

**AŞIRI GRADYAN ARTIRMA ALGORİTMASI
KULLANARAK SENTINEL-1 ZAMAN SERİSİ
GÖRÜNTÜLERİNDEN ÜRÜN SINIFLANDIRMA**

**CROP CLASSIFICATION FROM SENTINEL-1 TIME-
SERIES IMAGES USING EXTREME GRADIENT
BOOSTING ALGORITHM**

SERHAT ÇABUK

PROF. DR. MUSTAFA TÜRKER

Tez Danışmanı

Hacettepe Üniversitesi

Lisansüstü Eğitim-Öğretim ve Sınav Yönetmeliğinin

Geomatik Mühendisliği Anabilim Dalı için Öngördüğü

YÜKSEK LİSANS TEZİ olarak hazırlanmıştır.

ÖZET

AŞIRI GRADYAN ARTIRMA ALGORİTMASI KULLANARAK SENTINEL-1 ZAMAN SERİSİ GÖRÜNTÜLERİNDEN ÜRÜN SINIFLANDIRMA

Serhat ÇABUK

Yüksek Lisans, Geomatik Mühendisliği Bölümü

Tez Danışmanı: Prof. Dr. Mustafa TÜRKER

Şubat 2021, 92 Sayfa

Bu tez çalışmasında; Gediz ovası bölgesinde yetiştirilen yaz dönemine ait tarım ürünlerinin, Sentinel-1 SAR çok zamanlı uydu görüntüleri kullanılarak sınıflandırması için makine öğrenme algoritmaları kullanılmış ve karşılaştırmaları yapılmıştır. Sınıflandırmada, Aşırı Gradyan Arttırma (XGBoost) ve Rastgele Orman (RO) algoritmaları kullanılmıştır. Görüntü olarak, 2017 yılı yaz tarım ürünlerinin üretim dönemlerini kapsayan 10 Nisan, 3 Mayıs, 2 Haziran, 2 Temmuz, 1 Ağustos, 7 Eylül, 10 Ekim ve 16 Kasım 2017 tarihlerine ait 8 adet Sentinel-1 SAR uydu görüntüsü seçilmiştir. Sentinel-1 SAR uydu görüntüleri bantlarına ek olarak; DY (Düşey-Yatay)/DD (Düşey-Düşey) Oranı, Karşıtlık (Kontrast), Homojenlik (Homogeneity), Farklılık (Dissimilarity), Entropi (Entropy), Açısız İkinci Moment (Angular Secondary Moment), Ortalama, Varyans (Variance) ve Korelasyon (Correlation) doku özellik bantları da oluşturulmuş ve sınıflandırma işlemine dâhil edilmiştir. Sınıflandırmada, bütün zaman serisi görüntüleri ve doku özellik bantları küme olarak kullanılmıştır.

Sınıflandırma doğruluğunun hesaplanması için araziden toplanan yer gerçekleri verileri ve Çiftçi Kayıt Sistemi (ÇKS) sistemi verileri referans veri olarak kullanılmıştır. Sınıflandırma ile tespit edilen ürünler buğday, domates, mera, mısır, pamuk, üzüm, yonca

ve zeytindir. Referans veriler rastgele olarak yarı yarıya eğitim ve test veri setleri poligonları olarak ayrılmıştır.

Elde edilen sonuçlara göre, XGBoost algoritması genel doğruluk değeri (%83,67) ve $\hat{K}$ değeri (0,8059), RO algoritması genel doğruluk değeri (%83,55) ve $\hat{K}$ değerinden (0,8056), çok düşük oranda da olsa, daha yüksek çıkmıştır.

XGBoost algoritması ile sınıflandırmada ürün seviyesinde pamuk %99,13 ile en yüksek kullanıcı doğruluğu, %96,59 ile en yüksek üretici doğruluğu ve %98,25 ile en yüksek dengelenmiş doğruluğu sağlamıştır. RO algoritması ile sınıflandırmada ürün seviyesinde yine pamuk; %99,34 ile en yüksek kullanıcı doğruluğu, %96,87 ile en yüksek üretici doğruluğu ve %98,40 ile en yüksek dengelenmiş doğruluğu sağlamıştır.

XGBoost algoritması ile sınıflandırmada ürün seviyesinde yonca %23,96 ile en düşük kullanıcı doğruluğu, %10,39 ile en düşük üretici doğruluğu ve %55,03 ile en düşük dengelenmiş doğruluğu sağlamıştır. RO algoritması ile sınıflandırmada ürün seviyesinde yine yonca %25,17 ile en düşük kullanıcı doğruluğu, %5,74 ile en düşük üretici doğruluğu ve %52,79 ile en düşük dengelenmiş doğruluğu sağlamıştır.

Anahtar Kelimeler: Aşırı Gradyan Artırma, Rastgele Orman, Doku Özellikleri Bantları, Tarımsal Ürün Tespiti, Sentinel-1 SAR Uydu Görüntüsü,

ABSTRACT

CROP CLASSIFICATION FROM SENTINEL-1 TIME-SERIES IMAGES USING EXTREME GRADIENT BOOSTING ALGORITHM

Serhat ÇABUK

Master of Science, Department of Geomatics Engineering

Supervisor: Prof. Dr. Mustafa TÜRKER

February 2021, 92 Pages

In this thesis, machine learning algorithms were used and made comparison to make classification of summer crop types in Gediz plain by using multitemporal Sentinel-1 SAR satellite images. Extreme Gradient Boosting (XGBoost) and Random Forest (RF) machine learning algorithms were used in the classification. A series of eight Sentinel-1 SAR images acquired during the crop growth season of 2017 (April 10, May 3, June 2, August 1, September 7, October 10 and November 16) were selected as the images. In addition to original bands of Sentinel-1, the VH/VV ratio band, and the computed texture bands contrast, homogeneity, dissimilarity, entropy, angular second moment, variance, and correlation were included in the classification operation. All time series images and texture bands were used as a layer stack in the classification.

To compute classification accuracy, the data collected through field survey and the farmer's registry system database were used as reference data. The crop types used in classification are wheat, tomato, pasture, corn, cotton, vineyard, clover, and olive trees. The reference dataset was randomly divided into two halves, one half for training and the other half for validation.

Based on the results obtained, the overall accuracy (83.67%) and $\hat{\kappa}$ value (0.8059) of the XGBoost algorithm were slightly higher than the overall accuracy (83.55%) and $\hat{\kappa}$ value (0.8056) of the RF algorithm.

At crop type level, cotton provided the highest user's accuracy of 99.13%, the highest producer's accuracy of 96.59%, and the highest balanced accuracy of 98.25% in classification using the XGBoost algorithm. In classification using the RF algorithm, again cotton provided the highest user's accuracy of 99.34%, the highest producer's accuracy of 96.87%, and the highest balanced accuracy of 98.40%. On the other hand, clover provided the lowest user's accuracy of 23.96%, the lowest producer's accuracy of 10.39%, and the lowest balanced accuracy of 55.03% in classification using the XGBoost algorithm. In classification using the RF algorithm, again clover provided the lowest user's accuracy of 25.17%, the lowest producer's accuracy of 5.74%, and the lowest balanced accuracy of 52.79%.

Keywords: Extreme Gradient Boosting, Random Forest, Texture Feature Bands, Crop Products, Sentinel-1 SAR Satellite Images,

TEŞEKKÜR

Yüksek lisans tez konusunun seçiminden hazırlanmasına kadar her zaman yanımda olan, bilgi ve tecrübesiyle yardımlarını eksik etmeyen danışman hocam Prof. Dr. Mustafa TÜRKER'e,

Desteğini her zaman hissettiğim eşim Ayşegül YILMAZ ÇABUK'a,

Bugünlere gelirken emekleri olan babam Veli ÇABUK ve annem Havva ÇABUK'a,

Teşekkürlerimi sunarım.

Serhat ÇABUK

Şubat 2021

Ankara

İÇİNDEKİLER

ÖZET	i
ABSTRACT	iii
TEŞEKKÜR	v
İÇİNDEKİLER	vi
ŞEKİLLER	viii
TABLolar	ix
KISALTMALAR	x
1. GİRİŞ	1
1.1. Geçmiş Çalışmalar	3
1.2. Radar Görüntüleri ile Yapılan Geçmiş Çalışmalar	3
1.3. Tezin Amacı	5
1.4. Kullanılan Yazılımlar	6
1.5. Tezin Organizasyonu	6
2. YÖNTEM	6
2.1. Çalışma Alanı ve Veriler	6
2.1.1. Çalışma Alanı	6
2.2. Sentinel-1 SAR Uydu Görüntüleri	13
2.3. Makine Öğrenme Algoritmaları	15
2.3.1. Karar Ağacı (Decision Tree)	15
2.3.2. Gradyan Artırmalı Ağaç (Gradient Boosting Tree) Algoritmaları	16
2.3.3. Aşırı Gradyan Artırma (XGBoost)	17
2.3.4. Rastgele Orman (Random Forest)	20
2.4. Sentinel-1 Görüntülerin İşlenmesi	21
2.4.1. Görüntülerin Ön İşlemleri	21
2.4.2. Doku Özelliklerinin Çıkarılması	22
2.5. Sınıflandırma İşlemi	27
2.5.1. Eğitim ve Test Verilerinin Belirlenmesi	27
2.5.2. Sınıflandırma	29
2.5.3. Doğruluk Analizleri	31

3. BULGULAR VE analizler	34
3.1. Bulguların Analizleri	34
3.2. XGBoost ve RO Algoritmaları İşlem Süreleri.....	45
3.3. Özellik Önem Derecelerinin Seçilmesi ve Sınıflandırma Sonuçları	45
4. SONUÇLAR ve ÖNERİLER.....	51
KAYNAKLAR.....	53
5.EKLER	63
EK 1 - Sentinel-1 DD/DY Doku Bantları.....	63
EK 2 – Sınıflandırma Sonuçları Hata Matrisleri	77
EK 3 - Tez Çalışması Orjinallik Raporu.....	83
ÖZGEÇMİŞ.....	84

ŞEKİLLER

Şekil 1. Çalışma Alanı	7
Şekil 2. Sentinel-1 SAR Uydu Görüntüleri DY Bant Görüntüleri	9
Şekil 3. Sentinel-1 SAR Uydu Görüntüleri DD Bant Görüntüleri	10
Şekil 4. Arazi Çalışmaları Sırasında Yerinde Görülen Üzüm Bağlarından Birisi.....	11
Şekil 5. Arazi Çalışmaları Sırasında Yerinde Görülen Mısır Tarlalarından Birisi.....	11
Şekil 6. Arazi Yer Gerçeği Verilerinin Çalışma Bölgesindeki Dağılımı.....	13
Şekil 7. Bir Karar Ağacı Algoritması Örneği	16
Şekil 8. Radar Uydu Görüntülerine Uygulanan Ön İşlem Adımları.....	21
Şekil 9. 10 Nisan 2017 Tarihli Sentinel-1 DY Görüntüsünden Hesaplanmış Doku Bantları	25
Şekil 10. 10 Nisan 2017 Tarihli Sentinel-1 DD Görüntüsünden Hesaplanmış Doku Bantları.....	26
Şekil 11. Çalışma Alanının Bir Bölümüne Ait, Tampon Alan Oluşturma Sonrası Elde Edilen Yeni Parsel Sınırlarının Temmuz 2017 DY Polarizasyon Görüntüsü ile Birlikte Gösterimi	28
Şekil 12. Çalışmada Kullanılan İşlem Adımları Diagramı	29
Şekil 13. XGBoost Algoritması Sınıflandırma Çıktısı (Tematik Harita)	38
Şekil 14. RO Algoritması Sınıflandırma Çıktısı (Tematik Harita).....	39
Şekil 15. RO Algoritması Bant Önem Dereceleri.....	46
Şekil 16. XGBoost Algoritması Bant Önem Dereceleri	47

TABLÖLAR

Tablo 1. Tarım Ürünleri Gelişim Süreçleri-Fenolojik Özellikleri.....	8
Tablo 2. Çalışma İçin Seçilen Sentinel-1 SAR Görüntüleri Tarihleri.....	8
Tablo 3. ÇKS Verileri Örneği	12
Tablo 4. XGBoost Makine Öğrenme Algoritması Varsayılan Parametre Değerleri	19
Tablo 5. XGBoost Optimizasyon Parametreleri ve Değerleri.....	20
Tablo 6. Ürün Bazında Eğitim ve Test Verisi Parsel ve Piksel Sayıları	28
Tablo 7. Sınıflandırmada Kullanılan Sentinel-1 DY ve DD Polarizasyon Görüntüleri ve Bu Görüntülerden Çıkarılan Doku Özellikleri Bantları.....	30
Tablo 8. Çoklu Sınıfa Ait Hata Matrisi (Campbell, 1996'dan alınmıştır.).....	31
Tablo 9. İkili Hata Matrisine Bir Örnek	32
Tablo 10. XGBoost ve RO Algoritmaları ile Sınıflandırma Sonuçları İçin Hesaplanmış $\hat{\kappa}$ ve Genel Doğruluk Değerleri.....	34
Tablo 11. XGBoost Algoritması ile Sınıflandırma Sonucundan Hesaplanan Hata Matrisi .	35
Tablo 12. RO Algoritması ile Sınıflandırma Sonucundan Hesaplanan Hata Matrisi.....	36
Tablo 13. XGBoost ve RO Algoritmaları Ürün Seviyesinde Kullanıcı, Üretici ve Dengelenmiş Doğruluk Değerlerinin Karşılaştırması	36
Tablo 14. XGBoost ve RO Algoritmaları McNemar Test Sonuçları	36
Tablo 15. XGBoost ve RO Algoritmaları Sınıflandırma Sonuçları Hata Matrisleri Karşılaştırması.....	44
Tablo 16. En Önemli 50 Bant ve 70 Bant ile Sınıflandırma Sonuçları İçin Hesaplanmış $\hat{\kappa}$ ve Genel Doğruluk Değerleri.....	48
Tablo 17. En Önemli 50 Bant Kullanarak XGBoost Algoritması ile Sınıflandırma Sonucundan Hesaplanan Hata Matrisi	48
Tablo 18. En Önemli 50 Bant Kullanarak RO Algoritması ile Sınıflandırma Sonucundan Hesaplanan Hata Matrisi	49
Tablo 19. En Önemli 70 Bant Kullanarak XGBoost Algoritması ile Sınıflandırma Sonucundan Hesaplanan Hata Matrisi	49
Tablo 20. En Önemli 70 Bant Kullanarak RO Algoritması ile Sınıflandırma Sonucundan Hesaplanan Hata Matrisi	50
Tablo 21. En Önemli 50 Bant Kullanarak XGBoost ve RO Algoritmaları İle Sınıflandırma Sonucu Üretici ve Kullanıcı Doğruluklarının Karşılaştırılması.....	50
Tablo 22. En Önemli 70 Bant Kullanarak XGBoost ve RO Algoritmaları İle Sınıflandırma Sonucu Üretici ve Kullanıcı Doğruluklarının Karşılaştırılması.....	51

KISALTMALAR

Kısaltmalar

ESA	European Space Agency
AUA	Avrupa Uzay Ajansı
XGBoost	Extreme Gradient Boosting (Aşırı Gradyan Artırma)
RO	Rastgele Orman
Ntree	Rastgele Orman Algoritması Ağaç Sayısı
Mtry	Rastgele Orman Algoritması Rastgele Değişken Sayısı
SAR	Synthetic Aperture Radar (Sentetik Açıklıklı Radar)
SNAP	Sentinel Application Platform (Sentinel Uygulama Platformu)
UTM	Universal Transverse Mercator (Evrensel Transversal Merkator)
WGS	World Geodetic System (Dünya Jeodezi Sistemi)
NASA	National Aeronautics and Space Administration (Ulusal Havacılık ve Uzay Dairesi)
ASF	Alaska Satellite Facility (Alaska Uydu Tesisi)
DY	Vertical-Horizontal (Dikey-Yatay)
DD	Vertical- Vertical (Yatay -Yatay)
SM	StripMap (Şerit Haritası)
IW	Interferometric Wide Swath (İnterferometrik Geniş Alan)
EW	Extra Wide Swath (Ekstra Geniş Alan)
WV	Wave (Dalga)
GRD	Ground Range Detected (Yer Aralık Tespiti)
GBM	Gradient Boosting Machine (Gradyan Artırma Makinesi)
OOB	Out of Bag (Genelleştirilmiş Hata)
GLCM	Gray Level Co-occurrence Matrix

1. GİRİŞ

Tarımsal üretimin etkili bir şekilde yönetilmesi ve izlenmesi için tarımsal ürünler hakkında istatistik bilgilerin çıkarılmasında tematik haritalama oldukça önem arz etmektedir. Yetiştirilecek ürünlerin yönetilebilmesi ve karar verilebilmesi için, ürün türleri ve ekim yapılacak tarımsal arazi hakkında bilgi sahibi olmak gerekmektedir. Bunun yanında, üretilen mahsullerin kalitesi ve sağlığı hakkında elde edilecek veriler de sürdürülebilir tarımsal üretim açısından önem arz etmektedir. Önemli bir veri olan ürün türü bilgisine ve dolayısıyla tarımsal haritalamaya hızlı, ekonomik ve uygulanabilir bir yaklaşım sağlayan önemli teknolojilerden birisi uzaktan algılama teknolojisidir.

Tarımsal uygulamalar, uzaktan algılamanın en yaygın kullanıldığı alanlardan biridir. Uydu görüntülerinden ürün türü haritalaması[1][2][3][4], tarımda izleme ve değerlendirme çalışmaları için mekansal bilgi birikimini arttırmaktadır. Uzaktan algılama yöntemleriyle tarımsal ürün deseni tespiti ayrıca biyokütle, ürün verimi, evapotranspirasyon ve mahsul suyu verimliliği gibi tarımsal ekosistem değişkenlerin modellenmesini de sağlar.

Tarımda uzaktan algılama uygulamaları için optik ve SAR uydularının her ikisi de kullanılmaktadır[5][6][7][8]. Optik görüntülerden ürün türü tespitinde pek çok çalışma yapılmasına ve ilerleme kaydedilmesine rağmen bulut örtüsü ve zamansal çözünürlük gibi kısıtlamalar nedeniyle istenen tarihlere ait görüntüleri bulmak ve dolayısıyla gerekli bilgiyi çıkarmak her zaman mümkün olamamaktadır. Optik uyduların aksine, SAR uyduları görüntü alımında kendi enerji kaynaklarını kullanırlar. SAR uydularının önemli avantajları arasında hava durumlarından etkilenmemeleri ve gece gündüz görüntüleme yeteneğine sahip olmaları yer almaktadır.

Tarımsal alanlarda ürün türü tespiti radar-tabanlı görüntülerin yaygın kullanıldığı alanlardan birisidir. İlk çalışmalardan olan Bush ve Ulaby tarafından gerçekleştirilen çalışmada ürün tespiti için çok zamanlı görüntüler kullanılmış ve önemi vurgulanmıştır[9]. Hoogeboom tarafından gerçekleştirilen çalışmada havadan çekilen SAR görüntülerinden ürün türleri başarılı bir şekilde tespit edilmiştir[10]. Bu ilk çalışmaların yapıldığı tarihten itibaren, SAR görüntülerinden tarımsal ürün deseni tespiti için birçok çalışma gerçekleştirilmiştir[6][11][12][13][14][15][16].

Ücretsiz olarak Avrupa Uzay Ajansı (ESA)'ndan elde edilebilen Sentinel-1 SAR ve Sentinel-2 optik uydu verilerinin tarımsal ürün sınıflandırmasındaki kullanımının son

yıllarda yaygınlaştığı görülmektedir[17][18][19][20][21][22]. Sentinel-1 uyduları 1A ve 1B olmak üzere ikiz uydulardan oluşmaktadır ve eş zamanlı uydular olması nedeniyle zamansal çözünürlükleri 6 güne kadar düşmektedir[5][23]. Yüksek zamansal çözünürlükte hem optik görüntüler hem de SAR görüntüleri çekebilmesi Sentinel-1 ve -2 uydularının dinamik arazi örtüsüne sahip olan tarım alanları üzerindeki çalışmalar için çok değerli veri kaynağı haline getirmiştir. Küresel tarım alanlarının önemli bir bölümü sıklıkla bulutla kaplı olduğu için, Sentinel-1 SAR görüntülerinin kullanımı, net optik görüntü elde etmenin zor olduğu ya da yeterli sayıda görüntünün alınamadığı bölgelerde, tarımsal ürün deseni tespiti için ayrı bir öneme sahiptir. Diğer yandan, oldukça dinamik arazi örtüsüne sahip yıllık yetiştirilen ürünlerin uydu görüntülerinden sınıflandırılması genellikle zaman serisi görüntülerin kullanılmasını gerektirmektedir, çünkü tek tarihli görüntüler, farklı fenolojik süreçlere sahip ürünlerin ayırımında benzer spektral yansımalarından dolayı, yeterli başarı oranını sağlayamamaktadır[9][15][16][24][25][26].

Bu çalışmada, Avrupa Uzay Ajansı tarafından ücretsiz olarak kullanıma sunulan çoklu zamansal Sentinel-1 SAR uydu görüntülerinden, Aşırı Gradyan Artırma (Extreme Gradient Boosting-XGBoost) ve Rastgele Orman (Random Forest-RF) makine öğrenmesi algoritmaları kullanılarak tarım parselleri seviyesinde tarımsal yaz ürünlerinin tespiti gerçekleştirilmiştir. Çalışma alanı olarak, Türkiye'nin en verimli tarım alanlarından olan Gediz Ovası'ndan bir alan seçilmiştir. Ürün çeşitliliğinin fazla olması ve bir yıl içerisinde birden fazla ürün alınabilmesi nedeniyle Gediz Ovası, Sentinel-1 SAR zaman serisi görüntülerinden ürün türü tespiti için zorlu bir alan olup bu çalışma için önem arz etmektedir. Sınıflandırmada Nisan-Kasım 2017 tarihleri arasında C-bandında (5.6cm) ve çift polarizasyon modundaki (Düşey Düşey-DD ve Düşey Yatay-DY) Sentinel-1A ve -1B SAR zaman serisi görüntülerden seçilen görüntüler kullanılmıştır. Görüntü modu olarak 250 km tarama alanı ve 5m×20m uzaysal çözünürlüğe göre kaydedildiği bir mod olan İnterferometrik Geniş Alan (IW) modu ve GRD (Ground Range Detected) formatı tercih edilmiştir. Gerekli görüntü ön işlemleri (örn., radyometrik kalibrasyon, speckle filtresi, arazi düzeltmesi, yeniden örnekleme, projeksiyon sistemi tanımlama) için SNAP (The Sentinel Application Platform) programı kullanılmıştır. Görüntü sınıflandırması R programı ortamında Aşırı Gradyan Artırma (XGBoost) ve Rastgele Orman (RO) makine öğrenme algoritmaları kullanılarak gerçekleştirilmiştir. Doğruluk analizleri, elde edilen sınıflandırma sonuçları ile referans

verilerin (Yer gerçekleri verileri ve Tarım ve Orman Bakanlığı'ndan temin edilmiş Çiftçi Kayıt Sistemi (ÇKS) verileri) karşılaştırılması suretiyle gerçekleştirilmiştir.

1.1. Geçmiş Çalışmalar

Literatür araştırması yapıldığında uzaktan algılama yöntemleri kullanılarak tarım ürünleri tespitinde birçok çalışma yapılmıştır.

Zhang (vd. 2019) 5 farklı makine öğrenmesini karşılaştırdıkları çalışmada XGB ve RO algoritmalarının genel doğrulukta yüksek olduğundan bahsetmiştir. Bunun yanında XGB algoritmasının büyük veri setleri ile çalışırken daha iyi hesaplama yapabildiği ve kullanışlı olduğundan bahsetmiştir[27].

Van Tricht vd. 2018 yılında Belçika'da yaptığı çalışmada Sentinel-1 ve Sentinel-2 uydu görüntüleriyle iki aşamalı sınıflandırma uygulaması yapmıştır. İlk aşamasında yerleşim, tarım alanları, orman ve hidroğrafya alanlarını sınıflandırmıştır. İkinci aşamada ise tarım alanlarını ürün seviyesinde sınıflandırmıştır. Rastgele Orman algoritması kullanarak maksimum doğruluğu %82 oranı ile elde etmiştir[28].

Alex O.Onojeghuo vd. 2018 yılında Çin'de yaptığı çalışmada Sentinel-1 ve Landsat uydu görüntüleriyle çeltik pirinci tarım ürünü sınıflandırma uygulaması yapmıştır. Sınıflandırmada Rastgele Orman ve Destek Vektör Makineleri algoritmaları kullanılmıştır. Radar ve optik görüntülerin bütün olarak kullanıldığı çalışma sonucunda Rastgele Orman algoritmasının Destek Vektör Makineleri algoritmasına göre daha doğru ve anlamlı sonuçlar elde edilmiştir[29].

Narin Ö.G. (vd. 2018) tarafından Tokat bölgesinde yapılan çalışmada, Sentinel-2A uydu görüntüsü 11, 8, 4 bant kombinasyonu kullanılarak beş adet arazi örtüsü sınıflandırması yapılmıştır. Sınıflandırma için Maksimum Olabilirlik, Yapay Sinir Ağları ve DVM algoritmaları kullanılmıştır. Sınıflandırma sonuçlarına göre, genel doğruluk Yapay Sinir Ağları ile %94,62, DVM ile %92,32 ve Maksimum Olabilirlik ile %85,29 olarak hesaplanmıştır[30].

1.2. Radar Görüntüleri ile Yapılan Geçmiş Çalışmalar

Jose Bermudez Castro (vd. 2017) Brezilya'da yaptığı çalışmada tropik iklimin olduğu bölgelerde yıllık sulama, tarımı ürünü yetiştirme dönemlerinin karışıklığına çözüm bulmak için çok zamanlı SAR uydu görüntülerini kullanarak tarım ürünü sınıflandırması yapmıştır. Yapay Sinir Ağları ve otomatik kodlayıcı derin öğrenme tekniklerini

kullanmıştır. En iyi doğruluğu eğitilmiş ve çok zamanlı uygulaması ile yapay sinir ağları ile elde etmiştir[31].

Kumari (vd. 2019) Hindistan'da yaptığı çalışmada soya fasulyesi tarım ürünü sınıflandırma uygulaması için 11 adet çok zamanlı Sentinel-1 SAR uydu görüntülerini kullanmıştır. Sınıflandırma uygulamasında kural bazlı sınıflandırıcı ve destek vektör makineleri algoritmasını kullanmıştır. En iyi Genel Doğruluğu %83,1 ve kappa değeri 0,74 ile Destek Vektör Makineleri algoritması vermiştir[32].

Niculescu (vd. 2018) Romanya'da yaptığı çalışmada çok zamanlı Sentinel-1 görüntülerini kullanarak sahil bölgesinde tarım ürünleri sınıflandırma uygulaması yapmıştır. Uygulamada Geçitli Tekrarlayan Birimler derin öğrenme ve Rastgele Orman makine öğrenme algoritmalarını kullanmış ve karşılaştırmıştır. Sonuç olarak genel doğruluk göz önünde bulundurulduğunda derin öğrenme algoritmasının makine öğrenme algoritmasına göre daha iyi sonuçlar verdiğini tespit etmiştir[33].

Sahu (vd. 2018) Hindistan'da yaptığı çalışmada Sentinel-1 SAR uydu görüntülerini kullanarak meyve bahçelerinin 5 farklı özellikte meyvenin sınıflandırmasını yapmıştır. Yapılan çalışmada 3 farklı sınıflandırma uygulaması sırasıyla şunlardır; En Yakın Komşuluk, Karar Ağacı ve Rastgele Orman algoritmasıdır. Yapılan sınıflandırma sonuçlarına göre en büyük doğruluk %81,09 ile En Yakın Komşuluk sınıflandırmasıyla elde edilmiştir[34].

Robertson (vd. 2019) Kanada'da yaptığı çalışmada Dünya çapında 12 farklı bölgede 3 adet seçilen tarım ürününü (Mısır, Soya Fasulyesi, Buğday) 28 farklı çalışma alanında RADARSAT-2 ve Sentinel-1 uydu görüntülerini kullanarak sınıflandırmasını yapmıştır. Sınıflandırma için Karar Ağacı ve Rastgele Orman algoritmaları kullanılmıştır. İstatistiksel sonuçlara bakıldığında bütün bölgeler için Rastgele Orman algoritmasının Karar Ağacı algoritmasına göre daha yüksek genel doğruluğa sahip olduğu tespit edilmiştir. En yüksek sonuç ise Amerika Birleşik Devletleri Iowa eyaletindeki çalışma alanında %96,2 genel doğrulukla tespit edilmiştir. Aynı bölgede Karar Ağacı algoritmasının doğruluğu ise %95,6 olarak kaydedilmiştir[35].

Sentinel-1 SAR çok zamanlı görüntüleri kullanılarak XGB makine öğrenme algoritmasını Dobrinic (vd. 2020) kullanarak yaptığı çalışmada piksel seviyesinde sınıflandırma sonuçlarının iyi seviyede olduğu görülmüştür[36].

Gasparovic ve Dobrinic 2020 yılında yaptığı çalışmada tek ve çok zamanlı Sentinel-1 uydu görüntüleri kullanarak farklı makine öğrenme algoritmalarının doğruluklarını karşılaştırmıştır. Yapılan karşılaştırmalarda XGB algoritmasının çok zamanlı uydu görüntülerinin yapıldığı çalışmada daha iyi sonuçlar verdiği tespit edilmiştir[37].

Nguyen-Thanh Son vd. 2017 yılında Vietnam'ya yaptığı çalışmada Sentinel-1 uydu görüntüleriyle pirinç tarım ürünü sınıflandırma uygulaması yapmıştır. Sınıflandırmada Rastgele Orman ve Destek Vektör Makineleri algoritmaları kullanılmıştır. Sınıflandırma sonuçlarına göre genel doğruluk ve kappa değerleri karşılaştırıldığında Rastgele Orman algoritmasının Destek Vektör Makineleri algoritmasına göre daha iyi sonuçlar verdiği görülmüştür[38].

Nasirzadehdizaji R. (vd. 2019) Konya bölgesinde yaptığı çalışmada aynı hasat dönemine ait dört adet Sentinel-1 (DD-DY) görüntüleri kullanarak belirlenen 10 parametre ile geniş ve dar yapraklı mısır, ay çiçeği ve buğday tarım ürünlerinin boyları ve arazideki kapladığı alanların gölgelerinin geri yansıma değerleri ile ilişkileri ve etkilerini araştırmıştır. Sentinel-1 C bandı görüntülerinin tarım ürünleri büyüklüğü, fenoloji evreleri ve geniş yapraklı mısır ürününün arazide kapladığı alan ve gölgelerinin tespit edilmesinde iyi sonuçlar verdiği görülmüştür[39].

Üstüner M. (vd. 2019) tarafından Konya bölgesinde yapılan çalışmada, üç adet Rasarsat-2 PolSAR görüntüsü kullanılarak beş farklı tarım ürününün sınıflandırması yapılmıştır. Sınıflandırma için LightGBM, RO ve DVM algoritmaları kullanılmıştır. Sınıflandırma sonuçlarına göre genel doğruluk değerleri, LightGBM algoritması için %85,7, RO algoritması için %85,5 ve DVM algoritması için %83,4 olarak hesaplanmıştır[40].

Üstüner M. (vd. 2017) tarafından yapılan çalışmada, Sentinel-1 görüntüsünden İstanbul bölgesine ait beş adet arazi örtüsü sınıflandırma ile tespit edilmiştir. Sınıflandırma için üç adet ek bant (DD-DY, DD+DY+ DD-DY/DD+DY) oluşturulmuş ve algoritma olarak DVM, RO ve K-EYK algoritmaları kullanılmıştır. Sınıflandırma sonuçlarına göre en yüksek doğruluk değeri (%81,90) DVM algoritması ile elde edilmiştir. Spektral bant seçimi ve bilgi artışının sınıflandırma doğruluğunu artırmadığı görülmüştür[41].

1.3. Tezin Amacı

Bu tez çalışmasının ana amacı, Avrupa Uzay Ajansı tarafından ücretsiz olarak kullanıma sunulmuş olan Sentinel uydu görüntülerinden, sadece çoklu zamansal Sentinel-1 SAR uydu görüntüleri kullanarak XGBoost ve RO makine öğrenmesi algoritmaları ile tarım

parselleri seviyesinde tarımsal yaz ürünleri tespit etmek ve bu iki algoritmanın performanslarını karşılaştırmaktır. Çalışmanın diğer amaçları arasında, Sentinel-1 SAR zaman serisi DD ve DY polarizasyon modlarındaki görüntülerin ve bu polarizasyon modlarındaki görüntülerden hesaplanacak doku metriklerinin sınıflandırma doğruluğuna olan etkilerinin araştırılması yer almaktadır.

1.4. Kullanılan Yazılımlar

Çalışmada SNAP, R ve ArcGIS yazılımları kullanılmıştır. Sentinel-1 SAR uydu görüntülerinin ön işlemleri ve doku özelliklerinin çıkarılması SNAP programı kullanılarak yapılmıştır. ArcGIS, tarım parselleri sınır alanlarına yönelik hataların giderilmesi, eğitim-test alanlarının belirlenmesi ve sonuçların görsel olarak kontrol işlemleri için kullanılmıştır. XGBoost ve RO makine öğrenme algoritmalarıyla sınıflandırma işlemleri, R programı ortamında gerçekleştirilmiştir.

1.5. Tezin Organizasyonu

Tez çalışması 5 bölümden oluşmaktadır. Tezin 1.bölümü geçmiş çalışmalar, tezin amacı, kullanılan yazılımlar ve tezin organizasyonunu içermektedir. 2.bölümde yöntemler anlatılmıştır. Bu bölüm, çalışma alanı, veriler, kullanılan makine öğrenme algoritmaları, görüntülerin işlenmesi ve sınıflandırma uygulaması yöntemlerini içermektedir. 3. bölümde sınıflandırma sonucu elde edilen bulgular ve çıkan sonuçların değerlendirilmeleri ve karşılaştırmaları yer almaktadır. 4. bölümde bu tez çalışmasından çıkan sonuçlar verilmiştir. 5. bölümde ise çalışma sonuçlarının tabloları mevcuttur.

2. YÖNTEM

2.1. Çalışma Alanı ve Veriler

2.1.1. Çalışma Alanı

Çalışma alanı olarak tarım ürünlerinin incelenmesi söz konusu olduğu için ürün çeşitliliği yüksek bir bölge tercih edilmiştir. Tercih edilen bölge Ege Bölgesinin en verimli yerlerinden biri olan Gediz ovasıdır. Siyasi sınırları Manisa ili Salihli ve Ahmetli ilçelerinin sınırları içinde bulunmaktadır. Çalışma alanı yüz ölçümü yaklaşık 520 km²'dir. Kuzey-güney yönünde 18 km ve doğu-batı yönünde 30 km yaklaşık mesafe değerleridir (Sol Üst Enlem: 38,620° Boylam: 27,840°, Sağ Alt Enlem: 38,460°, Boylam: 28,180°).

Çalışma alanı çok çeşitli tarımsal ürünün yetiştiği bir alandır. Alanda yer alan temel ürünler çoğunlukla mısır, yonca, zeytin, domates, biber, patlıcan, pamuk, buğday ve üzumdür. Bu ürünlerin yanı sıra lahan, brokoli ve ıspanak tarım ürünlerinin de üretimi yapılmaktadır.

Yetiştirilen ürün çeşitliliği ve bir yıl içinde birden fazla hasat yapılabilmesi alanın tarım anlamında önemli bir bölge olmasını sağlamaktadır. Bölgede yetiştirilen ürün çeşitliliği gibi tarım parsellerinin büyüklükleri bakımından da çeşitlilik bulunmaktadır. Yaklaşık olarak en küçük tarım parseli 2 dekar ve en büyüğü 4600 dekardır. Çalışma alanı Şekil 1’de gösterilmiştir.



Şekil 1. Çalışma Alanı

2.1.2. Veriler

Çalışmada 8 farklı tarihe ait çoklu zamanlı Sentinel-1 SAR uydu görüntüleri kullanılmıştır. Sentinel-1 uydu görüntüleri Avrupa Uzay Ajansı Copernicus açık erişim internet sitesinden (<https://scihub.copernicus.eu/dhus/#/home>) ücretsiz olarak indirilmiştir[42]. Uydu görüntülerinin yanında arazide yerinde tespit edilmiş, görülmüş veriler ve Çiftçi Kayıt Sistemi (ÇKS) verileri kullanılmıştır.

2.1.2.1. Sentinel-1 SAR Uydu Görüntüleri ve Tarihlerin Seçimi

Çalışmada Sentinel 1 SAR zaman serisi uydu görüntüleri kullanılmıştır. Görüntü seçimi yapılırken çalışmanın gerçekleştirileceği bölgedeki tarım alanlarının fenolojik özelliklerinin bilinmesi önemlidir. Bu çalışmada, sınıflandırma ile tespit edilecek ürünlerin fenolojik özellikleri Türkiye Fenoloji Atlasından temin edilmiş ve Tablo 1’de gösterilmiştir[43].

Tablo 1. Tarım Ürünleri Gelişim Süreçleri-Fenolojik Özellikleri

	Mart	Nisan	Mayıs	Haziran	Temmuz	Ağustos	Eylül	Ekim	Kasım	Tarım Ürün Gelişim Süreci
Buğday										
Domates										Ekim Dönemi
Mısır										
Mısır2										Büyüme Dönemi
Pamuk										Yaprak Dökümü
Üzüm	Üzüm bitkisinin gelişiminde en önemli zamanı yaprak dökümüdür.									
Yonca	Yonca hasadı birkaç haftada bir yapıldığından tabloda gösterilmemektedir.									
Zeytin	Zeytin bitkisinin belirgin değişim göstermediği ve yaprak dökmediği için yansıma değeri değişim göstermemektedir									

Çalışma alanındaki tarım ürünlerinin fenolojik özellikleri göz önünde bulundurularak uydu görüntülerinin Nisan-Kasım ayları arasında olmasına karar verilmiştir. Dolayısıyla, bu aylar arasında 2017 yılına ait 8 adet uydu görüntüsü seçilmiştir. Fenolojik özelliklere bakıldığında ürünlerde aylık değişimlerin olduğu görülmektedir. Bu kapsamda görüntü seçimi, her ay için bir görüntü olacak şekilde, yapılmıştır. Seçilen tarihlere ait Sentinel-1 SAR görüntüleri DD ve DY olarak Şekil 2 ve Şekil 3’te gösterilmiştir.

Tablo 2. Çalışma İçin Seçilen Sentinel-1 SAR Görüntüleri Tarihleri

Görüntü Tarihleri	Sentinel-1 SAR Uydusu
	10 Nisan 2017
	03 Mayıs 2017
	02 Haziran 2017
	02 Temmuz 2017
	01 Ağustos 2017
	07 Eylül 2017
	10 Ekim 2017
	16 Kasım 2017

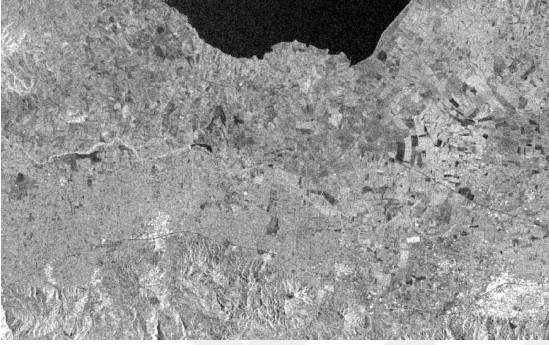
10 Nisan 2017



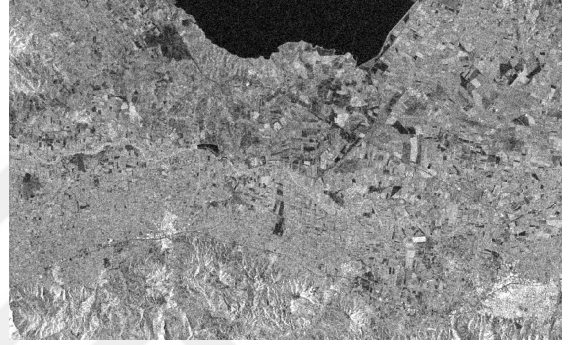
03 Mayıs 2017



02 Haziran 2017



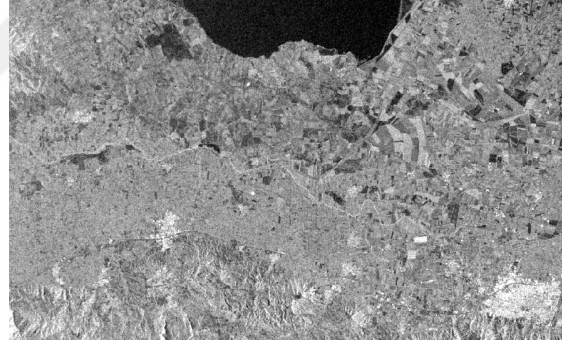
02 Temmuz 2017



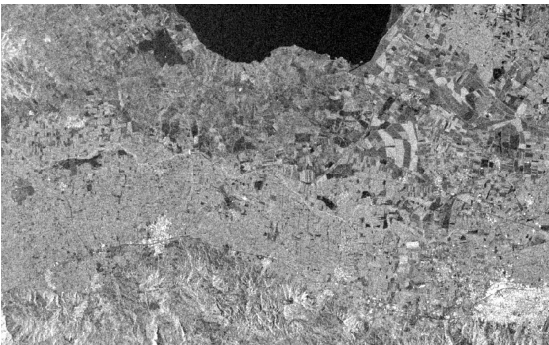
01 Ağustos 2017



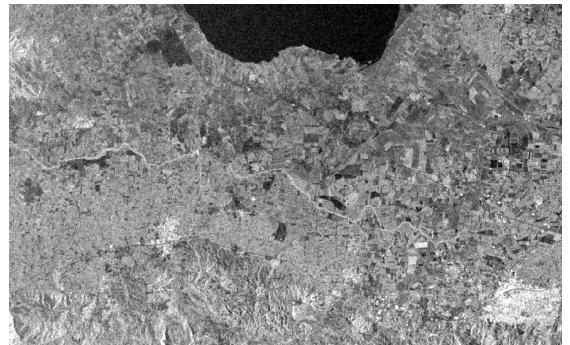
07 Eylül 2017



10 Ekim 2017

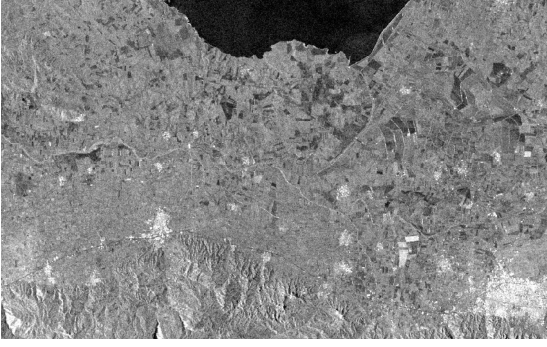


16 Kasım 2017

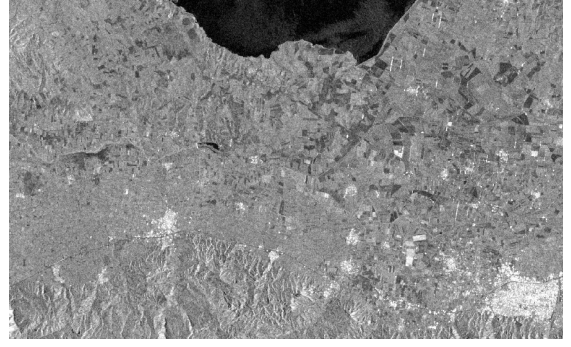


Şekil 2. Sentinel-1 SAR Uydu Görüntüleri DY Bant Görüntüleri

10 Nisan 2017



03 Mayıs 2017



02 Haziran 2017



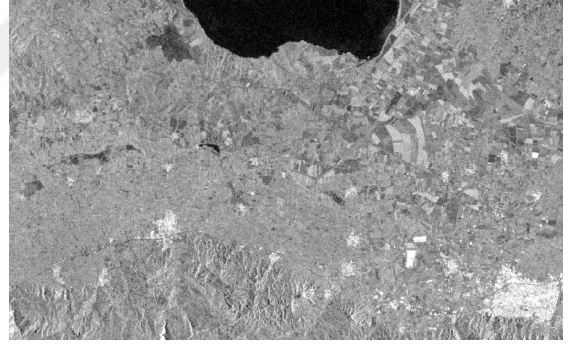
02 Temmuz 2017



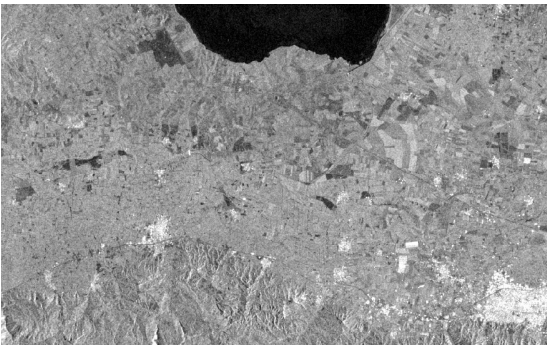
01 Ağustos 2017



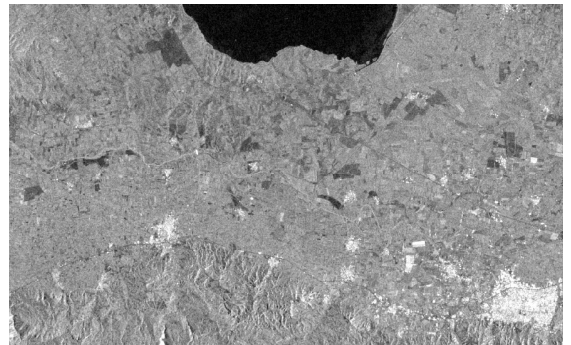
07 Eylül 2017



10 Ekim 2017



16 Kasım 2017



Şekil 3. Sentinel-1 SAR Uydu Görüntüleri DD Bant Görüntüleri

2.1.2.2. Arazi Yer Gerçeđi Verileri

Hacettepe Üniversitesi Geomatik Mühendisliđi Anabilim Dalı'nda önceki yıllarda yapılan bir lisansüstü tez çalışması için sahada toplanmış olan veriler arazi gerçeđi verileri (referans veriler) olarak kullanılmıştır[44].

Arazi gerçeđi verileri, 25.08.2017-05.09.2017 tarihleri aralığında gerçekleştirilen saha çalışmaları sırasında yerinde görülmüş ve tespit edilmiştir. Arazide yer gerçeđi verilerinin toplanması sırasında üzüm ve mısır için çekilmiş olan fotoğraflar sırasıyla Şekiller 4 ve 5'te gösterilmiştir[44].



Şekil 4. Arazi Çalışmaları Sırasında Yerinde Görülen Üzüm Bağlarından Birisi



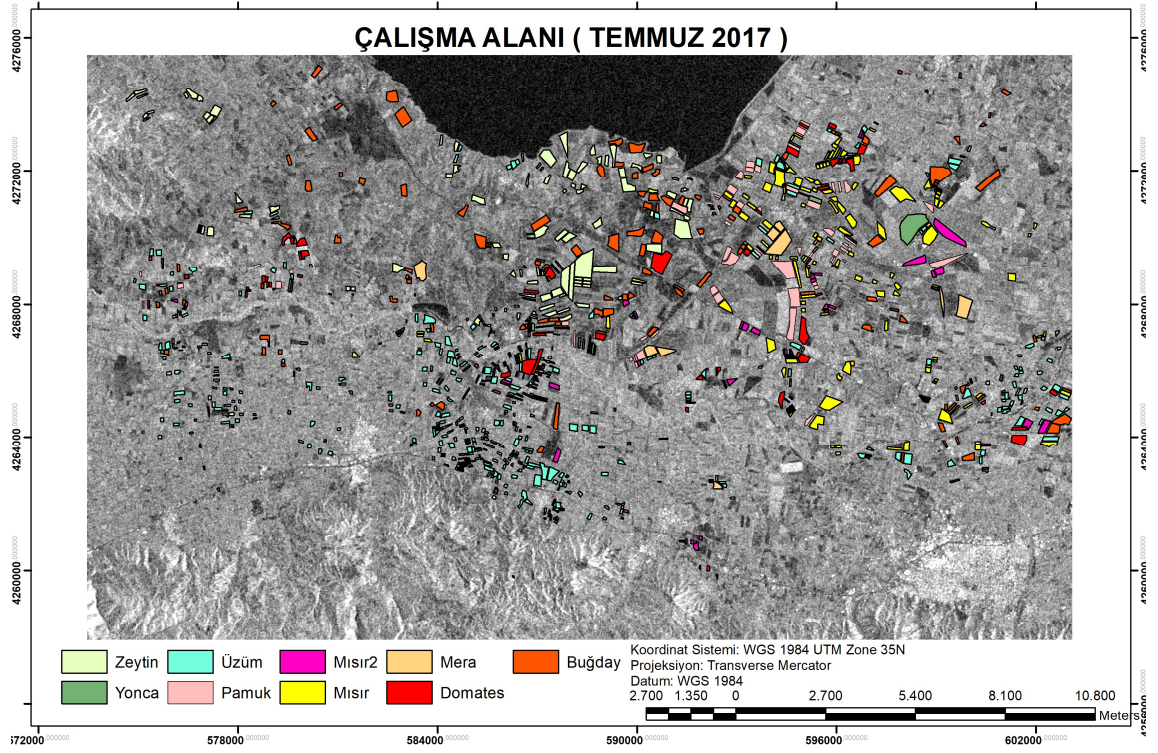
Şekil 5. Arazi Çalışmaları Sırasında Yerinde Görülen Mısır Tarlalarından Birisi.

Test ve eğitim veri setlerinin oluşturulması için arazi gerçeği verilerinin yanında tarım parselleri ürün bilgilerinin artırılması için ÇKS verileri de arazi gerçek verilerine eklenmiştir. ÇKS verileri, çalışma bölgesinde Tarım ve Orman Bakanlığı'nca arazide yapılan önceki çalışmalardaki verilerden elde edilmiştir[44][45]. Çalışma alanına ait ÇKS verileri öznitelik bilgilerinden seçilmiş bir ekran alıntısı Tablo 3'te gösterilmiştir.

Tablo 3. ÇKS Verileri Örneği

İl	İlçe	Köy	Ada No	Parsel No	Kullanılan Alan(da)	Parsel Alanı(da)	Ürün	Tarım Şekli	Ekim Tarihi	Hasat Tarihi	Ekili Alan (da)
MANİSA	AHMETLİ	KARGIN	0	1146	3.163	18.65	BUĞDAY (MUHTELİF)	Kuru	01/10/2016	01/04/2017	3.163
MANİSA	AHMETLİ	KARGIN	0	1148	4.268	25.166	ÜZÜM (KURUTMALIK ÇEKİRDEKSİZ)	Sulu	10/11/2015	20/08/2017	4.268
MANİSA	AHMETLİ	KARGIN	0	1151	4.823	28.434	ÜZÜM (KURUTMALIK ÇEKİRDEKSİZ)	Sulu	01/11/2015	21/08/2017	4.823
MANİSA	AHMETLİ	KARGIN	0	692	12.728	25.46	ÜZÜM (KURUTMALIK ÇEKİRDEKSİZ)	Sulu	01/11/2015	20/08/2018	4.318
MANİSA	AHMETLİ	KARGIN	0	1318	9.256	21.767	ÜZÜM (KURUTMALIK ÇEKİRDEKSİZ)	Sulu	31/12/1999	31/08/2018	9.256
MANİSA	AHMETLİ	KARGIN	0	1330	19.303	19.305	BUĞDAY (MUHTELİF)	Sulu	01/10/2016	01/05/2017	19.089
MANİSA	AHMETLİ	KARGIN	0	1330	19.303	19.305	MISIR (SILAJLIK)	Sulu	01/06/2017	01/10/2017	19.089
MANİSA	AHMETLİ	KARGIN	0	385	8.699	26.1	DOMATES (MUHTELİF)	Sulu	01/04/2017	01/07/2017	8.699
MANİSA	AHMETLİ	KARGIN	0	1035	2.73	2.73	ÜZÜM (KURUTMALIK ÇEKİRDEKSİZ)	Sulu	01/11/2016	20/08/2021	2.73
MANİSA	AHMETLİ	KARGIN	0	1140	8.668	8.668	ÜZÜM (KURUTMALIK ÇEKİRDEKSİZ)	Sulu	02/02/1993	21/08/2017	8.66
MANİSA	AHMETLİ	KARGIN	0	1149	50.334	50.334	ÜZÜM (KURUTMALIK ÇEKİRDEKSİZ)	Sulu	01/01/1998	20/08/2017	50.33

Arazi çalışmaları sırasında sahada incelenen 430 adet parselin arazi gerçeği verileri oluşturulmuştur. Ayrıca, ÇKS veri tabanından elde edilen bitki örtüsü alan bilgileri ile birlikte toplam 1195 adet parsele ait ürün bilgisi (örn.,buğday, domates, mera, mısır, mısır2, pamuk, üzüm, yonca, zeytin) elde edilmiştir. Alanı 1 dekardan küçük olan parseller sınıflandırma işlemine dâhil edilmemiştir. Dolayısıyla, toplam 1116 parsel referans veri olarak kullanılmıştır. Arazi yer gerçeği verilerinin çalışma bölgesindeki dağılımı Şekil 6'da gösterilmiştir.



Şekil 6. Arazi Yer Gerçeği Verilerinin Çalışma Bölgesindeki Dağılımı.

2.2. Sentinel-1 SAR Uydu Görüntüleri

Sentinel-1 SAR uyduları bütün hava şartlarında, gece ve gündüz C-bandında 4 farklı konumsal çözünürlük ve kapladıkları alanlarına göre görüntüler ortaya çıkaran bir uydudur. Öncelikle deniz, kara görüntüleme ve acil yönetim servisleri olmak üzere operasyon uygulamalarını destekleyen Copernicus Programındaki Avrupa Uzay Ajansı (AUA) tarafından geliştirilmiştir[46].

Sentinel-1 A ve B SAR uyduları, 3 Nisan 2014 ve 25 Nisan 2016 tarihlerinde fırlatılarak yörüngelerine gönderilmiştir. Güneş senkronize ve kutba yakın yörüngeye sahiptir. Uyduların yerden yüksekliği 693 km'dir[46].

Sentinel-1 uydularının görüntü çekim modları şu şekildedir:

- **Stripmap (SM) Modu:** 5x5 metre konumsal çözünürlük ve 80 km genişlikteki verileri toplar. Stripmap modu, temel olarak okyanus rüzgâr alanları ve yeryüzü radyal hızlarının tespiti faaliyetleri için kullanılır.

- **İnterferometrik Geniş Alan (Interferometric Wide Swath-IW) Modu:** 5x20 metre konumsal çözünürlük ve 250 km genişlikteki verileri toplar. İnterferometrik geniş

alan modu, okyanus rüzgâr alanları ve yeryüzü radyal hızlarının tespiti faaliyeti için kullanılır.

- **Ekstra Geniş Alan (Extra Wide Swath-EW) Modu:** 20x40 metre konumsal çözünürlük ve 400 kilometre genişlikteki verileri toplar.

- **Dalga (Wave-WV) Modu:** 5x5 metre konumsal çözünürlük ve 20x20 kilometre alandaki verileri toplar.

Sentinel-1 SAR uydusu öncelikli olarak yeryüzünde IW modunda, açık okyanuslarda ise WV modunda çalışır. Veri setleri WV modu için tek polarizasyon (DD veya HH) ve SM, IW, EW modları için çift polarizasyon (DD+DY veya HH+HV) ve tek polarizasyon (HH veya DD) olarak bulunmaktadır. Veri üretimi, her mod için, ham veri olarak ya da level-0, level-1 ve level-2 seviyelerinde üretilebilmektedir[46]. Level-0 SAR ürünleri sıkıştırılmış ve odaklanmamış SAR ham ürünlerinden oluşmaktadır. Diğer tüm level üretilen ürünlerin temelini teşkil etmektedir. Verilerin kullanılabilmesi için sıkıştırmadan çıkartılması ve üretilmesi gerekmektedir. Veriler uydu ve konum bilgisinin yanında gürültü, iç kalibrasyon ve yansıma kaynak paketlerini içermektedir. Uzun süreli arşivlerde tutulmaktadır. Ürünler görev süresinde ve uzay operasyonlarının bitmesinden 25 yıl sonrasında herhangi bir ürünün üretilmesi için işlenebilmektedir. Görüntüler, kullanıcılar için sadece SM, IW ve EW modlarında temin edilebilmektedir. Level-0 ürünler (Ham Veri) Level-1 ürünlere ham veri analizi, iç kalibrasyon, Doppler ağırlık merkezi tahmini, polinom uygunluğu, mesafe ve azimuth işlemleri ile dönüştürülmektedir. Level-1 verileri hem Single Look Complex (SLC) hem de Ground Range Detected (GRD) ürünleri olarak üretilebilmektedir. Dijital piksel değerlerini radyometrik olarak kalibrasyonu yapılmış geri yansıma değerlerine çevirmek için bütün gerekli bilgiler ürünlerde bulunabilmektedir. Level-2 verileri Level-1 verilerinden üretilen coğrafi koordinatları olan jeofiziksel ürünleri içermektedir. Level-2 Okyanus ürünleri rüzgâr, dalga ve günümüzdeki uygulamalar için SAR verisinden üretilen jeofiziksel bileşenleri içermektedir[47].

Bu tez çalışmasında, konumsal çözünürlüğü ve yeryüzü için kullanım uygunluğu göz önünde bulundurularak, İnterferometrik Geniş Alan modu tercih edilmiştir. Görüntüler GRD formatında ve Sentinel 1A/B uydularından C-bantı çift polarizasyon (DY-DD) görüntüleri olarak kullanılmıştır. Görüntülerin projeksiyon sistemi WGS1984, UTM Bölge 35 N olarak tanımlanmıştır.

2.3. Makine Öğrenme Algoritmaları

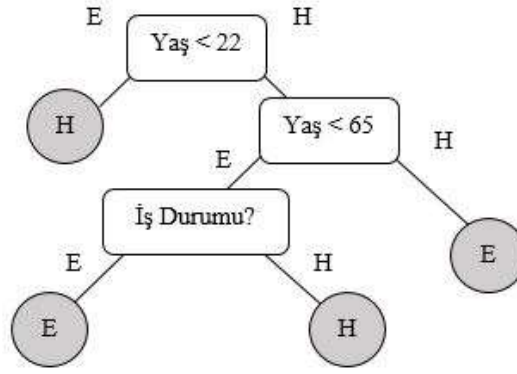
2.3.1. Karar Ağacı (Decision Tree)

Karar ağacı, eğitilmiş öğrenme tahminleri ve sınıflandırması için en güçlü ve yaygın öğrenme algoritmasıdır. Tahmin ve sınıflandırma için karar verme modeli oluşturma metodudur.[48] Karar ağacı öğrenmesi, belirli kriterlere göre eğitim verilerini tekrarlı olarak alt birimlere bölen tahmin modelleme metodudur[49].

Uygulaması kolay ve az bir ön bilgiye ihtiyaç duyar. Karar ağacı birçok öz nitelik değerine sahip her boyut veri seti ile oluşturulabilir. Karar ağacında eğitim örnekleri setinin başarılı bölünmesi bütün alt örnek setlerinde tek bir sınıf örneği kalana kadar devam eder[50].

Karar ağacı öğrenimi, eğitim örneklerini belirlenen kriterlere göre alt kümelerle tekrar tekrar bölerek modeli öğrenen bir tahmin modellemesi yöntemidir. Karar ağacı dalları etiketli eğitim örneklerini girdi olarak kabul eden ve yeni örneklerin etiketlerini tahmin eden modeli oluşturan denetimli öğrenenlerdir[49].

Karar ağacı 3 ana bileşene sahiptir. Bu bileşenler; düğümler (nodes), yapraklar (leaves) ve kenarlardır (edges). Her düğüm verinin ayrılacağı bir öznitelik değeri ile etiketlidir. Her düğüm öznitelik değerinin ihtimaline göre etiketlenmiş birden fazla kenarlara sahiptir. Her kenar ya 2 düğümü ya da düğüm ve yaprağı bağlar. Yapraklar verinin kategorilere ayrılması için karar değerleri ile etiketlenir. Karar ağacı, sonuç vermek için kök düğümünden başlar. Yaprak düğümünün temsil ettiği sınıfa erişene kadar ağacın dallarından aşağıya doğru takip eder. Her karar ağacı kural setlerini ifade eder. Kural setleri verilerin özelliklerine göre ayrılmasını sağlar[50]. Bir karar ağacı algoritması örneği Şekil 7’de gösterilmiştir.



Sıra	Yaş	İş Durumu	Ev Durumu
0	12	E	H
1	32	H	E
2	25	H	E
3	48	E	H
4	67	E	E
5	18	H	H

Şekil 7. Bir Karar Ağacı Algoritması Örneği

2.3.2. Gradyan Artırmalı Ağaç (Gradient Boosting Tree) Algoritmaları

Güçlendirme, bir öğrenme algoritmasının doğruluğunu yükseltmek için kullanılan genel bir metottür. Yüksek doğruluklu veya güçlü sınıflandırma aracının birçok doğru olmayan veya zayıf sınıflandırma aracının lineer birleşimi aracılığıyla oluşturulma prensibine dayanır[51].

Gradyan Artırmalı Ağaç yöntemi, eğitilmiş öğrenmedeki her yanlış sınıflandırma hatası ile tek bir güçlü sınıflandırma aracı üretmek için birçok sınıflandırma aracını birleştirme yaklaşımına dayanır. Bir veri örneği her tekrar serisinde oluşturulan temel sınıflandırma araçlarının ağırlıklı tahminleri oylayarak sınıflandırılır[50].

Güçlendirme ile her eğitim örneğine bir ağırlık ataması yapılır ve her artırma tekrarından sonra ağırlıklar yeni değere bağlı olarak değişir. Örnekleme, yeni eğitim setini elde etmek için eğitim örneklerinin dağılımına göre değer alır. Sınıflandırma aracı eğitim veri setinden oluşturulur ve orijinal veri setindeki eğitim örneklerinin tümünün sınıflandırılması için kullanılır. Eğitim örneklerinin ağırlıkları her güçlendirme serisinin sonunda güncellenir. Doğru sınıflandırılan eğitim örneklerinin ağırlıkları düşerken yanlış sınıflandırılan örneklerin ağırlıkları güçlendirilmiş olur. Bu durum sınıflandırma aracının devamlı iterasyonlarda sınıflandırılması zor olan eğitim örneklerine odaklanmasına

zorlar. Artırmanın en önemli engeli aşırı uyum gösterme (overfitting) durumunun oluşmasıdır[50].

2.3.3. Aşırı Gradyan Artırma (XGBoost)

XGBoost, kolay paralel işlem yapma ve yüksek tahmin doğruluğu gibi özelliklerinden dolayı yapay zekâ yöntemleri arasında yapılan yarışmalarda kullanılan araç olan, Gradient Boosting Machine (GBM) algoritmasının etkili ve ölçeklenebilir uygulamasıdır[52].

XGBoost, ağaç güçlendirmek için yüksek verimli, ölçeklenebilir bir makine öğrenme algoritmasıdır. İşlenmesi zor verilerle (Kaggle Yarışması gibi) bilinen en iyi sonuçları elde etmek için birçok alanda yaygın olarak kullanılmaktadır. Chen ve Guestrin tarafından geliştirilmiş olan XGBoost algoritması gradyen artırma çatısı altında en verimli hale getirilmiş, esnek ve taşınabilir olarak tasarlanmıştır[48][49][50]. Güçlendirmenin ana fikri, düşük doğruluğa sahip zayıf sınıflandırıcı serilerini birleştirerek daha iyi sınıflandırma performansıya güçlü sınıflandırıcı oluşturmaktır[52].

Veri setini şu şekilde varsayarsak; $D = \{(x_i, y_i) : i = 1 \dots n, x_i \in \mathbb{R}^m, y_i \in \mathbb{R}\}$, elimizde m özellikleriyle n örnekleri vardır. Eşitlik (1)'de $\hat{y}_i$ model tarafından tahmin edilen değer olarak alırsak:

$$\hat{y}_i = \sum_{k=1}^K f_k(x_i), f_k \in F \quad (1)$$

Eşitlik 1'de f_k bağımsız regresyon ağacını ifade etmekte ve Eşitlik (1)'de $f_k(x_i)$, k sayıdaki ağaçtan i sayıdaki örnekten verilen tahmin değerini belirtmektedir. Regresyon ağacı modelindeki f_k fonksiyon seti hedef fonksiyonu en aza indirerek öğrenilebilmektedir.

$$Obj = \sum_{i=1}^n l(y_i, \hat{y}_i) + \sum_{k=1}^K \Omega(f_k) \quad (2)$$

Eşitlik (2)'de l tahmin edilen $\hat{y}$ ve nesne olan y_i arasındaki farklı ölçen eğitilen kayıp fonksiyonudur. Aşırı uyumu önlemek için, Eşitlik (2)'de Ω terimi modelin karmaşıklığını engeller.

$$\Omega(f_k) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (3)$$

Eşitlik (3)'te γ ve λ düzeltmenin dereceleri. Eşitlik (3)'te T ve w sırasıyla yaprak sayısını ve her yapraktaki puanı ifade etmektedir.

Eşitlik (4)'te $\hat{y}_i^{(t)}$ t iterasyondaki i örneğinin tahmini olarak kabul edilirse, Eşitlik (4)'teki hedefi en aza indirmek için f_t eklemek gerekmektedir.

$$Obj^{(t)} = \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) \quad (4)$$

Eşitlik (5), Eşitlik (4) ü sadeleştirmek için ikinci derece Taylor açılımını kullanarak elde edilmekte ve sabit terimi ortadan kaldırmaktadır:

$$Obj^{(t)} = \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t(x_i)^2 \right] + \Omega(f_t) \quad (5)$$

Eşitlik (5)'te $g_i = \partial_{\hat{y}_i^{(t-1)}} l(y_i, \hat{y}_i^{(t-1)})$ ve $h_i = \partial_{\hat{y}_i^{(t-1)}}^2 l(y_i, \hat{y}_i^{(t-1)})$ l üzerinde birinci ve ikinci derece gradyanlarıdır. Sonrasında hedef Eşitlik (6)'da yeniden yazılır:

$$\begin{aligned} Obj^{(t)} &= \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t(x_i)^2 \right] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \\ &= \sum_{j=1}^T \left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T \end{aligned} \quad (6)$$

Eşitlik (6)'da $I_j = \{i \mid q(x_i) = j\}$ ifadesi j yaprağı setinin örneğini belirtmektedir. Sabit ağaç yapısı için q , j yaprağının en uygun ağırlığı w_j^* ve benzer en uygun değer Eşitlik (7) ve Eşitlik (8) ile hesaplanabilmektedir:

$$w_j^* = -\frac{G_j}{H_j + \lambda} \quad (7)$$

$$Obj^* = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \lambda T \quad (8)$$

Eşitlik (7) ve Eşitlik (8)'de $G_j = \sum_{i \in I_j} g_i$, $H_j = \sum_{i \in I_j} h_i$ ve obj değişkeni ağaç yapısı q 'nın

kalitesini ifade etmektedir. Değerler ne kadar küçük olursa, ağaç yapısı o kadar iyi olmaktadır. Tüm ağaç yapılarını saymak mümkün olmadığından, ağaca tekrarlanan dallar eklemek için açgözlü algoritma kullanılmaktadır. Eşitlik (9)'da I_L ve I_R , sol ve sağ düğüm noktaları ayrımı yapıldıktan sonraki örnek setleridir. Uygun olan bölünme

noktalarını numaralandırarak ve en az hedef fonksiyonu ile en çok kazanç bölümünü seçerek elde edilen formül Eşitlik(9)'daki gibi olmaktadır:

$$G = \frac{1}{2} \left[\frac{\left(\sum_{i \in I_L} g_i \right)^2}{\sum_{i \in I_L} h_i + \lambda} + \frac{\left(\sum_{i \in I_R} g_i \right)^2}{\sum_{i \in I_R} h_i + \lambda} - \frac{\left(\sum_{i \in I} g_i \right)^2}{\sum_{i \in I} h_i + \lambda} \right] - \gamma \quad (9)$$

XGBoost makine öğrenme algoritmasını tanımlarken belirli parametreler ve değerlerin girilmesi gerekmektedir. Bu değerler algoritmanın kullanılacağı uygulamaya göre değişiklik göstermektedir. Tablo 4'te gösterilen parametreler varsayılan değerler olarak kullanılmaktadır[55].

Tablo 4. XGBoost Makine Öğrenme Algoritması Varsayılan Parametre Değerleri

Parametre	Tanım	Aralık	Varsayılan Değer
nrounds	Modelde Kullanılacak Ağaç Sayısı	[1,∞[	10
Eta(η)	Öğrenme Oranı	[0,1]	0.3
Gamma(γ)	En Düşük Bölüm Kaybı Azaltma	[0,∞[	0
Max_depth	En Yüksek Ağaç Derinliği	[0,∞[	6
Minimum_Child_Weight	Yeni Ağaç İçin En Düşük Gözlem	[0,∞[	1
Subsample	Eğitim için Kullanılacak Rastgele Örnek Oranı	[0,1]	1
Column_Sample_Tree	Her Ağacı Eğitmek İçin Kullanılacak Olan Kolon Oranı	[0,1]	1
Column_Sample_Level	Her Ağacın Ayrımını Eğitmek İçin Kullanılacak Olan Kolon Oranı	[0,1]	1
Lambda(λ)	L ² Normu İçin Ağırlıklardaki Düzenleme Terimi	[0,∞[	1
Alpha(α)	L ¹ Normu İçin Ağırlıklardaki Düzenleme Terimi	[0,∞[	0

Optimum parametre değerlerinin tespiti herhangi bir makine öğrenme algoritmasında performansı en üst seviyeye çıkarmak için önemlidir. Bütün parametre olasılıklarını tanımlama sonrasında sınıflandırma sonuçlarının en iyi değerleri olan setleri seçmek için bütün kombinasyonları test etmeyi içermektedir.

Grid arama, algoritma için uygun olan parametre değerleri kullanılarak grid aralıklar oluşturmayı ifade etmektedir. Rastgele arama, farklı parametre değerlerini rastgele seçmek ve sonuçlarını test etmek anlamına gelmektedir. Bu çalışmada grid arama ile belirlenen aralıklar test edilmiştir. En uygun 5 parametre ve değerleri seçilmiş ve diğer parametreler varsayılan değerleri ile kullanılmıştır[56]. Seçilen en uygun parametreler için belirlenen değerler Tablo 5'te verilmiştir. Dolayısıyla, XGBoost algoritması sınıflandırma modeli bu parametre değerleri ile oluşturulmuştur.

Tablo 5. XGBoost Optimizasyon Parametreleri ve Değerleri

Parametre	Grid Değerleri	En Uygun Değerler
Eta	0.05, 0.2, 0.3, 0.4	0.2
Nrounds	1, 10, 100, 150, 170, 200	170
Max_depth	1, 4, 6, 10, 12, 20	12
Subsample	0.5, 0.6, 0.7, 0.9, 1	1
Col_sample_level	0,1	1

2.3.4. Rastgele Orman (Random Forest)

Rastgele Orman (RO), Breiman tarafından geliştirilen bir topluluk (ensemble) öğrenme algoritmasıdır[57]. Algoritmanın temeli daha doğru ve istikrarlı tahminleri elde etmek için birden çok karar ağacı oluşturulması ve bunların birleştirilmesine dayanmaktadır. [58]. Özellikle büyük veri setleri için farklı sınıflardaki veri dengesizliğinin üstesinden gelmesinden dolayı yaygın şekilde kullanılmaktadır. Paralel yapısından dolayı diğer sınıflandırma algoritmalarına göre kıyasla daha hızlıdır[59].

RO algoritması verilen girdi bilgisinden yeni örneği sınıflandırmak için sınıflandırma yapan ormandaki her ağaca aynı girdi değerini gönderir. Her ağaç sınıflandırma değerlerini belirler. Her sınıflandırma sonucu oy olarak adlandırılır. En yüksek değerde sınıflandırılan sonuç seçilir[60].

RO algoritmasında belirlenen eğitim verisinin 2/3 oranındaki kısmı karar ağaçları ve ormanın oluşturulması için kullanılır. Kalan 1/3 oranındaki kısım ise genelleştirilmiş hata (Out of Bag - OOB) hesaplaması için ayrılmaktadır. OOB algoritmanın sınıflandırma hatasının hesaplanmasında ve önemli özelliklerin belirlenmesinde kullanılmaktadır.

RO algoritmasında 2 ana parametrenin belirlenmesi gerekmektedir. Bu parametreler; ağaç sayısı (n_{tree}) ve düğümü bölmek için rastgele seçilen değişken sayısıdır (m_{try}).

Her ağaçtaki karar düğümünde seçilen rastgele özelliklerin sayısı (m_{try}) orman sınıflandırmasının hata oranını belirler. RO sınıflandırmasının hata oranı her 2 ağacın arasındaki korelasyon ve her bir ağacın sınıflandırma gücüne bağlıdır. m_{try} parametre değerini artırmak ağaçlar arasında korelasyonu artırır[61].

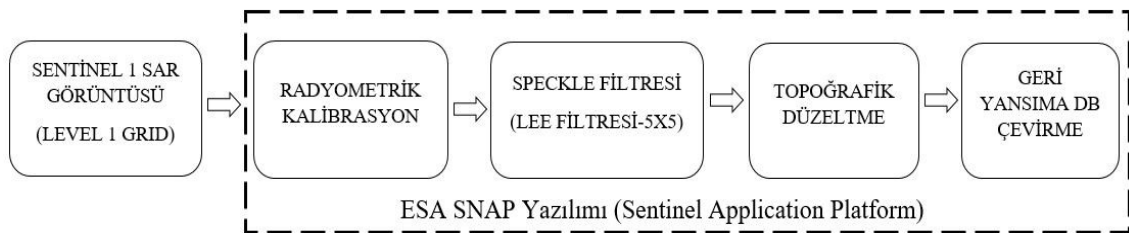
RO algoritmasının ağaç sayısının artması ile sınıflandırma performansı artmaktadır[61]. Fakat yapılan çalışmalar incelendiğinde önceden belirlenen ağaç sayısının üstünde değer alındığında rastgele orman algoritmasının sınıflandırma performansını önemli bir oranda artırmadığı görülmüştür[62].

2.4. Sentinel-1 Görüntülerin İşlenmesi

2.4.1. Görüntülerin Ön İşlemleri

Çalışmada çok zamanlı Sentinel-1A/1B SAR uydu görüntülerinin C bandı çift polarizasyon (DY/DD) olarak kullanılmıştır. Sentinel-1 SAR görüntüleri IW modu ve Level-1 Grid formatında temin edilmiştir.[19] Görüntüler Level-1 Grid formatında elde edildikten sonra ön işlemler uygulanmıştır. Ön işlemler ESA tarafından ücretsiz olarak sağlanan SNAP (Sentinel Application Platform) programında yapılmıştır.

Çalışmada yapılan ön işlemler ve adımları Şekil 8’de gösterilmiş olup sırasıyla şu şekildedir: Radyometrik Kalibrasyon, Benek(Speckle) Filtreleme, Topoğrafik Düzeltme, Geri Yansıma db Çevirme.



Şekil 8. Radar Uydu Görüntülerine Uygulanan Ön İşlem Adımları

- **Radyometrik Kalibrasyon:** Piksel değerlerini radyometrik olarak kalibrasyonu yapılmış olan SAR geri yansıma değerlerine çevirme işlemidir. Sentinel-1 GRD verilerinin görüntü yoğunluk değerlerini sigma nought değerine dönüştürür[63].

- **Benek (Speckle) Filtreleme:** SAR görüntülerinde taneli gürültü olarak görünen Speckle, birçok temel geri yansılardan yansıyan dalgaların kaynağa gelmesinden kaynaklanmaktadır.[64] Speckle filtresi, görüntüdeki speckle gürültüsünü azaltarak görüntüsünün kalitesini artırmak için kullanılmaktadır. İyileştirilmiş Lee filtresi kenar, çizgi detay, nokta hedefi ve doku bilgilerini belirleme özelliğinden dolayı görsel yorumlama için diğer speckle filtrelerine kıyasla daha üstün seviyededir[63].

- **Topografik Düzeltme:** SAR sistemlerinin yanal gözlemleme geometrisinden dolayı, yeryüzeyindeki engebeli topografya, SAR sinyallerinin geri saçılma değerlerini etkilemektedir. Bu etki, SAR verisi çekim açısındaki değişimlerden (artış ve azalış) dolayı çok engebeli arazilerde, örneğin tepelik alanlarda, çok yüksek miktarda olabilmektedir. Dolayısıyla, topografik düzeltme, bahsedilen bozulmaları düzeltmek ve görüntünün geometrik olarak gerçek yeryüzüne en yakın haline getirilmesini sağlamaktadır. Topografik düzeltme; topoğrafyadan kaynaklanan geometrik bozuklukları, detayların gerçeğe göre kısılması ve gölge gibi, her pikseli doğrulamak için sayısal yükseklik modeli kullanılarak yapılmaktadır[63].

- **Geri Yansımaya dB Çevirme:** Herhangi birim değeri olmayan geri yansımaya katsayılarını Eşitlik 10'daki logaritmik dönüşüm kullanarak dB olarak çevirmektedir[63].

$$\beta_{db}^o = 10 * \log_{10}(\beta^o) \quad (10)$$

2.4.2. Doku Özelliklerinin Çıkarılması

Sınıflandırma işlemi piksel tabanlı yapılmıştır. Çalışma bölgesinin geniş bir tarım arazisi olmasından dolayı sınıflandırma işleminde, SAR görüntülerinin geri yansımaya değerlerine ek olarak, görüntülerden elde edilen doku özellikleri bantları da kullanılmıştır.

Doku özellikleri, renk ve şeklin yanı sıra görüntünün görsel açısı hakkında önemli bir bilgi kaynağıdır[65]. Doku özellikleri tipik yeryüzü nesneleri ve yapılarından oluşan yoğunluk değişimlerini gösterir. Genellikle yeryüzü üzerindeki şekil veya nesnelerin düzenli tekrarları olarak tanımlanmaktadır. Doku analizleri yeryüzünde bulunan nesnelerin sınıflandırmasında önemli rol oynamaktadır. Doku analizi diğer bir tanımla doku görüntüsünden özellik bilgilerinin bir takım matematiksel işlemler aracılığıyla çıkartılması işlemidir. Konumsal olarak değişkenlik gösteren spektral değerler hakkında bilgi sağlar[66].

Gray Level Co-occurrence Matrix (GLCM) en iyi bilinen ve yaygın olarak kullanılan doku özellikleri çıkarma yöntemidir. GLCM ilk olarak 1973 yılında Haralick tarafından önerilmiştir[67][68]. GLCM ile görüntüden ikinci dereceden istatistiksel değerler hesaplanır ve görüntüdeki komşu piksellerin konumları ve değerleri hakkında bilgiler sağlar[69].

GLCM belirli bir d mesafe aralığı ve belirli bir açı θ değerine sahip gri değer çiftlerinin bağımlı frekanslarını ifade eden bir matristir. Mesafe değeri 1 den başlayarak bütün görüntü boyutunda 4 farklı yönde (0° , 45° , 90° , 135°) devam eder. Farklı açılar ve mesafeler çok farklı değerler vermektedir. Yön açısından farklı özellikler içeren görüntülerden doku özellik bilgilerinin çıkarılması büyük ölçüde seçilen açı değerine bağlıdır[65].

Bu çalışmada GLCM doku özellikleri hesaplama yönteminde kullanılan parametreler sırasıyla; Pencere Boyutu: 5×5 , Açılar: Hepsi, Nicemleme: Olasılığa Dayalı, Nicemleme Seviyesi: 32 ve Yer Değişimi: 4 olarak alınmıştır.

$N \times N$ boyutlarındaki bir I görüntüsünün GLCM formülü Eşitlik 11 ve 12’de gösterilmiştir.

$$L(x, y) = \sum_{x=1}^N \sum_{y=1}^N \{1, l(s, v) = x \text{ ve } l(s + \Delta s, v + \Delta v) = y\} \quad (11)$$

$$L(x, y) = \sum_{x=1}^N \sum_{y=1}^N \{0, Diğer\} \quad (12)$$

Eşitlikler 11 ve 12’de (s, v) görüntü üzerinde belirli bir noktanın konum bilgisini gösterir. Δs ve Δv verilen bir noktadaki $L(x, y)$ üzerinde i ve j yoğunluğunun birlikte oluşturulan birleştirilmiş olasılıklarıdır.

Bu çalışmada, karşıtlık (Contrast) (Eşitlik 13), homojenlik (homogeneity) (Eşitlik 14), farklılık (Dissimilarity) (Eşitlik 15), entropi (Entropy) (Eşitlik 16), açısal ikinci moment (Angular Secondary Moment) (Eşitlik 17), ortalama (Mean) (Eşitlik 18), varyans (Variance) (Eşitlik 19) ve korelasyon (Correlation) (Eşitlik 20) doku özellikleri kullanılmıştır.

- **Karşıtlık:** Herhangi bir nesne üzerinden diğer nesnelerden ayırt eden renk ve parlaklık farkıdır.

$$\sum_{k=0}^{N_g-1} k^2 \left\{ \sum_{x=1}^{N_g} \sum_{y=1}^{N_g} L(x, y) |x - y| = k \right\} \quad (13)$$

- **Homojenlik:** Nesnenin dağılımının tek tip yapıda olmasıdır.

$$\sum_x \sum_y \frac{1}{1+|x-y|} L(x,y) \quad (14)$$

- **Farklılık:** Ağırlık değerleri dışında karşılık ile aynıdır. İkisi de aynı bilgileri içerir fakat değer boyutları farklıdır.

$$\sum_x \sum_y |x-y| \cdot (L(x,y)) \quad (15)$$

- **Entropi:** Görüntüdeki doku düzensizliğini matematiksel değer olarak ifade etmektedir.

$$\sum_{x=1}^{N_g-1} \sum_{y=1}^{N_g-1} L(x,y) \cdot \log\{L(x,y)\} \quad (16)$$

- **Açısal İkinci Moment:** Herhangi bir görüntünün homojenlik değerinin ölçülmesidir.

$$\sum_{x=1}^{N_g-1} \sum_{y=1}^{N_g-1} (L(x,y))^2 \quad (17)$$

- **Ortalama:** Komşu piksel değerlerinin ortalamasıdır.

$$\sum_x \sum_y \frac{x+y}{2} L(x,y) \quad (18)$$

- **Varyans:** Heterojenlik durumunun ölçüsüdür. Gri ölçekli değerleri ortalamadan sapmaya başladığında varyans değeri artar.

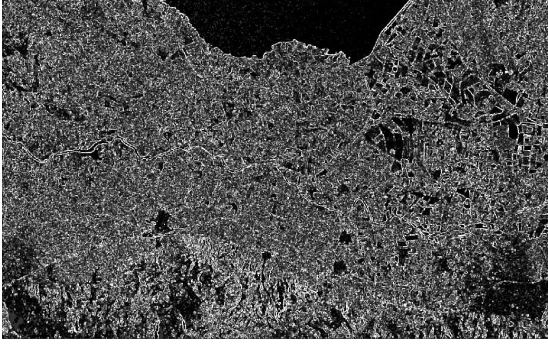
$$\sum_x \sum_y (x-\mu)^2 L(x,y) \quad (19)$$

- **Korelasyon:** Komşu piksel değerlerindeki gri seviye değerlerinin lineer bağımlılığı ölçümüdür.

$$\frac{\sum_x \sum_y (x-\mu_s)(y-\mu_v) * L(x,y)}{\sigma_s \sigma_v} \quad (20)$$

Bu çalışmada sınıflandırma için seçilen Sentinel-1 SAR görüntülerinin (toplam 8 ayrı tarihe ait görüntüler) her biri için yukarıda verilen 8 adet doku özellikleri bantları oluşturulmuştur. Sınıflandırmada her bir tarihe ait 8 adet Sentinel-1 SAR uydu görüntüsü DD ve DY polarizasyon görüntüleri, DY/DD oran görüntüleri ile DD ve DY polarizasyon görüntülerinden üretilmiş 128 adet doku özellikleri bantları (8 DY × 8 Doku + 8 DD × 8 Doku) kullanılmıştır. Böylece, sınıflandırmada kullanılan toplam bant sayısı 152 olmuştur. 10 Nisan 2017 tarihli Sentinel-1 DY görüntüsünden hesaplanmış doku bantları Şekil 9'da, Sentinel-1 DD görüntüsünden hesaplanmış doku bantları Şekil 10'da gösterilmiştir.

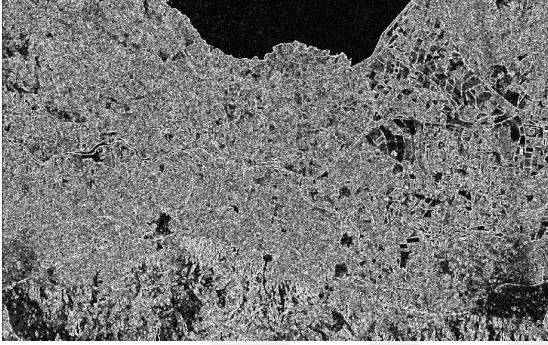
DY - Karşıtlık



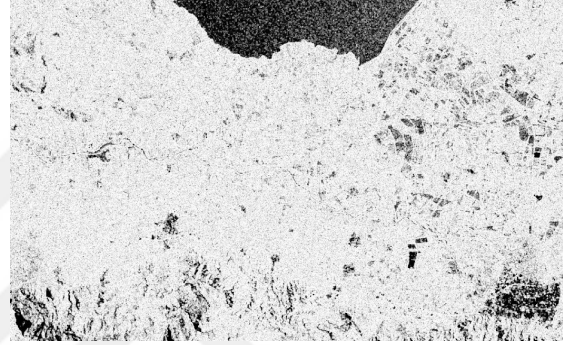
DY - Homojenlik



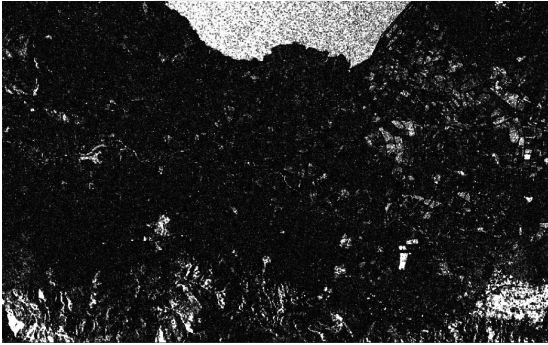
DY - Farklılık



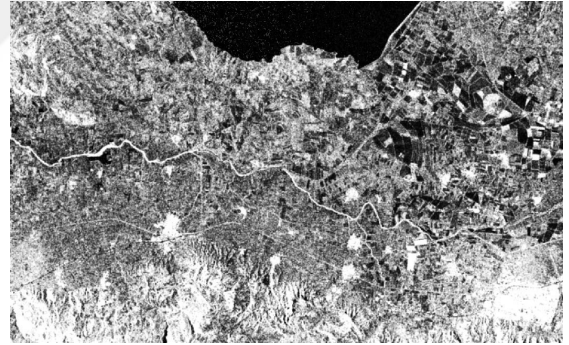
DY - Entropi



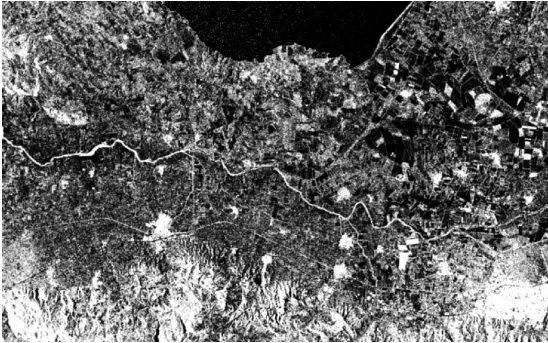
DY – Açısıl İkinci Moment



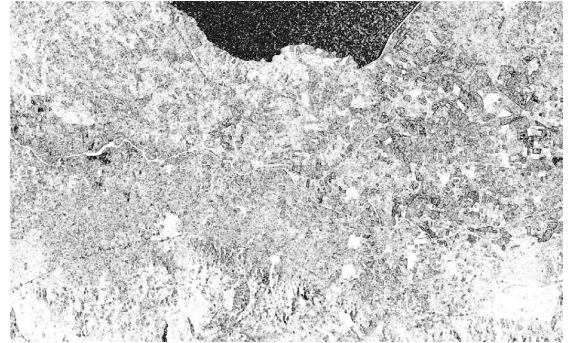
DY - Ortalama



DY - Varyans

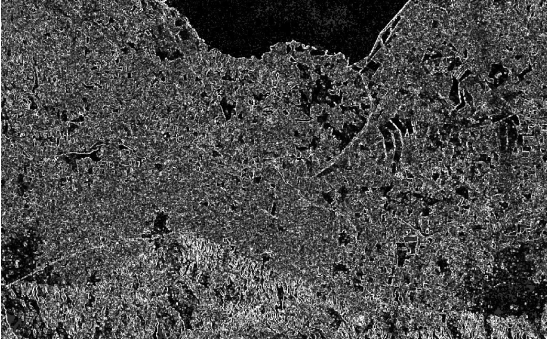


DY - Korelasyon



Şekil 9. 10 Nisan 2017 Tarihli Sentinel-1 DY Görüntüsünden Hesaplanmış Doku Bantları

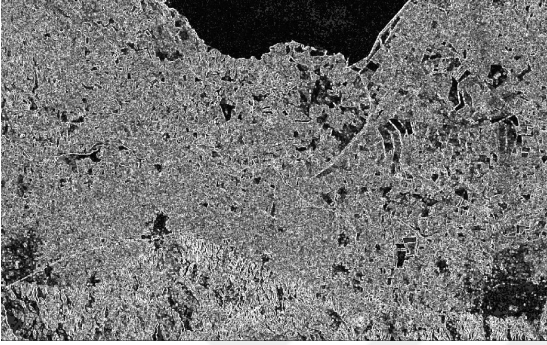
DD - Karşıtlık



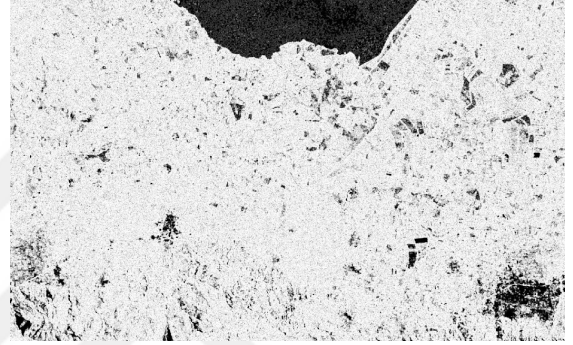
DD - Homojenlik



DD - Farklılık



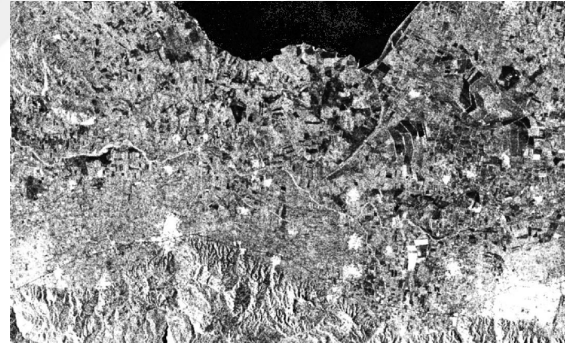
DD - Entropi



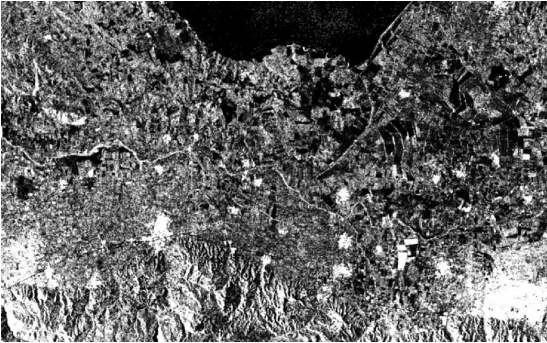
DD – Açısai İkinci Moment



DD - Ortalama



DD - Varyans



DD - Korelasyon



Şekil 10. 10 Nisan 2017 Tarihli Sentinel-1 DD Görüntüsünden Hesaplanmış Doku Bantları

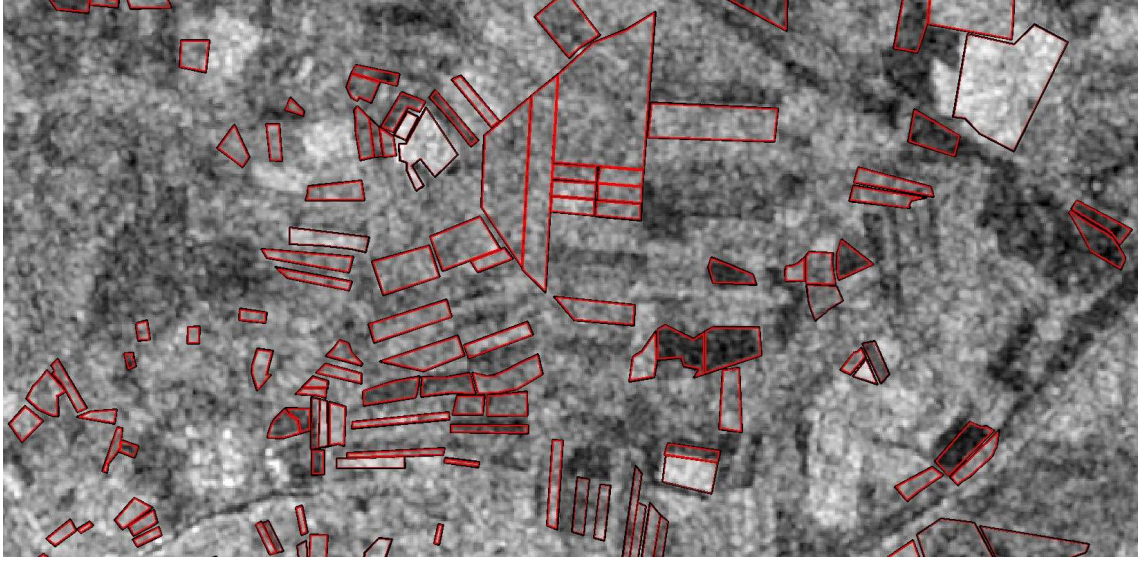
2.5. Sınıflandırma İşlemi

2.5.1. Eğitim ve Test Verilerinin Belirlenmesi

Çalışma alanı tarım ürünleri açısından zengin bir yapıya sahiptir. Bölgede yoğun olarak bulunan tarım ürünleri üzüm, mısır, zeytin, buğday, biber, pamuk, patlıcan ve yoncadır. Ayrıca, az miktarda brokoli, lahana ve ıspanak gibi kış ürünleri ekilmektedir. Az miktarda olmalarından dolayı bu ürünler sınıflandırmaya dâhil edilmemiştir.

Bu çalışmada, Hacettepe Üniversitesi Geomatik Mühendisliği Anabilim Dalı'nda aynı alanda daha önce bitirilen bir Yüksek Lisans tez çalışması için sahadan toplanmış arazi gerçeği verileri kullanılmıştır[44]. Arazi gerçeği verileriyle birlikte ÇKS verileri de sınıflandırmada referans veriler olarak kullanılmıştır.

Çalışmada alanda yoğun olarak bulunan 9 adet tarım ürününün tespiti yapılmıştır. Bu ürünler; buğday, domates, mera, mısır, mısır2, pamuk, üzüm, yonca ve zeytindir. Mısır2 olarak belirtilen tarım ürünü, buğday hasadından sonra körpe mısır ekilen alanlardır. Eğitim alanı örnekleri ve sınıflandırma algoritmalarının performans testleri için gerekli referans veriler, sahadan toplanan arazi gerçeği verileri ve ÇKS verisinde yer alan tarım parsellerinden seçilmiştir. Alanı 1 dekardan küçük parseller sınıflandırma işlemine dâhil edilmemiştir. Referans veri olarak toplam 1116 tarım parseli kullanılmıştır. Referans parsellerinin yaklaşık yarısı (560 adet) eğitim verisi olarak, diğer yarısı da (556 adet) test verisi olarak kullanılmıştır. Ürün bazında eğitim ve test verileri bilgileri Tablo 6.'da verilmiştir. Sınır piksellerinden kaynaklanabilecek hataları gidermek için bu pikseller sınıflandırmaya dâhil edilmemiştir. Bunun için, ArcMap yazılımında, parsel sınırlarının her iki tarafında yarım piksel (5m) olacak şekilde toplam 1 piksel genişliğinde (10m) tampon (*buffer*) alan oluşturulmuş ve sınıflandırmada tampon alan dışında kalan pikseller kullanılmıştır. Çalışma alanının bir bölümüne ait, tampon alan oluşturma sonrası elde edilen yeni parsel sınırları (tampon alan içinde kalan sınırlar) Şekil 11'de gösterilmiştir.



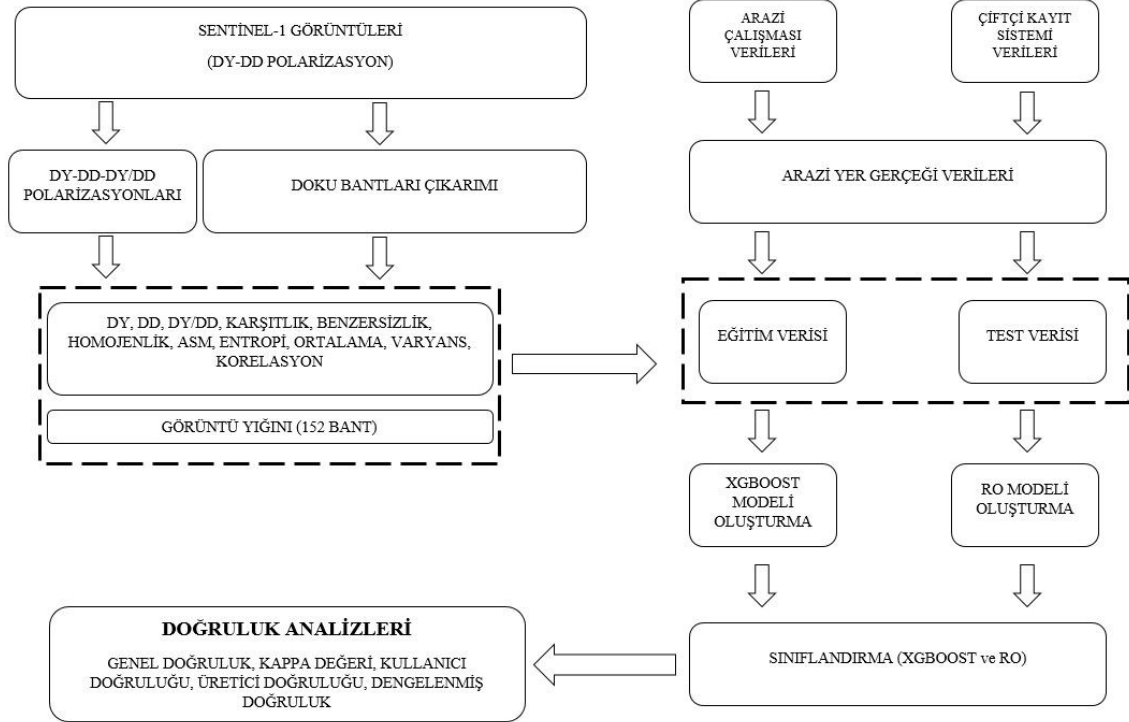
Şekil 11. Çalışma Alanının Bir Bölümüne Ait, Tampon Alan Oluşturma Sonrası Elde Edilen Yeni Parsel Sınırlarının Temmuz 2017 DY Polarizasyon Görüntüsü ile Birlikte Gösterimi

Tablo 6. Ürün Bazında Eğitim ve Test Verisi Parsel ve Piksel Sayıları

Tarım Ürünü	Eğitim Verisi (Parsel/Piksel)		Test Verisi (Parsel/Piksel)	
Buğday	97	38540	97	11548
Domates	31	14420	30	4875
Mera	5	8857	4	2537
Mısır	81	28417	81	10674
Mısır2	26	7740	25	3630
Pamuk	31	15953	31	7107
Üzüm	220	36289	219	15053
Yonca	10	6304	10	661
Zeytin	59	31502	59	12195
Toplam	560	188022	556	68280
Genel Toplam	1116/256302			

2.5.2. Sınıflandırma

Sınıflandırma işlemi açık kaynak kodlu ve ücretsiz olan R programında gerçekleştirilmiştir. R programında kullanılan kütüphaneler, yazılımın kendisi gibi açık kaynaklı ve ücretsizdir. Sınıflandırma için XGBoost ve RO makine öğrenme algoritmaları kullanılmıştır[66][67]. RO, son yıllarda en yaygın kullanılan makine öğrenme algoritmalarının başında gelmektedir ve hemen tüm sınıflandırma çalışmalarında önerilen yeni yaklaşımların/algoritmaların test sonuçlarının karşılaştırmasında referans olarak kullanılmıştır. Dolayısıyla, bu çalışmada da XGBoost algoritması ile elde edilen sonuçların başarı oranı değerlendirmesi, RO algoritması sonuçları ile karşılaştırılması suretiyle yapılmıştır. Çalışmada yapılan işlem adımları Şekil 12’de gösterilmiştir.



Şekil 12. Çalışmada Kullanılan İşlem Adımları Diagramı

Eğitim ve test veri setlerinin oluşturulması için referans parsel verisi %50 oranında rastgele olarak *trainshapefile* ve *testshapefile* adlarıyla iki ayrı gruba ayrılmıştır. Eğitim ve test parselleri verileri shapefile formatında *raster* kütüphanesinde mevcut olan *shapefile* fonksiyonu kullanılarak bilgisayara tanımlanmıştır. *Raster* kütüphanesinde olan *stack* fonksiyonu kullanılarak 8 farklı tarihe ait görüntüler ve bu görüntülerden elde edilen doku özellikleri bantları (2.4.2. Görüntülerin Doku Özelliklerinin Çıkarılması) çok zamanlı veri yığını (*stack*) olarak tanımlanmıştır.

Görüntülerden, parsellerin içinde yer alan piksellerin geri yansıma değerinin çıkartılması işlemi *extract* fonksiyonu kullanılarak gerçekleştirilmiştir. Sonuç olarak; çoklu zaman Sentinel-1 görüntülerin geri yansıma ve doku özelliklerine ait değerler her bir piksel için elde edilmiştir. Sınıflandırmada kullanılacak veri setinde her bir piksel için 8 adet DY, 8 adet DD, 8 adet DY/DD oranı, 64 adet DY doku özellik bantları, 64 adet DD doku özellik bantları olmak üzere toplam 152 bant yer almıştır. Sınıflandırmada kullanılan bantlara ait bilgiler Tablo 7.'de verilmiştir.

Tablo 7. Sınıflandırmada Kullanılan Sentinel-1 DY ve DD Polarizasyon Görüntüleri ve Bu Görüntülerden Çıkarılan Doku Özellikleri Bantları

DY Görüntüleri	DD Görüntüleri	DY/DD Görüntüleri	DY Doku Özellik Bantları	DD Doku Özellik Bantları	TOPLAM
8 Adet	8 Adet	8 Adet	64 Adet (8 Görüntü × 8 Doku Özelliği)	64 Adet (8 Görüntü × 8 Doku Özelliği)	152

Eğitim ve test verileri oluşturulan 152 katmanlı veri yığınının okutulduktan sonra, XGBoost algoritması ile sınıflandırma işlemi *xgboost* kütüphanesi kullanılarak gerçekleştirilmiştir. XGBoost algoritması sınıflandırma modelinin oluşturulması için, *Eta*: 0,2 – *Nrounds*: 170 – *Max_depth*: 12 – *Subsample*: 1 ve *Col_sample_level*: 1 parametre değerleri belirlenmiş ve bu değerler kullanılarak (2.3.3. Aşırı Gradyan Artırma (XGBoost)) sınıflandırma işlemi gerçekleştirilmiştir.

Çalışmada kullanılan diğer sınıflandırma algoritması RO ile sınıflandırma işlemi *RandomForest* kütüphanesi kullanılarak gerçekleştirilmiştir. RO algoritması için *ntree* ve *mtry* parametre değerlerinin belirlenmesi için farklı çalışmalar incelenmiştir[72][73][74][75][76][77][78]. RO algoritmasında hesaplamada etkili olana kadar *ntree* değerinin olabildiği kadar yüksek alınabileceği belirtilmiştir[78]. Lawrence tarafından yapılan bir çalışmada, *ntree* sınıflandırma ağacı sayısı 500 olarak belirlendikten sonra hata değeri sabit kalmış ve değişmemiştir[79]. Bu değer yaygın şekilde kullanılmasının diğer bir nedeni ise en yaygın RO algoritması olarak kullanılan R programının *randomForest* kütüphanesi varsayılan *ntree* değerinin 500 alınmasıdır[71]. *Ntree* değeri Adelabu (vd. 2014), Millard ve Richardson (2015), Nitze (vd. 2015) tarafından yapılmış çalışmalarda 5000 olarak, Colditz, Reese (vd. 2014), Sesnie (vd.

2010) tarafından yapılan çalışmalarda 1000 olarak, Belgiu (vd. 2016) tarafından yapılan çalışmada 500 olarak ve Guan (vd.2013) tarafından yapılan çalışmada 100 olarak alınmıştır[72][73][74][75][76][77][78]. Literatürdeki bu çalışmaları temel alarak RO sınıflandırması için varsayılan *ntree*: 500 değerinin kabul edilebilir olduğu görülmüş ve bu çalışmada bu değer kullanılmıştır. *mtry* parametresi değeri olarak değişken değerinin karekökü alınmıştır[80][81]. Sınıflandırma işleminde 152 katmanlı veri yığını kullanıldığı için, *mtry* parametre değeri 12 olarak hesaplanmıştır.

2.5.3. Doğruluk Analizleri

Görüntü sınıflandırması çıktılarının doğruluk analizinde hata matrisleri en yaygın kullanılan yöntemdir. Bu çalışmada kullanılan XGBoost ve RO algoritmaları sınıflandırma çıktıları hata matrisleri şeklinde değerlendirilmiştir. Hata matrisleri, sınıflandırma için de kullanılan R programı kullanılarak oluşturulmuştur.

Çoklu sınıfa ait bir hata matrisi sınıf sayısı kadar satır ve sütundan oluşmaktadır. Satır ve sütunlarda sınıflandırma için kullanılan referans veriler ve sınıflandırma sonuç verileri yer almaktadır. Hata matrisinde doğru sonuçlar satır ve sütun kesişimde bulunan diyagonal değerlerdir. Diğer değerler ise yanlış sınıflandırılanları göstermektedir. Çoklu sınıfa ait bir hata matrisi örnek olarak Tablo 8’de gösterilmiştir.

Tablo 8. Çoklu Sınıfa Ait Hata Matrisi (Campbell, 1996’dan alınmıştır.)

Çoklu Sınıf Hata Matrisi		Referans					Kullanıcı Doğruluğu	Satır Toplamı
		1	2	3	4	5		
Sınıflandırma	1	510	110	85	23	10	69%	738
	2	54	1155	235	253	35	67%	1732
	3	15	217	930	173	8	69%	1343
	4	37	173	238	864	27	65%	1339
	5	5	17	23	11	265	83%	321
	Üretici Doğruluğu	82%	69%	62%	65%	77%	68%	5473
	Sütun Toplamı	621	1672	1511	1324	345	5473	

Hata matrisi değerlerinden genel doğruluk, kappa ($\hat{\kappa}$) değeri, üretici ve kullanıcı doğrulukları hesaplanabilmektedir. Test verilerinin sonuçlarının doğruluğu üretici doğruluğu, eğitim verilerinin sonuçlarının doğruluğu ise kullanıcı doğruluğudur. Genel

doğruluk, doğru sınıflandırma sonucundaki verilerin, eğitim verilerine bölünmesiyle bulunur. $\hat{\kappa}$ değeri Eşitlik 21'e göre hesaplanır.

$$Kappa(\hat{\kappa}) = \frac{N \sum_{i=1}^r x_{ii} - \sum_{i=1}^r (x_{i+} * x_{+i})}{N^2 - \sum_{i=1}^r (x_{i+} * x_{+i})} \quad (21)$$

Eşitlik 21'de N değeri toplam eğitim verisini, r sınıf sayısını, x_{ii} doğru sınıflandırılmış eğitim verisi toplamını, x_{+i} i satır toplamını, x_{i+} i sütun toplamını göstermektedir[82].

Yukarıdaki doğruluk değerlerine ek olarak, her sınıf için, kullanıcı doğruluğu (Eşitlik 22), özgüllük (Eşitlik 23), yaygınlık (Eşitlik 24), üretici doğruluğu (Eşitlik 25), yanlış sınıf değeri (Eşitlik 26), tespit etme oranı (Eşitlik 27), tespit etme yaygınlığı (Eşitlik 28) ve dengelenmiş doğruluk (Eşitlik 2) değerleri de hesaplanabilmektedir. Bu değerlerin hesaplanmasında ikili hata matrisleri kullanılmaktadır. İkili hata matrisine bir örnek Tablo 9'da gösterilmiştir. İkili hata matrisi, sınıflandırmada kullanılan verilerin, var olan hali ile sınıflandırma sonuçlarına göre doğru veya yanlış olduğunu tablo halinde göstermektedir. Tablo 9'da; Doğru Pozitif (DP), Yanlış Pozitif (YP), Yanlış Negatif (YN) ve Doğru Negatif (DN) değerleri mevcuttur. DP, var olan halinin doğru sınıflandırılması ve sonucun doğru olmasıdır. YP, var olan halinin doğru sınıflandırılması ve sonucun yanlış olmasıdır. YN, var olan halinin yanlış sınıflandırılması ve sonucun yanlış olmasıdır. DN, var olan halinin yanlış sınıflandırılması ve sonucun doğru olmasıdır[83]. Bu çalışmada, elde edilen sonuçların analizleri, kullanıcı doğruluk değeri (Eşitlik 22), üretici doğruluk değeri (Eşitlik 25) ve dengelenmiş doğruluk değerine (Eşitlik 29) göre yapılmıştır.

Tablo 9. İkili Hata Matrisine Bir Örnek

İkili Hata Matrisi		Referans Veri	
		Sınıf C	Sınıf C Değil
Tahmin	Sınıf C	DP	YP
	Sınıf C Değil	YN	DN

$$KullanıcıDoğruluğu = \frac{DP}{DP + YN} \quad (22)$$

$$Özgüllük = \frac{DN}{YP + DN} \quad (23)$$

$$Yaygınlık = \frac{DP + YN}{DP + YP + YN + DN} \quad (24)$$

$$ÜreticiDoğruluğu = \frac{KullanıcıDoğruluğu \times Yaygınlık}{(KullanıcıDoğruluğu \times Yaygınlık) + ((1 - Özgüllük) \times (1 - Yaygınlık))} \quad (25)$$

$$YanlışSınıfDeğeri = \frac{KullanıcıDoğruluğu \times (1 - Yaygınlık)}{((1 - KullanıcıDoğruluğu) \times Yaygınlık) + ((Özgüllük) \times (1 - Yaygınlık))} \quad (26)$$

$$TespitEtmeOranı = \frac{DP}{DP + YP + YN + DN} \quad (27)$$

$$TespitEtmeYaygınlığı = \frac{DP + YP}{DP + YP + YN + DN} \quad (28)$$

$$DengelenmişDoğruluk = \frac{KullanıcıDoğruluğu + Özgüllük}{2} \quad (29)$$

XGBoost ve RO algoritmaları sınıflandırma sonuçları için hesaplanan hata matrisleri sırasıyla Tablo 10 ve 11’de verilmiştir (“Bölüm 3.1.Bulguların Analizleri”). Tablo 8’de sınıflandırma uygulamasının çoklu sınıf hata matrisi örneğinde gösterilen hesaplamalarla örnek olarak kullanıcı ve genel doğruluk değerlerinin hesaplaması şu şekildedir. Kullanıcı doğruluğu (user’s accuracy), sınıflandırma verisi üzerinden doğru sınıflandırması yapılan ürünün toplam sayıya bölünmesi ile elde edilir. Örneğin; 1. sınıfın sınıflandırma verisindeki toplam piksel sayısı, hata matrisi 1. sınıfa ait satır değerlerin toplamı, 738’dir. Doğru sınıflandırılan piksel sayısı, hata matrisi 1. sınıf satır ve sütunu kesişimi değeri, 510’dur. Dolayısıyla, 1. sınıf için kullanıcı doğruluk değeri, aşağıda gösterildiği gibi, 0.69 olarak hesaplanır.

$$\frac{510}{738(510+110+85+23+10)} = 0,69$$

Genel doğruluk, bütün değerlerin toplamına doğru sınıflandırma sonuçlarının bölünmesi ile elde edilir. Tablo 8’de toplam referans sayısı 5473’tür. Doğru sınıflandırılan piksel sayısı, hata matrisi 1. sınıf satır ve sütunu kesişimi değeri, 510’dur. Sırasıyla her bir sınıfın doğru sınıflandırılan piksel sayıları ve referans verideki toplam piksel sayıları toplanır. Doğru sınıflandırılan piksel sayılarının toplamının referans verileri piksel sayısı

toplamına bölümü ile elde edilir. Tablo 8’da verilen hata matrisine göre hesaplama şu şekildedir;

$$\frac{510+1155+930+864+265}{621+1672+1511+1324+345} = 0,68$$

3. BULGULAR VE ANALİZLER

3.1. Bulguların Analizleri

Sınıflandırma uygulaması, XGBoost ve RO makine öğrenme algoritmaları ile toplam 152 bant (8 adet DY polarizasyon bantı, 8 adet DD polarizasyon bantı, 8 adet DY/DD oran bantı, 64 adet DY bantlarından çıkarılmış doku özellikleri bantı ve 64 adet DD bantlarından çıkarılmış doku özellikleri bantı) kullanılarak gerçekleştirilmiştir. XGBoost ve RO makine öğrenme algoritmaları ile elde edilen sonuçların değerlendirmeleri ve bu iki algoritmanın sonuçlarının karşılaştırması $\hat{\kappa}$ ve genel doğruluk değerleri üzerinden yapılmıştır.

XGBoost ve RO algoritmaları ile elde edilen $\hat{\kappa}$ ve genel doğruluk değerleri Tablo 10’da gösterilmiştir. XGBoost algoritması ile sınıflandırma için hesaplanan $\hat{\kappa}$ değeri (0,8059) ve genel doğruluk değeri (%83,67), RO algoritması için hesaplanan $\hat{\kappa}$ değeri (0.8036) ve genel doğruluk değerinden (%83.55) daha yüksek çıkmıştır. Dolayısıyla, elde edilen sonuçlara göre, bu çalışma alanında, XGBoost algoritması RO algoritmasından, düşük oranda da olsa, daha iyi performans sağlamıştır.

Tablo 10. XGBoost ve RO Algoritmaları ile Sınıflandırma Sonuçları İçin Hesaplanmış $\hat{\kappa}$ ve Genel Doğruluk Değerleri

Doğruluklar	$\hat{\kappa}$		Genel Doğruluk	
Sınıflandırma Algoritmaları	XGBoost	RO	XGBoost	RO
Orijinal DY (8 Adet) Orijinal DD (8 Adet) DY/DD Oranı (8 Adet) DY Doku Bantları (64 Adet) DD Doku Bantları (64 Adet) Toplam: 152 Bant	0,8059	0,8036	%83,67	%83,55

Sınıflandırma sonuçlarından hesaplanmış olan ürün seviyesinde kullanıcı doğruluğu, üretici doğruluğu ve dengelenmiş doğruluk değerleri de değerlendirilmiştir. Her sınıf için, kullanıcı doğruluğu, üretici doğruluğu ve dengelenmiş doğruluk değerlerinin hesaplanması sırasıyla Eşitlikler 22, 25 ve 29 kullanılarak yapılmıştır.

XGBoost algoritması ile sınıflandırma için elde edilen kullanıcı, üretici ve dengelenmiş doğruluk değerleri Tablo 11’de, RO algoritması ile sınıflandırma için elde edilen kullanıcı, üretici ve dengelenmiş doğruluk değerleri ise Tablo 12’de verilmiştir. Pamuk, hem XGBoost hem de RO sınıflandırma sonuçlarında en yüksek kullanıcı, üretici ve dengelenmiş doğruluk değerlerini sağlamıştır. XGBoost sınıflandırma sonuçlarına göre (Tablo 11) pamuk için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %99,13, %96,59 ve %98,25 tir. RO sınıflandırma sonuçlarına göre (Tablo 12) pamuk için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri ise sırasıyla %99,34, %96,87 ve %98,40 tir.

Hem XGBoost hem de RO sınıflandırma sonuçlarına göre, en düşük kullanıcı, üretici ve dengelenmiş doğruluk değerlerine sahip ürün yonca olmuştur. Yonca için XGBoost sınıflandırma sonucundan (Tablo 11) hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %23,96, %10,39 ve %55,03; RO sınıflandırma sonucundan (Tablo 12) hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri ise sırasıyla %25,17, %5,74 ve %52,79 dır. XGBoost ve RO algoritmaları ile sınıflandırma sonucu elde edilen kullanıcı, üretici ve dengelenmiş doğruluk değerleri karşılaştırmalı olarak Tablo 13’te verilmiştir. Bu iki algoritmanın sonuçları için hesaplanan McNemar test sonuçları Tablo 14’te verilmiştir. XGBoost ve RO algoritmaları sınıflandırma çıktıları (tematik haritalar) sırasıyla Şekiller 13 ve 14’te gösterilmiştir.

Tablo 11. XGBoost Algoritması ile Sınıflandırma Sonucundan Hesaplanan Hata Matrisi

XGBOOST											
	Referans										
Sınıflandırma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin	Toplam	Kullanıcı Doğruluğu
Buğday	9704	26	372	294	1931	98	36	129	169	12759	76,06
Domates	141	3909	0	283	103	48	47	4	7	4542	86,06
Mera	235	9	2068	8	324	1	5	59	9	2718	76,09
Mısır	120	438	12	9279	96	7	74	15	27	10068	92,16
Mısır2	911	44	13	140	782	22	34	5	41	1992	39,26
Pamuk	12	32	0	0	6	7090	12	0	0	7152	99,13
Üzüm	235	341	4	534	68	55	13841	68	1219	16365	84,58
Yonca	8	1	1	42	88	1	0	69	78	288	23,96
Zeytin	197	70	55	114	218	18	1078	315	10667	12732	83,78
Toplam	11563	4870	2525	10694	3616	7340	15127	664	12217	68616	
Üretici Doğruluğu	83,92	80,27	81,90	86,77	21,63	96,59	91,50	10,39	87,31	Genel Doğruluk	%83,67
Dengelenmiş Doğruluk	89,28	89,64	90,46	92,70	59,88	98,25	93,39	55,03	91,83	Kappa	0,8059

Tablo 12. RO Algoritması ile Sınıflandırma Sonucundan Hesaplanan Hata Matrisi

RO											
	Referans										
Sınıflandırma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin	Toplam	Kullanıcı Doğruluğu
Buğday	10267	4	522	332	2005	91	52	128	151	13552	75,76
Domates	86	3829	0	230	87	12	21	0	0	4265	89,78
Mera	228	11	1906	0	60	0	0	70	12	2287	83,34
Mısır	143	559	10	9240	455	11	85	9	31	10543	87,64
Mısır2	177	34	3	40	538	8	2	0	22	824	65,29
Pamuk	13	22	0	0	9	7097	3	0	0	7144	99,34
Üzüm	355	276	2	556	77	97	13732	56	1411	16562	82,91
Yonca	7	0	1	22	19	0	0	38	64	151	25,17
Zeytin	227	118	71	248	354	10	1179	361	10469	13037	80,30
Toplam	11503	4853	2515	10668	3604	7326	15074	662	12160	68365	
Üretici Doğruluğu	89,25	78,90	75,79	86,61	14,93	96,87	91,10	05,74	86,09	Genel Doğruluk	%83,55
Dengelenmiş Doğruluk	91,74	89,11	87,60	92,18	57,24	98,40	92,89	52,79	90,76	Kappa	0,8059

Tablo 13. XGBoost ve RO Algoritmaları Ürün Seviyesinde Kullanıcı, Üretici ve Dengelenmiş Doğruluk Değerlerinin Karşılaştırması

XGBoost ve RO Algoritmaları Ürün Seviyesinde Doğrulukları Karşılaştırması											
%	Algoritma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin	
Kullanıcı Doğruluğu	XGBoost	76,06	86,06	76,09	92,16	39,26	99,13	84,58	23,96	83,78	
	RO	75,76	89,78	83,34	87,64	65,29	99,34	82,91	25,17	80,30	
Üretici Doğruluğu	XGBoost	83,92	80,27	81,90	86,77	21,63	96,59	91,50	10,39	87,31	
	RO	89,25	78,90	75,79	86,61	14,93	96,87	91,10	5,74	86,09	
Dengelenmiş Doğruluk	XGBoost	89,28	89,64	90,46	92,70	59,88	98,25	93,39	55,03	91,83	
	RO	91,74	89,11	87,60	92,18	57,24	98,40	92,89	52,79	90,76	

Tablo 14. XGBoost ve RO Algoritmaları McNemar Test Sonuçları

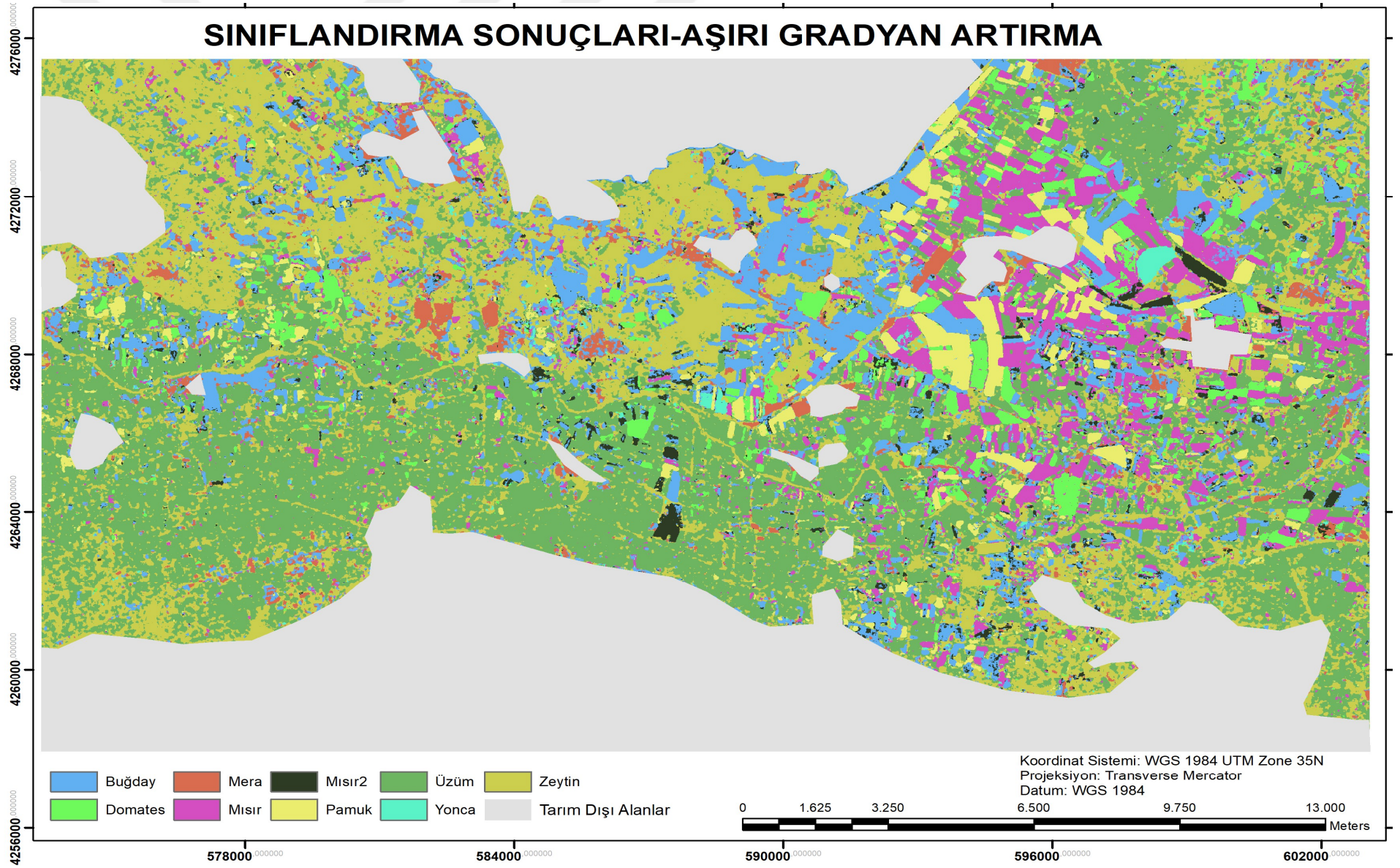
McNemar Testi	Yanlış Sınıf	Doğru Sınıf
Yanlış Sınıf	4260	11500
Doğru Sınıf	11207	64356

Tablo 14'te 1/1 hücresi her iki algoritma tarafından yanlış sınıflandırma sayısını, 1/2 hücresi sadece XGBoost algoritması ile yanlış sınıflandırma sayısını, 2/1 hücresi sadece RO algoritması ile yanlış sınıflandırma sayısını ve 2/2 hücresi her iki algoritma tarafından doğru sınıflandırma sayısını göstermektedir.

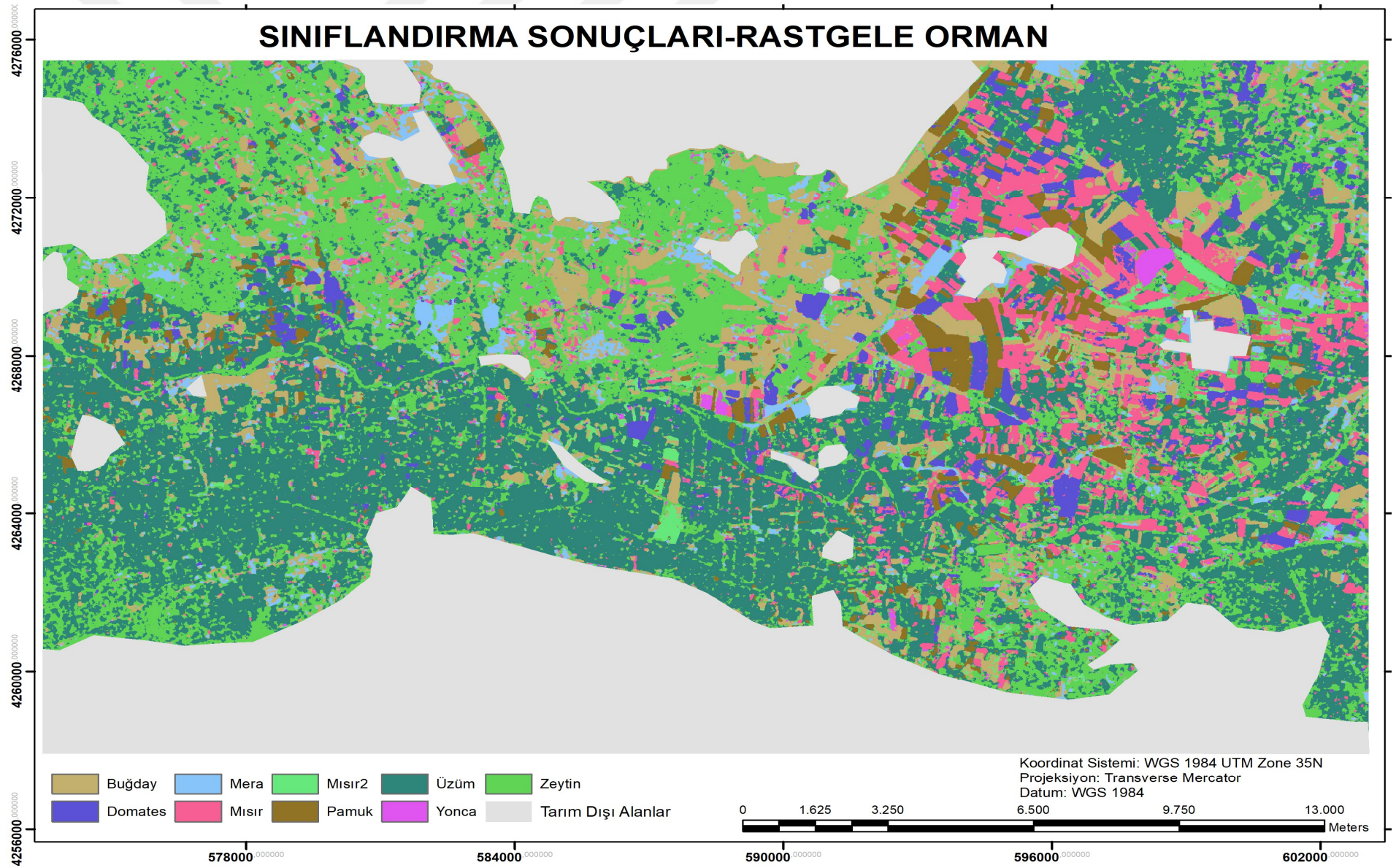
$$\frac{(|e_{01} - e_{10}| - 1)^2}{e_{01} + e_{10}} \sim X_1^2 \quad (30)$$

McNemar Ki Kare test sonucu (Eşitlik 30) çıkan değer (3,755) serbestlik derecesi 1 için istatistiksel anlamlılık sınır değeri olan 3.84'ten küçük olduğundan, iki algoritma arasında istatistiksel olarak anlamlı fark olmadığı sonucu elde edilir. $p = 0.05265$ ($p > 0.05$) olduğu için bu sonucun yorumu iki değerlendirme yöntemine ait sonuçlar istatistiksel olarak birbirinden farklıdır veya birbirileri ile uyumlu değildir şeklindedir.[84]





Şekil 13. XGBoost Algoritması Sınıflandırma Çıktısı (Tematik Harita)



Şekil 14. RO Algoritması Sınıflandırma Çıktısı (Tematik Harita)

Tablo 11'a göre XGBoost algoritmasıyla tarım ürünlerinin ürün bazında sınıflandırma sonuçlarının değerlendirmeleri şu şekildedir;

Buğday için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %76,06, %83,92 ve %89,28 tür. Buğday sınıfına ait toplam 11563 referans pikselden 9704 adedi doğru olarak sınıflandırılmıştır. Buğdayın pamuk ve yonca haricindeki diğer tarım ürünleri ile karıştırıldığı görülmektedir. En yüksek karışma 911 piksel ile mısır2 sınıfı ile olmuştur. Karışma miktarı mera ve üzüm ile 235 piksel, domates ile 141 piksel, mısır ile 120 piksel ve zeytin ile 197 piksel seviyesindedir. Karışan ürünlerin büyüme döneminde buğdayla benzer özellikler gösterdiği için buğdayın yansımaya yakın değerler vermiştir.

Domates için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %86,06, %80,27 ve %89,64 tür. Domates sınıfına ait toplam 4870 referans pikselden 3909 adedi doğru olarak sınıflandırılmıştır. Domatesin mera ve yonca haricindeki diğer ürünler ile karıştırıldığı görülmektedir. En yüksek karıştırma 438 piksel ile mısır ile olmuştur. Karışma miktarı buğday ile 26 piksel, mısır2 ile 44 piksel, pamuk ile 32 piksel, üzüm ile 341 piksel ve zeytin ile 70 piksel düzeyindedir. Domatesin ekim döneminde; buğday büyüme, mısır, mısır2 sınıfında ekilen iki farklı üründen biri olan buğday ve pamuk ekim dönemlerindedir. Ürünlerin aynı tarihlerde ekim döneminde olmasından dolayı sınıflandırmada karıştırılmıştır.

Mera için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %76,09, %81,90 ve %90,46 dır. Mera sınıfına ait toplam 2525 referans pikselden 1906 adedi doğru olarak sınıflandırılmıştır. Meranın buğday ve zeytin ile karıştırıldığı görülmektedir. En yüksek karıştırma 372 piksel ile buğdayla olmuştur. Zeytin ile karışma 55 piksel düzeyindedir. Meralar hayvan otlaması için boş bırakılan arazilerin olduğu bölgelerdir. Karışmaların, bölgedeki ürünlerinin ekim ve büyüme dönemlerinde benzer yansıma değerleri vermesinden dolayı olduğu düşünülmektedir.

Mısır için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %92,16, %86,77 ve %92,70 tir. Mısır sınıfına ait toplam 10694 referans pikselden 9279 adedi doğru olarak sınıflandırılmıştır. Mısırın mera, pamuk ve yonca haricindeki diğer tarım ürünleri ile karıştırıldığı görülmektedir. En yüksek karışma 534 piksel ile üzümle olmuştur. Karışma miktarı buğday ile 294 piksel, domates ile 283 piksel, mısır2 ile 140 piksel ve zeytin ile 114 piksel olmuştur. Mısır, domatesle aynı dönem ve mısır2 sınıfında

ekilen iki farklı üründen biri olan mısır ile 2-3 aylık dönemde ekilmektedir. Dolayısıyla, ekim ve büyüme dönemlerindeki benzer yansıma özelliklerinden dolayı sınıflandırmada karışma olmaktadır.

Mısır2 için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %39,26, %21,63 ve %59,88 dir. Mısır2 sınıfına ait toplam 3616 referans pikselden 782 adedi doğru olarak sınıflandırılmıştır. Mısır2nin pamuk haricindeki diğer tarım ürünleri ile karıştığı görülmektedir. Mısır2, buğday ve buğday hasadından sonra ekilen körpe mısırı içermektedir. Dolayısıyla, aynı tarım ürününü içeren dönemlere ait sınıflandırmalarda yüksek oranda karıştırma gerçekleşmiştir. En yüksek karışma 1931 piksel ile buğdayla olmuştur. Karışma miktarı, domates ile 103 piksel, mera ile 324 piksel, mısır ile 96 piksel, üzüm ile 68 piksel, yonca ile 88 piksel ve zeytin ile 114 piksel düzeyindedir.

Pamuk için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %99,13, %96,59 ve %98,25 tir. Pamuk sınıfına ait toplam 7340 referans pikselden 7090 adedi doğru olarak sınıflandırılmıştır. Pamuğun buğday, domates ve üzüm tarım ürünleri ile düşük oranda karıştırıldığı görülmektedir. Karışma miktarı, buğday ile 98 piksel, domates ile 48 piksel ve üzüm ile 55 piksel düzeyindedir.

Üzüm için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %84,58, %91,50 ve %93,39 dur. Üzüm sınıfına ait toplam 15127 referans pikselden 13841 adedi doğru olarak sınıflandırılmıştır. Üzümün zeytin ile 1078 piksel karıştırıldığı görülmektedir. Karıştırıldığı tarım ürünlerinin hepsi aynı dönemde ekimi yapılan ürünlerdir. Büyüme dönemleri farklılık gösterse de geri yansıma değerlerinde etkili olmuş ve sınıflandırmayı etkilemiştir.

Ürünler içinde en düşük doğruluk değerlerine sahip olan yonca için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %23,96, %10,39 ve %55,03 tür. Yonca sınıfına ait toplam 664 referans pikselden 69 adedi doğru olarak sınıflandırılmıştır. Yoncanın domates, mısır2 ve pamuk haricindeki diğer tarım ürünleri ile karıştırıldığı görülmektedir. En yüksek karışma 315 piksel ile zeytinle olmuştur. Karışma miktarı, buğday ile 129 piksel, mera ile 59 piksel, mısır ile 15 piksel ve üzüm ile 68 piksel düzeyinde olmuştur. Yonca ekildikten sonra yaklaşık olarak 5 yıl boyunca arazide kalmaktadır. Hasat her 2-3 haftada bir sefer olacak şekilde yapılmaktadır. Düzenli

ve kısa süreli hasatların yapılmasının tarım ürünü sınıflandırmasının karıştırılmasında çok etkili olduğu değerlendirilmektedir.

Zeytin için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %83,78, %87,31 ve %91,83 tür. Zeytin sınıfına ait toplam 12217 referans pikselden 10667 adedi doğru olarak sınıflandırılmıştır. Zeytinin buğday, üzüm ve yonca tarım ürünleri ile karıştırıldığı görülmektedir. En yüksek karışma 1219 piksel ile üzümle olmuştur. Karışma miktarı, buğday ile 169 piksel ve yonca ile 78 piksel olmuştur. Zeytin dikim ve büyüme dönemi gerçekleştikten sonra değişim göstermemekte yaprak bile dökmemektedir. Dolayısıyla, ekildikten sonra büyüme döneminde olan diğer tarım ürünleri ile karışma göstermiştir.

Tablo 12'ye göre RO algoritmasıyla tarım ürünlerinin ürün bazında sınıflandırma sonuçlarının değerlendirmeleri şu şekildedir;

Buğday için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %75,76, %89,25 ve %91,74 tür. Buğday sınıfına ait toplam 11503 referans pikselden 10267 adedi doğru olarak sınıflandırılmıştır. Buğdayın pamuk ve yonca haricindeki diğer tarım ürünleri ile karıştırıldığı görülmektedir. En yüksek karışma 355 piksel ile üzümle olmuştur. Karışma miktarı, domates ile 86 piksel, mera ile 228 piksel, mısır ile 143 piksel, mısır2 ile 177 piksel ve zeytin ile 227 piksel düzeyindedir.

Domates için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %89,78, %78,90 ve %89,11 dir. Domates sınıfına ait toplam 4853 referans pikselden 3829 adedi doğru olarak sınıflandırılmıştır. Domatesin karıştığı ürünler arasında mısır, mısır2, üzüm ve zeytin yer almaktadır. Bu ürünlerden en yüksek karışma 559 piksel ile mısırla olmuştur. Karışma miktarı, mısır2 ile 34 piksel, üzüm ile 276 piksel ve zeytin ile 118 piksel düzeyindedir.

Mera için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %83,34, %75,79 ve %87,60 tır. Mera sınıfına ait toplam 2515 referans pikselden 1906 adedi doğru olarak sınıflandırılmıştır. Meranın karıştığı iki ürün buğday ve zeytindir. Bu ürünler arasında en yüksek oranda karışma 522 piksel ile buğdayla olmuştur. Zeytinle olan karışma miktarı 71 piksel düzeyindedir.

Mısır için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %87,64, %86,61 ve %92,18 dir. Mısır sınıfına ait toplam 10688 referans pikselden 9240 adedi doğru olarak sınıflandırılmıştır. Mısırın mera, mısır2, pamuk ve yonca haricindeki

diğer tarım ürünleri ile karıştırıldığı görülmektedir. En yüksek karışma 556 piksel ile üzümle olmuştur. Karışma miktarı, buğday ile 332 piksel, domates ile 230 piksel, üzüm ile 556 piksel ve zeytin ile 248 piksel düzeyindedir.

Mısır2 için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %65,29, %14,93 ve %57,24 tür. Mısır2 sınıfına ait toplam 3604 referans pikselden 538 adedi doğru olarak sınıflandırılmıştır. Mısır2nin pamuk ve yonca haricindeki diğer tarım ürünleri ile karıştırıldığı görülmektedir. En yüksek karışma 2005 piksel ile buğdayla olmuştur. Karışma miktarı, domates ile 87 piksel, mera ile 60 piksel, mısır ile 455 piksel, üzüm ile 77 piksel ve zeytin ile 354 piksel düzeyindedir.

Pamuk için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %99,34, %96,87 ve %98,40 tır. Pamuk sınıfına ait toplam 7326 referans pikselden 7097 adedi doğru olarak sınıflandırılmıştır. Pamuğun daha ziyade buğday ve üzümle karıştırıldığı görülmektedir. Buğday ile 91 piksel ve üzüm ile 97 piksel karışma göstermiştir.

Üzüm için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %82,91, %91,10 ve %92,89 dur. Üzüm sınıfına ait toplam 15074 referans pikselden 13732 adedi doğru olarak sınıflandırılmıştır. Üzümün karışma gösterdiği ürünler mısır ve zeytindir. Bu ürünler arasında en yüksek karışma 1078 piksel ile zeytinle olmuştur. Mısır ile karışma miktarı 85 piksel düzeyindedir.

Yine, ürünler içinde en düşük doğruluk değerlerine sahip olan yonca için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %25,17, %5,74 ve %52,79 dur. Yonca sınıfına ait toplam 662 referans pikselden 38 adedi doğru olarak sınıflandırılmıştır. Yoncanın domates, mısır2 ve pamuk haricindeki diğer tarım ürünleri ile karıştırıldığı görülmektedir. Yoncanın en yüksek karışma gösterdiği ürün 361 piksel ile zeytin olmuştur. Karışma miktarı, buğday ile 128 piksel, mera ile 70 piksel, mısır ile 9 piksel ve üzüm ile 56 piksel düzeyindedir.

Zeytin için hesaplanmış kullanıcı, üretici ve dengelenmiş doğruluk değerleri sırasıyla, %80,30, %86,09 ve %90,76 dır. Zeytin sınıfına ait toplam 12160 referans pikselden 10469 adedi doğru olarak sınıflandırılmıştır. Zeytinin karıştığı ürünler arasında buğday, üzüm ve yonca yer almaktadır. En yüksek karışma 1411 piksel ile üzümle olmuştur. Karışma miktarı, buğday ile 151 piksel ve yonca ile 64 piksel düzeyindedir.

XGBoost ve RO algoritmaları sınıflandırma sonuçları doğruluk değerleri karşılaştırmalı olarak Tablo 15’te verilmiştir.

Tablo 15. XGBoost ve RO Algoritmaları Sınıflandırma Sonuçları Hata Matrisleri Karşılaştırması

	REFERANS																	
	XGB	RO	XGB	RO	XGB	RO	XGB	RO	XGB	RO	XGB	RO	XGB	RO	XGB	RO	XGB	RO
Sınıflandırma	Buğday		Domates		Mera		Mısır		Mısır2		Pamuk		Üzüm		Yonca		Zeytin	
Buğday	9704	10267	26	4	372	522	294	332	1931	2005	98	91	36	52	129	128	169	151
Domates	141	86	3909	3829	0	0	283	230	103	87	48	12	47	21	4	0	7	0
Mera	235	228	9	11	2068	1906	8	0	324	60	1	0	5	0	59	70	9	12
Mısır	120	143	438	559	12	10	9279	9240	96	455	7	11	74	85	15	9	27	31
Mısır2	911	177	44	34	13	3	140	40	782	538	22	8	34	2	5	0	41	22
Pamuk	12	13	32	22	0	0	0	0	6	9	7090	7097	12	3	0	0	0	0
Üzüm	235	355	341	276	4	2	534	556	68	77	55	97	13841	13732	68	56	1219	1411
Yonca	8	7	1	0	1	1	42	22	88	19	1	0	0	0	69	38	78	64
Zeytin	197	227	70	118	55	71	114	248	218	354	18	10	1078	1179	315	361	10667	10469

3.2. XGBoost ve RO Algoritmaları İşlem Süreleri

Her iki algoritma ile sınıflandırma işlemi Intel® Core™ İ7 6700HQ 2.60GHz işlemci, 16 GB RAM, 1 TB HDD ve Windows 10 işlemi sistemi bulunan dizüstü bilgisayar ortamında gerçekleştirilmiştir. Sınıflandırma işlemleri için Sentinel-1 görüntülerinden eğitim ve test verileri belirlendikten sonra algoritmalar arasında model oluşturma süreleri kaydedilmiştir. Her iki algoritma da aynı verileri kullanmasına rağmen işlemci çekirdeklerini farklı olarak kullandığı görülmüştür. XGBoost algoritması, parametre olarak belirlenen işlemci çekirdek sayısını, aynı anda ve paralel olarak kullanmaktadır. Diğer taraftan, RO algoritması için işlemci çekirdekleriyle ilgili parametre değeri girilemediği tespit edilmiştir. XGBoost modeli oluşturulurken bilgisayar üzerinde başka bir program çalıştırılmamış, çekirdek sayısı parametre değeri olarak 7 girilmiş ve model oluşturmak için kullanılmıştır. RO modeli oluşturulurken bilgisayar üzerinde yine başka bir program çalıştırılmamış, varsayılan çekirdek sayısı olarak 2 alınmış ve model oluşturulmuştur. Bu parametrelerle model oluşturma süresi XGBoost algoritması için 25 dakika, RO algoritması için ise 120 dakika olarak ölçülmüştür.

3.3. Özellik Önem Derecelerinin Seçilmesi ve Sınıflandırma Sonuçları

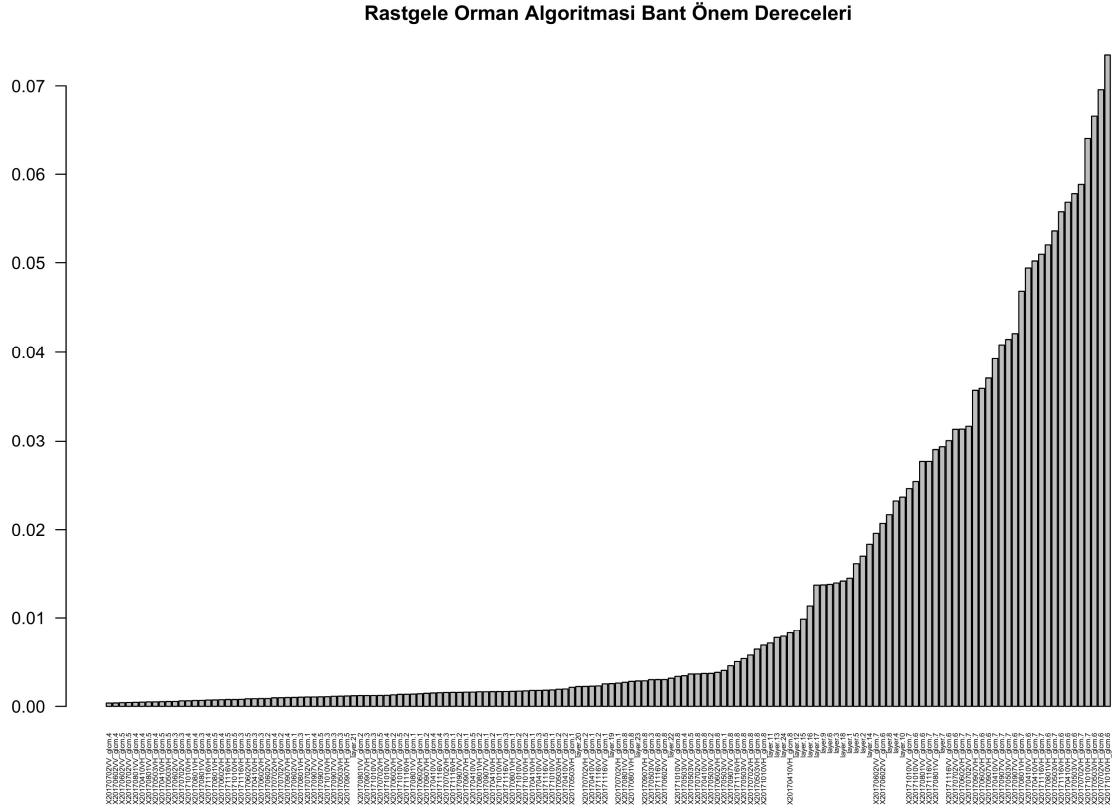
Özellik seçimi makine öğrenme algoritmaları uygulamalarında önemli bir adım olmuştur. Modern veri setleri pratik model oluşturma için birçok değişken ile tanımlanmaktadır. Genellikle, değişkenlerin çoğunun sınıflandırmada önemi düşük ve ilerlemede önemleri bilinmemektedir. Çok büyük veri setleri ile çalışmak algoritmaları yavaşlatmakta ve fazlasıyla gereksiz kaynak tüketmektedir. [85]

Özellik seçimi veri analizi, makine öğrenmesi ve veri madenciliğinde gittikçe artan öneme sahiptir. Büyük veri setleri için önemsiz ve gereksiz özellikteki verilerin, aşırı uyum göstermeyi (overfitting) engellemek ve büyük veri setlerinin üstesinden gelmek için alt kümesi seçilerek filtreden geçirilmesi önem arz etmektedir.[86]

XGBoost ve RO algoritmaları değişken önem dereceleri iki ölçü değeri tanımlar. XGBoost algoritması için ilk ölçü değeri, modeldeki her ağaç için her özelliğin katkısıyla hesaplanan, ilgili özelliğin modele göreceli katkısını (Gain) ifade eder. İkinci ölçü değeri ise, özelliğin ağaçlarda kullanılma bağıl sayısı (Frequency) hesaplanır. Daha yüksek oranda kazanç (Gain) ve frekans (Frequency) daha önemli tahmin değişkenini ifade eder. RO algoritması için ilk ölçü değeri (Mean Decrease Accuracy), her değişken çıkarıldığında modelin ne kadar doğruluk kaybettiğini ifade eder. İkinci ölçü değeri

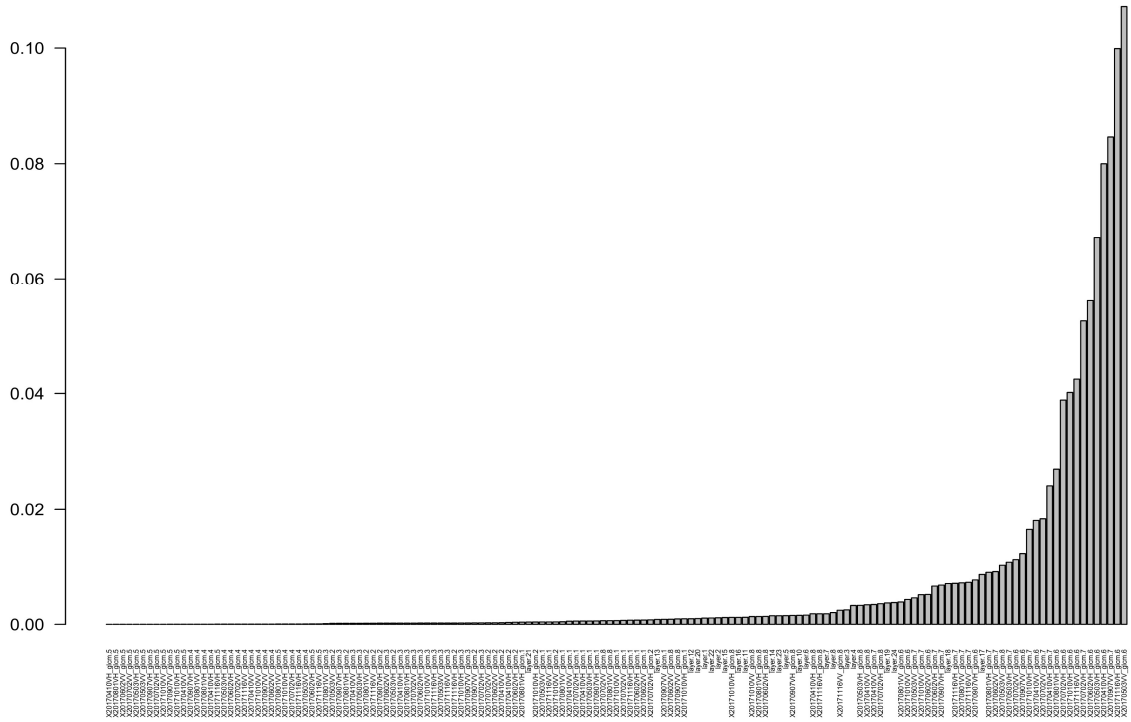
(Mean Decrease Gini) ise, rastgele orman sonuçlarında ortaya çıkan düğüm ve yapraklardaki homojenliğe her değişkenin nasıl katkı sağladığını ölçer. Her iki ölçü değerinin yüksek olması, tahmin değişkeninin daha önemli olduğunu gösterir. [87]

Bu çalışmada kullanılan RO algoritması için Doğrulukta Ortalama Düşüş (Mean Decrease Accuracy) ve XGBoost algoritması için Kazanç (Gain) kullanarak hesaplanmış bant önem dereceleri sırasıyla Şekil 15 ve 16’da verilmiştir.



Şekil 15. RO Algoritması Bant Önem Dereceleri

XGBoost Algoritması Bant Önem Dereceleri



Şekil 16. XGBoost Algoritması Bant Önem Dereceleri

Hesaplanan bant önem derecelerinin analizi yapıldığında eşik değerleri olarak 50 bant ve 70 bant seçilmiştir. Dolayısıyla, sınıflandırma işlemi en önemli 50 bant ve 70 bantla gerçekleştirilmiş olup elde edilen $\hat{\kappa}$ ve genel doğruluk değerleri Tablo 16’da verilmiştir. En önemli 50 bant kullanarak XGBoost algoritması ile sınıflandırma sonucu hesaplanan $\hat{\kappa}$ değeri (0,6925) ve genel doğruluk değeri (%74,13), RO algoritması için hesaplanan $\hat{\kappa}$ değeri (0,6988) ve genel doğruluk değerinden (%74,82), çok az oranda, daha düşük çıkmıştır. Benzer şekilde, en önemli 70 bant kullanarak XGBoost algoritması ile sınıflandırma sonucu hesaplanan $\hat{\kappa}$ değeri (0,7357) ve genel doğruluk değeri (%77,83), RO algoritması için hesaplanan $\hat{\kappa}$ değeri (0,7395) ve genel doğruluk değerinden (%78,21), çok az oranda, daha düşük çıkmıştır.

XGBoost algoritması için, 70 bant ile yapılan sınıflandırma sonucu hesaplanan genel doğruluk değeri (%77.83) ve $\hat{\kappa}$ değeri (0.7357), 50 bant ile yapılan sınıflandırma sonucu hesaplanan genel doğruluk değeri (%74.13) ve $\hat{\kappa}$ değerine (0.6925) göre daha yüksek çıkmıştır. Genel doğruluk değeri ve $\hat{\kappa}$ değeri için artış miktarı sırasıyla, %3.7 ve 0.0432 dir.

RO algoritması için, 70 bant ile yapılan sınıflandırma sonucu hesaplanan genel doğruluk değeri (%78.21) ve $\hat{\kappa}$ değeri (0.7395), 50 bant ile yapılan sınıflandırma sonucu hesaplanan genel doğruluk değeri (%74.82) ve $\hat{\kappa}$ değerine (0.6988) göre daha yüksek çıkmıştır. Genel doğruluk ve $\hat{\kappa}$ değerleri için artış miktarı sırasıyla, %3.39 ve 0.0407 dir

Tablo 16. En Önemli 50 Bant ve 70 Bant ile Sınıflandırma Sonuçları İçin Hesaplanmış $\hat{\kappa}$ ve Genel Doğruluk Değerleri

Doğruluklar	$\hat{\kappa}$		Genel Doğruluk	
Sınıflandırma Algoritmaları	XGBoost	RO	XGBoost	RO
50 Bant	0,6925	<u>0,6988</u>	%74,13	<u>%74,82</u>
70 Bant	0,7357	<u>0,7395</u>	%77,83	<u>%78,21</u>

Bant önem derecelerine göre en önemli 50 bant için XGBoost algoritması ile sınıflandırma için elde edilen kullanıcı, üretici ve dengelenmiş doğruluk değerleri Tablo 17’de, RO algoritması ile sınıflandırma için elde edilen kullanıcı, üretici ve dengelenmiş doğruluk değerleri Tablo 18’de verilmiştir.

Tablo 17. En Önemli 50 Bant Kullanarak XGBoost Algoritması ile Sınıflandırma Sonucundan Hesaplanan Hata Matrisi

XGB-50											
	Referans										
Sınıflandırma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin	Toplam	Kullanıcı Doğruluğu
Buğday	9152	51	553	700	2186	512	22	159	386	13721	66,70
Domates	75	3266	17	562	237	194	97	1	44	4493	72,69
Mera	353	2	1786	198	4	3	8	1	17	2372	75,29
Mısır	236	864	139	7906	224	24	260	48	96	9797	80,70
Mısır2	1281	11	27	110	748	1281	33	97	90	3678	20,34
Pamuk	5	139	2	9	77	4880	225	0	0	5337	91,44
Üzüm	356	427	0	763	105	340	14170	166	2554	18881	75,05
Yonca	29	6	1	325	15	56	110	44	116	702	6,27
Zeytin	76	104	0	121	20	50	202	148	8914	9635	92,52
Toplam	11563	4870	2525	10694	3616	7340	15127	664	12217	68616	
Üretici Doğruluğu	79,15	67,06	70,73	73,93	20,69	66,49	93,67	06,63	72,96	Genel Doğruluk	%74,13
Dengelenmiş Doğruluk	85,57	82,57	84,92	85,33	58,09	82,87	92,43	52,83	85,84	Kappa	0,6925

Tablo 18. En Önemli 50 Bant Kullanarak RO Algoritması ile Sınıflandırma Sonucundan Hesaplanan Hata Matrisi

RO-50											
	Referans										
Sınıflandırma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin	Toplam	Kullanıcı Doğruluğu
Buğday	9497	50	480	716	2060	680	23	174	387	14067	67,51
Domates	61	2948	0	198	171	119	60	0	35	3592	82,07
Mera	251	2	1897	142	37	10	2	0	9	2350	80,72
Mısır	242	1007	81	7954	253	69	357	30	147	10140	78,44
Mısır2	604	2	0	57	719	281	4	36	12	1715	41,92
Pamuk	11	152	3	35	120	5247	222	0	16	5806	90,37
Üzüm	422	450	3	830	96	678	13765	163	3324	19731	69,76
Yonca	22	3	6	124	15	13	59	35	96	373	9,38
Zeytin	235	123	0	277	67	46	168	209	7976	9101	87,64
Toplam	11345	4737	2470	10333	3538	7143	14660	647	12002	66875	
Üretici Doğruluğu	83,71	62,23	76,80	76,98	20,32	73,46	93,89	5,41	66,46	Genel Doğruluk	%74,82
Dengelenmiş Doğruluk	87,74	80,60	88,05	86,56	59,38	86,26	91,23	52,45	82,20	Kappa	0,6988

Bant önem derecelerine göre en önemli 70 bant için XGBoost algoritması ile sınıflandırma için elde edilen kullanıcı, üretici ve dengelenmiş doğruluk değerleri Tablo 19’da, RO algoritması ile sınıflandırma için elde edilen kullanıcı, üretici ve dengelenmiş doğruluk değerleri Tablo 20’de verilmiştir.

Tablo 19. En Önemli 70 Bant Kullanarak XGBoost Algoritması ile Sınıflandırma Sonucundan Hesaplanan Hata Matrisi

XGB-70											
	Referans										
Sınıflandırma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin	Toplam	Kullanıcı Doğruluğu
Buğday	9326	99	815	447	2300	126	17	198	402	13730	67,92
Domates	76	3398	1	423	148	225	58	0	16	4345	78,20
Mera	554	0	1587	62	37	1	21	0	17	2269	69,94
Mısır	173	682	53	8749	153	15	149	34	95	10103	86,60
Mısır2	936	10	48	135	680	493	57	66	54	2479	27,43
Pamuk	3	140	3	8	12	6275	46	0	0	6487	96,73
Üzüm	362	444	2	786	99	176	14585	156	2764	19374	75,28
Yonca	60	5	11	21	172	1	30	42	107	449	9,35
Zeytin	83	92	5	63	15	28	164	168	8762	9380	93,41
Toplam	11563	4870	2525	10694	3616	7340	15127	664	12217	68616	
Üretici Doğruluğu	80,65	69,77	62,85	81,81	18,80	85,49	96,41	6,33	71,72	Genel Doğruluk	%77,83
Dengelenmiş Doğruluk	86,47	84,14	80,91	89,74	58,02	92,57	93,73	52,86	85,31	Kappa	0,7357

Tablo 20. En Önemli 70 Bant Kullanarak RO Algoritması ile Sınıflandırma Sonucundan Hesaplanan Hata Matrisi

RO-70											
	Referans										
Sınıflandırma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin	Toplam	Kullanıcı Doğruluğu
Buğday	9499	47	660	588	1938	85	24	195	373	13409	70,84
Domates	62	3100	0	226	51	298	13	0	1	3751	82,64
Mera	268	0	1698	35	19	0	9	0	6	2035	83,44
Mısır	233	769	46	8325	193	21	165	33	98	9883	84,24
Mısır2	366	11	31	34	1024	13	7	32	17	1535	66,71
Pamuk	5	136	1	5	13	6131	38	0	1	6330	96,86
Üzüm	465	515	3	752	112	458	13696	118	3350	19469	70,35
Yonca	31	5	6	30	108	2	40	23	103	348	6,61
Zeytin	224	94	0	185	65	46	199	234	7866	8913	88,25
Toplam	11153	4677	2445	10180	3523	7054	14191	635	11815	65673	
Üretici Doğruluğu	85,17	66,28	69,45	81,78	29,07	86,92	96,51	3,62	66,68	Genel Doğruluk	%78,21
Dengelenmiş Doğruluk	89,00	82,61	84,46	89,49	64,12	93,29	92,65	51,56	82,32	Kappa	0,7395

En önemli 50 bant kullanarak XGBoost ve RO algoritmaları ile sınıflandırma sonucu elde edilen kullanıcı, üretici ve dengelenmiş doğruluk değerleri karşılaştırmalı olarak Tablo 21’de verilmiştir. Yine en önemli 70 bant kullanarak XGBoost ve RO algoritmaları ile sınıflandırma sonucu elde edilen kullanıcı, üretici ve dengelenmiş doğruluk değerleri karşılaştırmalı olarak Tablo 22’de verilmiştir.

Tablo 21. En Önemli 50 Bant Kullanarak XGBoost ve RO Algoritmaları İle Sınıflandırma Sonucu Üretici ve Kullanıcı Doğruluklarının Karşılaştırılması

%	Algoritma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Kullanıcı Doğruluğu	XGB-50	66,70	72,69	75,29	80,70	20,34	91,44	75,05	6,27	92,52
	RO-50	67,51	82,07	80,72	78,44	41,92	90,37	69,76	9,38	87,64
Üretici Doğruluğu	XGB-50	79,15	67,06	70,73	73,93	20,69	66,49	93,67	6,63	72,96
	RO-50	83,71	62,23	76,80	76,98	20,32	73,46	93,89	5,41	66,46
Dengelenmiş Doğruluk	XGB-50	85,57	82,57	84,92	85,33	58,09	82,87	92,43	52,83	85,84
	RO-50	87,74	80,60	88,05	86,56	59,38	86,26	91,23	52,45	82,20

Tablo 22. En Önemli 70 Bant Kullanarak XGBoost ve RO Algoritmaları İle Sınıflandırma Sonucu Üretici ve Kullanıcı Doğruluklarının Karşılaştırılması

%	Algoritma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Kullanıcı Doğruluğu	XGB-70	67,92	78,20	69,94	86,60	27,43	96,73	75,28	9,35	93,41
	RO-70	70,84	82,64	83,44	84,24	66,71	96,86	70,35	6,61	88,25
Üretici Doğruluğu	XGB-70	80,65	69,77	62,85	81,81	18,80	85,49	96,41	6,33	71,72
	RO-70	85,17	66,28	69,45	81,78	29,07	86,92	96,51	3,62	66,58
Dengelenmiş Doğruluk	XGB-70	86,47	84,14	80,91	89,74	58,02	92,57	93,73	52,86	85,31
	RO-70	89,00	82,61	84,46	89,49	64,12	93,29	92,65	51,56	82,32

4. SONUÇLAR VE ÖNERİLER

- Bu çalışmada elde edilen sonuçlara göre, XGBoost algoritması ile elde edilen genel doğruluk ve kappa değerleri (%83,67, 0,8059), RO algoritması ile elde edilen doğruluk değerlerine göre (%83,55, 0,8036), düşük oranda da olsa, daha yüksek çıkmıştır.

- Yalnız Sentinel-1 görüntüleri kullanılarak yapılan sınıflandırma işleminde her iki algoritmanın genel doğruluk değerleri %83'ün üzerinde çıkmıştır. Sentinel-1 görüntülerinin gece, gündüz ve kötü hava şartlarında alınabilmesi her döneme ait görüntüleri elde etme imkânı sağlamaktadır. Sentinel-1 görüntülerinin, tarımsal ürün tespiti amaçlı sınıflandırma uygulamalarında istenilen doğruluk değerlerine ulaşabildiği değerlendirilmektedir.

- Hem XGBoost hem de RO sınıflandırma sonuçlarına göre, pamuk en yüksek kullanıcı doğruluğu, üretici doğruluğu ve dengelenmiş doğruluk değerlerini, yonca en düşük kullanıcı doğruluğu, üretici doğruluğu ve dengelenmiş doğruluk değerlerini sağlamıştır. Sınıflar arası karışmaların her iki algoritma için aynı ürünlerde fakat farklı oranlarda olduğu görülmüştür. Genel olarak, buğdayın pamuk ve yonca haricindeki ürünlerle, domatesin mera ve yonca haricindeki ürünlerle, meranın buğday ve zeytinle, mısırın mera, pamuk ve yonca haricindeki ürünlerle, mısır2nin pamuk haricindeki ürünlerle, pamuğun buğday, domates ve üzümle, üzümün zeytinle, yoncanın domates,

mısır² ve pamuk haricindeki ürünlerle, zeytinin ise buğday, üzüm ve yoncayla karıştırıldığı görülmüştür.

- Sınıflandırma sonuçlarının yanı sıra model oluşturma zamanlarının da önemli olduğu görülmüştür. Sınıflandırmada 152 bantlı veri yığının kullanıldığı bu çalışmada, XGBoost algoritması ile model oluşturulması için 25 dakika, RO algoritması ile model oluşturulması için 120 dakika süre gerekmiştir. XGBoost algoritması model oluştururken işlemci çekirdeklerinin paralel kullanılması imkânı sağlamaktadır. RO algoritması model oluştururken bu imkânı sağlamamaktadır. Daha büyük alanlar ve daha fazla görüntünün kullanılacağı uygulamalarda XGBoost algoritması RO algoritmasına göre önemli miktarda zaman tasarrufu sağlayacaktır.

- Bu çalışmada, XGBoost algoritması çok yaygın kullanılan RO algoritmasına göre, düşük oranda da olsa, daha iyi sonuçlar vermiştir. Bu da XGBoost algoritmasının Sentinel-1 görüntülerinden tarımsal ürün tespiti çalışmalarında tercih edilebilecek bir algoritma olduğunu göstermektedir. Ancak, daha kesin bir sonuca ulaşmak için benzer çalışmanın farklı birkaç alanda daha uygulanması gerekir.

- Bant önem derecelerine göre sınıflandırma, genel doğruluk ve kappa değerlerinde çok az oranda düşüş olmasına rağmen, makine öğrenme algoritmalarından sınıflandırma modeli oluşturma sürelerinde önemli miktarda zaman tasarrufu sağlamaktadır.

- En önemli 50 bant kullanarak XGBoost ve RO algoritmaları ile sınıflandırma sonuçları için hesaplanmış genel doğruluk değeri %74'ün üzerinde, en önemli 70 bant kullanarak XGBoost ve RO algoritmaları ile sınıflandırma sonuçları için hesaplanmış genel doğruluk değeri %77'nin üzerinde çıkmıştır.

- Tarımsal ürünlerin uydu görüntülerinden sınıflandırmasında ürünlerin ekim, büyüme, yaprak dökümü ve hasat dönemleri önem arz etmektedir. Bu fenolojik dönemlere göre belirleyici özelliklerdeki tarihler dikkate alınarak uydu görüntülerini seçmek ürünlerdeki zamansal değişimlerin sınıflandırmaya olan olumlu etkisini artıracaktır.

- Tarım ürünlerinin birbirinden ayırt edilebilmesi için ekimden sonraki büyüme ve hasat süreçleri çok önemlidir. Ürünlerin ekim sonrası büyüme ve gelişme sürecinde gösterdikleri değişimler geri yansıma değerlerini etkilemektedir. Ancak, geri yansıma değerlerindeki bu değişimler ürünlerin birbirleri ile karıştırılmasını önlemekte ve sınıflandırma sonuçlarını olumlu etkilemektedir. Özetle, tarımsal ürün tespitinde çoklu

tarihli görüntülerin birlikte kullanılmasının sınıflandırma doğruluk değerlerine olumlu etkisi vardır. Ancak, çoklu tarihli görüntülerin veri yığını olarak sınıflandırmaya dâhil edilmesi, makine öğrenme algoritmalarından sınıflandırma modeli oluşturma süresini olumsuz yönde etkilemektedir.

- ÇKS verileri tarım alanlarında üretim yapan çiftçilerin beyanlarından oluşturulmaktadır. Gerek çiftçiler tarafından verilen bilgilerde gerekse sisteme aktarılması sırasında yanlışlıklar olabilmektedir. Dolayısıyla, bu verilerin sınıflandırmada referans veriler olarak kullanılabilmesi için öncelikle analiz edilmesi ve hataların ayıklanması ve/veya düzeltilmesi gerekir. Diğer taraftan, hatalı ÇKS verileri, örn. parseldeki ürün türü, sınıflandırma sonuçlarına göre tespit edilip düzeltilebilir.

- XGBoost algoritması ile Sentinel-1 görüntüleri kullanılarak farklı ürünler ve zamansal değişimlerin olduğu çalışma alanlarına ait veri setleri ile sınıflandırma sonuçlarının karşılaştırması önerilmektedir.

- XGBoost algoritması kullanılarak farklı tip uydu görüntüleri kullanılarak tarım ürünleri sınıflandırma sonuçlarının karşılaştırması önerilmektedir.

- Sentinel-1 görüntüleri kullanılarak diğer makine öğrenme algoritmalarıyla tarım ürünlerinin tespit edilmesindeki sonuçların karşılaştırılması önerilmektedir.

- Sentinel-1 görüntülerinin yanında farklı radar görüntülerinin sınıflandırmaya dâhil edilmesinin sonuçları olumlu yönde etkileyeceği ve sınıflandırma doğruluk değerlerini yükselteceği değerlendirilmektedir.

- Radar görüntülerinin yanında sınıflandırmada radar bitki örtüsü indeksi tutarlılık verilerinin de kullanılmasının sonuçları olumlu yönde etkileyeceği düşünülmektedir.

- Sentinel-1 görüntüleri sınıflandırması ile farklı fenolojik özelliklere ve coğrafi bölgelere ait tarım ürünlerinin tespitinin yapılması önerilmektedir.

KAYNAKLAR

- [1] M. Turker and M. Arikan, “Sequential masking classification of multi-temporal Landsat7 ETM+ images for field-based crop mapping in Karacabey, Turkey,” *Int. J. Remote Sens.*, vol. 26, no. 17, pp. 3813–3830, 2005.
- [2] M. Turker and A. Ozdarici, “Field-based crop classification using SPOT4, SPOT5, IKONOS and QuickBird imagery for agricultural areas: A comparison study,” *Int.*

- J. Remote Sens.*, vol. 32, no. 24, pp. 9735–9768, 2011.
- [3] M. W. Liu, M. Ozdogan, and X. Zhu, “Crop type classification by simultaneous use of satellite images of different resolutions,” *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3637–3649, 2014.
 - [4] L. Viskovic, I. N. Kosovic, and T. Mastelic, “Crop classification using multi-spectral and multitemporal satellite imagery with machine learning,” in *2019 27th International Conference on Software, Telecommunications and Computer Networks, SoftCOM 2019*, 2019.
 - [5] G. Lemoine and O. Leo, “Crop mapping applications at scale: Using Google Earth Engine to enable global crop area and status monitoring using free and open data sources,” in *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2015, vol. 2015-Novem, pp. 1496–1499.
 - [6] U. Lussem, C. Hütt, and G. Waldhoff, “Combined analysis of Sentinel-1 and RapidEye data for improved crop type classification: An early season approach for rapeseed and cereals,” *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch.*, vol. 41, no. July, pp. 959–963, 2016.
 - [7] R. Nasirzadehdizaji, F. B. Sanli, Z. Cakir, and E. Sertel, “Crop mapping improvement by combination of optical and SAR datasets,” in *2019 8th International Conference on Agro-Geoinformatics, Agro-Geoinformatics 2019*, 2019.
 - [8] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, “Deep Learning Classification of Land Cover and Crop Types Using Remote Sensing Data,” *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 778–782, 2017.
 - [9] T. F. Bush and F. T. Ulaby, “An evaluation of radar as a crop classifier,” *Remote Sens. Environ.*, vol. 7, no. 1, pp. 15–36, 1978.
 - [10] P. Hoogeboom, “Classification of Agricultural Crops in Radar Images,” *IEEE Trans. Geosci. Remote Sens.*, vol. GE-21, no. 3, pp. 329–336, 1983.
 - [11] D. Bargiel, “A new method for crop classification combining time series of radar images and crop phenology information,” *Remote Sens. Environ.*, vol. 198, pp. 369–383, 2017.
 - [12] P. Taylor, G. M. Foody, M. B. McCulloch, and W. B. Yates, “International Journal

- of Remote Sensing Crop classification from C-band polarimetric radar data,” no. October 2012, pp. 37–41, 2007.
- [13] C. Hütt, W. Koppe, Y. Miao, and G. Bareth, “Best accuracy land use/land cover (LULC) classification to derive crop types using multitemporal, multisensor, and multi-polarization SAR satellite images,” *Remote Sens.*, vol. 8, no. 8, 2016.
 - [14] H. McNairn and J. Shang, “A review of multitemporal synthetic aperture radar (SAR) for crop monitoring,” *Remote Sens. Digit. Image Process.*, vol. 20, pp. 317–340, 2016.
 - [15] H. Skriver, “Crop classification by multitemporal C- and L-band single- and dual-polarization and fully polarimetric SAR,” *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 6, pp. 2138–2149, 2012.
 - [16] R. Sonobe, H. Tani, X. Wang, N. Kobayashi, and H. Shimamura, “Random forest classification of crop type using multioral TerraSAR-X dual-polarimetric data,” *Remote Sens. Lett.*, vol. 5, no. 2, pp. 157–164, 2014.
 - [17] O. Csillik and M. Belgiu, “Cropland Mapping from Sentinel-2 Time Series Data Using Object-based Image Analysis,” *20th Agil. Conf. Geogr. Inf. Sci.*, vol. 20, no. June 2019, p. 5, 2017.
 - [18] M. Drusch *et al.*, “Sentinel-2: ESA’s Optical High-Resolution Mission for GMES Operational Services,” *Remote Sens. Environ.*, vol. 120, pp. 25–36, 2012.
 - [19] T. Whelen and P. Siqueira, “Time-series classification of Sentinel-1 agricultural data over North Dakota,” *Remote Sens. Lett.*, vol. 9, no. 5, pp. 411–420, 2018.
 - [20] X. Song *et al.*, “Evaluation of Sentinel-2A satellite imagery for mapping cotton root rot,” *Remote Sens.*, vol. 9, no. 9, pp. 1–17, 2017.
 - [21] R. Sonobe, Y. Yamaya, H. Tani, X. Wang, N. Kobayashi, and K. ichiro Mochizuki, “Assessing the suitability of data from Sentinel-1A and 2A for crop classification,” *GIScience Remote Sens.*, vol. 54, no. 6, pp. 918–938, 2017.
 - [22] F. Vuolo, M. Neuwirth, M. Immitzer, C. Atzberger, and W. T. Ng, “How much does multi-temporal Sentinel-2 data improve crop type classification?,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 72, no. April, pp. 122–130, 2018.
 - [23] G. Satalino, A. Balenzano, F. Mattia, and M. Davidson, “Sentinel-1 SAR data for

- mapping agricultural crops not dominated by volume scattering,” in *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2012, pp. 6801–6804.
- [24] B. Brisco and R. J. Brown, “Multidate SAR/TM synergism for crop classification in western Canada,” *Photogramm. Eng. Remote Sensing*, vol. 61, no. 8, pp. 1009–1014, 1995.
- [25] S. Siachalou, G. Mallinis, and M. Tsakiri-Strati, “A hidden markov models approach for crop classification: Linking crop phenology to time series of multi-sensor remote sensing data,” *Remote Sens.*, vol. 7, no. 4, pp. 3633–3650, 2015.
- [26] T. G. Van Niel and T. R. McVicar, “Determining temporal windows for crop discrimination with remote sensing: A case study in south-eastern Australia,” *Comput. Electron. Agric.*, vol. 45, no. 1–3, pp. 91–108, 2004.
- [27] M. Zhang, W. Shi, and Z. Xu, “Systematic comparison of five machine-learning models in classification and interpolation of soil particle size fractions using different transformed data,” *Hydrol. Earth Syst. Sci.*, vol. 24, no. 5, pp. 2505–2526, 2020.
- [28] K. Van Tricht, A. Gobin, S. Gilliams, and I. Piccard, “Synergistic use of radar sentinel-1 and optical sentinel-2 imagery for crop mapping: A case study for Belgium,” *Remote Sens.*, vol. 10, no. 10, 2018.
- [29] A. O. Onojeghuo, G. A. Blackburn, Q. Wang, P. M. Atkinson, D. Kindred, and Y. Miao, “Mapping paddy rice fields by applying machine learning algorithms to multi-temporal sentinel-1A and landsat data,” *Int. J. Remote Sens.*, vol. 39, no. 4, pp. 1042–1067, 2018.
- [30] Ö. G. NARİN, A. DELEN, and S. ABDİKAN, “Sentinel-2a Verisi Kullanarak Tarımsal Ürün Deseninin Belirlenmesi,” 2018.
- [31] J. D. B. Castro, R. Q. Feitoza, L. C. La Rosa, P. M. A. Diaz, and I. D. A. Sanches, “A Comparative Analysis of Deep Learning Techniques for Sub-Tropical Crop Types Recognition from Multitemporal Optical/SAR Image Sequences,” in *Proceedings - 30th Conference on Graphics, Patterns and Images, SIBGRAPI 2017*, 2017, pp. 382–389.
- [32] M. Kumari, C. S. Murthy, V. Pandey, and G. D. Bairagi, “Soybean cropland mapping using multi-temporal sentinel-1 data,” in *International Archives of the*

Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives, 2019, vol. 42, no. 3/W6, pp. 109–114.

- [33] S. Niculescu, D. Ienco, and J. Hanganu, “Application of deep learning of multi-temporal Sentinel-1 images for the classification of coastal vegetation zone of the danube delta,” in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 2018, vol. 42, no. 3, pp. 1311–1318.
- [34] H. Sahu, D. Haldar, A. Danodia, and S. Kumar, “Classification of orchard crop using sentinel-1A synthetic aperture radar data,” in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 2018, vol. 42, no. 5, pp. 335–338.
- [35] L. Di. Robertson *et al.*, “Using Dense Time-Series of C-Band Sar Imagery for Classification of Diverse, Worldwide Agricultural Systems,” in *International Geoscience and Remote Sensing Symposium (IGARSS)*, 2019, pp. 6231–6234.
- [36] D. Dobrinić, D. Medak, and M. Gašparović, “Integration of multitemporal sentinel-1 and sentinel-2 imagery for land-cover classification using machine learning methods,” in *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*, 2020, vol. 43, no. B1, pp. 91–98.
- [37] M. Gašparović and D. Dobrinić, “Comparative assessment of machine learning methods for urban vegetation mapping using multitemporal Sentinel-1 imagery,” *Remote Sens.*, vol. 12, no. 12, 2020.
- [38] N. T. Son, C. F. Chen, C. R. Chen, and V. Q. Minh, “Assessment of Sentinel-1A data for rice crop classification using random forests and support vector machines,” *Geocarto Int.*, vol. 33, no. 6, pp. 587–601, 2018.
- [39] R. Nasirzadehdizaji, F. B. Sanli, S. Abdikan, Z. Cakir, A. Sekertekin, and M. Ustuner, “Sensitivity analysis of multi-temporal Sentinel-1 SAR parameters to crop height and canopy coverage,” *Appl. Sci.*, vol. 9, no. 4, 2019.
- [40] M. Üstüner and F. Balık Şanlı, “Çok zamanlı polarimetrik SAR verileri ile tarımsal ürünlerin sınıflandırılması,” *J. Geod. Geoinf.*, 2019.
- [41] M. Ustuner, F. B. Sanli, G. Bilgin, and S. Abdikan, “Land use and cover

- classification of Sentinel-1A SAR imagery: A case study of Istanbul,” in *2017 25th Signal Processing and Communications Applications Conference, SIU 2017*, 2017, pp. 1–4.
- [42] European Space Agency (ESA), “Copernicus Open Access Hub,” *Copernicus Open Access Hub (previously known as Sentinels Scientific Data Hub)*, 2020. [Online]. Available: <https://scihub.copernicus.eu/>.
- [43] O. Şimşek, Y. Nadaroğlu, G. Yücel, Ö. Dokuyucu, and Ş. Gökdağ, “Türkiye Fenoloji Atlası,” in *Türkiye Fenoloji Atlası*, 2014, pp. 18–103.
- [44] M. Yilmaz, “Görüntülerinden Rastgele Orman Algoritması ile Nesne-Tabanlı Sınıflandırması Object-Based Classification of Summer Crops From Sentinel-2 Satellite Images Using Random Forest Algorithm,” 2018.
- [45] T. Dizdaroğlu, “Comparison of Machine Learning Algorithms for Agricultural Crop Classification From Sentinel-1 and Sentinel-2,” 2019.
- [46] European Commission and ESA, “Sentinel-5 Precursor: DATA ACCESS AND PRODUCTS,” 2015. [Online]. Available: <https://sentinel.esa.int/web/sentinel/sentinel-5p-tropomi-wiki>.
- [47] European Space Agency, “Sentinel-1 SAR User Guide,” *Esa*, 2020. [Online]. Available: <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-1-sar>.
- [48] Y. Zhong, “İ» ß²¿’§-·- ±° Ý¿-»- Þ¿-»¼ ±² Ü»½-·±² İ®»»,” 2016.
- [49] R. Mitchell and E. Frank, “Accelerating the XGBoost algorithm using GPU computing,” *PeerJ Comput. Sci.*, vol. 2017, no. 7, 2017.
- [50] D. M. Farid, G. M. Maruf, and C. M. Rahman, “A new approach of Boosting using decision tree classifier for classifying noisy data,” *2013 Int. Conf. Informatics, Electron. Vision, ICIEV 2013*, pp. 0–3, 2013.
- [51] E. J. Ong and R. Bowden, “A boosted classifier tree for hand shape detection,” *Proc. - Sixth IEEE Int. Conf. Autom. Face Gesture Recognit.*, pp. 889–894, 2004.
- [52] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” *Proc. ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, vol. 13-17-Aug, pp. 785–794, 2016.
- [53] M. Luckner, B. Topolski, and M. Mazurek, “Application of XGboost algorithm in

- fingerprinting localisation task,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2017, vol. 10244 LNCS, pp. 661–671.
- [54] I. Babajide Mustapha and F. Saeed, “Bioactive Molecule Prediction Using Extreme Gradient Boosting,” *Molecules*, vol. 21, no. 8, 2016.
 - [55] I. L. Cherif and A. Kortebi, “On using eXtreme Gradient Boosting (XGBoost) Machine Learning algorithm for Home Network Traffic Classification,” *IFIP Wirel. Days*, vol. 2019-April, pp. 1–6, 2019.
 - [56] C. B. Escabias, “Tree Boosting Data Competitions with XGBoost,” 2017.
 - [57] M. S. Alam and S. T. Vuong, “Random forest classification for detecting android malware,” *Proc. - 2013 IEEE Int. Conf. Green Comput. Commun. IEEE Internet Things IEEE Cyber, Phys. Soc. Comput. GreenCom-iThings-CPSCOM 2013*, pp. 663–669, 2013.
 - [58] Y. Zhai and X. Zheng, “Random Forest based Traffic Classification Method in SDN,” *Int. Conf. Cloud Comput. Big Data Blockchain, ICCBB 2018*, pp. 5–9, 2018.
 - [59] A. Paul, D. P. Mukherjee, P. Das, A. Gangopadhyay, A. R. Chintla, and S. Kundu, “Improved Random Forest for Classification,” *IEEE Trans. Image Process.*, vol. 27, no. 8, pp. 4012–4024, 2018.
 - [60] S. V. Patel and V. N. Jokhakar, “A random forest based machine learning approach for mild steel defect diagnosis,” *2016 IEEE Int. Conf. Comput. Intell. Comput. Res. ICCIC 2016*, 2017.
 - [61] L. Breiman, “Random forests,” *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
 - [62] T. M. Oshiro, P. S. Perez, and J. A. Baranauskas, “How many trees in a random forest?,” in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2012, vol. 7376 LNAI, pp. 154–168.
 - [63] F. Filippini, “Sentinel-1 GRD Preprocessing Workflow,” *Proceedings*, vol. 18, no. 1, p. 11, 2019.
 - [64] J. S. Lee, I. Jurkevich, P. Dewaele, P. Wambacq, and A. Oosterlinck, “Speckle

- filtering of synthetic aperture radar images: a review,” *Remote Sens. Rev.*, vol. 8, no. 4, pp. 313–340, 1994.
- [65] S. Singh, D. Srivastava, and S. Agarwal, “GLCM and its application in pattern recognition,” in *5th International Symposium on Computational and Business Intelligence, ISCBI 2017*, 2017, pp. 20–25.
 - [66] G. A. S. Saroja and C. H. Sulochana, “Texture analysis of non-uniform images using GLCM,” in *2013 IEEE Conference on Information and Communication Technologies, ICT 2013*, 2013, pp. 1319–1322.
 - [67] A. Monadjemi, “Towards Efficient Texture Classification and Abnormality Detection,” University of Bristol, 2004.
 - [68] R. M. Haralick, K. Shanmugam, and I. Dinstein, “TexturalFeatures.pdf,” *IEEE Trans. Syst. Man Cybern.*, pp. 610–621, 1973.
 - [69] A. Ali, X. Jing, and N. Saleem, “GLCM-based fingerprint recognition algorithm,” in *Proceedings - 2011 4th IEEE International Conference on Broadband Network and Multimedia Technology, IC-BNMT 2011*, 2011, pp. 207–211.
 - [70] T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016. [Online]. Available: <https://github.com/dmlc/xgboost/issues>. [Accessed: 19-Dec-2020].
 - [71] A. Liaw and M. Wiener, “RandomForest,” *R news*, vol. XXXIX, no. 1, pp. 54.1–54.10, 2014.
 - [72] S. Adelabu, O. Mutanga, and E. Adam, “Evaluating the impact of red-edge band from Rapideye image for classifying insect defoliation levels,” *ISPRS J. Photogramm. Remote Sens.*, vol. 95, pp. 34–41, 2014.
 - [73] K. Millard and M. Richardson, “On the importance of training data sample selection in Random Forest image classification: A case study in peatland ecosystem mapping,” *Remote Sens.*, vol. 7, no. 7, pp. 8489–8515, 2015.
 - [74] I. Nitze, B. Barrett, and F. Cawkwell, “Temporal optimisation of image acquisition for land cover classification with random forest and MODIS time-series,” *Int. J. Appl. Earth Obs. Geoinf.*, vol. 34, no. 1, pp. 136–146, 2015.

- [75] R. R. Colditz, “An evaluation of different training sample allocation schemes for discrete and continuous land cover classification using decision tree-based algorithms,” *Remote Sens.*, vol. 7, no. 8, pp. 9655–9681, 2015.
- [76] S. E. Sesnie, B. Finegan, P. E. Gessler, A. M. S. Smith, R. B. Zayra, and S. Thessler, “The multispectral separability of Costa Rican rainforest types with support vector machines and Random Forest decision trees,” *Int. J. Remote Sens.*, vol. 31, no. 11, pp. 2885–2909, 2010.
- [77] M. Belgiu and L. Drăgu, “Random forest in remote sensing: A review of applications and future directions,” *ISPRS J. Photogramm. Remote Sens.*, vol. 114, pp. 24–31, Apr. 2016.
- [78] H. Guan, J. Li, M. Chapman, F. Deng, Z. Ji, and X. Yang, “Integration of orthoimagery and lidar data for object-based urban thematic mapping using random forests,” *Int. J. Remote Sens.*, vol. 34, no. 14, pp. 5166–5186, 2013.
- [79] R. L. Lawrence, S. D. Wood, and R. L. Sheley, “Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (randomForest),” *Remote Sens. Environ.*, vol. 100, no. 3, pp. 356–362, 2006.
- [80] P. O. Gislason, J. A. Benediktsson, and J. R. Sveinsson, “Random forests for land cover classification,” in *Pattern Recognition Letters*, 2006, vol. 27, no. 4, pp. 294–300.
- [81] M. Sheykhmousa, M. Mahdianpari, H. Ghanbari, F. Mohammadimanesh, P. Ghamisi, and S. Homayouni, “Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 6308–6325, 2020.
- [82] Ö. Akar, “Rastgele Orman Sınıflandırıcısına Doku Özellikleri Entegre Edilerek Benzer Spektral Özellikteki Tarımsal Ürünlerin Sınıflandırılması, Trabzon (Doktora Tezi , 2013),” Trabzon: Karadeniz Teknik Üniversitesi, 2013.
- [83] N. Deepa, M. Z. Khan, B. Prabadevi, D. R. P. M. Vincent, P. K. R. Maddikunta, and T. R. Gadekallu, “Multiclass model for agriculture development using multivariate statistical method,” *IEEE Access*, vol. 8, pp. 183749–183758, 2020.
- [84] E. Alpaydin, “Introduction to Machine Learning Third Edition,” *Introd. to Mach.*

Learn., 2010.

- [85] M. B. Kursa and W. R. Rudnicki, “Feature selection with the boruta package,” *J. Stat. Softw.*, 2010.
- [86] A. Bommert, X. Sun, B. Bischl, J. Rahnenführer, and M. Lang, “Benchmark for filter methods for feature selection in high-dimensional classification data,” *Comput. Stat. Data Anal.*, 2020.
- [87] A. Orynbaikyzy, U. Gessner, B. Mack, and C. Conrad, “Crop type classification using fusion of sentinel-1 and sentinel-2 data: Assessing the impact of feature selection, optical data availability, and parcel sizes on the accuracies,” *Remote Sens.*, 2020.



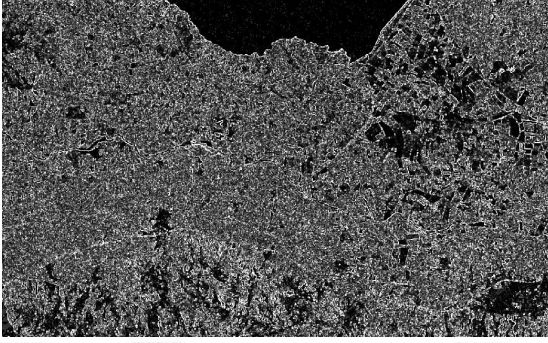
5.EKLER

EK 1 - Sentinel-1 DD/DY Doku Bantları

03 Mayıs 2017 Tarihli Sentinel-1 DY Görüntüsünden Hesaplanmış Doku Bantları

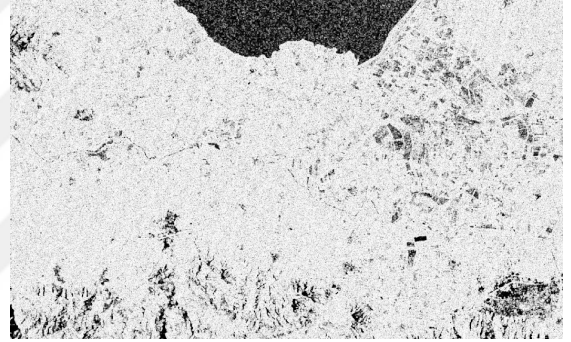
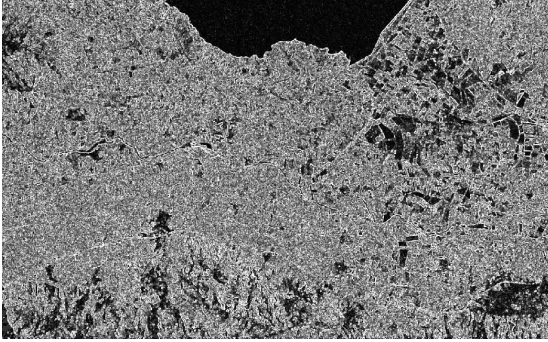
DY - Karşıtlık

DY - Homojenlik



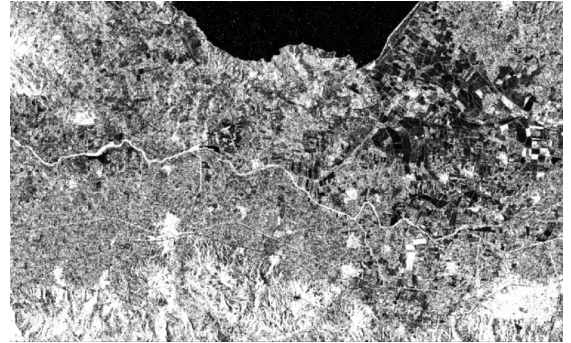
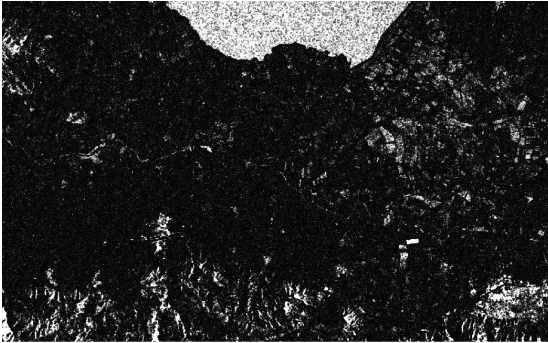
DY - Farklılık

DY - Entropi



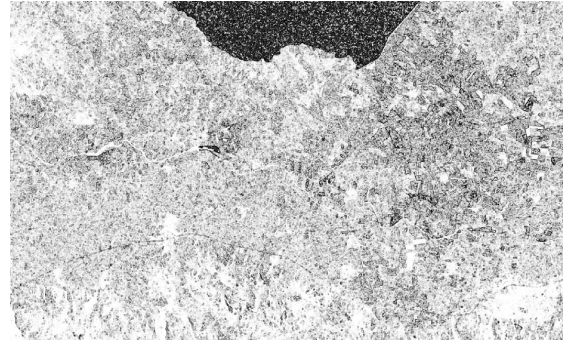
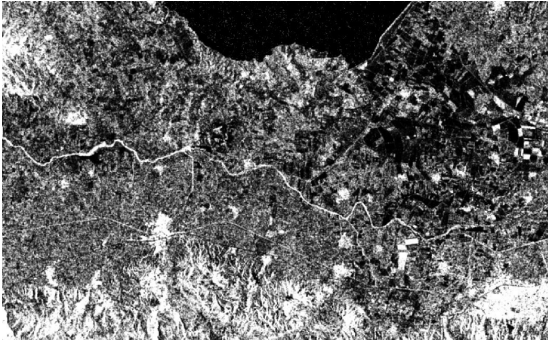
DY – Açısal İkinci Moment

DY - Ortalama



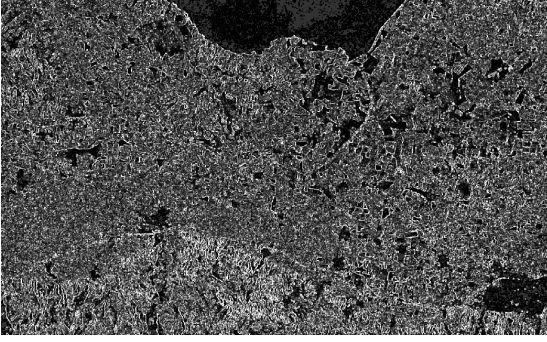
DY - Varyans

DY - Korelasyon



03 Mayıs 2017 Tarihli Sentinel-1 DD Görüntüsünden Hesaplanmış Doku Bantları

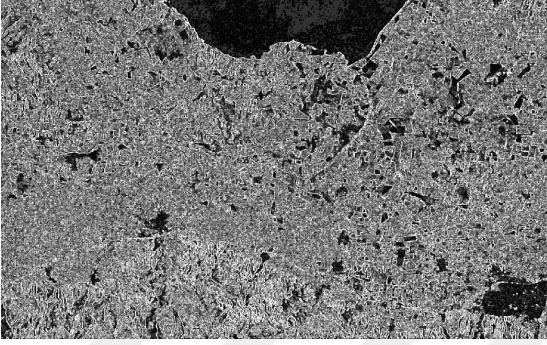
DD - Karşıtlık



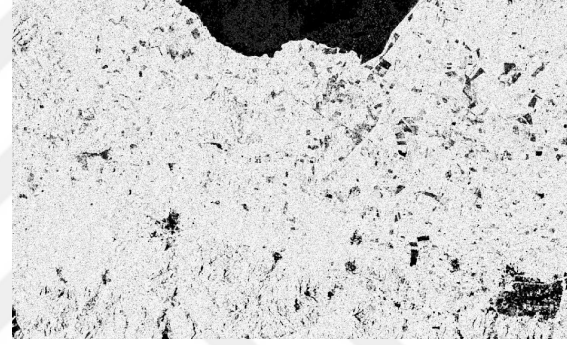
DD - Homojenlik



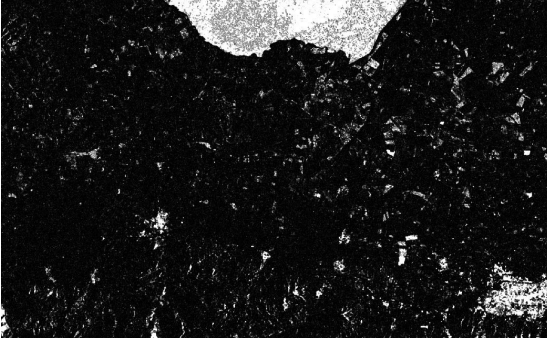
DD - Farklılık



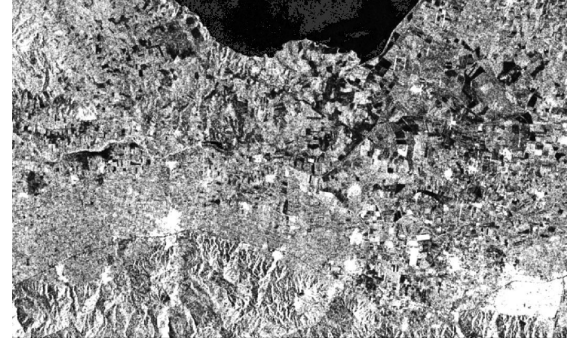
DD - Entropi



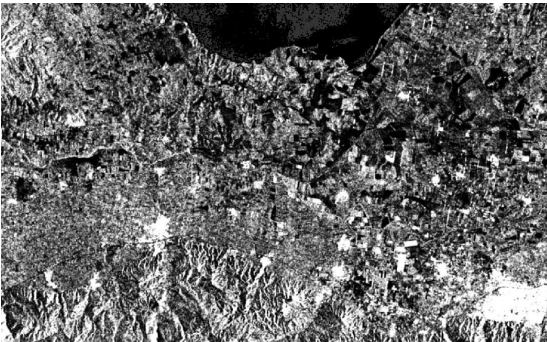
DD – Açısal İkinci Moment



DD - Ortalama



DD - Varyans

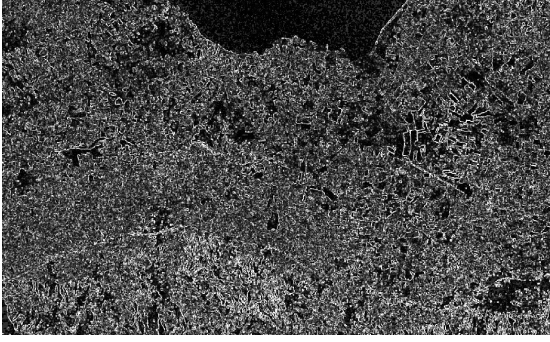


DD - Korelasyon



02 Haziran 2017 Tarihli Sentinel-1 DY Görüntüsünden Hesaplanmış Doku Bantları

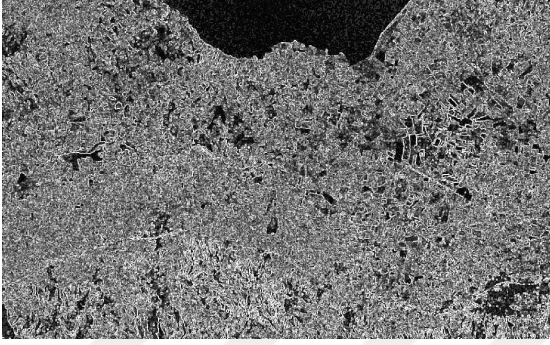
DY - Karşıtlık



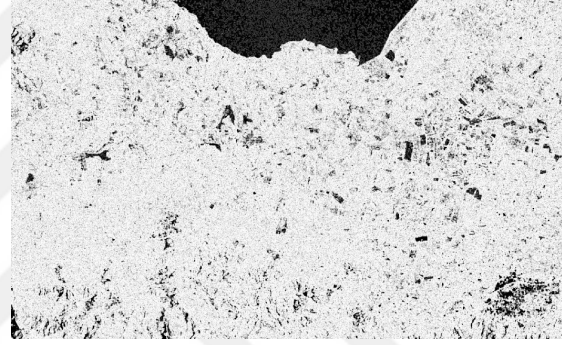
DY - Homojenlik



DY - Farklılık



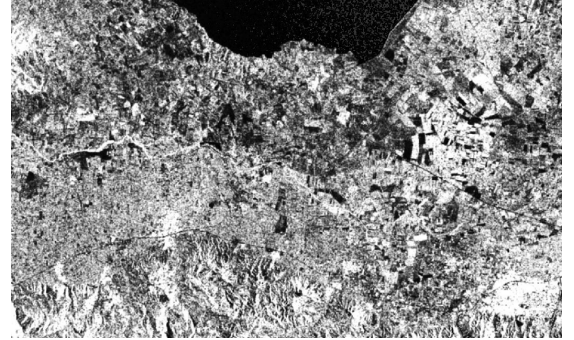
DY - Entropi



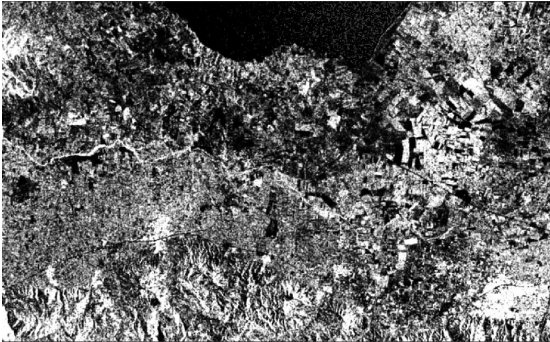
DY – Açısal İkinci Moment



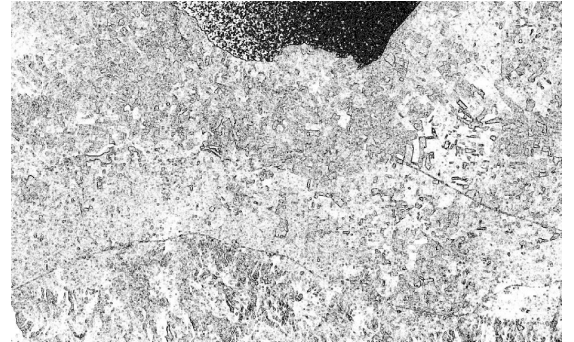
DY - Ortalama



DY - Varyans

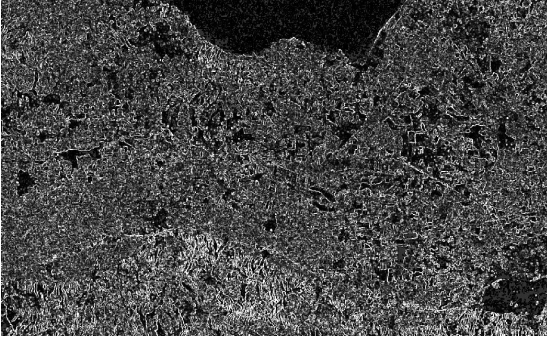


DY - Korelasyon



02 Haziran 2017 Sentinel-1 DD Görüntüsünden Hesaplanmış Doku Bantları

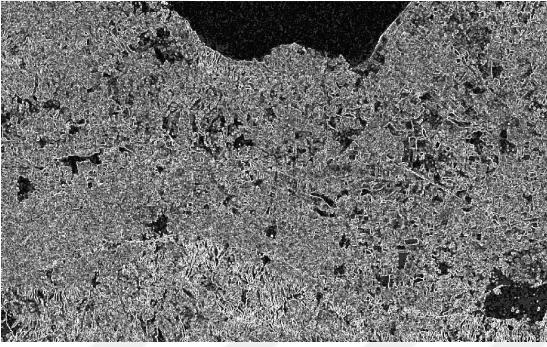
DD - Karşıtlık



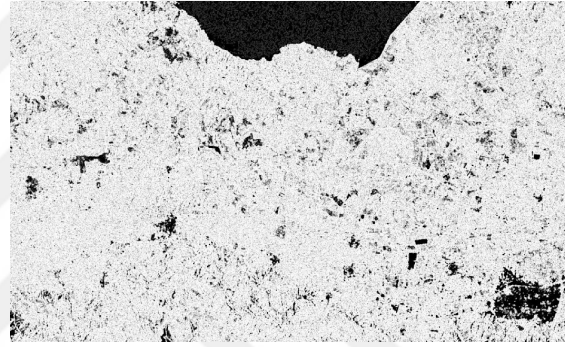
DD - Homojenlik



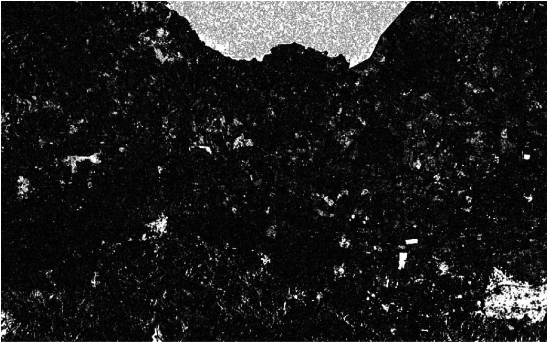
DD - Farklılık



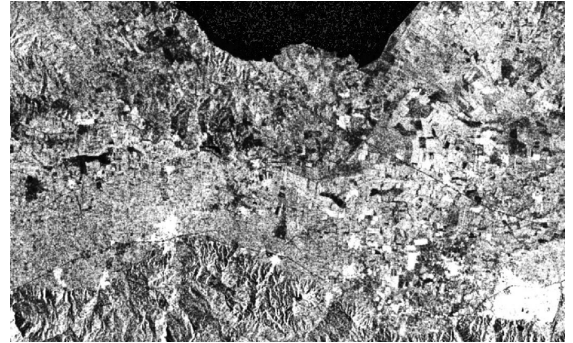
DD - Entropi



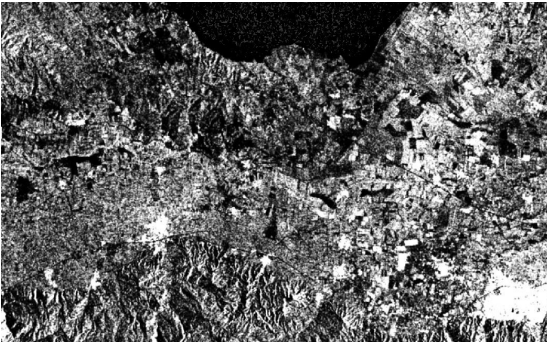
DD – Açısal İkinci Moment



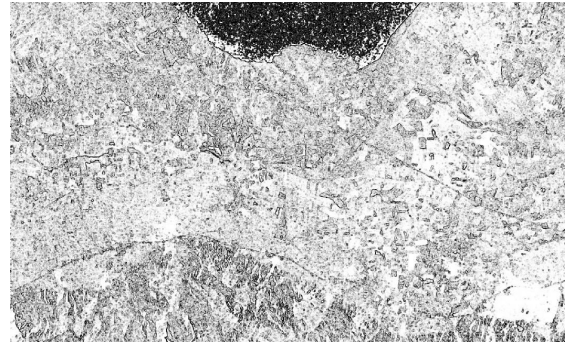
DD - Ortalama



DD - Varyans

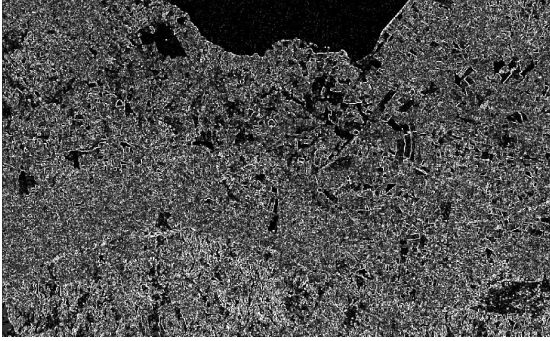


DD - Korelasyon



02 Temmuz 2017 Sentinel-1 DY Görüntüsünden Hesaplanmış Doku Bantları

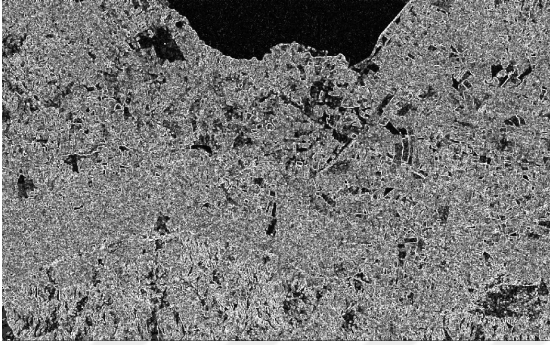
DY - Karşıtlık



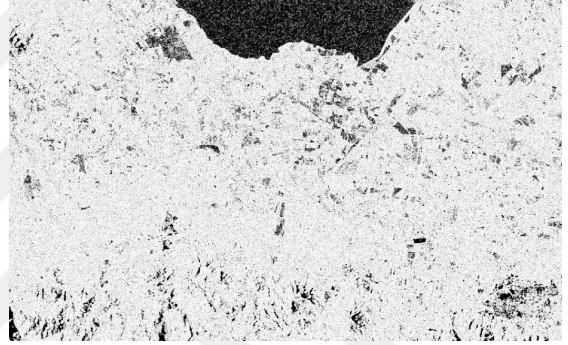
DY - Homojenlik



DY - Farklılık



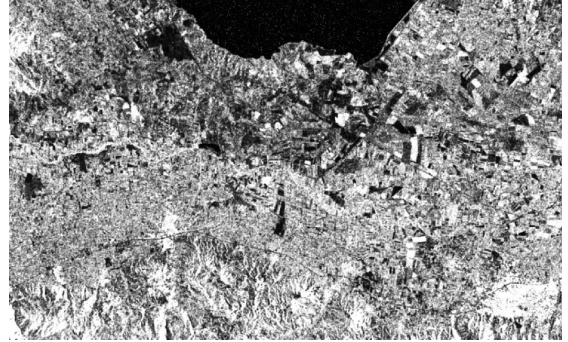
DY - Entropi



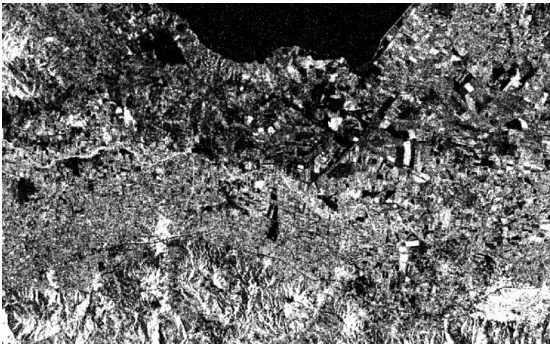
DY – Açısıl İkinci Moment



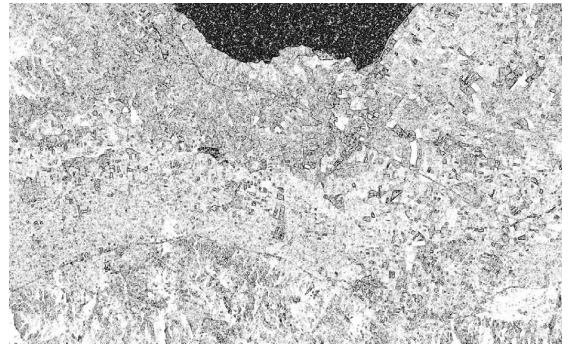
DY - Ortalama



DY - Varyans

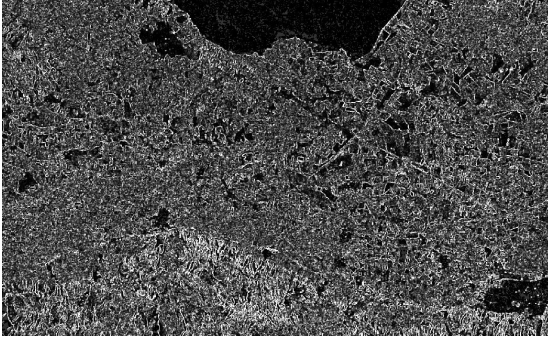


DY - Korelasyon



02 Temmuz 2017 Sentinel-1 DD Görüntüsünden Hesaplanmış Doku Bantları

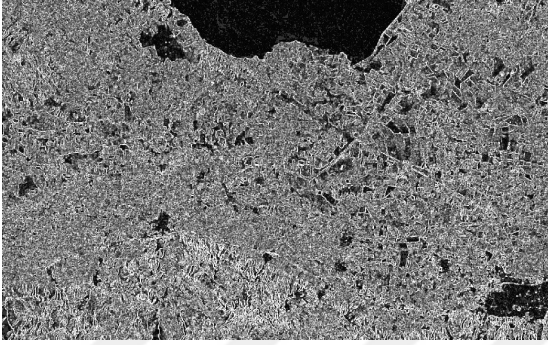
DD - Karşıtlık



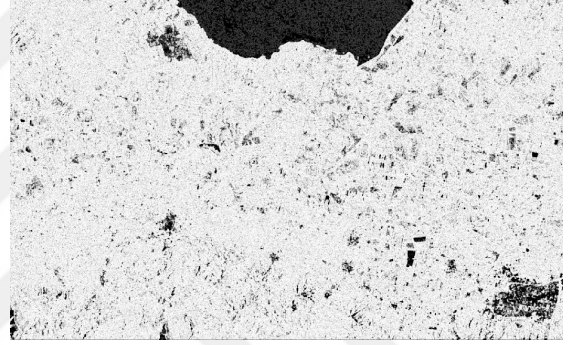
DD - Homojenlik



DD - Farklılık



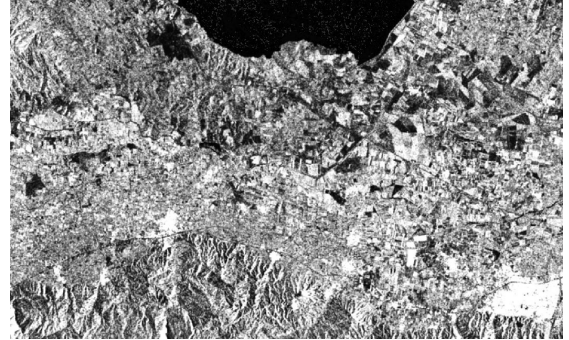
DD - Entropi



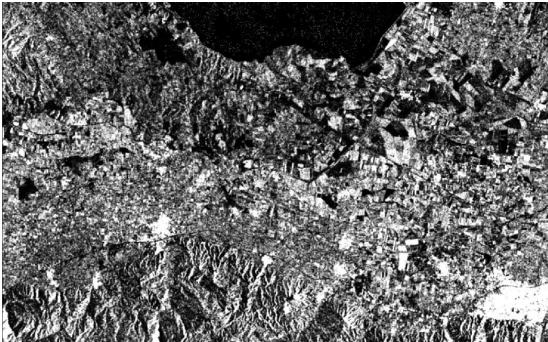
DD – Açısal İkinci Moment



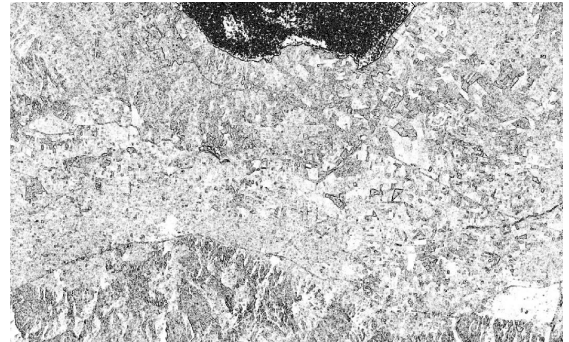
DD - Ortalama



DD - Varyans

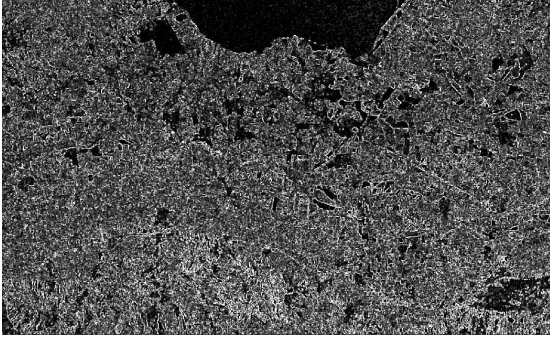


DD - Korelasyon



01 Ağustos 2017 Sentinel-1 DY Görüntüsünden Hesaplanmış Doku Bantları

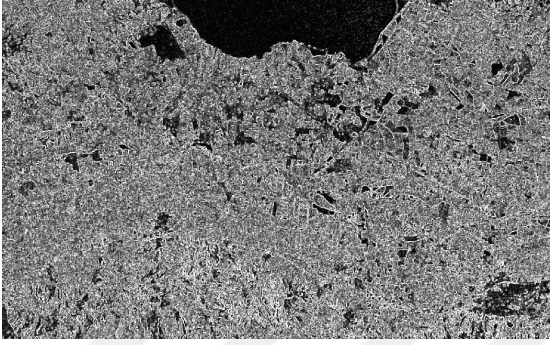
DY - Karşıtlık



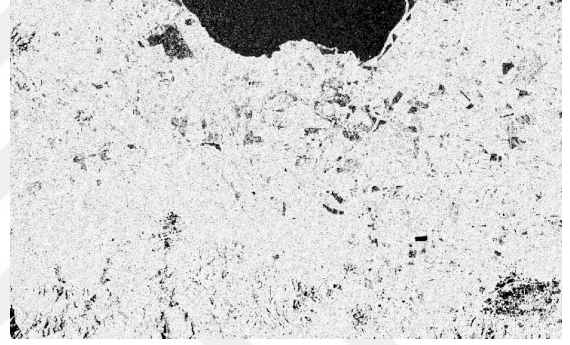
DY - Homojenlik



DY - Farklılık



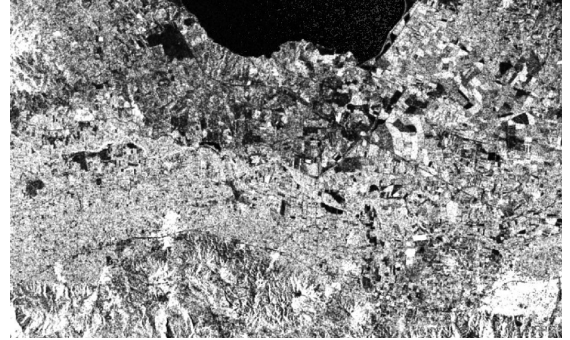
DY - Entropi



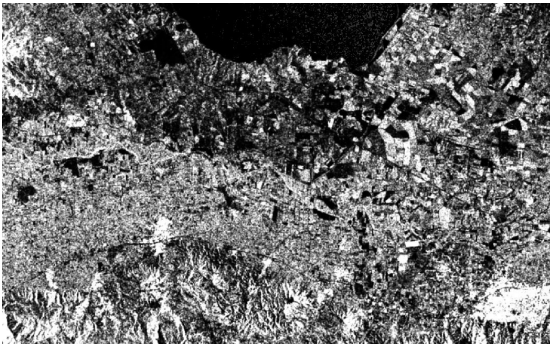
DY – Açısal İkinci Moment



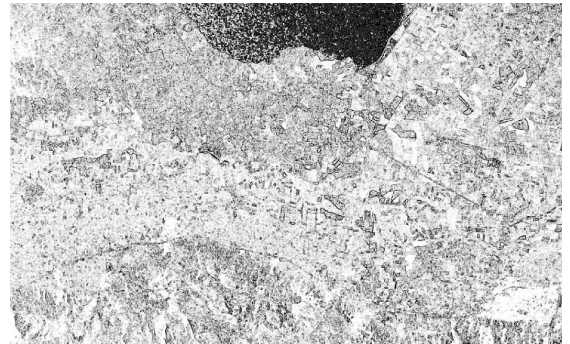
DY - Ortalama



DY - Varyans

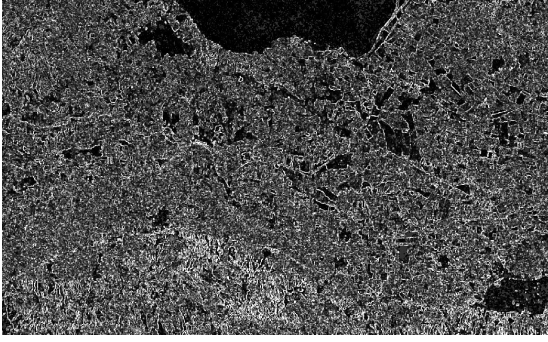


DY - Korelasyon



01 Ağustos 2017 Sentinel-1 DD Görüntüsünden Hesaplanmış Doku Bantları

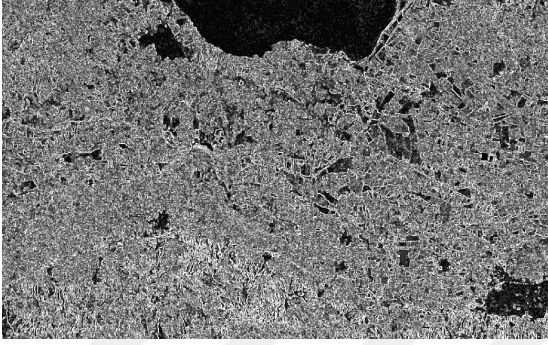
DD - Karşıtlık



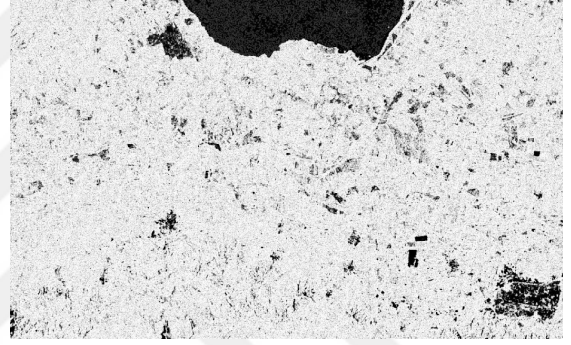
DD - Homojenlik



DD - Farklılık



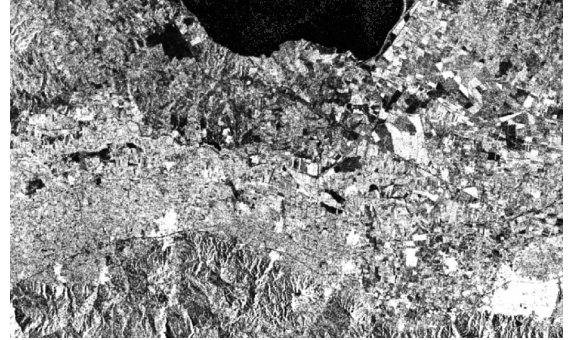
DD - Entropi



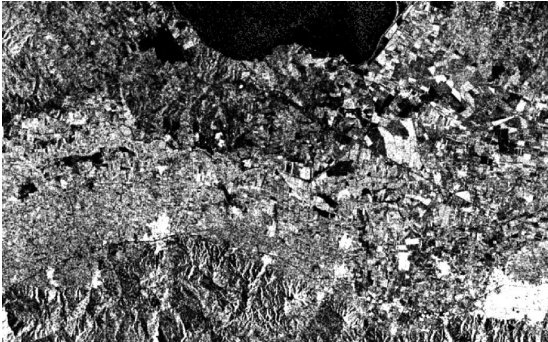
DD – Açısal İkinci Moment



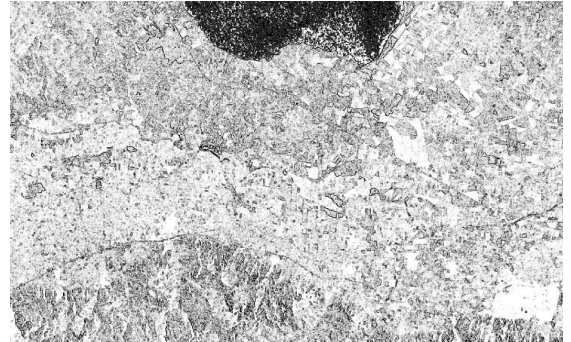
DD - Ortalama



DD - Varyans

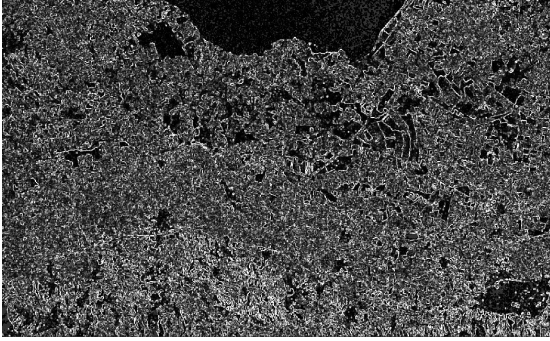


DD - Korelasyon



07 Eylül 2017 Sentinel-1 DY Görüntüsünden Hesaplanmış Doku Bantları

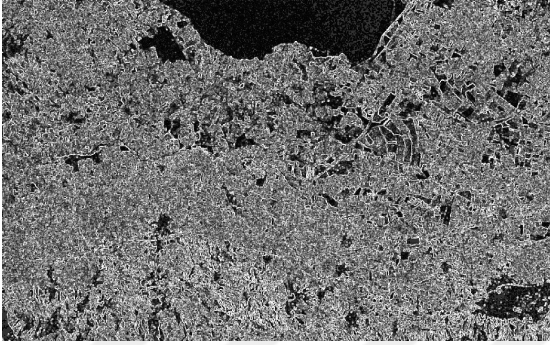
DY - Karşıtlık



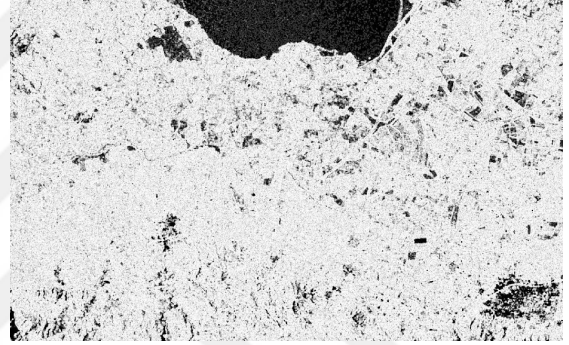
DY - Homojenlik



DY - Farklılık



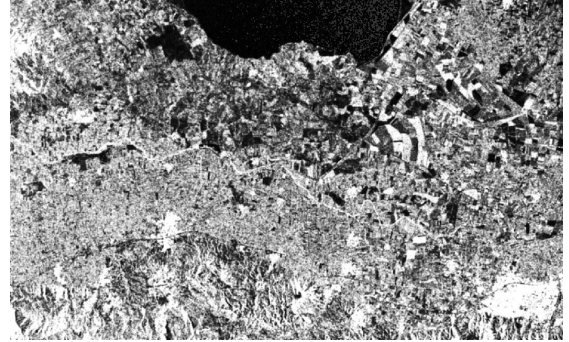
DY - Entropi



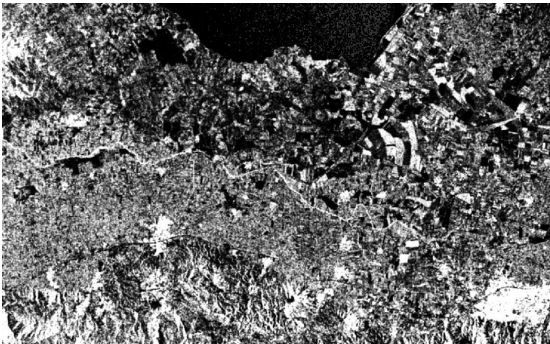
DY – Açısal İkinci Moment



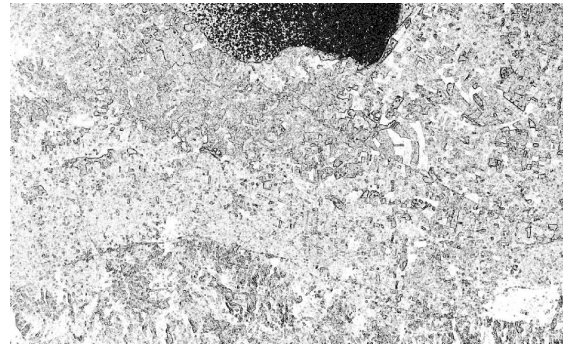
DY - Ortalama



DY - Varyans

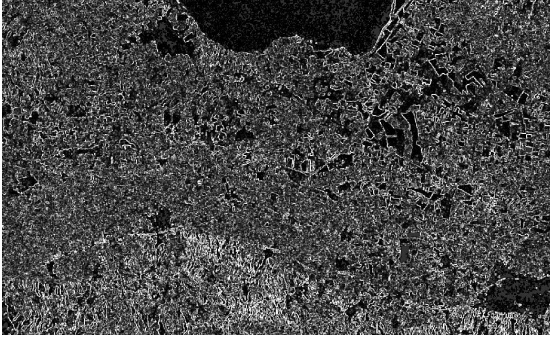


DY - Korelasyon



07 Eylül 2017 Sentinel-1 DD Görüntüsünden Hesaplanmış Doku Bantları

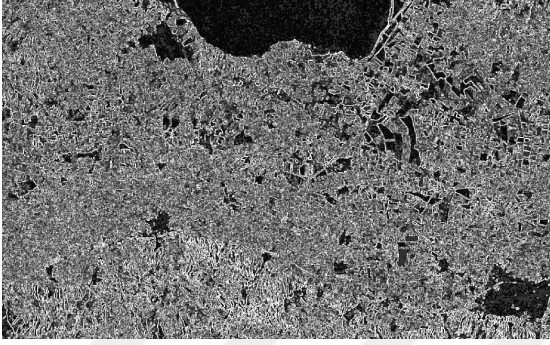
DD - Karşıtlık



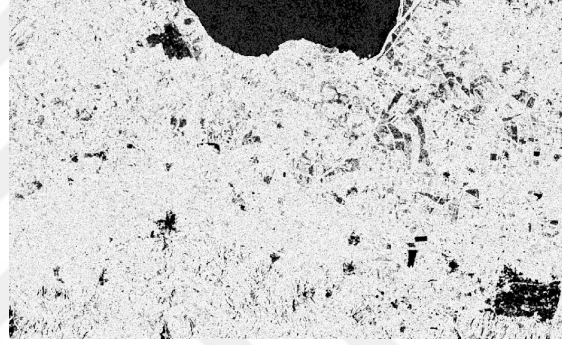
DD - Homojenlik



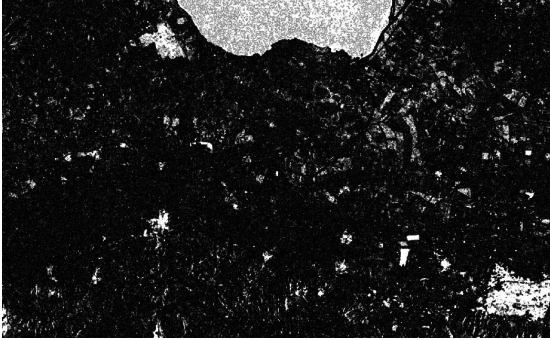
DD - Farklılık



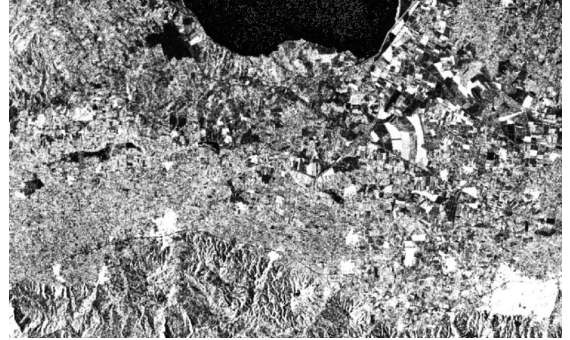
DD - Entropi



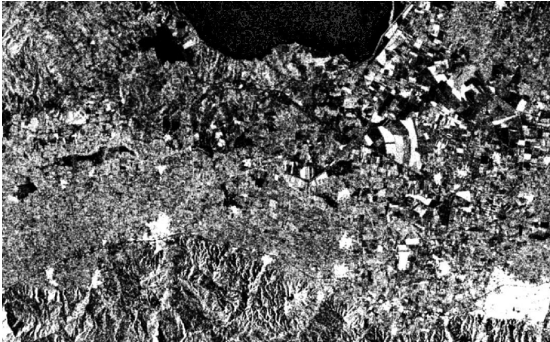
DD – Açısal İkinci Moment



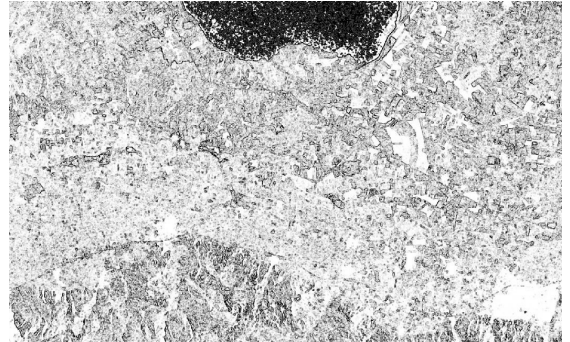
DD - Ortalama



DD - Varyans

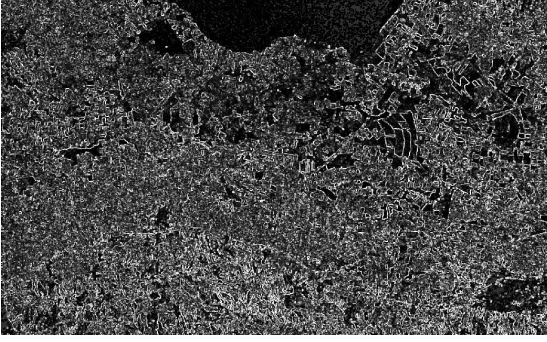


DD - Korelasyon



10 Ekim 2017 Sentinel-1 DY Görüntüsünden Hesaplanmış Doku Bantları

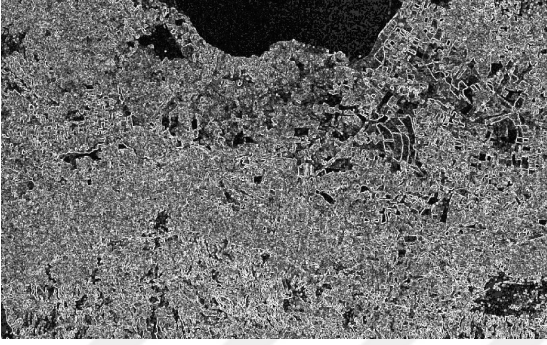
DY - Karşıtlık



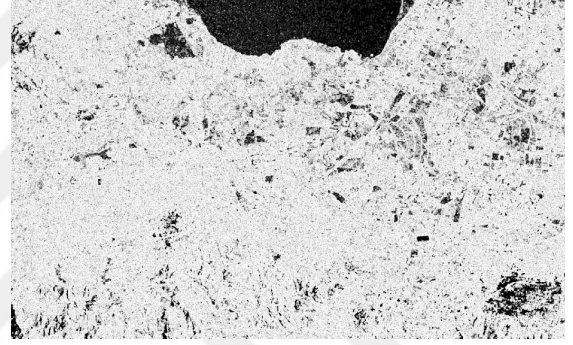
DY - Homojenlik



DY - Farklılık



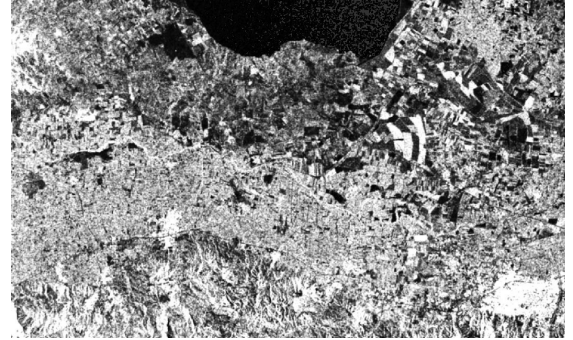
DY - Entropi



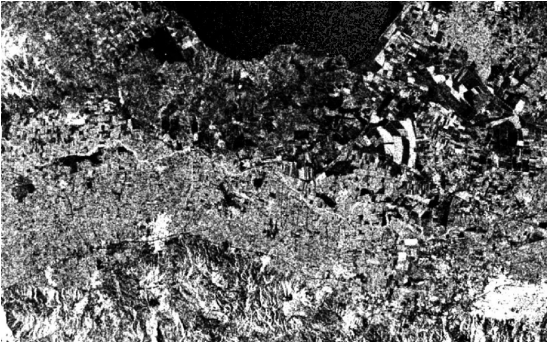
DY – Açısal İkinci Moment



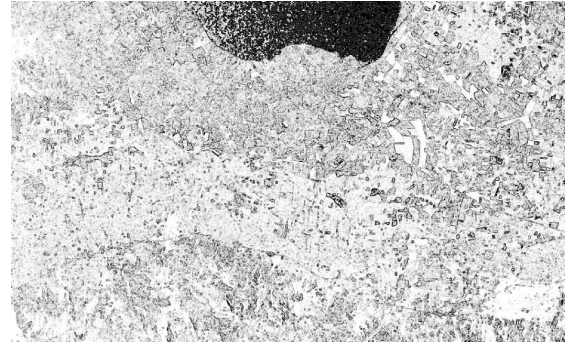
DY - Ortalama



DY - Varyans

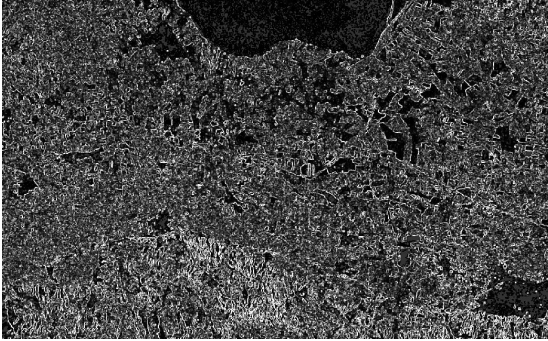


DY - Korelasyon

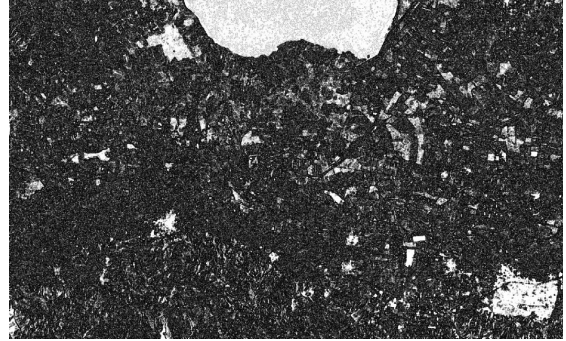


10 Ekim 2017 Sentinel-1 DD Görüntüsünden Hesaplanmış Doku Bantları

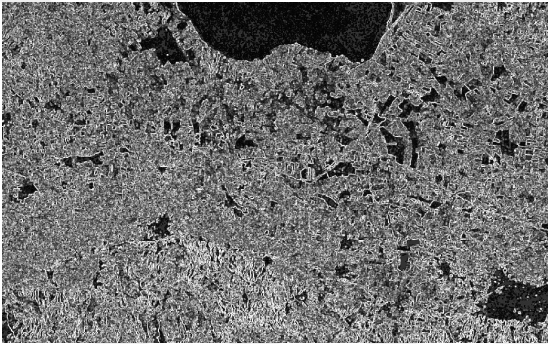
DD - Karşıtlık



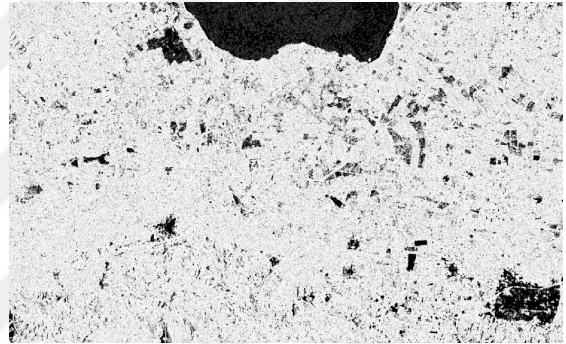
DD - Homojenlik



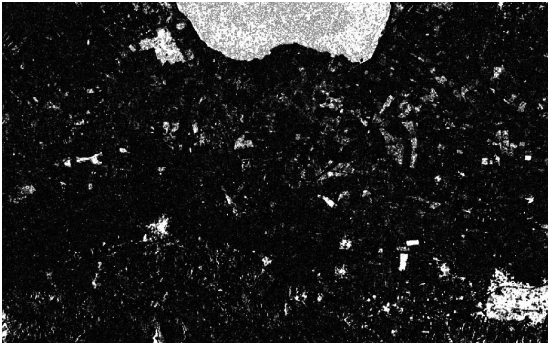
DD - Farklılık



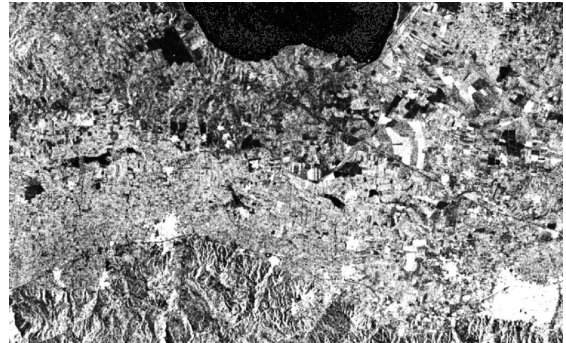
DD - Entropi



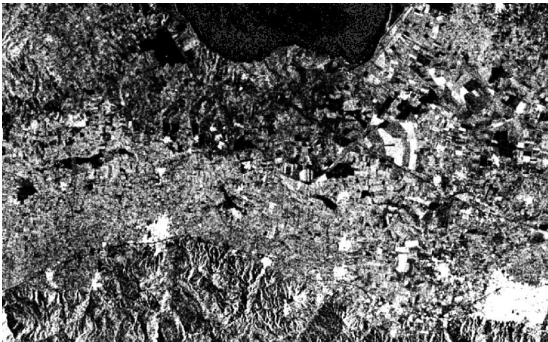
DD – Açısal İkinci Moment



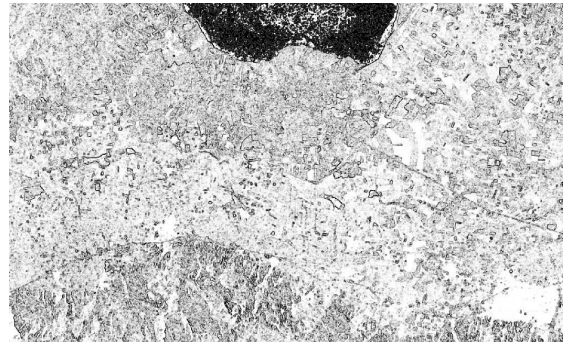
DD - Ortalama



DD - Varyans

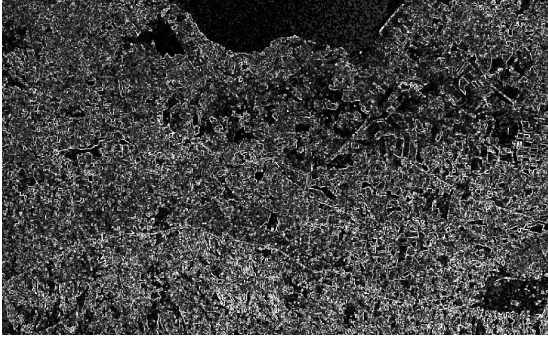


DD - Korelasyon



16 Kasım 2017 Sentinel-1 DY Görüntüsünden Hesaplanmış Doku Bantları

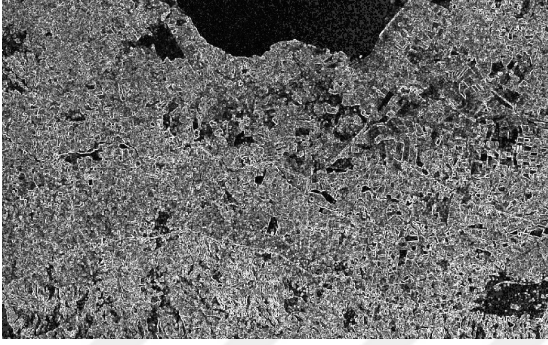
DY - Karşıtlık



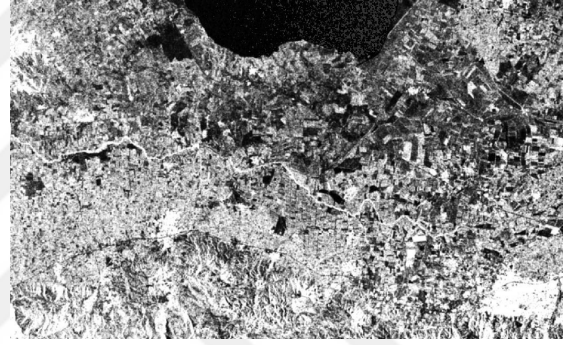
DY - Homojenlik



DY - Farklılık



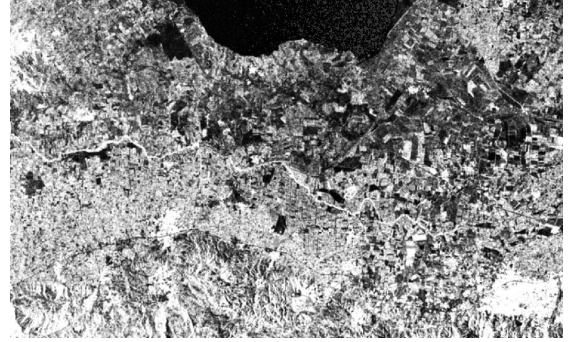
DY - Entropi



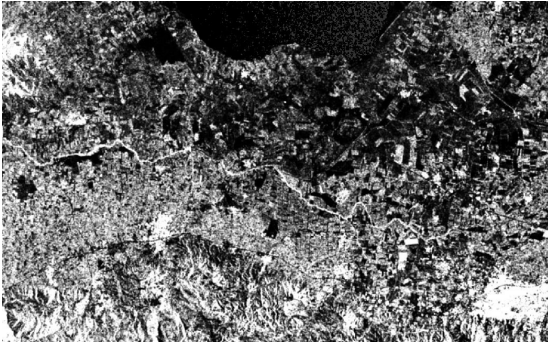
DY – Açısıl İkinci Moment



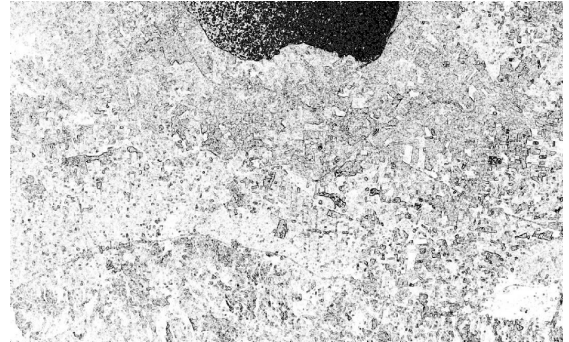
DY - Ortalama



DY - Varyans

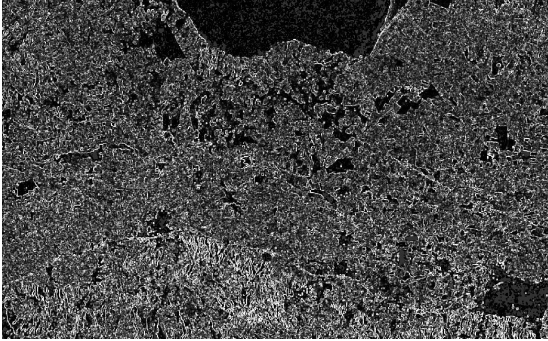


DY - Korelasyon

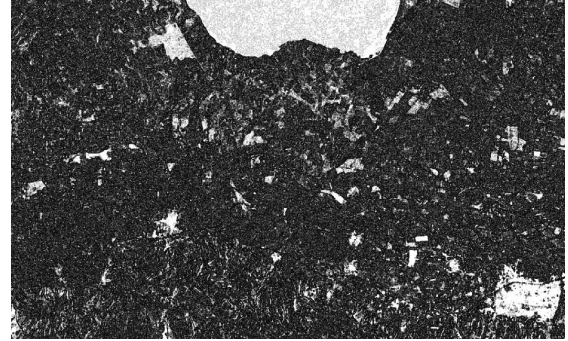


16 Kasım 2017 Sentinel-1 DD Görüntüsünden Hesaplanmış Doku Bantları

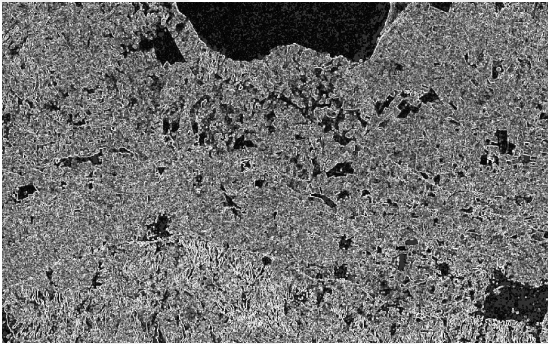
DD - Karşıtlık



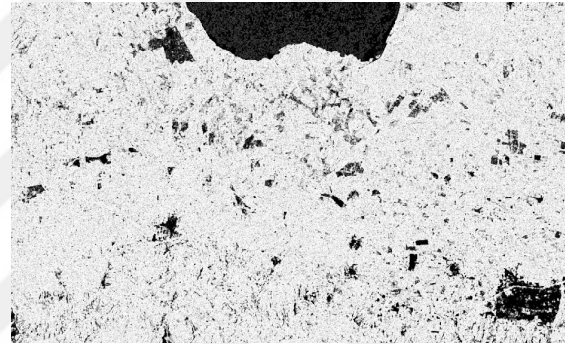
DD - Homojenlik



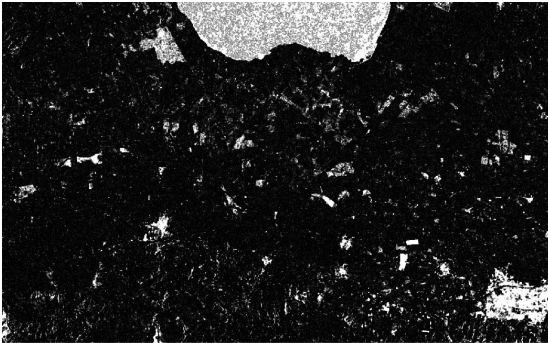
DD - Farklılık



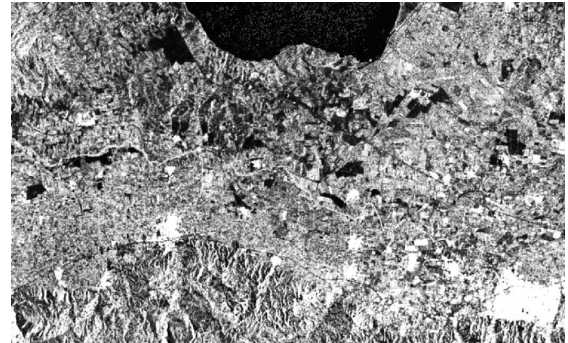
DD - Entropi



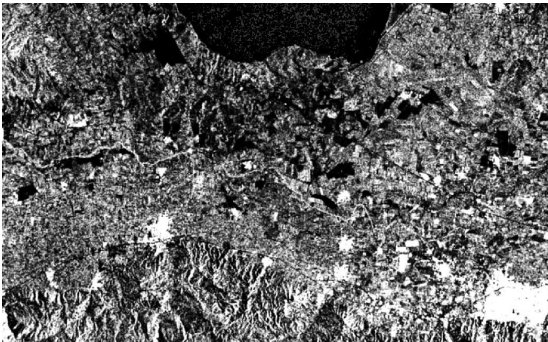
DD – Açısal İkinci Moment



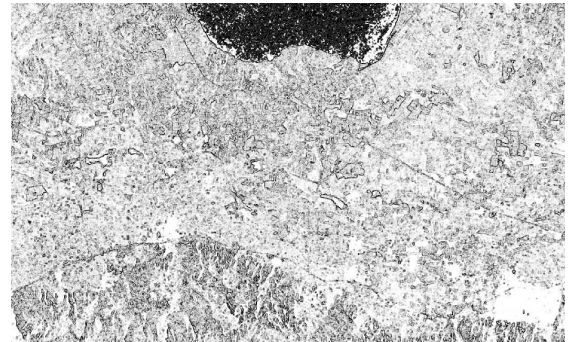
DD - Ortalama



DD - Varyans



DD - Korelasyon



EK 2 – Sınıflandırma Sonuçları Hata Matrisleri

XGBoost Algoritması Sınıflandırma Sonuçları Hata Matrisi

XGBOOST									
	Referans								
Sınıflandırma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Buğday	9704	26	372	294	1931	98	36	129	169
Domates	141	3909	0	283	103	48	47	4	7
Mera	235	9	2068	8	324	1	5	59	9
Mısır	120	438	12	9279	96	7	74	15	27
Mısır2	911	44	13	140	782	22	34	5	41
Pamuk	12	32	0	0	6	7090	12	0	0
Üzüm	235	341	4	534	68	55	13841	68	1219
Yonca	8	1	1	42	88	1	0	69	78
Zeytin	197	70	55	114	218	18	1078	315	10667

Genel Doğruluk: %83,67

Kappa: 0,8059

	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Üretici Doğruluğu	0.8392	0.8027	0.8190	0.8677	0.2163	0.9659	0.9150	0.1039	0.8731
Özgüllük	0.9465	0.9900	0.9902	0.9864	0.9814	0.9990	0.9528	0.9968	0.9634
Kullanıcı Doğruluğu	0.7606	0.8606	0.7609	0.9216	0.3926	0.9913	0.8458	0.2396	0.8378
Yanlış Tahmin Değeri	0.9667	0.9850	0.9931	0.9758	0.9575	0.9959	0.9754	0.9913	0.9723
Yaygınlık	0.1685	0.0710	0.0368	0.1559	0.0527	0.1070	0.2205	0.0097	0.1780
Tespit Etme Oranı	0.1414	0.0570	0.0301	0.1352	0.0114	0.1033	0.2017	0.0010	0.1555
Tespit Etme Yaygınlığı	0.1859	0.0662	0.0396	0.1467	0.0290	0.1042	0.2385	0.0042	0.1856
Dengelenmiş Doğruluk	0.8928	0.8964	0.9046	0.9270	0.5988	0.9825	0.9339	0.5504	0.9183

Rastgele Orman Algoritması Sınıflandırma Sonuçları Hata Matrisi

RO									
	Referans								
Sınıflandırma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Buğday	10267	4	522	332	2005	91	52	128	151
Domates	86	3829	0	230	87	12	21	0	0
Mera	228	11	1906	0	60	0	0	70	12
Mısır	143	559	10	9240	455	11	85	9	31
Mısır2	177	34	3	40	538	8	2	0	22
Pamuk	13	22	0	0	9	7097	3	0	0
Üzüm	355	276	2	556	77	97	13732	56	1411
Yonca	7	0	1	22	19	0	0	38	64
Zeytin	227	118	71	248	354	10	1179	361	10469

Genel Doğruluk: %83,55

Kappa: 0,8036

	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Üretici Doğruluğu	0.8925	0.7890	0.7579	0.8661	0.1493	0.9687	0.9110	0.0574	0.8609
Özgüllük	0.9422	0.9931	0.9942	0.9774	0.9956	0.9992	0.9469	0.9983	0.9543
Kullanıcı Doğruluğu	0.7576	0.8978	0.8334	0.8764	0.6529	0.9934	0.8291	0.2517	0.8030
Yanlış Tahmin Değeri	0.9775	0.9840	0.9908	0.9753	0.9546	0.9963	0.9741	0.9909	0.9694
Yaygınlık	0.1683	0.0710	0.0368	0.1560	0.0527	0.1072	0.2205	0.0097	0.1779
Tespit Etme Oranı	0.1502	0.0560	0.0279	0.1352	0.0079	0.1038	0.2009	0.0006	0.1531
Tespit Etme Yaygınlığı	0.1982	0.0623	0.0335	0.1542	0.0121	0.1045	0.2423	0.0022	0.1907
Dengelenmiş Doğruluk	0.9174	0.8911	0.8760	0.9218	0.5724	0.9840	0.9289	0.5279	0.9076

En Önemli 50 Bant Kullanarak XGBoost Algoritması Sınıflandırma Sonuçları Hata Matrisi

XGB-50									
	Referans								
Sınıflandırma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Buğday	9152	51	553	700	2186	512	22	159	386
Domates	75	3266	17	562	237	194	97	1	44
Mera	353	2	1786	198	4	3	8	1	17
Mısır	236	864	139	7906	224	24	260	48	96
Mısır2	1281	11	27	110	748	1281	33	97	90
Pamuk	5	139	2	9	77	4880	225	0	0
Üzüm	356	427	0	763	105	340	14170	166	2554
Yonca	29	6	1	325	15	56	110	44	116
Zeytin	76	104	0	121	20	50	202	148	8914

Genel Doğruluk: %74,13

Kappa: 0,6925

	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Üretici Doğruluğu	0.7915	0.6706	0.7073	0.7393	0.2069	0.66485	0.9367	0.0663	0.7296
Özgüllük	0.9199	0.9808	0.9911	0.9674	0.9549	0.99254	0.9119	0.9903	0.9872
Kullanıcı Doğruluğu	0.6670	0.7269	0.7530	0.8070	0.2034	0.91437	0.7505	0.0627	0.9252
Yanlış Tahmin Değeri	0.9561	0.9750	0.9888	0.9526	0.9558	0.96112	0.9808	0.9909	0.9440
Yaygınlık	0.1685	0.0710	0.0368	0.1559	0.0527	0.10697	0.2205	0.0098	0.1780
Tespit Etme Oranı	0.1334	0.0476	0.0260	0.1152	0.0109	0.07112	0.2065	0.0006	0.1299
Tespit Etme Yaygınlığı	0.2000	0.0655	0.0346	0.1428	0.0536	0.07778	0.2752	0.0102	0.1404
Dengelenmiş Doğruluk	0.8557	0.8257	0.8492	0.8533	0.5809	0.82870	0.9243	0.5283	0.8584

En Önemli 50 Bant Kullanarak RO Algoritması Sınıflandırma Sonuçları Hata Matrisi

RO-50									
	Referans								
Sınıflandırma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Buğday	9497	50	480	716	2060	680	23	174	387
Domates	61	2948	0	198	171	119	60	0	35
Mera	251	2	1897	142	37	10	2	0	9
Mısır	242	1007	81	7954	253	69	357	30	147
Mısır2	604	2	0	57	719	281	4	36	12
Pamuk	11	152	3	35	120	5247	222	0	16
Üzüm	422	450	3	830	96	678	13765	163	3324
Yonca	22	3	6	124	15	13	59	35	96
Zeytin	235	123	0	277	67	46	168	209	7976

Genel Doğruluk: %74,82

Kappa: 0,6988

	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Üretici Doğruluğu	0.8371	0.6223	0.7680	0.7698	0.2032	0.7346	0.9389	0.0541	0.6646
Özgüllük	0.9177	0.9896	0.9930	0.9613	0.9843	0.9906	0.8857	0.9949	0.9795
Kullanıcı Doğruluğu	0.6751	0.8207	0.8072	0.7844	0.4192	0.9037	0.6976	0.0938	0.8764
Yanlış Tahmin Değeri	0.9650	0.9717	0.9911	0.9581	0.9567	0.9690	0.9810	0.9908	0.9303
Yaygınlık	0.1696	0.0708	0.0369	0.1545	0.0529	0.1068	0.2192	0.0097	0.1795
Tespit Etme Oranı	0.1420	0.0441	0.0284	0.1189	0.0108	0.0785	0.2058	0.0005	0.1193
Tespit Etme Yaygınlığı	0.2103	0.0537	0.0351	0.1516	0.0256	0.0868	0.2950	0.0056	0.1361
Dengelenmiş Doğruluk	0.8774	0.8060	0.8805	0.8656	0.5938	0.8626	0.9123	0.5245	0.8220

En Önemli 70 Bant Kullanarak XGBoost Algoritması Sınıflandırma Sonuçları Hata Matrisi

XGB-70									
	Referans								
Sınıflandırma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Buğday	9326	99	815	447	2300	126	17	198	402
Domates	76	3398	1	423	148	225	58	0	16
Mera	544	0	1587	62	37	1	21	0	17
Mısır	173	682	53	8749	153	15	149	34	95
Mısır2	936	10	48	135	680	493	57	66	54
Pamuk	3	140	3	8	12	6275	46	0	0
Üzüm	362	444	2	786	99	176	14585	156	2764
Yonca	60	5	11	21	172	1	30	42	107
Zeytin	83	92	5	63	15	28	164	168	8762

Genel Doğruluk: %77,83

Kappa: 0,7357

	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Üretici Doğruluğu	0.8065	0.6977	0.6285	0.8181	0.1881	0.8549	0.9642	0.0633	0.7172
Özgüllük	0.9228	0.9851	0.9897	0.9766	0.9723	0.9965	0.9105	0.9940	0.9890
Kullanıcı Doğruluğu	0.6792	0.7821	0.6994	0.8660	0.2743	0.9673	0.7528	0.0935	0.9341
Yanlış Tahmin Değeri	0.9592	0.9771	0.9859	0.9668	0.9556	0.9829	0.9890	0.9909	0.9417
Yaygınlık	0.1685	0.0710	0.0368	0.1559	0.0527	0.1070	0.2205	0.0097	0.1781
Tespit Etme Oranı	0.1359	0.0495	0.0231	0.1275	0.0099	0.0915	0.2126	0.0006	0.1277
Tespit Etme Yaygınlığı	0.2001	0.0633	0.0331	0.1472	0.0361	0.0945	0.2824	0.0065	0.1367
Dengelenmiş Doğruluk	0.8647	0.8414	0.8091	0.8974	0.5802	0.9257	0.9373	0.5286	0.8531

En Önemli 70 Bant Kullanarak RO Algoritması Sınıflandırma Sonuçları Hata Matrisi

RO-70									
	Referans								
Sınıflandırma	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Buğday	9499	47	660	588	1938	85	24	195	373
Domates	62	3100	0	226	51	298	13	0	1
Mera	268	0	1698	35	19	0	9	0	6
Mısır	233	769	46	8325	193	21	165	33	98
Mısır2	366	11	31	34	1024	13	7	32	17
Pamuk	5	136	1	5	13	6131	38	0	1
Üzüm	465	515	3	752	112	458	13696	118	3350
Yonca	31	5	6	30	108	2	40	23	103
Zeytin	224	94	0	185	65	46	199	234	7866

Genel Doğruluk: %78,21

Kappa: 0,7395

	Buğday	Domates	Mera	Mısır	Mısır2	Pamuk	Üzüm	Yonca	Zeytin
Üretici Doğruluğu	0.8517	0.6628	0.6945	0.8178	0.2907	0.8692	0.9651	0.0362	0.6658
Özgüllük	0.9283	0.9893	0.9947	0.9719	0.9918	0.9966	0.8879	0.9950	0.9806
Kullanıcı Doğruluğu	0.7084	0.8265	0.8344	0.8424	0.6671	0.9686	0.7035	0.0661	0.8825
Yanlış Tahmin Değeri	0.9684	0.9745	0.9883	0.9668	0.9610	0.9845	0.9893	0.9906	0.9304
Yaygınlık	0.1698	0.0712	0.0372	0.1550	0.0537	0.1074	0.2161	0.0097	0.1799
Tespit Etme Oranı	0.1446	0.0472	0.0259	0.1268	0.0156	0.0934	0.2086	0.0004	0.1198
Tespit Etme Yaygınlığı	0.2042	0.0571	0.0310	0.1505	0.0234	0.0964	0.2965	0.0053	0.1357
Dengelenmiş Doğruluk	0.8890	0.8261	0.8446	0.8949	0.6412	0.9329	0.9265	0.5156	0.8232