 - 0.5
Destek Vektör Makineleri	Çekirdek Fonksiyonu: Linear – Gaussian – Polynomial - RTF
	Kernel Ölçeği: 1 – 5 – 10 – 30 – 50 - 100
	Çoklu Sınıflandırma: Bire karşı bir - Bire karşı hepsi
Karar Ağaçları	Bölünme Kriteri: Gini İndeksi - Twoing Algoritması
	Maksimum Bölünme Sayısı: 10 – 20 – 30 – 50 - 100
Torbalama Karar Ağaçları	Maksimum Bölünme Sayısı: 30 – 50 - 100
K-En Yakın Komşuluk	Mesafe Kriteri: Euclidean – Mahalanobis - Minkowski, NumNeighbors: 1 -5 - 10
Lineer Diskriminant	Gama Katsayısı: 0, Diskriminant Tipi: Lineer
Alt Uzay En Yakın Komşuluk	Öğrenme Döngü Sayısı: 30 - 50 - 100, Alt Uzay Sayısı: Öznitelik Sayısı / 2

İlk parametrik analiz Oto-Kodlayıcı tekniği ile gerçekleştirilmiştir. Tablo 2’de görüldüğü gibi Salinas-A veri seti üzerinde Oto-Kodlayıcı ile yapılan sınıflandırmada en yüksek test doğruluğunu veren ‘10’ düzenleme değerinde, ‘50’ saklı katman sayısında, ‘0.05’ Sparsity oranında encoder transfer fonksiyonunun ‘purelin’, decoder transfer fonksiyonunun ‘satlin’ olduğu uygulamadır. Indian Pines veri seti üzerinde Oto-Kodlayıcı ile yapılan sınıflandırmada en yüksek eğitim ve test doğruluğunu veren ‘20’ düzenleme değerinde, ‘30’ saklı katman sayısında, ‘0.05’ Sparsity oranında encoder transfer fonksiyonunun ‘purelin’, decoder transfer fonksiyonunun ‘satlin’ olduğu uygulamadır.

Tablo 2. Salinas-A ve Indian Pines veri seti için Oto-Kodlayıcı ile yapılan sınıflandırmada en yüksek doğruluğu veren parametreler ve doğruluk değerleri

Veri Seti	Encoder Transfer Fonksiyonu	Decoder Transfer Fonksiyonu	Düzenleme Değeri	Sparsity Oranı	Saklı Katman Sayısı	Eğitim Doğruluğu	Test Doğruluğu
Salinas-A	Logsig	Logsig	10	0.05	30	0.9264	0.9099
	Logsig	Purelin	20	0.5	30	0.9093	0.9028
	Logsig	Satlin	10	0.05	50	0.9073	0.8986
	Purelin	Logsig	1	0.5	50	0.951	0.9103
	Purelin	Purelin	20	0.5	50	0.9288	0.9047
	Purelin	Satlin	10	0.05	50	0.9336	0.9108
	Satlin	Logsig	20	0.5	30	0.9109	0.9024
	Satlin	Purelin	20	0.05	50	0.8741	0.8674
	Satlin	Satlin	1	0.5	30	0.8803	0.8809
Indian Pines	Logsig	Logsig	1	0.05	50	0.8154	0.7909
	Logsig	Purelin	20	0.5	30	0.7959	0.7792
	Logsig	Satlin	10	0.5	50	0.8001	0.7804
	Purelin	Logsig	10	0.05	50	0.9159	0.8275
	Purelin	Purelin	20	0.05	30	0.9064	0.8251
	Purelin	Satlin	20	0.05	30	0.9225	0.8253
	Satlin	Logsig	10	0.05	50	0.8146	0.7984
	Satlin	Purelin	10	0.05	30	0.8019	0.7798
	Satlin	Satlin	1	0.05	30	0.8133	0.7899

İkinci uygulama ise Destek Vektör Makineleri tekniğinin parametrelere bağlı sınıflandırma performansı incelenmiştir. Tablo 3'te görüldüğü gibi Salinas-A veri seti üzerinde Destek Vektör Makineleri ile yapılan sınıflandırmada en yüksek eğitim ve test doğruluğunu sağlayan 'bire karşı bir' çoklu sınıflandırmada, çekirdek fonksiyonunun 'polynomial', kernel ölçeğinin '5' olduğu uygulamadır. Indian Pines veri seti üzerinde Destek Vektör Makineleri ile yapılan sınıflandırmada en yüksek eğitim ve test doğruluğunu sağlayan 'Bire Karşı Bir' çoklu sınıflandırmada, çekirdek fonksiyonunun 'polynomial', kernel ölçeğinin '10' olduğu uygulamadır.

Tablo 3. Salinas-A ve Indian Pines veri seti için DVM ile yapılan sınıflandırmada en yüksek doğruluğu veren parametreler ve doğruluk değerleri

Veri Seti	Çoklu Sınıflandırma	Çekirdek Fonksiyonu	Kernel Ölçeği	Eğitim Doğruluğu	Test Doğruluğu
Salinas-A	Bire Karşı Bir	Lineer	1	0.8961	0.8842
Salinas-A	Bire Karşı Hepsi	Lineer	1	0.8597	0.8459
	Bire Karşı Bir	Gaussian	5	0.9025	0.872
	Bire Karşı Hepsi	Gaussian	5	0.8871	0.8655
	Bire Karşı Bir	Polynomial	5	0.9564	0.8898
	Bire Karşı Hepsi	Polynomial	5	0.9472	0.887
	Bire Karşı Bir	RTF	5	0.9025	0.872
	Bire Karşı Hepsi	RTF	5	0.8871	0.8655
Indian Pines	Bire Karşı Bir	Lineer	5	0.7983	0.7782
	Bire Karşı Hepsi	Lineer	1	0.7913	0.7684
	Bire Karşı Bir	Gaussian	10	0.8222	0.7823
	Bire Karşı Hepsi	Gaussian	5	0.8731	0.7901
	Bire Karşı Bir	Polynomial	10	0.9068	0.8305
	Bire Karşı Hepsi	Polynomial	5	0.984	0.8193
	Bire Karşı Bir	RTF	10	0.8222	0.7823
	Bire Karşı Hepsi	RTF	5	0.8731	0.7901

Üçüncü uygulamada ise Karar Ağaçları'nın parametrik analizi yapılmıştır. Tablo 4'te görüldüğü gibi Salinas-A ve Indian Pines veri seti üzerinde Karar Ağaçları ile yapılan sınıflandırmada en yüksek eğitim ve test doğruluğunu sağlayan çalışma 'Twoing Algoritması'nda '100' maksimum bölünme sayısı olan uygulamadır.

Tablo 4. Salinas-A ve Indian Pines veri seti için Karar Ağaçları ile yapılan sınıflandırmada en yüksek doğruluğu veren parametreler ve doğruluk değerleri

Veri Seti	Bölünme Kriteri	Maksimum Bölünme Sayısı	Eğitim Doğruluğu	Test Doğruluğu
Salinas-A	Gini İndeksi	100	0.9139	0.8608
	Twoing Algoritması	100	0.9204	0.8683
Indian Pines	Gini İndeksi	100	0.6992	0.6615
	Twoing Algoritması	100	0.7139	0.6799

Dördüncü uygulamada ise Torbalama Karar Ağaçları tekniğinin maksimum bölünme sayısına göre analiz yapılmıştır. Tablo 5'te görüldüğü gibi Salinas-A veri seti üzerinde Torbalama Karar Ağaçları ile yapılan sınıflandırmada en yüksek

eğitim ve test doğruluğunu sağlayan uygulamanın parametresi maksimum bölünme sayısı '100' dür. Indian Pines veri seti üzerinde Torbalama Karar Ağaçları ile yapılan sınıflandırmada en yüksek test doğruluğunu sağlayan uygulamanın parametresi maksimum bölünme sayısı diğer veri setinde olduğu gibi '100' dür. Çalışmada kapsamında kullanılan parametrelere göre her iki veri seti için de Torbalama Karar Ağaçları ile yapılan sınıflandırma sonucu elde edilen test doğruluğunun, Karar Ağaçları ile yapılan sınıflandırma sonucu elde edilen test doğruluğundan daha yüksek olduğu görülmektedir.

Tablo 5. Salinas-A ve Indian Pines veri seti için Torbalama Karar Ağaçları ile yapılan sınıflandırmada en yüksek doğruluğu veren parametreler ve doğruluk değerleri

Veri Seti	Maksimum Bölünme Sayısı	Eğitim Doğruluğu	Test Doğruluğu
Salinas-A	30	0.9182	0.8814
	50	0.9192	0.8842
	100	0.9202	0.8856
Indian Pines	30	0.722	0.7024
	50	0.7237	0.7001
	100	0.7224	0.7032

Beşinci uygulamada ise Lineer Diskriminant Analiz tekniğinin sınıflandırma performansı incelenmiştir. Tablo 6'da görüldüğü gibi Salinas-A ve Indian Pines veri seti üzerinde Lineer Diskriminant Analiz ile yapılan sınıflandırmada uygulanan parametreler ile başarı ölçütleri belirtilmiştir.

Tablo 6. Salinas-A ve Indian Pines veri seti için Lineer Diskriminant Analiz ile yapılan sınıflandırmada en yüksek doğruluğu veren parametreler ve doğruluk değerleri

Veri Seti	Gama Katsayısı	Sınıflandırma Tipi	Eğitim Doğruluğu	Test Doğruluğu
Salinas-A	0	Lineer	0.8321	0.8267
Indian Pines	0	Lineer	0.7039	0.6861

Altıncı uygulamada ise K-En Yakın Komşuluk tekniğinin parametrik analizi yapılmıştır. Tablo 7'de görüldüğü gibi Salinas-A ve Indian Pines veri seti üzerinde K-En Yakın Komşuluk ile yapılan sınıflandırmada mesafe kriterinin 'Euclidean' ve 'Minkowski' olduğu uygulamalarda 'Mahalanobis' mesafe kriterine kıyasla çok daha iyi sınıflandırma doğruluk değerleri elde edilmiştir.

Tablo 7. Salinas-A ve Indian Pines veri seti için K-En Yakın Komşuluk ile yapılan sınıflandırmada en yüksek doğruluğu veren parametreler ve doğruluk değerleri

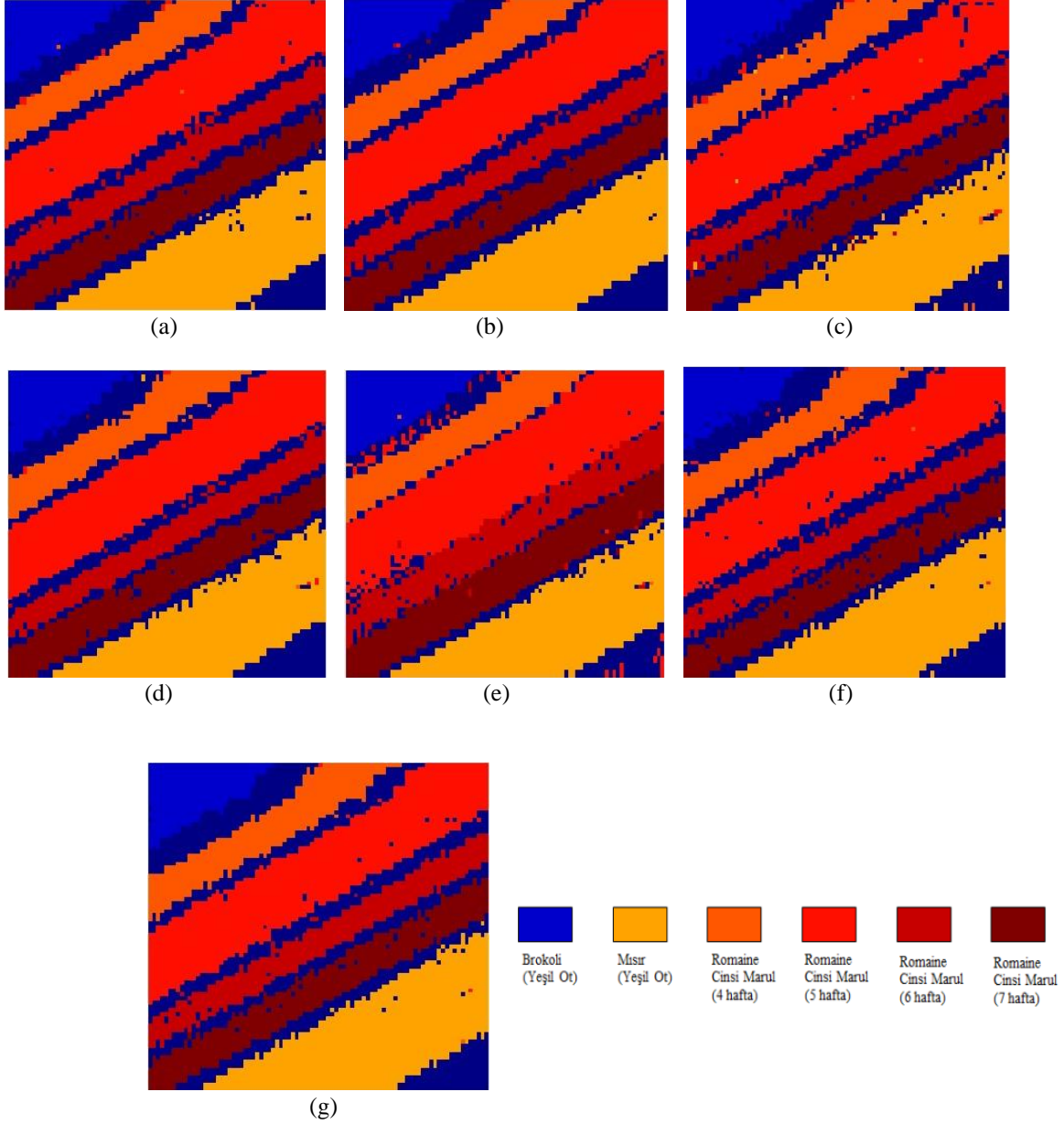
Veri Seti	Mesafe Kriteri	NumNeighbors	Eğitim Doğruluğu	Test Doğruluğu
Salinas-A	Euclidean	1	0.9758	0.8636
	Mahalanobis	1	0.9758	0.6254
	Minkowski	1	0.9758	0.8636
Indian Pines	Euclidean	10	0.7829	0.7316
	Mahalanobis	1	1	0.6582
	Minkowski	10	0.7829	0.7316

Son olarak yedinci uygulamada ise Alt Uzay En Yakın Komşuluk tekniğinin öğrenme döngü sayısına göre analiz yapılmıştır. Tablo 8'de görüldüğü gibi Salinas-A ve Indian Pines veri seti üzerinde Alt Uzay En Yakın Komşuluk ile yapılan sınıflandırmada en yüksek eğitim ve test doğruluğunu sağlayan öğrenme döngü sayısı '50' olan uygulamadır. Çalışmada kapsamında kullanılan parametrelere göre Alt Uzay En Yakın Komşuluk ile yapılan sınıflandırma sonucu elde edilen test doğruluğunun, K-En Yakın Komşuluk ile elde edilen test doğruluğuna kıyasla daha başarılı sonuçlar verdiği görülmektedir.

Tablo 8. Salinas-A ve Indian Pines veri seti için Alt Uzaklık En Yakın Komşuluk ile yapılan sınıflandırmada en yüksek doğruluğu veren parametreler ve doğruluk değerleri

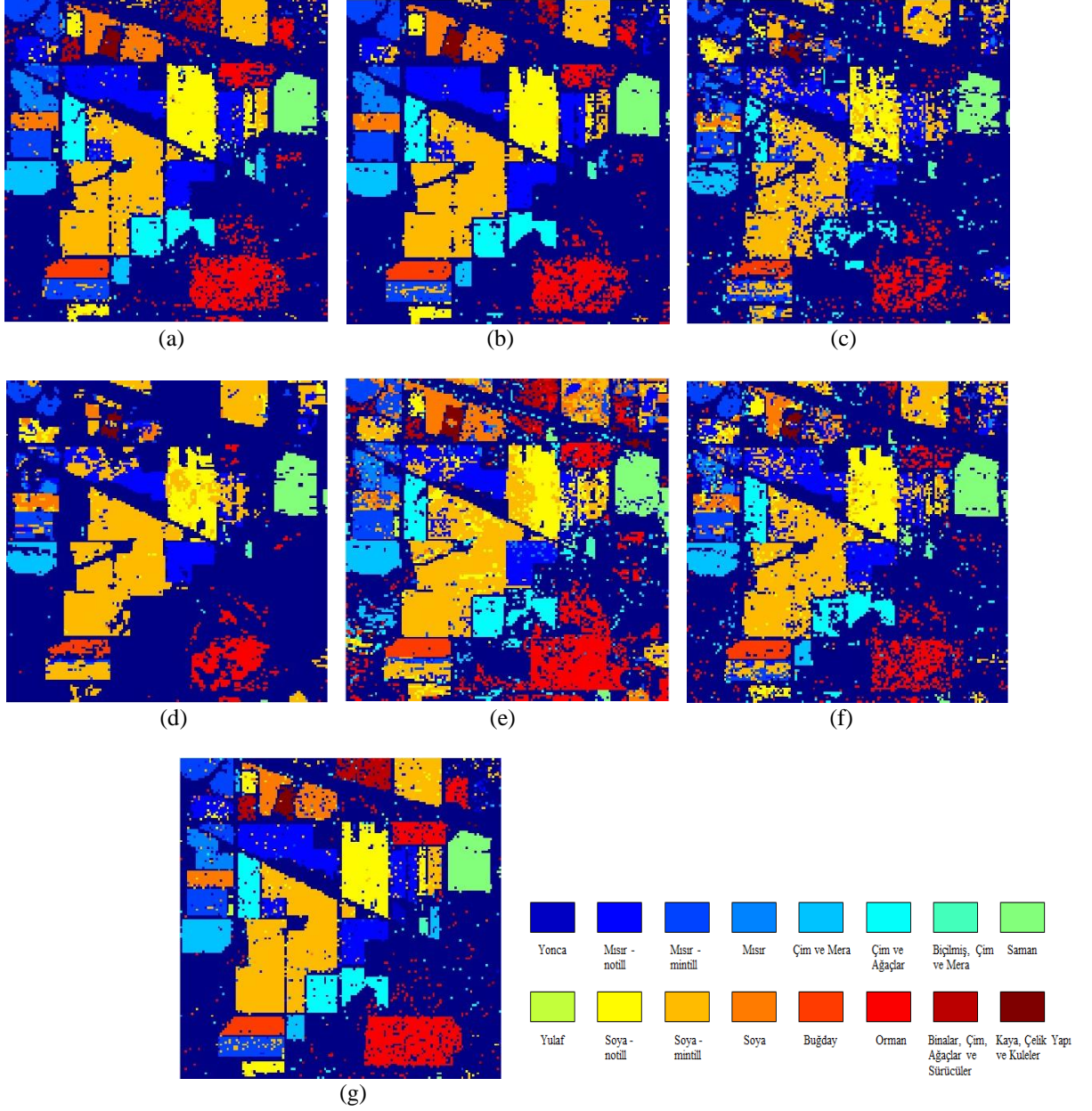
Veri Seti	Öğrenme Döngü Sayısı	Eğitim Doğruluğu	Test Doğruluğu
Salinas-A	50	0.9758	0.8916
Indian Pines	50	1	0.7649

Şekil 4'te belirtilen Salinas-A veri seti üzerinde tüm sınıflandırma teknikleri ile yapılan sınıflandırma uygulamalarında, her tekniğin kendi parametrelerindeki en yüksek test doğruluğuna sahip görüntülerdir.



Şekil 4. Sınıflandırma Tekniklerinin Salinas-A Verisi İçin En Yüksek Performansa Sahip Çıktıları: a)Oto-Kodlayıcı, b) Destek Vektör Makineleri, c) Karar Ağaçları, d) Torbalama Karar Ağaçları, e)Lineer Diskriminant Analiz, f) K-En Yakın Komşuluk, g)Alt Uzaklık En Yakın Komşuluk

Şekil 5'te belirtilen Indian Pines veri seti üzerinde tüm sınıflandırma teknikleri ile yapılan sınıflandırma uygulamalarında, her tekniğin kendi parametrelerindeki en yüksek test doğruluğuna sahip görüntülerdir.



Şekil 5. Sınıflandırma Tekniklerinin Indian Pines Verisi İçin En Yüksek Performansa Sahip Çıktıları: a) Oto-Kodlayıcı, b) Destek Vektör Makineleri, c) Karar Ağaçları, d) Torbalama Karar Ağaçları, e) Lineer Diskriminant Analiz, f) K-En Yakın Komşuluk, g) Alt Uzay En Yakın Komşuluk

4. Sonuçlar

Günümüzde birçok disiplin tarafından sıklıkla kullanılan hiperspektral görüntülerde, multispektral görüntülerdeki gibi tematik bilgiye erişim sınıflandırma ile yapılmaktadır. Fakat hiperspektral görüntülerin bant sayısı çok fazla olması sınıflandırma sürecini zorlaştırmaktadır. Bu çalışmada, Salinas-A ve Indian Pines veri setleri üzerinde Oto-Kodlayıcının performansı farklı parametreler belirlenerek değerlendirilmiştir. Deneysel uygulama sırasında sınıflandırma doğruluklarını karşılaştırabilmek için her biri farklı parametrelerde olacak şekilde altı farklı sınıflandırma yöntemi değerlendirilmedi göz önünde bulundurulmuştur. Ortaya çıkan sonuçlar incelendiğinde her iki veri seti için Oto-Kodlayıcı tekniğinin encoder transfer fonksiyonunun ‘purelin’, decoder transfer fonksiyonunun ‘logsig’ değişkenlerinde yapılan sınıflandırmanın diğer yöntemlere göre daha başarılı sonuçlar verdiği görülmektedir.

Kaynaklar

- Atasever, Ü.H. (2018) Investigation of the effectiveness of random subspace technique in classifying hyperspectral satellite images. T. Taşkanat, M.A. Günen, E. Beşdok (ed.) *II. International Scientific and Vocational Studies Congress – Engineering and Natural Sciences*, Nevşehir, Turkey, 70-78.
- Breiman, L. (1996) Bagging predictors. *Mach Learn* 24(2), 123-140.
- Cunningham, P., Delany, S. (2007) k-Nearest neighbour classifiers. *School of Computer Science and Informatic*, 15(4), 84-92.
- Günen, M.A. (2018) Hiperspektral görüntülerin otokodlayıcı derin öğrenme ile sınıflandırılması. Ü.H. Atasever, E. Beşdok (ed.) *VII. UZAL - CBS*, Eskişehir, Turkey, 51-59.
- Ho, T. K. (1998) The random subspace method for constructing decision forests. *IEEE transactions on pattern analysis and machine intelligence* 20(8), 832-844.
- Huang, C., Davis, L. S., Townshed, JRG. (2002) An assessment of support vector machines for land cover classification. *International Journal of Remote Sensing* 23(4), 725–749.
- Safavian, S. R., Landgrebe, D. (1991) A survey of decision tree classifier methodology. *IEEE Transactions on Systems, Man, and Cybernetics* 21(3), 660–674.
- Shippert, P. (2004) Why use hyperspectral imagery. *Photogrammetric Engineering & Remote Sensing* 70(4), 377-380.
- Teke, M. (2013) Classification on hyperspectral images using enhanced Fisher discriminant criterion. U. Sakarya (ed.) *Signal Processing and Communications Applications Conference (SIU) 21st*, Haspolat, Turkey, 1-4.