

SINIFLANDIRMA SONRASI DOĐRULUK ANALİZİ

Doç. Dr. Ođuz GÜngör

Karadeniz Teknik Üniversitesi
Jeodezi ve Fotogrametri Mühendisliđi Bölümü
61080 Trabzon
ogungor@ktu.edu.tr

Doğruluk Deęerlendirmesi

- Sınıflandırma işleminin doğruluęu hakkında bilgi sahibi olmak için, yapılan sınıflandırmanın doğru kabul edilen coęrafik bir veri ile karşılaştırılmasıdır.
- Genelde doğru kabul edilen veri yer kontrol verisinden çıkartılır.
- Sınıflandırılan görüntüdeki tüm pikselleri test etmek veya doğrulamak normalde çok pratik bir işlem deęildir.
- Bu yüzden bir grup referans pikseller kullanılmaktadır.
- Referans pikseller, sınıflandırılmış görüntü üzerindeki bilinen yada bilinecek verilerdir.
- Referans pikseller rastgele seçilmektedirler. (Congalton, R. 1991. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. Remote Sensing of Environment 37: 35-46.)

Doğruluk Deęerlendirmesi

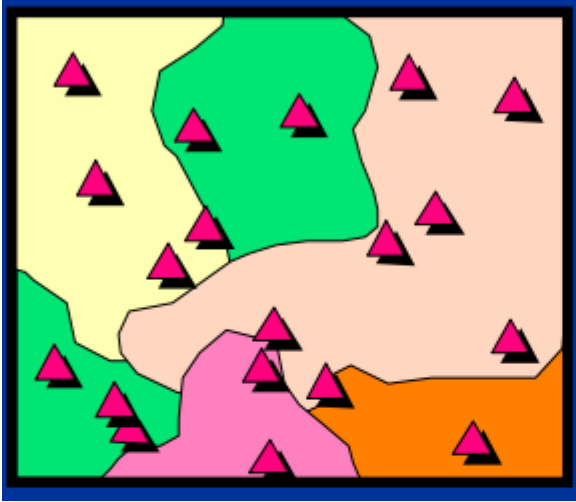
- Doğruluk Deęerlendirmesi ünitesi analizciye tematik raster harita üzerindeki bazı pikseller ile sınıfı bilinen referans piksellerini karşılaştırma imkanı sunmaktadır.
- Bu sınıflandırma sonuçlarını yer gözlem verileri ile, önceden test edilmiş haritalar ile, hava fotoęrafları ile veya diğer veriler ile organizeli bir şekilde karşılaştırma yöntemidir.
- Erdas Imagine programı sizin sınıflandırdığınız görüntünüz üzerinde rastgele noktalar üretir. Noktalar üretildikten sonra, referans noktaları olacak olan bu noktalar için sınıf etiketleri girmeniz gerekmektedir.
- Bu referans deęerlerinizi sizin sınıflandırılmış görüntünüzdeki sınıf deęerleriniz ile karşılaştırılacaktır.

Ne kadar referans pikseli gereklidir?

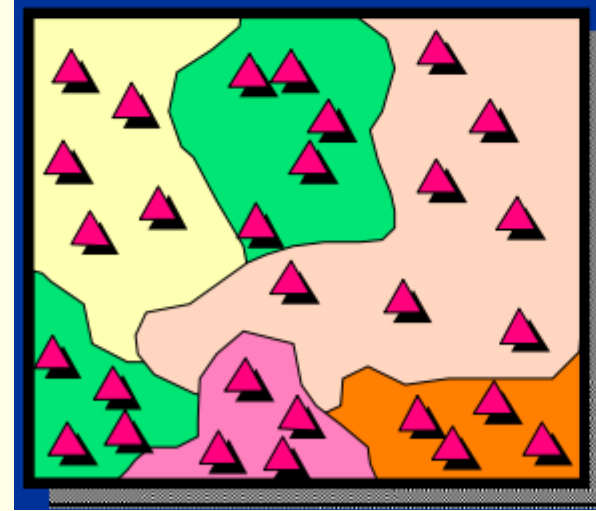
- Sınıflandırmanın doğruluğunu hassas bir şekilde belirleyebilmek için referans piksel sayısı çok önemlidir.
- Bir sınıfın doğruluk ortalamasını +%5 ve -%5 aralığında tahmin edebilmek için 250'den fazla referans piksele ihtiyaç olduğu bilinmektedir. (Congalton, R. 1991. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. Remote Sensing of Environment 37: 35-46.)
- Bu sayı çok yüksek çözünürlükteki uydu görüntülerinde çok daha fazla olmalıdır.

- ERDAS IMAGINE referans pikselleri seçmek için kare pencere kullanmaktadır.
- Pencerenin boyutu kullanıcı tarafından belirlenir.
- Rastgele pikselleri seçmek için 3 farklı dağılım yöntemi vardır.
 - **Rastgele—Herhangi bir kural kullanılmaz**
 - **Hiyerarşik Rastgele— Her bir tematik sınıf için rastgele noktaları sınıfların görüntü üzerinde kapladığı alanlarla orantılı olacak şekilde dağıtır.**
 - **Eşitlenmiş Rastgele — Sınıflara eşit sayıda rastgele noktalar dağıtılır.**

Rastgele



Hiyerarşik Rastgele



- Kaynak:<http://web.pdx.edu/~nauna/resources/9-accuracyassessment.pdf>

Hata Matrisi (Confusion or Error Matrix)

Point ID	Class	User																
1	4	4																
2	1	1	1	Su	13	0	0	0	0	0	0	0	13	68%	100%			
3	1	1	2	Yeşil Alan	0	52	0	0	0	0	0	0	52	100%	100%			
4	1	1	3	Toprak	0	0	19	0	0	0	0	0	19	100%	100%			
5	5	5	4	Kentsel Yapı1	0	0	0	28	0	0	0	0	28	97%	100%			
6	5	5	5	Kentsel Yapı2	0	0	0	0	79	1	0	0	80	100%	99%			
7	5	5	6	Kentsel Yapı3	0	0	0	1	0	32	0	0	33	97%	97%			
8	5	5	7	Gölge	6	0	0	0	0	0	0	19	25	100%	76%			
9	5	5		Sütun Toplam	19	52	19	29	79	33	19	250						
10	5	5		genel Doğruluk		96.80%												
11	5	5																
12	4	4																
13	2	2		Cohen's Kappa Statistic														
14	5	5																
15	2	2		Pr(a)=	0.9680													
16	2	2																
17	2	2		A(satır için)	Olasılıklar													
18	6	5	x	Su	0.0520													
19	6	6		Yeşil Alan	0.2080													
20	3	3		Toprak	0.0760													
21	2	2		Kentsel Yapı1	0.1120													
22	5	5		Kentsel Yapı2	0.3200													
23	5	5		Kentsel Yapı3	0.1320													
24	5	5		Gölge	0.1000													
25	7	7																

Üretici Doğruluğu %	Kullanıcı Doğruluğu %
$\frac{13}{19} \times 100 = 68$	$\frac{13}{13} \times 100 = 100$
$\frac{52}{52} \times 100 = 100$	$\frac{52}{52} \times 100 = 100$
$\frac{19}{19} \times 100 = 100$	$\frac{19}{19} \times 100 = 100$
$\frac{28}{28} \times 100 = 100$	$\frac{28}{28} \times 100 = 100$
$\frac{79}{80} \times 100 = 99$	$\frac{79}{80} \times 100 = 99$
$\frac{32}{33} \times 100 = 97$	$\frac{32}{33} \times 100 = 97$
$\frac{19}{25} \times 100 = 76$	$\frac{19}{25} \times 100 = 76$

$$Üretici_d. = \frac{Köşegen_elemanı}{Sütun_Toplamı} \times 100$$

$$Kullanıcı_d. = \frac{Köşegen_elemanı}{Satır_Toplamı} \times 100$$

$$Genel_doğruluk = \frac{Köşegen_elemanlar}{ToplamPikselSayısı} \times 100 = \frac{242}{250} \times 100 = 96,80$$

A(satır için)	Olasılıklar	B(sütun için)	Olasılıklar	Olasılık
Su	0.0520	Su	0.0760	0.0040
Yeşil Alan	0.2080	Yeşil Alan	0.2080	0.0433
Toprak	0.0760	Toprak	0.0760	0.0058
Kentsel Yapı1	0.1120	Kentsel Yapı1	0.1160	0.0130
Kentsel Yapı2	0.3200	Kentsel Yapı2	0.3160	0.1011
Kentsel Yapı3	0.1320	Kentsel Yapı3	0.1320	0.0174
Gölge	0.1000	Gölge	0.0760	0.0760

Olasılıklar hesaplarken:	Olasılıklar hesaplarken:
$\frac{Satır_Toplamı}{ToplamPikselSayısı}$	$\frac{Sütun_Toplamı}{ToplamPikselSayısı}$
$13 : 250 = 0,0520$	$19 : 250 = 0,0760$
$52 : 250 = 0,2080$	$52 : 250 = 0,2080$

$$Pr(e) = 0.1977$$

$$Kappa = \frac{(Pr(a) - Pr(e))}{(1 - Pr(e))} = 0.9601$$

Hata Matrisi referans noktalarını cxc matrisi içindeki sınıflandırılmış noktalar ile karşılaştırmaktadır. (c sınıf sayısıdır (Sınıf 0'ı da içine alacak şekilde))

■ Üretici doğruluğu

- Bir spektral sınıfı eğitmek veya test etmek için toplanan piksel topluluğundan kaç tanesinin doğru olarak sınıflandırıldığını gösterir. Örneğin, orman sınıfını eğitmek için N adet piksel toplanmış olsun. Bu piksellerden kaç tanesinin sınıflandırma sonrasında da orman olarak sınıflandırıldığını gösterir

■ Kullanıcı doğruluğu

- Bu sizin ürettiğiniz sınıflandırma haritasının doğruluğunu / güvenilirliğini ölçer. Yani sınıflandırma haritasındaki bir pikselin yeryüzünde de o pikseli temsil edip etmediğini ölçer. Örneğin, eğitim amaçlı toplanan piksellerin tamamından M tanesi orman olarak sınıflandırılmış olsun. Bu sınıfını eğitmek için M adet piksel toplanmış olsun. Bu M tane pikselden kaç tanesinin yeryüzünde de orman sınıfını temsil edip etmediğini ölçer.

Kappa Coefficient

- The Kappa coefficient expresses the proportionate reduction in error generated by a classification process compared with the error of a completely random classification.
- For example, a value of .82 implies that the classification process is avoiding 82 percent of the errors that a completely random classification generates (Congalton, R. 1991. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. Remote Sensing of Environment 37: 35-46.).