

SENTINEL-2 GÖRÜNTÜLERİNİN YAPAY ZEKA ALGORİTMALARI İLE SINIFLANDIRMASI VE SINIFLANDIRMAYA ETKİ EDEN FAKTÖRLERİN SHAP TEKNİĞİ İLE AÇIKLANABİLİRLİĞİ

Doğu İLMAK^{1*}, Selin UYAR¹, Muzaffer Can İBAN²,

¹Mersin Üniversitesi, Uzaktan Algılama ve CBS ABD, 33343, Yenişehir, Mersin, doguilmak, slnuyar@gmail.com

²Doç. Dr., Mersin Üniversitesi, Harita Mühendisliği Bölümü, 33343, Yenişehir, Mersin, caniban@mersin.edu.tr

ÖZET

Son yıllarda, arazi örtüsü ve arazi kullanımı (AÖAK) sınıflandırmasında yapay zekâ tabanlı yöntemlerde önemli ilerlemeler kaydedilmiştir. Bu yöntemler sınıflandırma performansını artırsa da hangi bant ve indislerin sınıflandırma sürecine katkı sağladığı konusu yeterince araştırılmamıştır. Bu boşluğu doldurmak amacıyla, SHapley Additive Explanations (SHAP) gibi açıklanabilir yapay zekâ yöntemleri, sınıflandırma modellerini değerlendirmede ve karar süreçlerini açıklamada önemli araçlar sunmaktadır. Bu çalışmada, Çukurova Havzası'nın AÖAK sınıflandırması için Sentinel-2 uydu görüntüsünden alınan dokuz adet bant, dokuz adet indis (ör. NDVI, NDWI, NDTI) ile Sayısal Yükseklik Modeli (SYM) kullanılmıştır. Eğitim veri seti, 18 Ağustos 2023 tarihi Sentinel-2 görüntüsüyle oluşturulmuş ve toplamda 1,575,249 piksel etiketlenmiştir. Bu veri seti, yedi adet makine öğrenmesi ve üç adet derin öğrenme modeli ile sınıflandırılmış; sonuçlar F-1 Skoru, Doğruluk ve Kappa performans metrikleriyle değerlendirilmiştir. Sonuçlara göre, RF ve 1D CNN modelleri en iyi performansı göstermiştir. RF modelinin tüm görüntü üzerine uygulanması sonucunda, çalışma sahasının %40.23'ü ormanlar, %31.74'ü tarım alanları, %21.18'i çıplak araziler, %4.81'i yapay yüzeyler ve %2.04'ü su yüzeyleri olarak sınıflandırılmıştır. SHAP tekniği kullanılarak, RF modeli girdi özellikler bazında açıklanmaya çalışılmıştır. Mutlak ortalama SHAP değerlerine göre, sonuca en çok katkı sağlayan bağımsız değişkenler sırasıyla SYM, B11 ve NDTI olmuştur. Sınıflar bazında değerlendirildiğinde, tarım alanlarında en etkili değişkenler SYM, NDWI ve NDTI; yapay yüzeylerde NDTI, SYM ve NDWI olmuştur. Ormanlar için SYM, B12 ve NDTI; su yüzeyleri için B11, B8 ve B6; çıplak araziler için ise SYM, B11 ve B12 en etkili bağımsız değişkenler olarak belirlenmiştir. Bu çalışma, SHAP tekniğinin görüntü sınıflandırmasında model açıklanabilirliği ve performans değerlendirmesi açısından önemli bir katkı sunduğunu ispatlamaktadır. SHAP, her bir bağımsız değişkenin model sonucuna olan katkısını net bir şekilde ortaya koyarak, hangi özelliklerin sınıflandırma sürecinde etkili olduğunu belirlemeye yardımcı olmuştur. Bu, modelin karar mekanizmalarının anlaşılmasını ve optimize edilmesini sağlar.

Anahtar Sözcükler: açıklanabilir yapay zekâ, AÖAK, makine öğrenmesi, Sentinel-2, SHAP

ABSTRACT

CLASSIFICATION OF SENTINEL-2 IMAGES WITH ARTIFICIAL INTELLIGENCE ALGORITHMS AND EXPLAINABILITY OF FACTORS AFFECTING CLASSIFICATION USING SHAP TECHNIQUE

In recent years, significant advancements have been made in artificial intelligence-based methods for land use and land cover (LULC) classification. While these methods have improved classification performance, the contribution of specific bands and indices to the classification process remains underexplored. To fill this gap, explainable AI techniques, such as SHapley Additive Explanations (SHAP), provide valuable tools for evaluating classification models and explaining decision-making processes. In this study, nine bands from Sentinel-2 satellite imagery, nine indices (e.g., NDVI, NDWI, NDTI), and a Digital Elevation Model (DEM) were used for the LULC classification of the Çukurova Basin, Türkiye. The training dataset was created using Sentinel-2 imagery from August 18, 2023, with a total of 1,575,249 labeled pixels. This dataset was classified using seven machine learning and three deep learning models, with results evaluated using F-1 Score, Accuracy, and Kappa performance metrics. According to the results, RF and 1D CNN models demonstrated the best performance. Applying the RF model to the entire image resulted in the classification of 40.23% of the study area as forests, 31.74% as agricultural lands, 21.18% as barren lands, 4.81% as artificial surfaces, and 2.04% as water bodies. SHAP analysis was used to explain the RF model based on input features. According to the absolute mean SHAP values, the most influential features were DEM, B11, and NDTI. When evaluated by classes, the most influential features were DEM, NDWI, and NDTI for agricultural lands; NDTI, DEM, and NDWI for artificial surfaces; DEM, B12, and NDTI for forests; B11, B8, and B6 for water bodies; and DEM, B11, and B12 for barren lands. This study demonstrates that SHAP significantly contributes to model explainability and performance evaluation in image classification by clearly identifying the most impactful features in the classification process.

Keywords: explainable AI, LULC, machine learning, Sentinel-2, SHAP

1. GİRİŞ

Arazi Örtüsü ve Arazi Kullanımı (AÖAK) bilgisi, insan faaliyetlerinin çevre ile olan fiziksel etkileşimlerini anlamak için önemli bir bileşendir. Uzaktan algılama verilerinin işlenmesiyle üretilen AÖAK haritaları, yeryüzü hakkında kıymetli bilgiler sağladıkları için, kuraklığın takibi, ormansızlaşma ve kentsel saçaklanmanın izlenmesi, bitki hastalıklarının tespiti, yangın izleme, su kalitesi analizi ve taşkın haritalaması gibi birçok çalışmada hayati bir rol

oyunmaktadır. Ayrıca, hükümetler, karar vericiler, uluslararası kuruluşlar, araştırmacılar ve vatandaşlar, arazi ile ilgili sosyo-ekonomik ve mekânsal değişimleri AÖAK bilgisi ile takip etmektedir (Kavzoğlu ve Bilucan, 2023). Bu nedenle, görüntü sınıflandırma işlemiyle yüksek doğrulukta ve güvenilir AÖAK haritaları üretmek son derece önemlidir. Bu işlem, bir görüntüdeki tüm piksellerin belirli bir etiket ya da AÖAK türüne göre sınıflandırılması olarak tanımlanır. Bu amaçla, literatürde birçok sınıflandırma tekniği geliştirilmiş ve başarıyla uygulanmıştır (Avcı vd., 2023).

Son yıllarda AÖAK alanında yapay zekâ tabanlı sınıflandırma yöntemlerinde önemli gelişmeler kaydedilmiştir. Yapay zekânın uzaktan algılama alanında güç kazandıran uygulamalarına rağmen, uydu görüntüleri görsel spektrumun ötesinde, belirli nesnelerin yansıma davranışlarına dayalı faydalı bilgiler içermektedir. Bu nedenle, çok bantlı görüntülerin işlenmesi, özellikle büyük miktarda görüntü içeren veri setleriyle çalışan yapay zekâ modellerinde, yüksek hesaplama gücü ve ek eğitim süresi gerektirir. Bu yüzden, bir sınıflandırma modeline verilen bilgi miktarını optimize etmek için bantların doğru seçilmesi önemlidir. Ayrıca, bu bantların neden önemli olduğunu anlamak ve modellerin tahmin davranışını hangi nesnelerin etkilediğini araştırmak amacıyla mekânsal açıklamaların görselleştirilmesi gerekmektedir (Temenos vd., 2023). Bir başka deyişle, yapay zekâ yöntemleri sınıflandırma performansını artırmış olsa da sınıflandırma sürecinde hangi bant ve indislerin nasıl kullanıldığı, bu faktörlerin sınıflandırma sürecine katkıları ve önem dereceleri konusunda belirgin bir araştırma boşluğu bulunmaktadır (Atik ve Atik, 2024). Bu boşluğu doldurmak amacıyla açıklanabilir yapay zekâ teknikleri, örneğin SHapley Additive Explanations (SHAP) gibi yöntemler, sınıflandırma performansını değerlendirmenin yanı sıra sınıflandırma kararlarını açıklamak için önemli araçlar sunmaktadır.

DeneySEL ve gözlemsel veri analizi, fiziksel olayların doğru şekilde modellenmesi ve anlaşılması için temel oluşturur. Ancak, sadece modelin oluşturulması genellikle tek başına yeterli olmaz ve sistemin davranışını anlama ve tanımlama arzusu doğar. Yapay zekâ modelleriyle elde edilen sonuçların bilimsel tutarlılığı, güvenilirliği ve açıklanabilirliği, karmaşık sistemlerle çalışırken hayati öneme sahiptir (Reichstein vd., 2019). Dolayısıyla, görselleştirilebilen, sorgulanabilen ve yorumlanabilen güçlü modeller geliştirmek önem arz etmektedir. Uzaktan algılama alanındaki yapay zekâ ve makine öğrenmesi modellerinde şeffaflık, yorumlanabilirlik ve açıklanabilirlik elde etmek, modelin performans tutarlılığını, genellenebilirliğini ve sağlamlığını sağlamak ve bilim insanları arasında daha geniş kabul ve güven oluşturmak açısından gereklidir (Tuia vd., 2024).

Bu çalışmada, Çukurova Havzası'nda AÖAK sınıflandırması gerçekleştirilmiştir. Sentinel-2 uydu görüntülerinin çeşitli bantları, ilgili indisler ve SYM bağımsız değişkenler olarak kullanılmıştır. 18 Ağustos 2023 tarihli Sentinel-2 görüntüsü ile 8486 km²'lik bir alanda kapsamlı bir eğitim veri seti oluşturulmuş ve farklı makine öğrenmesi (Karar Ağaçları (DT), Gradyan Artırma (Gradient Boosting), Aşırı Gradyan Artırma (XGBoost), Rastgele Orman (RF), Destek Vektör Makineleri (SVM), k-en Yakın Komşuluk (KNN), Gauss Toy Bayes (Gaussian NB)) ve derin öğrenme (Çok Katmanlı Algılayıcı (MLP), Tekrarlayan Sinir Ağları (RNN), Bir Boyutlu Konvolüsyonel Sinir Ağı (1D CNN)) algoritmalarıyla sınıflandırma deneyleri yapılmıştır. Performans değerlendirmesi F-1 Skoru, Doğruluk ve Kappa katsayısı metrikleriyle gerçekleştirilmiştir. Çalışma, CORINE 1. düzey sınıflarına dayanmakta olup, toplamda 1,575,249 piksel etiketlenmiştir. Yerel doğruluk sağlamak için Google Earth'ten çok yüksek çözünürlüklü görüntüler kullanılmıştır.

AÖAK sınıflandırmasının ardından, sonuçlar SHAP algoritmasına aktarılmıştır. Bu sayede, sınıflandırma doğruluğunu artıran bantlar ve indisler belirlenmiştir. Açıklanabilir yapay zekâ (XAI) algoritması olan SHAP, pozitif ve negatif korelasyonlar üreterek bu bantların ve indislerin sınıflandırma üzerindeki etkilerini ortaya koymaktadır. Önerilen çerçeve hem spektral hem de mekânsal düzeyde sınıflandırma sürecine ilişkin açıklamalar sağlamaktadır. Spektral düzeyde, her bir sınıf için tüm bantların ve indislerin özellik önem derecesi belirlenmiştir. Mekânsal düzeyde ise, nihai tahmin edilen sınıflara en fazla katkı sağlayan piksel kümeleri görselleştirilmiştir. Böylece, sınıflandırma sürecinin hem doğruluğu hem de şeffaflığı artırılmıştır.

2. MATERYAL VE YÖNTEM

2.1 Kullanılan Veriler

Bu çalışmada, 18 Ağustos 2023 tarihli bulutsuz Sentinel-2 MSI uydu görüntüsü kullanılmıştır. Görüntü, Adana Metropol Alanı'nın kuzeybatı kısmını da içeren ve 8,486 km²'lik Çukurova Havzası'nın bir kısmını kapsamaktadır. Bu bölge, kırsal ve ormanlık alanlara doğru önemli bir kentsel yayılım olgusuna sahiptir. Sentinel-2 MSI uydu verileri, 2015 yılından bu yana ücretsiz olarak erişilebilmekte olup, Avrupa Uzay Ajansı tarafından geliştirilen orta çözünürlüklü multispektral ve atmosferik olarak düzeltilmiş görüntüler sağlamaktadır. Sentinel-2A uydusu, 13 spektral banda sahip olup, bu bantların mekânsal çözünürlüğü 10 ile 60 metre arasında değişmektedir. Bu çalışmada kullanılan bantlar ve indisler hakkında detaylı bilgi Çizelge 1 ve Çizelge 2'de sunulmuştur. Sentinel-2 bantları ve bu bantlardan türetilen indisler, AÖAK sınıflandırmasında, her biri farklı yüzey özelliklerini ve spektral yanıtları yansıttıkları için büyük önem taşır. Örneğin, B2 (Mavi), B3 (Yeşil) ve B4 (Kırmızı) bantları, özellikle bitki örtüsü, su kütleleri ve çiplak toprak gibi farklı yüzey özelliklerini ayırt etmek için kullanılır. Yakın Kızılötesi (B8) bitki sağlığını

değerlendirmek için kritik öneme sahiptir, çünkü bitkiler bu spektrumda yüksek yansıtıcılık gösterir. Kısa Dalga Kızılötesi (SWIR) bantları (B11 ve B12), toprak nemi, su içeriği ve bitki örtüsünün su stresi gibi özelliklerin tespitinde etkilidir.

Çizelge 1. Çalışmada Kullanılan Sentinel-2 Bantları

Tanım	Sentinel-2 Bant İsmi	Spektral Aralık (nm)	Mekânsal Çözünürlük (m)
Mavi (Blue)	B2	496.6 - 492.1	10
Yeşil (Green)	B3	560 - 559	10
Kırmızı (Red)	B4	664.5 - 665	10
Bitki Kırmızı Kenar 1 (Red Edge 1)	B5	703.9 - 703.8	20
Bitki Kırmızı Kenar 2 (Red Edge 2)	B6	740.2 - 739.1	20
Bitki Kırmızı Kenar 3 (Red Edge 3)	B7	782.5 - 779.7	20
Yakın Kızılötesi (Near-Infrared –NIR)	B8	835.1 - 833	10
Kısa Dalga Kızılötesi 1 (SWIR1)	B11	1613.7 - 1610.4	20
Kısa Dalga Kızılötesi 2 (SWIR2)	B12	2202.4 - 2185.7	20

Çizelge 2. Çalışmada Kullanılan Sentinel-2 Bantları

Tanım	İngilizcesi	Formül
Çıplak Toprak İndeksi	Bare Soil Index (BSI)	$((B11 + B4) - (B8 + B2)) / ((B11 + B4) + (B8 + B2))$
Geliştirilmiş Bitki Örtüsü İndeksi	Enhanced Vegetation Index (EVI)	$2.5 \times (B8 - B4) / (B8 + 6 \times B4 - 7.5 \times B2 + 1)$
Normalize Edilmiş Fark Bitki Örtüsü İndeksi	Normalized Difference Vegetation Index (NDVI)	$(B8 - B4) / (B8 + B4)$
Normalize Edilmiş Fark Su İndeksi	Normalized Difference Water Index (NDWI)	$(B3 - B8) / (B3 + B8)$
Normalize Edilmiş Fark Yapılaşma İndeksi	Normalized Difference Built-Up Index (NDBI)	$(B11 - B8) / (B11 + B8)$
Normalize Edilmiş Fark Bulanıklık İndeksi	Normalized Difference Turbidity Index (NDTI)	$(B3 - B6) / (B3 + B6)$
Yeşil Yaprak İndeksi	Green Leaf Index (GLI)	$(2 \times B3 - B4 - B2) / (2 \times B3 + B4 + B2)$
Toprak Düzeltmiş Bitki Örtüsü İndeksi	Soil-Adjusted Vegetation Index (SAVI)	$(B8 - B4) \times (1 + L) / (B8 + B4 + L)$, $L = 0.5$
Arazi Yüzeyi Su İndeksi	Land Surface Water Index (LSWI)	$(B8 - B11) / (B8 + B11)$

Bu bantlardan türetilen indisler, farklı AÖAK türlerinin daha özellikli ve hassas bir şekilde belirlenmesini sağlar. Örneğin, NDVI bitki örtüsünün yoğunluğunu ve sağlığını ölçerken, NDWI su kütlelerini ve nemi tespit eder. SAVI, bitki örtüsü yoğunluğunun düşük olduğu alanlarda toprak etkisini azaltarak daha doğru sonuçlar sağlar. BSI, çıplak toprak alanlarını belirlemeye yardımcı olurken, NDBI yapılaşmış alanların tespitinde kullanılır. Bu bant ve indisler, geniş bir spektral aralıkta farklı yüzeylerin özelliklerini yansıttığı için AÖAK sınıflandırmasında tercih edilmektedir. Dolayısıyla daha yüksek doğruluk oranlarına ulaşılır ve bölgenin AÖAK haritalaması detaylı bir şekilde yapılabilir.

SYM olarak, yaklaşık 30 metre çözünürlükte sunulan Shuttle Radar Topography Mission (SRTM) V3 ürünü (SRTM Plus) kullanılmıştır.

2.2 Sınıflandırma Çerçevesi

Bu çalışmada, CORINE 1. düzey sınıflandırma şemasına dayalı olarak beş farklı AÖAK sınıfı kullanılmıştır: (1) **yapay yüzeyler**, yerleşim alanları, betonla kaplı araziler ve her türlü yapılaşmış alanları içerir; (2) **tarım alanları**, yıllık ve çok yıllık ekinler, çayırlar ve seraları kapsar; (3) **ormanlar**; (4) **su kütleleri**, denizler ve akarsular gibi su yüzeylerini içerir; ve (5) **çıplak araziler**, bitki örtüsünün seyrek olduğu, kumlu veya kayalık alanları içerir (İban ve

Şahin, 2022). Uydu görüntülerinin sınıflandırılması, her pikseli otomatik olarak bu AÖAK sınıflarından birine atamayı amaçlar. Ancak denetimli görüntü sınıflandırması, sınıflandırılacak AÖAK sınıflarına dair önceden bilgi gerektirir. Bu nedenle, seçilen AÖAK sınıflarının görüntüler üzerinde işaretlenmesi ve sınıflandırma algoritmalarının eğitilmesi için, her beş AÖAK sınıfını temsil eden toplamda 1,575,250 piksel içeren, rastgele dağıtılmış 99 poligon oluşturulmuştur. Bu örnek veri seti, %80'i eğitim, %20'si ise doğrulama için kullanılmak üzere rastgele ikiye bölünmüştür. Eğitim pikselleri sınıflandırma algoritmalarının eğitimi için kullanılırken, kalan doğrulama pikselleri sınıflandırmanın performansını değerlendirmek amacıyla kullanılmıştır. Poligonların seçimi, yüksek çözünürlüklü Bing ve Google Earth görüntülerinden destek alınarak görsel bir yaklaşımla yapılmıştır. Bu sayede, uydu görüntülerinin doğru şekilde etiketlenmesi sağlanmış ve eğitim veri setinin kalitesi artırılmıştır. Bu poligonlara ilişkin bilgiler Çizelge 3'te sunulmaktadır.

Çizelge 3. Çalışmada Kullanılan Eğitim ve Doğrulama Piksellerine İlişkin Bilgiler

AÖAK Sınıfı	Eğitim Pikseli Sayısı	Doğrulama Pikseli Sayısı	Toplam
Yapay Yüzeyler	79740	20171	99911
Tarım Alanları	445159	110422	555581
Ormanlar	198754	49808	248562
Su Kütleleri	385493	96770	482263
Çıplak Araziler	151054	37879	188933
Toplam	1260200	315050	1575250

Bu çalışmada, AÖAK sınıflandırması için farklı makine öğrenmesi ve derin öğrenme algoritmaları kullanılmıştır. Makine öğrenmesi modelleri arasında Karar Ağaçları (DT), Gradyan Artırma (Gradient Boosting), Aşırı Gradyan Artırma (XGBoost), Rastgele Orman (RF), Destek Vektör Makineleri (SVM), k-en Yakın Komşuluk (KNN) ve Gauss Toy Bayes (Gaussian NB) algoritmaları yer almaktadır. Derin öğrenme tarafında ise Çok Katmanlı Algılayıcı (MLP), Tekrarlayan Sinir Ağları (RNN) ve Bir Boyutlu Konvolüsyonel Sinir Ağları (1D CNN) kullanılmıştır. Bu modellerin hiperparametre uyumlandırması GridSearchCV yöntemi ile gerçekleştirilmiştir. Her bir sınıflandırma algoritmasının performansı üç temel metrik ile değerlendirilmiştir: Bunlar Cohen's Kappa, Doğruluk ve F-1 skoru şeklindedir. Çalışmanın bulgular kısmında, en iyi performans gösteren model üzerinde harita üretimi ve model açıklanabilirliği sağlanmıştır.

Makine öğrenimi modellerinin açıklanabilirliği, XAI'nin ana konularından biridir XAI yöntemlerinden olan SHAP, her örnek (piksel) için hangi özelliklerin sınıflandırma sonuçları etkilediğini gösterir. SHAP, Shapley değerlerine dayalıdır ve her özelliğin katkısını adil bir şekilde ölçer. Bu çalışmada, SHAP Python kütüphanesi kullanılarak en iyi performans gösteren modelin açıklanabilirliği sağlandı. SHAP özet grafikleri ve bağlılık grafikleri, modeldeki özelliklerin etkilerini ve kritik eşiklerini anlamaya yardımcı olmuştur.

3. BULGULAR VE TARTIŞMA

3.1 Sınıflandırma Performansı

Yedi makine öğrenmesi ve üç derin öğrenme modeli arasında en iyi performansı, RF ve 1D CNN modelleri göstermiştir. RF modeli, 0.9945 Kappa değeri, %99.60 doğruluk ve 0.9960 F-1 skoru ile en yüksek sonuçları elde etmiştir. 1D CNN modeli ise 0.9920 Kappa, %99.42 doğruluk ve 0.9921 F-1 skoru ile ikinci en yüksek performansı sergilemiştir. RF modeli, özellikle karar ağaçlarının topluluğu sayesinde yüksek performans sağlamış, 1D CNN ise verinin mekânsal ve sıralı özelliklerini yakalamada başarılı olmuştur.

Çizelge 5 ve Çizelge 6'daki karmaşıklık matrisleri sırasıyla, RF ve 1D CNN modellerinin AÖAK sınıflandırmasındaki performansını karşılaştırmaktadır. Her iki model de yüksek doğruluk seviyelerine sahip olsa da RF modeli genel olarak daha dengeli ve tutarlı sonuçlar vermiştir. 1D CNN ise özellikle tarım alanları ve orman sınıfında bir miktar daha yüksek doğruluk sağlamıştır. Tarım alanları ve ormanların genellikle yüksek doğrulukla sınıflandırılmasının nedenlerinden biri, bu sınıfların belirgin spektral ve tekstürel özelliklere sahip olmasıdır. Örneğin, tarım alanları genellikle geniş ve homojen alanlar sunarken, ormanlar yoğun bitki örtüsüne sahip olduğundan, modeller bu iki sınıfı ayırt etmede başarılı olmuştur. Yapay yüzeyler sınıfında her iki model de yüksek doğruluk sağlamış olsa da kullanıcı doğruluğu açısından RF (%95.82) modelinde bir miktar daha fazla hata yapılmıştır. Çıplak araziler sınıfında hem RF hem de 1D CNN yüksek doğruluk elde etmiştir, ancak üretici doğruluğu açısından RF %97.38, 1D CNN ise %98.46 değerlerine sahiptir. Bu fark, çıplak arazilerin yapay yüzeylerle karıştırılma eğilimi ile açıklanabilir. Çıplak araziler genellikle düzensiz doku ve spektral özelliklere sahip olduğundan, her iki modelin de bu sınıfta bazı hatalar yapması olasıdır. Su yüzeyleri sınıfı, her iki model tarafından da neredeyse mükemmel doğrulukla sınıflandırılmıştır (%100'e

yakın). Su yüzeylerinin, diğer sınıflardan oldukça farklı spektral yansıtıcılığa sahip olması, modellerin bu sınıfı yüksek doğrulukla ayırt edebilmesini kolaylaştırmıştır.

Çizelge 4. Algoritmaların Sınıflandırma Performansı (En iyi makine öğrenmesi ve derin öğrenme yöntemleri koyu renkle gösterilmiştir.)

Model	Kappa	Doğruluk	F-1 Skoru
DT	0.9764	0.9826	0.9826
Gradient Boosting	0.9877	0.9909	0.9909
XGBoost	0.9847	0.9887	0.9887
RF	0.9945	0.9960	0.9960
SVM	0.9732	0.9802	0.9802
KNN	0.9869	0.9903	0.9903
Gaussian NB	0.7542	0.8139	0.8088
1D CNN	0.9920	0.9942	0.9921
MLP	0.9710	0.9787	0.9698
RNN	0.9890	0.9921	0.9893

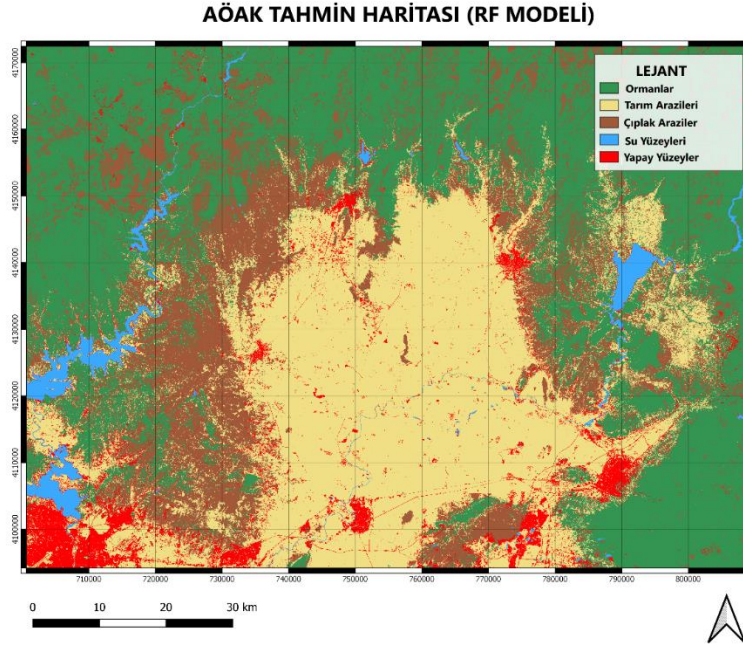
Çizelge 5. RF Modelinin Karmaşıklık Matrisi

	Tarım Alanları	Yapay Yüzeyler	Çıplak Araziler	Ormanlar	Su Yüzeyleri	Toplam	Kullanıcı Doğruluğu
Tarım Alanları	109421	110	815	76	0	110422	0.9909
Yapay Yüzeyler	661	19327	131	52	0	20171	0.9582
Çıplak Araziler	839	28	36985	27	0	37879	0.9764
Ormanlar	205	57	51	49495	0	49808	0.9937
Su Yüzeyleri	0	0	0	0	96770	96770	1
Toplam	111126	19522	37982	49650	96770	315050	
Üretici Doğruluğu	0.9847	0.9900	0.9738	0.9969	1		

Çizelge 6. 1D CNN Modelinin Karmaşıklık Matrisi

	Tarım Alanları	Yapay Yüzeyler	Çıplak Araziler	Ormanlar	Su Yüzeyleri	Toplam	Kullanıcı Doğruluğu
Tarım Alanları	109736	195	456	35	0	110422	0.9938
Yapay Yüzeyler	347	19754	60	10	0	20171	0.9793
Çıplak Araziler	527	21	37303	27	1	37879	0.9848
Ormanlar	55	26	68	49659	0	49808	0.9970
Su Yüzeyleri	0	0	0	0	96770	96770	
Toplam	110422	20171	37879	49808	96770	315050	
Üretici Doğruluğu	0.9916	0.9879	0.9846	0.9986	0.9999		

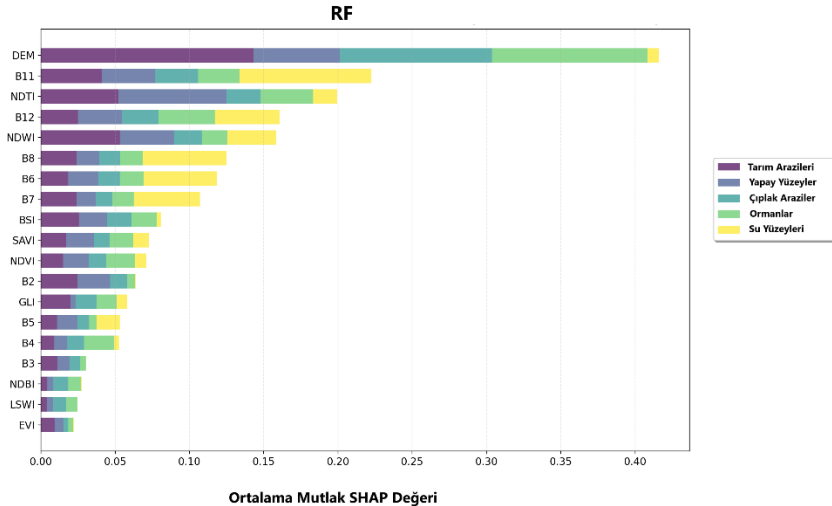
Eğitilen RF modeli, Sentinel-2 görüntüsünün tüm pikselleri için sınıf tahmininde bulunarak bir AÖAK tematik haritası oluşturmuştur. Tahmin sonuçlarına göre, çalışma alanının %40.23'ü orman, %31.74'ü tarım alanı, %21.18'i çıplak arazi, %4.81'i yapay yüzey ve %2.04'ü su yüzeyi olarak sınıflandırılmıştır.



Şekil 1. AÖAK tahmin haritası (RF modeli)

3.2 Sınıflandırma Modellerinin Açıklanabilirliği

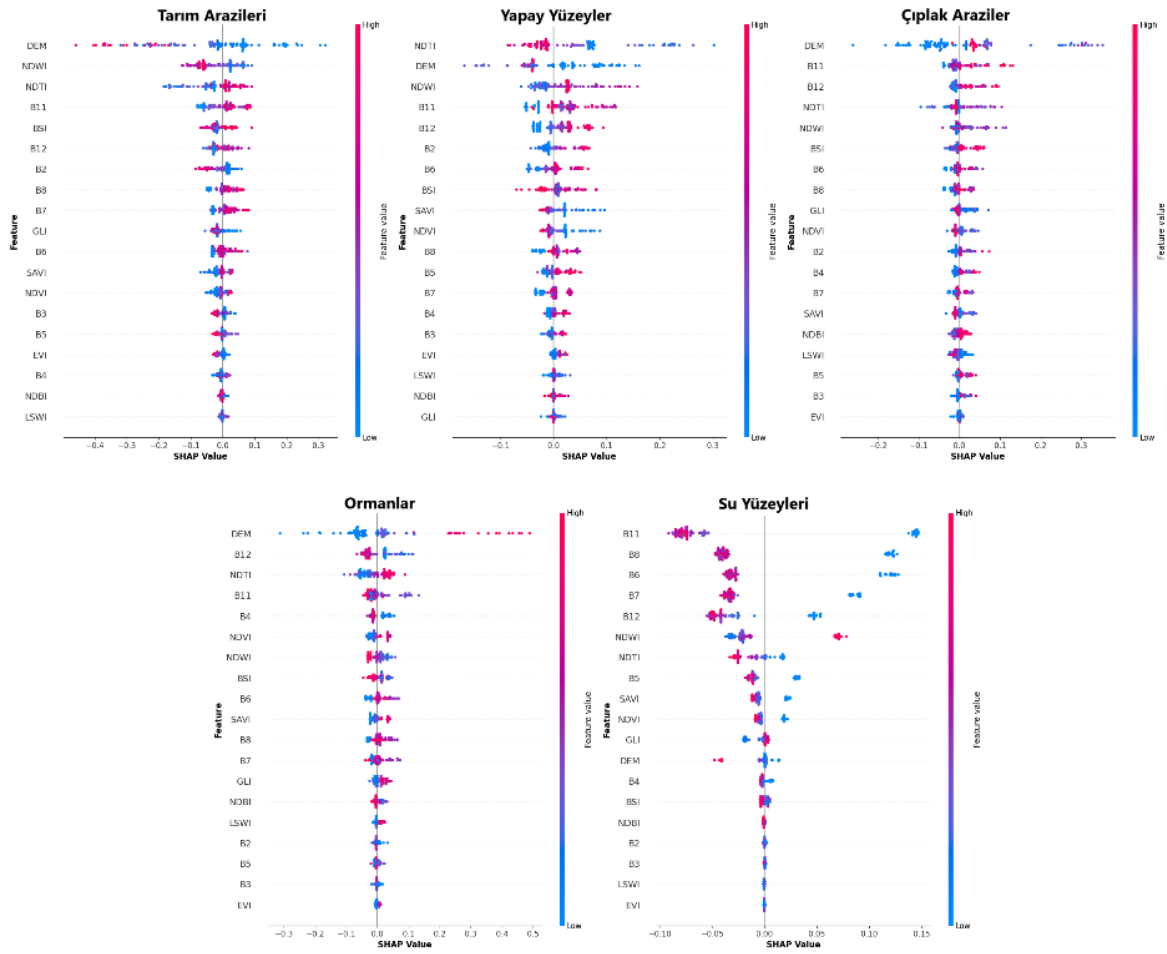
Şekil 2'de, RF modelinin farklı AÖAK sınıflarına yönelik özelliklerin (*features*) önem dereceleri ortalama mutlak SHAP değerleri ile görselleştirilmiştir. Bu şekil, her bir sınıf için her özelliğin ne derece katkı sağladığını ve bu katkıların mutlak değerlerini hesaplayarak ortaya koymaktadır. Burada, modelin eğitimi sırasında en fazla etkiye sahip özellikleri ve bu etkilerin hangi AÖAK sınıflarında daha belirgin olduğunu gözlemlemek olanaklıdır. Bu bağlamda, RF modeli eğitildikten sonra, her bir AÖAK sınıfı için tüm girdi özelliklerinin (SYM, farklı bantlar ve indeksler) SHAP değerleri hesaplanmıştır. Her bir özelliğin mutlak SHAP değerlerinin ortalaması alınarak, her bir sınıf için özellikler sıralanmıştır ve hangi girdi verilerin modelin karar süreçlerini daha fazla etkilediğini anlamaya yardımcı olmuştur. Her özelliğin hangi AÖAK sınıflarında daha fazla etkiye sahip olduğunu gösteren bir yığılmış çubuk grafik oluşturulmuştur. Grafikte her renk, farklı bir AÖAK sınıfını temsil ederken, çubuklar belirli bir özelliğin her sınıf için toplam etki değerini göstermektedir.



Şekil 2. RF modelinin tüm sınıflarda ürettiği ortalama mutlak SHAP değerleri.

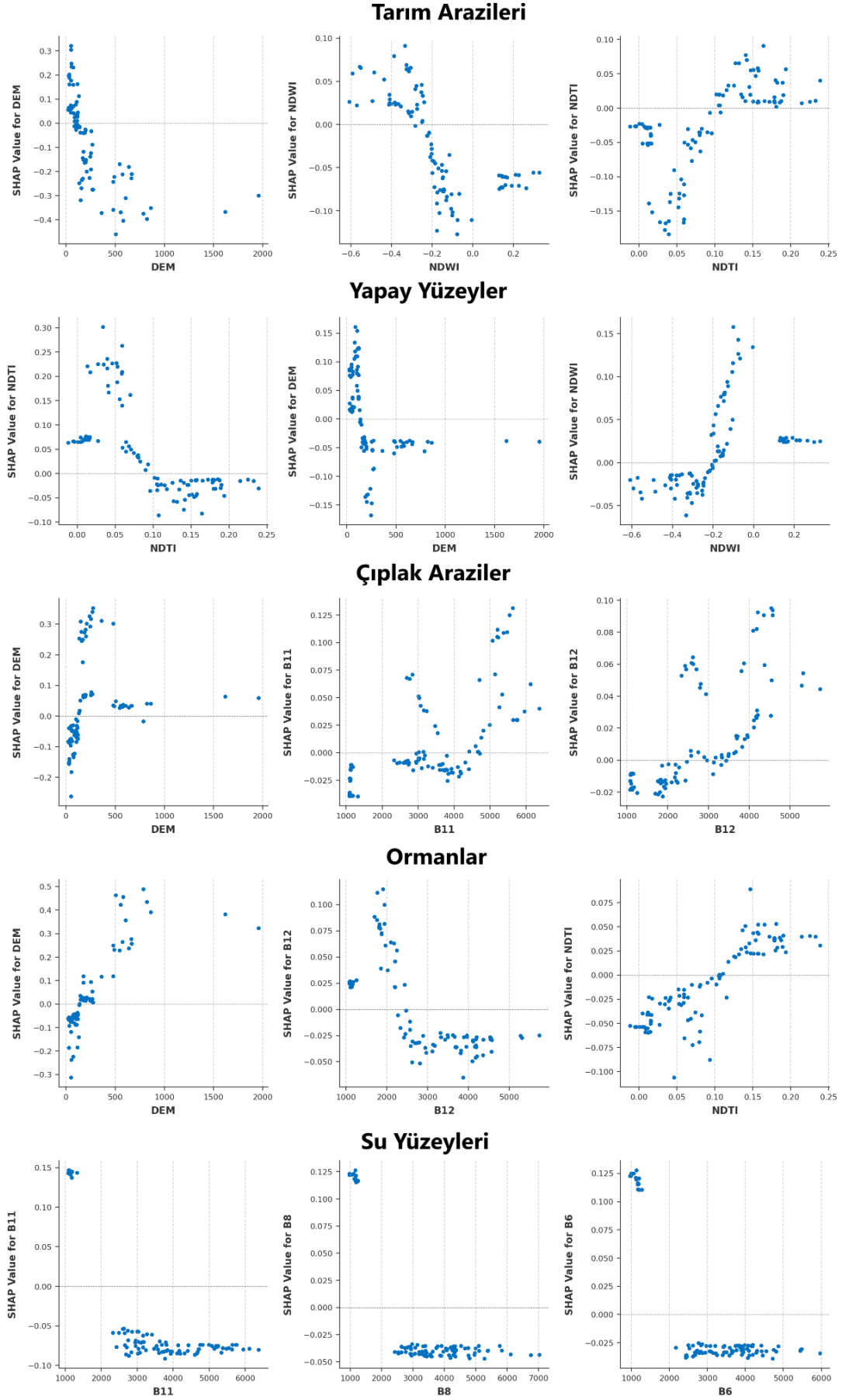
Bu kapsamda SYM (DEM), model için en önemli özellik olup, özellikle çıplak araziler ve orman sınıflarında yüksek etkiye sahip olduğu gözlemlenmiştir. Bu durum, topografik özelliklerin orman ve çıplak arazi sınıflandırılmasında önem arz ettiğini göstermektedir. İkinci sırada gelen B11 (SWIR1), özellikle orman ve tarım alanlarının sınıflandırılmasında belirleyici bir rol oynarken, yapay yüzeyler için daha düşük bir etkiye sahiptir. Tarım ve orman sınıfları, genellikle bu bandın yansıma özellikleriyle iyi ayrılmıştır. Üçüncü sırada gelen NDTI, çıplak arazilerde en yüksek etkiyi gösterirken, diğer sınıflar üzerinde orta dereceli bir etkiye sahiptir.

Şekil 2’de modelin geneli (global) üzerinde bir açıklanabilirlik söz konusudur. Ancak, modele girdi olarak sunulan özelliklerin model performansına katkısını numune pikseller özelinde (lokal) incelemek de mümkündür. Bu kapsamda, test verilerinden her bir sınıf için rastgele seçilen 100’er adet pikselin tahmin sonucuna etki eden özelliklerin ürettiği SHAP değerlerinin nokta grafikleri Şekil 3’te sunulmuştur. Bu grafikler, her özelliğin değerinin (yüksekten düşüğe) modelin çıktısına olan etkisini görselleştirmektedir. Yüksek SHAP değerleri (magenta renkli noktalar), belirli bir sınıfın tahmininde daha fazla olumlu etkiye sahipken, düşük değerler (mavi renkli noktalar) olumsuz bir etkiyi gösterir.



Şekil 3. RF modelinin seçili test piksellerinde ürettiği SHAP değerleri

Tarım arazilerinde, düşük değere sahip olan SYM ve NDWI özellikleri pozitif SHAP değerlerine sahip olup, tarım alanlarının sınıflandırılmasında en belirleyici faktörlerdir. NDTI özelliğinin değeri yükseldikçe, SHAP değerleri yükselmektedir. Bir başka deyişle, düşük kot, düşük NDWI ve yüksek NDTI değerleri, bir pikselin tarım arazisi olarak sınıflandırılması olasılığını artırır. Bu durum Şekil 4’teki SHAP bağılılık grafikleri tarafından da desteklenmektedir. Bağılılık grafikleri, yaklaşık 100 metreden az kota sahip piksel numunelerinin çoğunlukla tarım arazisi olarak sınıflandırıldığını ispatlamaktadır. Benzer şekilde, NDWI değerlerinin -0.2’den küçük olması, NDTI değerlerinin 0.10’dan büyük olması da bu piksellerin tarım arazisi olarak sınıflandırılması olasılığını artırmaktadır. Diğer sınıflar için de belirlenen en önemli üç özelliğin eşik değerleri de bu grafiklerden gözlemlenebilir ve Çizelge 7’de özetlenmektedir.



Şekil 4. RF modelinin her sınıfa ait en önemli üç özelliğinin ürettiği SHAP değerlerinin bağlılık grafikleri

Çizelge 7. RF modelinin farklı sınıflar en önemli bulduğu üç faktörün ürettiği pozitif SHAP değerleri

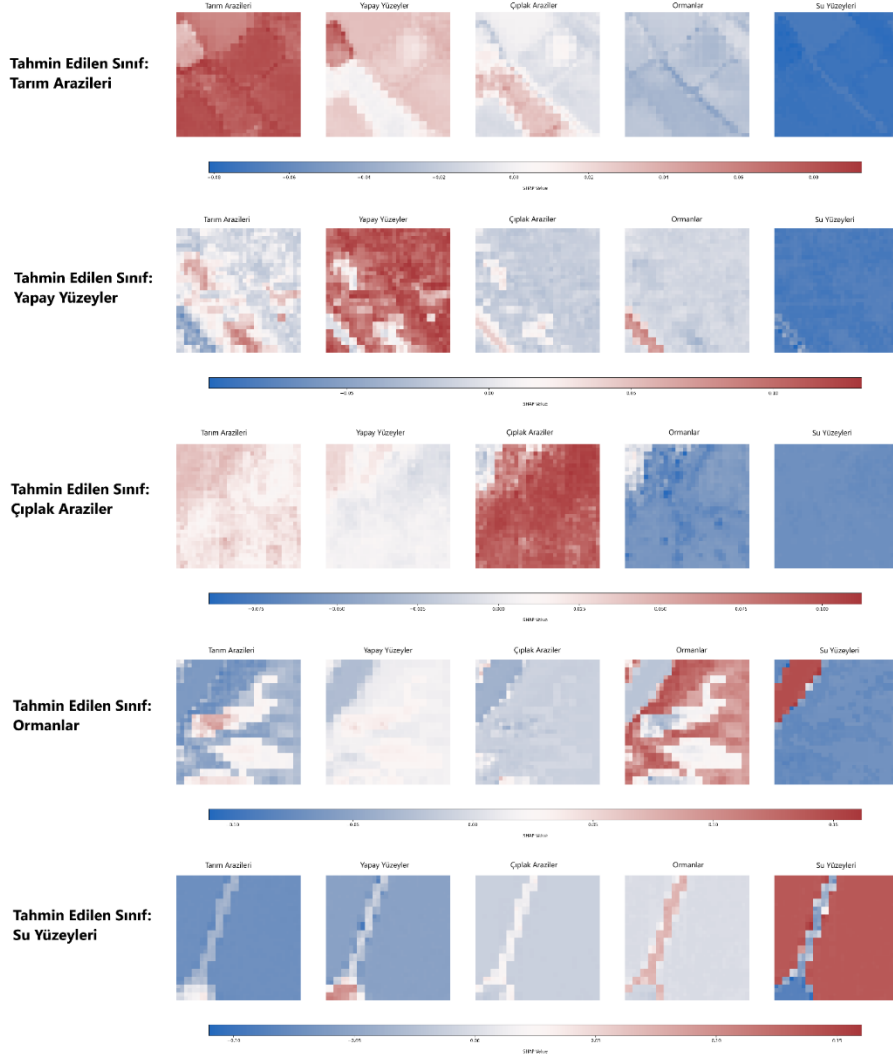
Sınıf	En Önemli Üç Özellik	Pozitif SHAP Değeri Üreten Pksel Değerleri
Tarım Arazileri	DEM NDWI NDTI	<100 metre < -0.2 > 0.10
Yapay Yüzeyler	NDTI DEM NDWI	< 0.10 > 200 metre > -0.2
Çıplak Araziler	DEM B11 B12	> 200 metre > 3500 < 2200
Ormanlar	DEM B12 NDTI	> 200 metre < 2200 > 0.11
Su Yüzeyleri	B11 B8 B6	~ 1000 ~ 1000 ~ 1000

Şekil 5'te her bir sınıf özelinde seçilen örnek alanlarda açıklanabilirliğin nasıl olduğu görselleştirilmiştir. Örneğin tarım arazisi sınıfında SHAP değerleri pozitif (kırmızı tonlar) olarak oluşmuştur. Bir başka deyişle, model tarım arazisi olarak etiketlenen bölgelerde yüksek bir güvenle bu sınıfı tanımlamaktadır. Su yüzeyleri ve çıplak araziler gibi diğer sınıflar negatif SHAP değerleriyle temsil edilmiştir. Dolayısıyla, modelin bu sınıfları tarım arazisi olarak yanlış sınıflandırmadığı anlaşılmaktadır. Bu bağlamda, grafiklerle spektral olarak açıklanan RF modelinin sınıflandırma performansı, mekânsal olarak da açıklanabilmektedir. Genel olarak, her sınıf için modelin doğru sınıflandırmalar yaptığı ve SHAP değerlerinin, sınıf özelinde farklı özelliklere nasıl katkıda bulunduğunu açıkça gösterdiğini söylemek mümkündür. Modelin özellikle su yüzeyleri ve ormanlar gibi belirgin sınıfları doğru bir şekilde ayırt edebildiği gözlemlenmiştir.

4. SONUÇLAR

Bu çalışma, SHAP yöntemi kullanılarak Çukurova Havzası'na ait Sentinel-2 görüntüsü kullanılarak AÖAK sınıflandırmasının daha iyi açıklanabilmesini sağlamaktadır. RF ve 1D CNN modelleri %99'u aşan bir doğruluk ile en yüksek performansı sergilemiştir. SHAP analizi, özellikle SYM, B11 ve NDTI değişkenlerinin tarım alanları gibi belirli sınıfların sınıflandırılmasında kritik öneme sahip olduğunu göstermiştir. Tarım alanlarında, düşük kot ve belirli NDWI ile NDTI eşik değerleri tarım arazisi olarak sınıflandırma olasılığını artırmaktadır. Bu tür detaylı açıklanabilirlik analizleri, hem mekânsal hem spektral bağlamda yapılmış olup, her AÖAK sınıfının belirlenmesinde modelin hangi bant ve indisleri göz önünde daha çok bulundurduğu ve bu özelliklerin sınıflandırmaya daha çok katkı sağlayan değer aralıkları tespit edilmiştir.

Gelecek çalışmalarda SHAP gibi açıklanabilir yapay zeka yöntemlerinin kullanımı yaygınlaştırılmalı, model performansının yanı sıra karar süreçleri de derinlemesine incelenmelidir. SHAP tekniği ile belirlenen en önemli girdi özellikler ayrıştırılabilir; önemsiz olan girdi özellikler veri setinden çıkarılarak daha hızlı ve etkili sınıflandırma çerçeveleri oluşturulabilir. Bu bağlamda, amaca yönelik olarak yapay zeka teknikleri ile görüntü sınıflandırması çalışmalarını gerekli olan bant ve indisleri kullanarak gerçekleştirmek daha kolay olacaktır. Ayrıca, farklı coğrafi bölgelerde benzer analizlerin farklı uydu görüntüleri ve daha fazla sınıf sayısı ile gerçekleştirilmesi, AÖAK sınıflandırmasında başarıyı artırabilir.



Şekil 5. RF modelinin her sınıfa ait en önemli üç özelliğinin ürettiği SHAP değerlerinin bağlılık grafikleri

KAYNAKLAR

- Atik, S. O. ve Atik, M.E.,** 2024. Optimal band selection using explainable artificial intelligence for machine learning-based hyperspectral image classification. *Journal of Applied Remote Sensing*, 18(4), 042604.
- Avcı, C., Budak, M., Yağmur, N. ve Balçık, F.,** 2023. Comparison between random forest and support vector machine algorithms for LULC classification. *International Journal of Engineering and Geosciences*, 8(1), 1-10.
- Iban, M.C. ve Sahin, E.,** 2022. Monitoring land use and land cover change near a nuclear power plant construction site: Akkuyu case, Turkey. *Environmental Monitoring and Assessment*, 194, 724.
- Kavzoglu, T. ve Bilucan, F.** 2023. Effects of auxiliary and ancillary data on LULC classification in a heterogeneous environment using optimized random forest algorithm. *Earth Science Informatics*, 16, 415-435.
- Reichstein, M., Camps-Valls, G., Stevens, B. vd.,** 2019. Deep learning and process understanding for data-driven Earth system science. *Nature*, 566, 195-204.
- Temenos, A., Temenos, N., Kaselimi, M., Doulamis, A. ve Doulamis, N.,** 2023 Interpretable Deep Learning Framework for Land Use and Land Cover Classification in Remote Sensing Using SHAP, *IEEE Geoscience and Remote Sensing Letters*, 20, pp. 1-5.
- Tuia, D. vd.** 2024, Artificial Intelligence to Advance Earth Observation: A review of models, recent trends, and pathways forward, *IEEE Geoscience and Remote Sensing Magazine* (in-press).