

MrSID Görüntü Sıkıştırma Formatının Sınıflandırma Doğruluğuna Etkisi (The Impact of MrSID Image Compression Format on Accuracy of Classification)

Gökhan ARASAN¹, Mustafa ERDOĞAN¹

¹Harita Genel Komutanlığı 06100 Cebeci, ANKARA
gokhan.arasan@hgk.msb.gov.tr

ÖZET

Bilgisayar teknolojisinin günlük yaşantımızın bir parçası haline gelmesi sonucunda verilerin bir ortamdan başka bir ortama iletilmesi ve bu verilerin depolanması en büyük problemlerden biri haline gelmiştir. Verilerin iletim ve depolanma problemi, veri sıkıştırma yöntemleri kullanılarak giderilmeye çalışılmaktadır.

Görüntü, ses ve video verileri sıkıştırmanın sıklıkla kullanıldığı veri kaynaklarıdır. Analog ve analitik fotogrametrik sistemler, bilgisayar teknolojisindeki gelişmelere paralel olarak yerini dijital fotogrametrik sistemlere bırakmıştır. Fotogrametrik işlemlerin tamamı, yaşanan süreç içerisinde bilgisayar sistemlerinde uygulanmaya başlamıştır. Fotogrametrik üretim zincirinin temelini oluşturan görüntülerin işlenmesi ve depolanması aşamalarında, görüntü sıkıştırma yoğun bir şekilde fotogrametrik uygulamalarda kullanılmaya başlanmıştır.

Bu çalışmada, Ankara'nın güneyinde Gölbaşı bölgesi seçilmiştir. Yaklaşık 300 km² alanı kapsayan bölge, detay çeşitliliği açısından oldukça zengindir. Bu nedenle, sınıflandırma ile tespit edilecek detay çeşitliliği açısından bölgenin uygun olduğu değerlendirilmektedir. GeoTIFF formatında 30 m çözünürlüklü 10 bant Landsat 8 uydu görüntüsü MrSID görüntü sıkıştırma formatında Lossless ve 10, 20, 30, 40, 60 oranında sıkıştırılmıştır. Piksel tabanlı sınıflandırma yöntemiyle görüntü sınıflandırması yapılmış ve hata matrisi yöntemiyle görüntülerin sınıflandırma doğruluğu belirlenmiştir. Elde edilen sonuçlar orijinal sıkıştırılmamış görüntü ve kayıpsız sıkıştırma ile elde edilen görüntünün sınıflandırma sonuçlarının benzer olduğunu göstermektedir. Ancak sıkıştırma oranları arttıkça sınıflandırma doğruluğu da düşmektedir.

Anahtar Kelimeler: Görüntü Sıkıştırma, Kayıpsız ve Kayıplı Sıkıştırma Teknikleri, Görüntü Sınıflandırma, Görüntü Sıkıştırmada Doğruluk, MrSID.

ABSTRACT

Computer technology is a part of daily life; for this reason, the most common issue is data transfer from one source to another and data storage. To overcome data transfer and storage issue, different data compression methods can be used.

Generally, data compression methods are used for compression image, audio and video files. Digital photogrammetric systems are replaced to analog and

analytic photogrammetric systems as the computer technology has been developed day by day. As a result of this, all photogrammetric processes are started to be applied to computer systems. In the photogrammetric applications, image compression frequently is used for image processing and storage which are basics of photogrammetric production.

In this article, Gölbaşı region, at south side of Ankara, has been studied. The region is approximately has 300 km² area and the area is very reach about detail variation. The area is evaluated as a suitable place because of its detail variation established by classification. 10 band Landsat 8 satellite images, 30 m resolution and in GeoTIFF format, have been compressed with lossless and lossy with 10, 20, 30, 40 and 60 rates in MrSID image compression format. Image classifications have been done by using resolution based classification technique and the accuracy of image classifications has been determined by confusion matrix method. Results of this study show that classification results of original uncompressed imagery and lossless compressed imagery are similar. But increase in compression ratio results as decrease in classification accuracy.

Keywords: Image Compression, Lossless and Lossy Compression Techniques, Accuracy on Image Compression, MrSID

1. GİRİŞ

Sürekli olarak kullandığımız ve çoğu zaman farkında olmadığımız veriler, aslında sıkıştırılmış veri kümeleridir. Görüntü, ses ve video verileri sıkıştırmanın sıklıkla kullanıldığı veri kaynaklarıdır. Verilerin bir ortamdan diğer bir ortama iletilmesi ve depolanma problemi, veri sıkıştırma yöntemleri kullanılarak giderilmeye çalışılmaktadır.

Analog ve analitik fotogrametrik sistemler, bilgisayar teknolojisindeki gelişmelere paralel olarak yerini dijital fotogrametrik sistemlere bırakmıştır. Fotogrametrik işlemlerin tamamı, yaşanan süreç içerisinde bilgisayar sistemlerinde uygulanmaya başlamıştır. Fotogrametrik üretim zincirinin temelini oluşturan görüntülerin işlenmesi ve depolanması aşamalarında, görüntü sıkıştırma yoğun bir şekilde fotogrametrik uygulamalarda kullanılmaya başlanmıştır.

Kullanılan sıkıştırma yöntemi ve oranı, elde edilecek olan doğruluk ve hassasiyeti doğrudan etkileyecek olması nedeni ile, gerçekleştirilecek fotogrametrik uygulamaya yönelik çözümlerin belirlenmesi önemlidir. Kullanılan sıkıştırma yöntem ve oranlarının, ardışık olarak yapılan işlemlerin tamamında ne şekilde sonuç vereceği konusu bu çalışmada ortaya konmuştur.

2. GÖRÜNTÜ SINIFLANDIRMA VE SIKIŞTIRMA TEKNİKLERİ

a. Görüntü Sınıflandırma

Uzaktan algılamada görüntü sınıflandırması, uydu görüntülerinden arazi örtüsüne yönelik bilgi çıkarımında en sık kullanılan yöntemlerden biridir (Borak ve Strahler, 1999; Chintan ve ark., 2004; Quattara ve ark., 2004). Farklı yöntemler kullanarak arazi yüzeyine ilişkin veri çıkarımı, bu verilerin analizi, yorumlanması, bu işlemlerin güncel ve çok kısa süreler içinde yapılabilirliğinin sağlanması ise uzaktan algılama verileri ile görüntü sınıflandırmanın temel konuları arasındadır.

Uzaktan algılamada görüntü sınıflandırması, görüntüdeki her pikselin ait olduğu sınıfın belirlenmesi işlemidir. Görüntü sınıflandırma işleminde amaç, nesnelerin farklı spektral bantlardaki yansıma değerleri kullanılarak, belirlenen bir matematiksel esasa göre pikselin ait olduğu sınıfı bulmaktır. Böylece görüntüdeki bütün pikseller arazide karşılık geldikleri sınıflar içine otomatik olarak atanır (Ayhan, 2003).

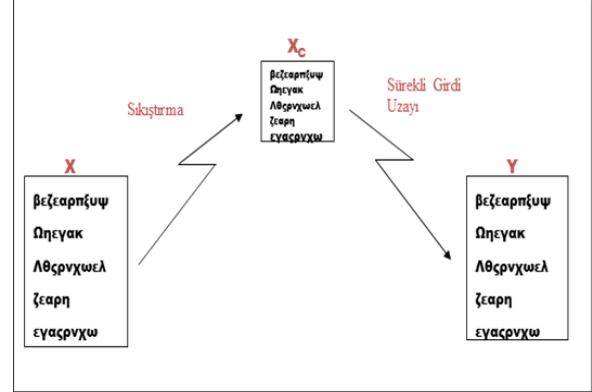
Boyacı; 2012 CBS-Uzaktan algılama entegrasyon ve örnek uygulama: Uydu görüntülerinden detay ve otomatik öznitelik tespiti çalışmasında piksel ve nesne tabanlı sınıflandırma örnekleriyle detaylı bir çalışma yaptığı değerlendirilmiştir.

b. Sıkıştırma Tekniği

Bilgisayar biliminde ve bilgi teorisinde veri sıkıştırma veya kaynak kodlama, özel kodlama algoritmaları kullanarak kodlanmamış kullanımından daha az boyutta olacak şekilde verinin kodlanması işlemidir. Herhangi bir veri iletişimde gönderici ve alıcının ikisi birden kodlama algoritmasını okuyabiliyorsa sıkıştırılmış veriler kullanılabilir.

Sıkıştırma tekniği veya algoritması denildiğinde gerçekte iki farklı algoritma kastedilmektedir. Birincisi X orijinal verisini girdi olarak alan ve bunu sıkıştırarak daha küçük boyutta bir X_C verisi oluşturan sıkıştırma

algoritması, ikincisi de X_C sıkıştırılmış verisini girdi olarak alan ve bunu açarak bir Y verisi oluşturan yeniden oluşturma algoritmasıdır. Bu işlemler şematik olarak Şekil 1' de gösterilmektedir (Sayood 2005)



Şekil 1. Sıkıştırma ve yeniden oluşturma

Yeniden oluşturulan verinin özellikleri, sıkıştırma tekniklerinin temel olarak iki bölümde sınıflandırılmasında belirleyici olmaktadır. Birincisi yüksek sıkıştırma uygulanan ve Y verisinin X verisi ile aynı olduğu kayıpsız sıkıştırma yöntemleri, diğeri de Y verisinin X verisi ile farklılıklar gösterdiği kayıplı sıkıştırma yöntemleridir.

Sıkıştırma yöntemlerinin karşılaştırılması konusunda Uçar;2011 Sıkıştırılmış raster görüntülerin fotogrametrik otomasyonda kalite ve doğruluk üzerindeki etkilerinin araştırılması çalışmasında farklı sıkıştırma formatında ve oranında görüntüler üreterek sonuçları değerlendirmiştir.

c. MrSID (Multiresolution Seamless Image Database)

Ortofotolar gibi yer konuşlu raster görüntülerin kodlanması için LizardTech firması tarafından geliştirilen ve patentli bir dosya formatıdır. Kayıplı veya kayıpsız görüntü sıkıştırma yapabilmekte olup şu anda kullanılmakta olan en son sürümü 4'üncü nesil MrSID'dir. Görüntü içeriğine ve renk derinliğine bağlı olarak farklı sıkıştırma oranları elde edilebilmektedir.

3. UYGULAMA

Bu çalışmada kaynak veri olarak 30 m çözünürlüklü Landsat 8 (10 bant) uydu görüntüsü kullanılmıştır. Bu görüntü farklı oranlarda MrSID formatında sıkıştırılmış ve sıkıştırılmış görüntüler kullanılarak piksel tabanlı sınıflandırma gerçekleştirilmiştir. Piksel tabanlı sınıflandırma

için en büyük olasılık sınıflandırma yöntemi kullanılmıştır. Elde edilen sınıflandırma sonuçları 120 referans noktada değerlendirilerek doğruluk araştırması yapılmıştır.

a. Çalışma Bölgesi

Çalışma bölgesi olarak Ankara'nın güneyinde Gölbaşı bölgesinde yaklaşık 300 km² alanı kapsayan bölge, detay çeşitliliği açısından oldukça zengindir. Bölgede özellikle büyük yerleşim alanları, tarım alanları ve iki adet göl detayı dikkati çekmektedir. Bu nedenle, sınıflandırma ile tespit edilecek detay çeşitliliği açısından bölgenin uygun olduğu değerlendirilmektedir. Çalışma alanı Şekil 2' de gösterilmiştir.



Şekil 2. Çalışma bölgesinin Landsat 8 Uydu Görüntüsü

b. Kullanılan Yazılımlar ve Veriler

ERDAS Imagine 2014 Yazılımı: Uydu görüntüleri ve hava fotoğraflarının görüntülenmesi, işlenmesi, zenginleştirilmesi ve analiz edilmesi işlemlerinin gerçekleştirilmesini sağlayan bir görüntü işleme yazılımıdır. Uzaktan algılama uygulamaları için elverişlidir (Erdas, 2014).

Uygulama çalışmasında, sınıflandırma için kaynak veri olarak Landsat uydu görüntüsü, sınıflandırma sonuçlarının doğruluğunu test etmek amacı ile 30 cm çözünürlüklü sayısal hava fotoğrafı kullanılmıştır (Şekil 3).

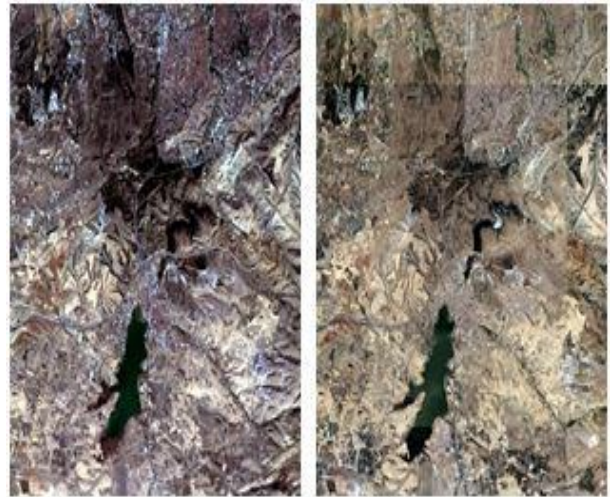
Landsat uydu görüntüsü geometrik düzeltmesi yapılmış olarak "www.earthexplorer.usgs.gov" internet adresinden temin edilmiştir. Uygulamada

15 m çözünürlüklü pankromatik bant dışında tüm bantlar kullanılmıştır. Tablo 1 Landsat 8 uydu görüntüsü bant çözünürlüğü. Görüntü alım tarihi 29 Ekim 2013'dur.

Tablo 1. Landsat 8 Uydu Görüntüsü Bant Çözünürlüğü

Spektral Aralık	Dalgaboyu (µm)	Çözünürlük
Band 1-Kıyı/Aerosol	0,433-0,453	30 m
Band 2- Mavi	0,450-0,515	30 m
Band 3- Yeşil	0,525-0,600	30 m
Band 4- Kırmızı	0,630-0,680	30 m
Band 5-Yakın Infrared	0,845-0,885	30 m
Band 6- Kısa Dalga Infrared	1,560- 1,660	30 m
Band 7- Kısa Dalga Infrared	2,100-2,300	30 m
Band 8- Pankromatik	0,500-0,680	15 m
Band 9- Sırrus	1,360- 1390	30 m

Hava Fotoğrafı: 100,5 mm. odak uzaklığına sahip Vexcel UltracamX kamera ile elde edilen 1/40.000 ölçekli hava fotoğrafları 4 bantlı ve 30 cm çözünürlüğe sahiptir. Bu görüntülerden üretilen ortofotolar Harita Genel Komutanlığından temin edilmiş olup, görüntü alım tarihi 28 Ağustos 2013'dir.



(a)

(b)

Şekil 3: Uygulamada kullanılan veriler a. Landsat 8 Uydu Görüntüsü b. Hava Fotoğrafı

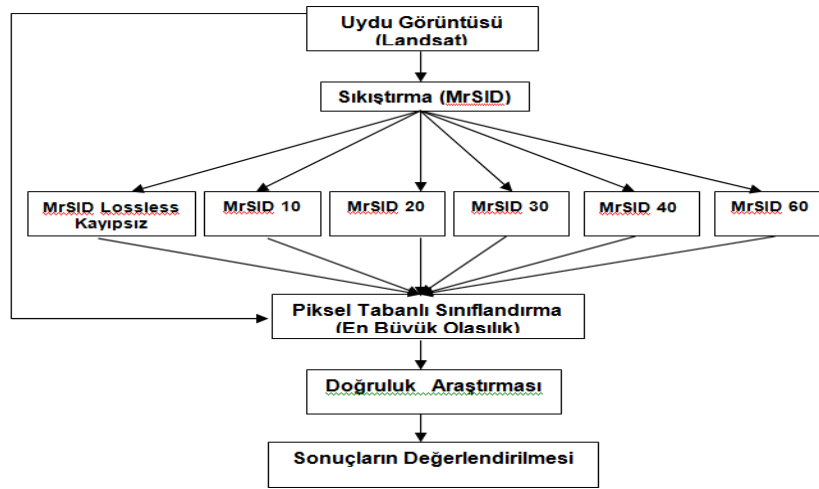
c. Metodoloji (İş Akışı)

Uygulama kapsamında Landsat uydu görüntüsü kullanılarak piksel tabanlı sınıflandırma için en büyük olasılık algoritması ile sınıflandırma yapılmıştır. Çalışmada 5 sınıf için sınıflandırma yapılmış ve sonuçlar 120 referans noktada yapılan doğruluk araştırması ile değerlendirilmiştir (Şeki 4 Metodoloji).

Piksel tabanlı sınıflandırma için ilk işlem adımı uydu görüntüsü üzerinden eğitim alanlarının toplanmasıdır. 5 detay sınıfı için Landsat uydu

görüntüsü üzerinden eğitim alanları alan vektör detaylar olarak belirlenmiş (Şekil 5) ve bu alanlar 7 görüntüde de (kayıpsız orijinal görüntü, ve kayıpsız, 10, 20, 30, 40 ve 60 oranında sıkıştırılmış mrsid görüntü) eğitim alanlarının toplanması için kullanılmıştır (Şekil 6).

Toplanan eğitim alanları girdi veri olarak kullanılmış ve görüntülerin sınıflandırması yapılmıştır. En büyük olasılık sınıflandırmasında ek bir parametre belirtilmemiştir.



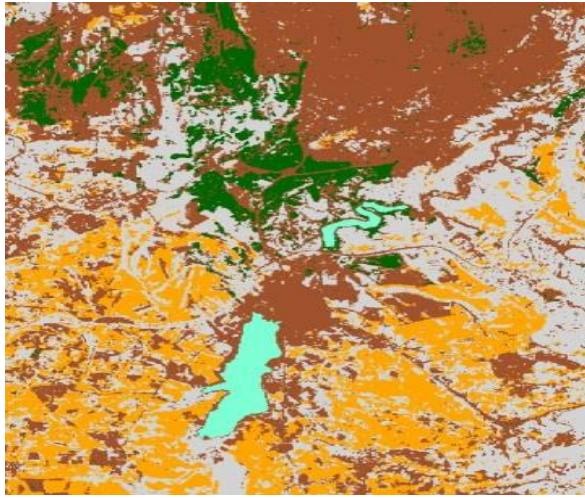
Şekil 4. Metodoloji



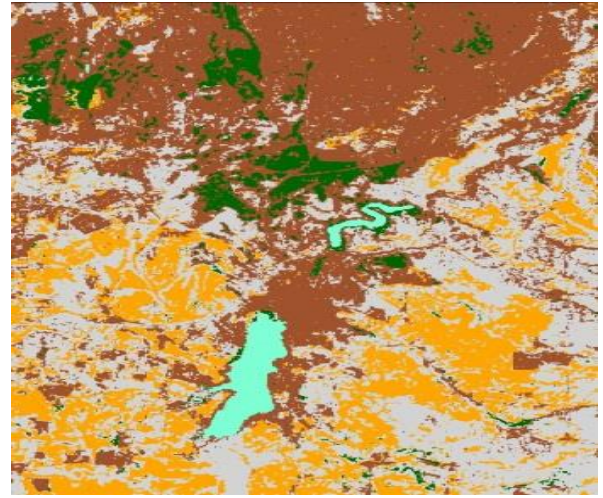
Şekil 5. Sınıf Alanları

Class #	>	Signature Name	Color	Red	Green	Blue	Value	Order	Count	Prob.	P	I	H	A	FS
1		ciplak_arazi		0.827	0.827	0.827	4	4	508	1.000	✓	✓	✓	✓	✓
2		oman		0.000	0.392	0.000	8	11	1316	1.000	✓	✓	✓	✓	✓
3		su		0.498	1.000	0.831	10	19	4720	1.000	✓	✓	✓	✓	✓
4		tarim		1.000	0.647	0.000	11	27	873	1.000	✓	✓	✓	✓	✓
5		yerlesim		0.627	0.322	0.176	9	34	1082	1.000	✓	✓	✓	✓	✓

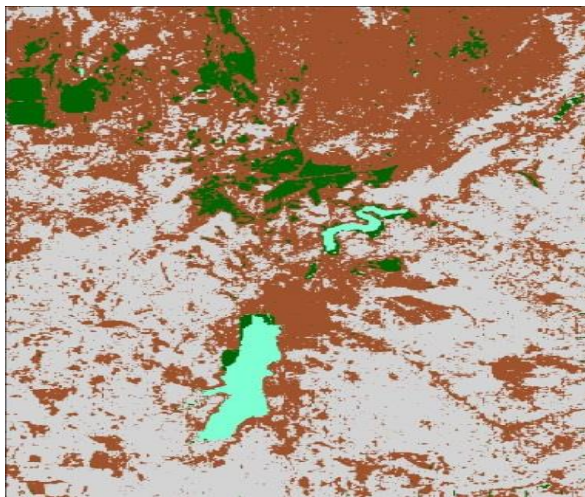
Şekil 6. Sınıf Dosyaları



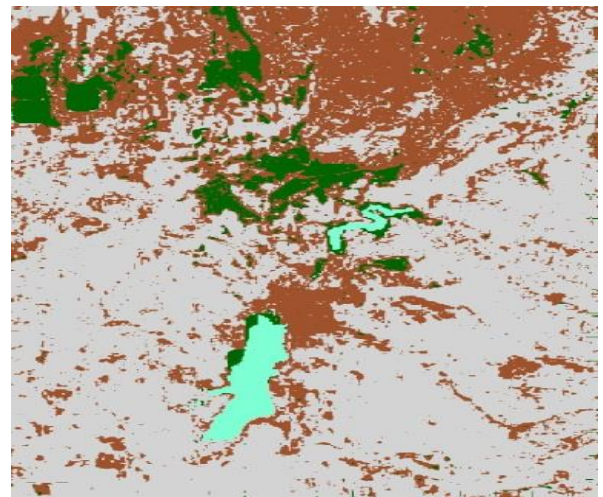
(a)



(b)



(c)



(ç)

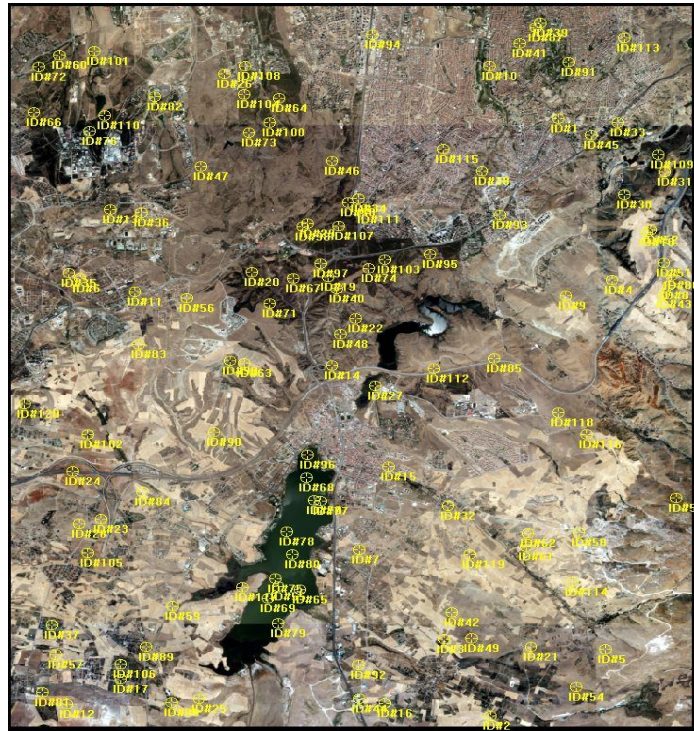
Şekil 7. Sınıflandırma sonuçları (a.Orijinal görüntü b. MrSID 20 c. MrSID 30 ç. MrSID 60 kat sıkıştırılmış görüntü)

Orijinal görüntü ile 20, 30 ve 60 oranında sıkıştırma ile elde edilen MrSID formatlı görüntülerin sınıflandırma sonuçları Şekil 7'de gösterilmiştir. Görüntüler incelendiğinde sıkıştırma oranı arttıkça sınıflandırma sonucunda tarım alanlarının çıplak araziye karıştığı görülmektedir.

Elde edilen sınıflandırma sonuçlarının doğruluğu Şekil 8'de gösterilen 120 referans noktada test edilmiştir. Yapılan tüm doğruluk araştırmalarında burada belirlenen aynı 120 nokta kullanılmıştır. Böylece farklı noktalara bağlı olarak yapılabilecek değerlendirme farklarının en

aza indirgenmesi sağlanmış ve daha nesnel bir değerlendirme yapılmıştır

5 sınıf için yapılan doğruluk değerlendirme sonuçları Tablo 2'de gösterilmiştir. Burada, her bir görüntü için elde edilen kullanıcı ve üretici doğrulukları ile yöntemin ortalama doğruluğu ve kapa değeri yer almaktadır. Doğruluk araştırması sonuçlarında çıplak arazinin üretici doğruluğu (ÜD) tarım araziyle karışması sonucu artarken kullanıcı doğruluğunun (KD) düştüğü gözlemlenmektedir. Sonuçlar incelendiğinde sıkıştırma oranı arttıkça sınıflandırma doğruluğunun da düştüğü görülmektedir



Şekil 8. Doğruluk araştırması için kullanılan noktalar

Tablo 2. Landsat ve MrSID kayıpsız, 10, 20, 30, 40, 60 oranında sıkıştırılmış görüntülerde sınıflandırma sonuçları

Sınıflar	Landsat		MrSID Kayıpsız		MrSID 10		MrSID 20		MrSID 30		MrSID 40		MrSID 60	
	ÜD (%)	KD (%)	ÜD (%)	KD (%)	ÜD (%)	KD (%)	ÜD (%)	KD (%)	ÜD (%)	KD (%)	ÜD (%)	KD (%)	ÜD (%)	KD (%)
Çıplak Arazi	69.44	82.29	75.00	87.10	77.78	73.68	58.33	63.64	66.67	50.00	80.56	52.73	80.56	49.15
Orman	73.08	86.36	61.54	88.89	53.85	87.50	53.85	93.33	42.31	84.62	42.31	84.62	46.15	85.71
Yerleşim	100.00	61.54	100.00	57.14	95.83	54.76	87.50	48.84	95.83	46.94	91.67	52.38	79.17	51.35
Su	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
Tarım	83.33	95.24	75.00	94.74	58.33	100.00	58.33	73.68	-	-	-	-	-	-
Genel Doğruluk Kapa	81.67	0.7660	79.17	0.7334	74.17	0.6670	66.67	0.5727	56.67	0.4357	60.00	0.4751	58.33	0.4508

4. SONUÇ VE ÖNERİLER

Literatürde çeşitli sıkıştırma tekniklerinin özellikle de yaygın olarak kullanılan jpeg sıkıştırma tekniğinin sınıflandırma üzerine etkisi konusunda çeşitli çalışmalar mevcuttur. Danoedoro (2013) tarafından yapılan çalışmada; ALOS-AVNIR2 orijinal sıkıştırılmamış görüntünün, JPEG formatında %10'dan %90'a kadar sıkıştırılması ile elde edilen görüntülerden daha yüksek sınıflandırma doğruluğuna sahip olduğu belirlenmiştir. Lau vd. (2001) tarafından yapılan çalışmada ise SPOT multispektral görüntü kullanılarak JPEG sıkıştırma formatında 35 kattan fazla yapılan sıkıştırma ile elde edilen görüntünün sınıflandırma doğruluğunun önemli ölçüde düştüğü tespit edilmiştir. Bu çalışma da elde edilen sonuçlar literatürdeki çalışmaları destekler niteliktedir. Ancak literatürdeki çalışmalardan farklı olarak bu çalışmada MrSID sıkıştırma tekniğinin sınıflandırma doğruluğu üzerindeki etkileri araştırılmıştır.

Bu çalışmada Landsat uydu görüntüsü MrSID sıkıştırma formatında 6 farklı oranda (kayıpsız, 10:1, 20:1, 30:1, 40:1, 60:1) sıkıştırılarak sınıflandırma sonuçları karşılaştırılmıştır. Sıkıştırma ve sonrasında gerçekleştirilen piksel tabanlı sınıflandırma işlemi ile ilgili elde edilen sonuçlar ve öneriler aşağıda belirtilmiştir.

- Sıkıştırma oranları nispetinde dosya boyutları küçülerek, genel olarak sıkıştırma probleminin temelini teşkil eden veri depolama ve veri iletimi kapsamında çözümlerin elde edilebileceği görülmüştür.

- Sıkıştırma oranının artması ile birlikte orijinal görüntüdeki dokuların, komşu pikseller ile gerçekleştirilen görüntü dengelemesi sonrasında yavaş yavaş kaybolduğu gözlenmiştir.

- Sıkıştırma oranı arttıkça tarım alanı detayı ile çıplak arazi detayının bazı alanlarda karıştığı görülmektedir. Uygulama çalışmasında kullanılan uydu görüntüsü Ekim ayına ait görüntüdür. Bu dönemde tarım alanlarının geneli hasat sonrası olduğundan çıplak arazi alanları ile karışabilmektedir. Bu nedenle bahar döneminde alınmış görüntüler kullanılarak yapılacak bir çalışma ile farklı sonuçlar elde edilebileceği faydalı olacağı değerlendirilmektedir.

- Kayıpsız sıkıştırma ile görüntüler bölgenin karakteristiğine bağlı olarak yaklaşık 5 kat küçültülebilmektedir. Bu görüntünün sınıflandırma için kullanımı orijinal görüntüye benzer doğruluk sonuçları vermektedir.

- MrSID formatında 10 kat sıkıştırmada büyük bir doğruluk kaybı olamamaktadır. Ancak 20 ve daha fazla kat yapılan sıkıştırmalarda doğruluk belirgin şekilde düşmektedir.

Sonuç olarak, özellikle kayıpsız ve düşük oranlarda sıkıştırılmış görüntülerin sınıflandırma için kullanılabileceği değerlendirilmektedir.

KAYNAKLAR

- Ayhan, E., Karslı, F., Tunç, E., 2003, **Uzaktan Algılanmış Görüntülerde Sınıflandırma ve Analiz**, Harita Dergisi, Sayı:130,32-46
- Borak, J.S. and Strahler, A.H., 1999, **Feature Selection and Land Cover Classification of a MODIS-like Dataset** for a Semi-arid Environment, International Journal of Remote Sensing, 20, 919-938.
- Boyacı, D., 2012,. **CBS-Uzaktan algılama entegrasyon ve örnek uygulama: Uydu görüntülerinden detay ve otomatik öznelik tespiti**, Doktora Tezi, *Selçuk Üniversitesi Fen Bilimleri Enstitüsü*, Konya.
- Chintan, A.S., Arora, M.K., Paramod, K.V., 2004, **Unsupervised Classification of Hyperspectral Data : an ICA Mixture Model Based Approach**, International Journal of Remote Sensing, 25, 481-487.
- Danoedoro, P., 2013, **The Effect Of Image Compression Level On The Land-Cover Classification Accuracy Of ALOS-AVNIR2 data Using Per-Pixel And Object Based Approaches**, Conference:34th Asian Conference on Remote Sensing, At Bali, Indonesia
- Erdas Field Guide, 2014, ERDAS Inc., Atlanta, USA.
- Quattara, T., Gwyn, Q.H.J., Dubois, J.-M.M., 2004, **Evalition of the Runoff Potential in High Relief Semi-Arid Regions Using Remote Sensing Data: Application to Bolivia**, International Journal of Remote Sensing, 25, 423-435.
- Sayood, K., 2005, **Introduction to data compression, 3rd ed.**, Morgan Kaufmann Publishers, San Diego, USA.

Uçar, E., 2011,. **Sıkıştırılmış raster görüntülerin fotogrametrik otomasyonda kalite ve doğruluk üzerindeki etkilerinin araştırılması**, Doktora Lisans Tezi, *Selçuk Üniversitesi Fen Bilimleri Enstitüsü*, Konya

W. L. LAU, Z. L. LI and W. K. LAM, 2001, **Effect Of JPEG Compression On Image Classification**, The Hong Kong Polytechnic University, Hong Kong SAR, Department of Land Surveying and Geo-Informatics, Hong Kong, PR China