

YAĞ TABAKALARININ SAPTANMASINDA DESTEK VEKTÖR MAKİNELERİ VE RASTGELE ORMAN YÖNTEMLERİNİN KULLANIMI

Ümit Haluk ATASEVER¹, Coşkun ÖZKAN¹, Filiz SUNAR²

¹Erciyes Üniversitesi, Harita Mühendisliği Bölümü, 38039, Melikgazi, Kayseri, uhatasever@erciyes.edu.tr, cozkan@erciyes.edu.tr

²Istanbul Teknik Üniversitesi, Geomatik Mühendisliği Bölümü, 34469, Maslak, İstanbul, fsunar@itu.edu.tr

ÖZET

Günümüzde enerji üretiminde en fazla kullanılan kaynakların başında petrol (yağ) gelmekte ve gelişmiş ülkelerin büyük bir kısmı petrol ihtiyaçlarını tankerler aracılığı ile karşılamaktadır. Ancak yoğun petrol taşımacılığı zaman zaman petrol kaynaklı deniz kirliliğine neden olmaktadır. Günümüzde çevre uygulamaları ve sorunlarının çözümünde etkin bir araç olan uzaktan algılama ve Yapay Açıklıklı Radar (SAR) teknolojisi, yağ tabakalarının izlenmesi ve saptanmasında başarıyla kullanılmaktadır. SAR görüntülerinde yağ tabakaları düşük geri saçılım nedeniyle koyu renkte görülmektedir. Ancak radar görüntülerinde koyu renkte görülen diğer oşinografik objelerden yağ tabakalarının ayırt edilmesi önemli bir sorun teşkil etmektedir.

Bu çalışmada Lübnan kıyılarında 2006 yılında meydana gelen yağ tabakası kirliliğinin Radarsat-1 uydu verisi ile saptanmasında istatistiksel öğrenme tabanlı Destek Vektör Makinesi ve karar ağaçları temelli Rastgele Orman Yöntemi kullanılmıştır. Bu amaca yönelik olarak ortak yazarların daha önceki yaptıkları çalışmalarda kullandıkları yapay sinir ağları ve yapay arı kolonisi algoritması, diğer bir deyişle farklı algoritma ve yöntemler karşılaştırılarak, söz konusu yöntemlerin yağ tabakalarının saptanmasındaki etkinliği ve performansı irdelenmiştir.

Anahtar Sözcükler: Deniz Yağ (petrol) Kirliliği, SAR Uydu Görüntüleri, Yapay Sinir Ağları, Destek Vektör Makinesi, Rastgele Orman Yöntemi

ABSTRACT

USE OF SUPPORT VECTOR MACHINES AND RANDOM FOREST METHODES FOR OIL SPILL DETECTION

Nowadays, oil is the most commonly used resource for energy production and majority of developed countries meet needs of oil via oil tanker. However, intensive oil transportation causes marine pollution from time to time. Today, remote sensing and Synthetic Aperture Radar (SAR) technology, effective tools to solve the environmental problems and applications, are successfully used in the detection and monitoring of oil layer. Oil layers appear dark due to low backscattering in SAR image. However distinguishing of oil layers from oceanographic features which appear dark in SAR images is an important issue.

In this study, statistical learning based Support Vector Machine and Decision Tree Based Random Forest method were used for the detection of oil pollution by using Radarsat-1 data at shores of Lebanon in 2006. For this purpose, effectiveness and performance of these methods for oil detection were examined by comparing neural networks and artificial bee colony algorithm which are used in earlier studies done by co-authors.

Keywords: Oil Spill, SAR Satellite Images, Neural Networks, Support Vector Machine, Random Forest Method

1. GİRİŞ

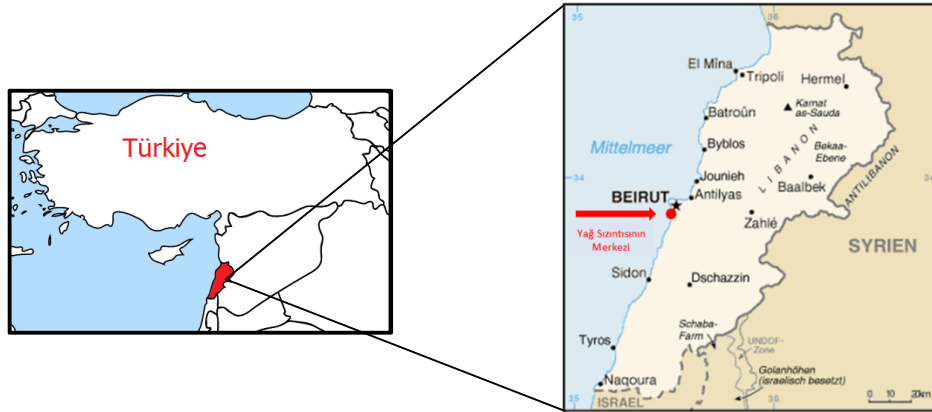
Petrol günümüzde en çok kullanılan fosil enerji kaynaklarından biridir. 2011'deki günlük dünya toplam petrol üretimi yaklaşık 87 milyon varildir (İnternet-1). Ülkelerin ihtiyaç duydukları petrol, tankerler ve boru hatlarıyla temin edilmekte ve belirtilen üretimin yaklaşık yarısı tankerlerle çeşitli deniz rotalarından iletilmektedir (İnternet-1). Bu denli büyük deniz taşımacılığı deniz kirliliğine neden olmaktadır (Sabins, 1997). Dünya genelindeki sızmış petrol miktarının 4.5 milyon tondan daha fazla olduğu tahmin edilmektedir (Bava, vd., 2002). Petrole bağlı bu kirlilik, başta memeliler ve kuşlar olmak üzere biyolojik hayatı olumsuz etkilemektedir. Bu artıkların tekrar doğada dönüşümü ise çok uzun süre almaktadır. Özellikle rafine edilmiş petrol ürünleri ham petrol kadar yapışkan olmamasına rağmen, denizler için daha zehirli ve tehlikelidir. Günümüzde uzaktan algılama teknolojisi, denizlerde oluşan yağ tabakasının tespitinde kullanılan en önemli teknolojilerden biridir. Yapay Açıklıklı Radar(SAR) ise her türlü hava şartlarında çalışabilme kabiliyetinden dolayı yağın tespitinde optik algılayıcılara nazaran daha başarılı sonuçlar vermektedir. Yağ alanlarının geri saçılım enerji seviyeleri, yağın deniz yüzeyindeki kılcak dalgaları engellenmesinden dolayı oldukça düşüktür. Bu durum doğal olarak yağ olan alanların radar görüntüsünde koyu renkte gözükmesine neden olur. Bununla beraber yağın tespiti için radar görüntülerindeki koyu renkte gözükken bölgelerin tespiti yeterli olmamaktadır. Bunun nedeni ise yağa ek olarak içsel dalgalar, denizlerdeki düşük rüzgar alanları, kara tarafından korunan kısımlarında geri saçılım enerjilerinin oldukça düşük olmasıdır. Dolayısıyla yağ tabakalarının doğal olaylardan ayırt edilebilmesi yağ tabakası tespitinde en önemli konulardan biridir (Brekke, vd.,

2005a; Sabins, 1997; Alpers, vd.,1991; Hovland, vd., 1994). Yapay Açıklıklı Radar (SAR) verisi ile yağ tabakasına bağlı kirliliğin belirlenmesinde izlenecek adımlar *i*)segmentasyon, *ii*)özellik çıkartımı ve *iii*)sınıflandırma olarak tanımlanmaktadır (Solberg, vd., 1997; Del Frate, vd.,2000; Karathanassi, vd., 2006; Brekke ve Solberg., 2005b).

Bu çalışmada, yağ ve yağ benzeri koyu renkli görüntü alanları segmente edilmiş ve yağ tabakaları tespit edilmeye çalışılmıştır. Bu amaçla istatistiksel bir öğrenme yöntemi olan Destek Vektör Makineleri ve karar ağaçları temelli bir grup (ensemble) öğrenme yöntemi olan Rastgele Orman yöntemi kullanılmıştır. Destek Vektör Makineleri yöntemi, çok boyutlu uzayda sınıfları en iyi ayıran hiper düzlemin hesaplanmaya çalışıldığı, birçok mühendislik probleminde başarılı sonuçlar veren bir yaklaşımdır. Rastgele Orman yöntemi ise, giriş verilerinin oluşturduğu özellik vektörlerinden bağımsız olarak rastgele vektörlerin oluşturulması ve bu vektörler ile elde edilen karar ağaçlarının birleştirilmesinden ibarettir. Bu çalışmada Ozkan ve arkadaşları (2011) tarafından kullanılan aynı eğitim ve test verileri dikkate alınmış ve ilgili çalışmadaki YSA bazı sonuçlar, bu çalışmada kullanılan Destek Vektör Makineleri ve Rastgele Orman yöntemlerinden elde edilen sonuçlarla karşılaştırılmıştır.

2. ÇALIŞMA ALANI VE ÖZELLİK ÇIKARTIMI

Günümüzde ENVISAT, Radarsat-1, Radarsat-2 vb. birçok radar uydusu yağ tabakalarının belirlenmesinde etkin olarak kullanılmaktadır. Bu çalışmada SAR uydu verisi olarak İstanbul Teknik Üniversitesi, Uydu Haberleşmesi ve Uzaktan Algılama Merkezi'nden temin edilen Temmuz 2007 tarihli, Lübnan kıyılarını kapsayan Radarsat-1 görüntüsü kullanılmıştır. Beyrut'un yaklaşık 19 km güneyindeki Jiyeh Dolu İstasyonunun İsrail devleti tarafından bombalanması sonrasında denize dökülen yağ tabakası büyük bir çevre felaketine neden olmuş ve Lübnan kıyı şeridinin yaklaşık 1/3'ü etkilenmiştir (Şekil 1). Ayrıca, akıntı ve hava koşullarına bağlı olarak, komşu ülkelerden biri olan Türkiye'nin de yağ tabakaları sebebiyle olumsuz etkilenmesinden endişe duyulmuştur (Ozkan, vd., 2011). Bu tip olaylara ve/veya benzeri felaketlere gecikmeden etkin bir şekilde müdahale edilebilmesi için kirliliğe maruz kalan alanların hızlı ve doğru bir şekilde saptanması gerekmektedir.



Şekil 1. Çalışma bölgesinin haritası.

Yağın tespitinde kullanılacak özellik bilgileri dört farklı kategoriye aşağıda belirtildiği şekilde ayrılmaktadır (Solberg, vd., 1997; Del Frate, vd., 2000; Karathanassi, vd., 2006; Brekke ve Solberg, 2005b).

1. Yağ tabakası alanı gibi geometrik karakteristikleri,
2. Yağın geri saçılım değerlerine ait standart sapma, ortalama, minimum ve maksimum değerleri gibi fiziksel özellikler,
3. Bağlamsal bilgi,
4. Doku bilgisi.

Yukarıda belirtilen karakteristik özellikler pikseller yerine objeler ile temsil edildiğinden piksel tabanlı sınıflandırma yerine daha çok obje tabanlı sınıflandırma yöntemleri kullanılmaktadır.

Topouzelis ve arkadaşları yukarıda belirtilen kategorilere ait toplam 25 alt özelliğin yağın tespitinde kullanılabileceğini göstermişlerdir (Topouzelis, vd., 2009). Çalışmalarında özellikle 10 özelliğin diğer belirlenen yöntemlere göre sınıflandırma doğruluğunu arttırdığını vurgulamışlardır. Bu çalışmada kullanılan 9 özellik bilgisi aşağıda sıralanmaktadır:

- Lokal Alan Kontrast Oranı (ConLa): Objenin geri saçılım değerinin ortalamasının, bölge merkezli çevrenin ortalama değerine oranıdır.

- Ortalama Orana Objeye Etkisi (Opm): Objeye değerlerinin standart sapmasının ortalama değere oranıdır.
- Çevre Alan Oranı (P/A): Objenin çevre uzunluğunun alanına oranıdır.
- Arka Plan Standart Sapması (BSd): Objeyi çevreleyen ve kullanıcı tarafından seçilen ilgi bölgesindeki piksellere ait yoğunluk değerlerinin standart sapmasıdır.
- Objeye Standart Sapması (OSd): Objenin standart sapma değeridir.
- Nesne Karmaşıklığı (C): Geometrik objenin karmaşıklığını ya da basitliğini tanımlayan değerdir.
- Ortalama Haralick Doku Değeri (THm): Alt-nesnelerinin eş oluşum matrislerinin gri seviyelerinin ortalamasıdır.
- Opm/Bpm Oranı: Ortalama orana objeye etkisinin ortalama değere arka plan etkisinin oranıdır.
- Şekil Faktörü (SP2): Objenin şeklini tanımlar. Asimetri veya form faktörü olarak da adlandırılır.

Yukarıda tanımlanan özelliklerden SP2, P/A ve C geometrik karakteristikler; THm dokusal karakteristik; BSd, ConLa, Opm, Opm/Bpm ve OSd fiziksel karakteristiklerdir (Topouzelis, vd., 2009). Segmentasyon işlemiyle oluşturulan koyu renkli nesnelere 68'i yağ tabakası, 53'ü ise yağ benzeri koyu renkli tabaka olarak UHUZAM tarafından tespit edilmiştir. Yağ ve yağ benzeri tabakalara ait istatistiksel veriler Çizelge 1'de sunulmuştur.

Çizelge 1. Yağ ve yağ benzeri verinin istatistiksel bilgileri (Özkan, vd., 2011).

Özellikler	Yağ				Yağ Benzeri			
	Min.	Maks.	Ort.	S.Sap.	Min.	Maks.	Ort.	S.Sap.
SP2	0.053	1	0.547	0.202	0.269	1	0.666	0.173
BSd	13.635	68.04	40.453	14.677	23.475	68.066	41.780	9.551
ConLa	0.373	0.85	0.547	0.088	0.440	0.663	0.570	0.062
Opm	0.177	0.457	0.257	0.066	0.169	0.297	0.207	0.025
Opm/Bpm	0.454	5.424	1.071	0.685	0.433	1.410	0.808	0.181
OSd	7.009	53.165	20.275	9.309	8.972	29.838	16.791	4.349
P/A	0.095	0.881	0.486	0.221	0.173	0.673	0.448	0.110
C	2.020	51.252	13.250	9.873	9.213	74.762	23.898	11.588
THm	11.400	36.202	22.346	5.890	12.400	35.577	21.437	5.312

3. KULLANILAN YÖNTEMLER

3.1 Destek Vektör Makineleri (DVM)

Destek Vektör Makineleri istatistiksel öğrenme alanında Vapnik ve Cortes tarafından geliştirilmiş bir öğrenme biçimidir (Cortes ve Vapnik, 1995). Uydu görüntülerinin sınıflandırmasında, DVM lineer olarak ayıramayan verinin lineer olarak ayrılabilmesi için yüksek boyuta aktararak farklı örnekler arasında maksimum sınırın bulunması işlemidir (Cortes ve Vapnik, 1995). Bu yöntemde verilerin lineer olarak ayrılabilmesi ve ayıramamaları olmak üzere iki temel durum söz konusudur. Lineer olarak ayrılabilen verilerin maksimum sınırlarının bulunması oldukça kolay olmasına rağmen, lineer olarak ayıramayan verilerin maksimum sınırının bulunması için ise, verilerin lineer olarak ayrılacakları bir üst uzaya aktararak çözüm yapılması gerekmektedir (Demirci, 2007).

3.1.1 Lineer Olarak Ayrılabilir Durumu

Sistemin eğitimi için kullanılacak verinin N elemandan oluşan $\chi = \{x_i, y_i\}, i=1,2,\dots,N$, sınıflara ait etiket değerlerinin $y_i \in \{-1, 1\}$ olduğu düşünüldüğünde, hesaplanmak istenen aşırı düzlemin üst tarafında bulunan noktalar aşağıda verilen eşitsizliğe uyar:

$$W^T X + b > 0, y_1 = +1 \quad (1)$$

Benzer şekilde aşırı düzlemin alt kısmında kalan noktalar ise aşağıda belirtilen eşitliğe uymaktadır:

$$W^T X + b < 0, y_2 = -1 \quad (2)$$

Bu iki eşitsizlik kullanılarak $y_i(W^T X + b) - 1 \geq 0$ eşitsizliği elde edilir. Bu koşulu sağlayan aşırı düzlemin iki kümenin en yakın elemanlarına olan dik uzaklıkları toplamı sınır olarak isimlendirilir. Dolayısıyla aşırı düzlemin bulunabilmesi için W (ağırlıklar) ve b (orijine olan uzaklık) değerlerinin bulunması gerekmektedir. 3 numaralı eşitlikte W değerlerinin minimum olması gerekir ve problem, optimizasyon problemi haline gelir. Bu optimizasyon probleminin çözümünde, ilk olarak eşitsizliğin çözümüne yönelik olarak Lagrange fonksiyonundan yararlanır. Problemin Lagrange denklemi ise,

$$L_p = \frac{1}{2} \|w\|^2 - \sum_{i=1}^N \alpha_i y_i (w^T x_i + b) + \sum_{i=1}^N \alpha_i \quad (3)$$

şekindedir. Eşitlikte $\alpha_i > 0$ değerleri destek vektörlerini tanımlamaktadır. Denklemin çözümünün karmaşık olmasından dolayı Karush-Kuhn-Tucker koşulları kullanılarak dual forma dönüştürülen problem tekrar Lagrange formülasyonunda yerine konulduğunda,

$$\sum_i \alpha_i - \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j y_i y_j x_i^T x_j \quad \alpha_i \geq 0, \forall i \quad (4)$$

eşitliği elde edilir. Bu eşitlik ile tüm eğitim verileri için bir Lagrange çarpanı hesaplanır ve $\alpha_i > 0$ olan x_i örnekleri destek vektörlerini tanımlamaktadır. Bu destek vektörler aracılığı ile aşırı düzlemler ve ardından da optimum aşırı düzlem hesaplanır (Özkan, 2008).

3.1.2 Lineer Olarak Ayrılamama Durumu ve Çekirdek Fonksiyonlar

Birçok gerçek dünya problemde olduğu gibi uydu verilerinin sınıflandırılmasında da verideki gürültü ve veri sınıflarının doğrusal olmaması gibi nedenlerden dolayı veri setinin doğrusal olarak ayrılabilmesi mümkün olmamaktadır. Bu problem ise optimizasyona hataları temsil eden bir değişkenin ($\xi_i, i=1, 2, \dots, N$) eklenmesi ile çözülmeye çalışılır. Yeni eşitsizlikler ise,

$$w^T x_i + b \geq 1 - \xi_i, y_i = +1 \quad (5)$$

$$w^T x_i + b \leq -1 + \xi_i, y_i = -1 \quad \xi_i \geq 0, \forall i \quad (6)$$

şeklinde elde edilir.

Denklemden $\xi_i = 0$ ise veri doğru sınıflandırılmış, $0 < \xi_i < 1$ durumunda veri doğru sınıflandırılmış, H_1 ve H_2 aşırı düzlemleri arasında yer alıyor; son olarak $\xi_i \geq 1$ ise sınıflandırma işlemi yanlış yapılmış demektir.

Lineer olarak ayrılamama durumunda sınıflandırma hatasını minimum yaparken, aşırı düzlemler arasındaki uzaklığı maksimum yapacak, Lagrange çarpanlarının üst sınırını gösteren bir C parametresi eklenir. Lagrange denkleminin yeni formu ise aşağıdaki eşitlik ile gösterilmektedir:

$$L_p = \frac{1}{2} \|w\|^2 + C \sum_i \xi_i - \sum_i \alpha_i \{y_i (w^T x_i + b) - 1 + \xi_i\} - \sum_i \mu_i \xi_i \quad (7)$$

Yukarıdaki eşitliğe KKT koşulları uygulandığında ise,

$$L_d = \sum_i \alpha_i - \frac{1}{2} \sum_{i,j} x_i^T x_j y_i y_j \alpha_i \alpha_j \quad 0 \leq \alpha_i \leq C, \forall i \quad (8)$$

elde edilir. Bu fonksiyondaki $x_i^T x_j$ çarpımı daha büyük boyutlu özellik uzayında tekrar tanımlanması gerekmektedir. x_i^T nin dönüşümü için $\phi(x_i)^T$ ve x_j dönüşümü için $\phi(x_j)$ fonksiyonları kullanılır. Diğer bir deyişle eşitlik aşağıdaki forma dönüştürülmüş olur (Atasever, 2011):

$$L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j \phi(x_i)^T \phi(x_j) \quad (9)$$

Yukarıdaki eşitlikte $\phi(x_i)^T \phi(x_j)$ çarpımı için çekirdek fonksiyonları kullanılabilir. Çekirdek fonksiyonları ile verilerin yüksek boyutlu ve doğrusal olarak ayrılacakları uzaya taşınarak ve çözümün bu yüksek boyutlu uzayda aranması için uzaylar arası dönüşüm yapılabilir (Atasever, 2011). En sık kullanılan çekirdek fonksiyonlar Çizelge 2'de verilmiştir (Özkan, 2008):

Çizelge 2. Destek Vektör Makineleri'nde sıkça kullanılan çekirdek fonksiyonlar

Kernel Fonksiyonu	Matematiksel İfadesi	Parametre
Polinom Kerneli	$K(x, y) = ((x \cdot y) + 1)^d$	Polinom derecesi (d)
Lineer Kernel	$K(x, y) = (x \cdot y)$	Genel Kernel İfadesi
Kuadratik Kernel	$K(x, y) = x \cdot y(1 + x \cdot y)$	Genel Kernel İfadesi
Radyal Tabanlı Fonksiyon Kerneli	$K(x, y) = e^{-\gamma \ x-y\ ^2}$	Kernel boyutu (γ)

3.2. Rastgele Orman Yöntemi

Rastgele Orman yöntemi Breiman tarafından karar ağaçları (decision trees), torbalama (bagging) ve rastgele alt-uzay (random subspace) yöntemlerinin birleştirilmesi ile elde edilmiş bir yöntemdir (Breiman, 2001). Torbalama ile eğitim verilerinden rastgele yeni örneklem oluşturulur. Rastgele alt uzay ile de karar ağacına ait her bir düğümde en iyi dallara ayrılan değişken, tüm değişkenler arasından rastgele seçilen değişkenler arasından seçilir. Rastgele orman yönteminde de başka yöntemlerde olduğu gibi kullanıcı tarafından belirlenmesi gereken parametreler vardır. Bunlar oluşturulacak ağaç sayısı ve kullanılacak örneklem miktarıdır (Pal, 2005). Rastgele orman sınıflandırıcısının oluşturulmasında CART yöntemi esas alınır. Dallanma kriteri olarak da sınıfların safsızlığını ölçen Gini indeksi dikkate alınır. Herhangi bir T düğümü için Gini İndeksi aşağıda verilen eşitlik ile hesaplanır:

$$Gini(T) = 1 - \sum_j [p(j|T)]^2 \quad (10)$$

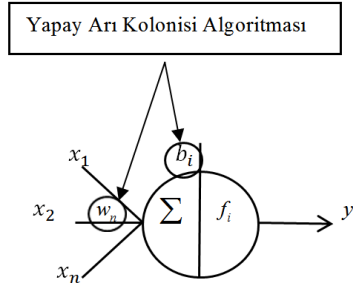
Bu eşitlikte $p(j|T)$, T düğümündeki J sınıfına ait bağıl olasılığı göstermektedir (İnternet-3).

3.3. Yapay Arı Kolonisi(ABC) Algoritması İle Eğitilmiş Yapay Sinir Ağları

Yapay sinir ağlarının kullanıldığı bir çok çalışmada ağırlıkların ve bias (eşik) değerlerinin hesaplanmasında, klasik optimizasyon değerleri yerine, yapay zeka optimizasyon değerleri kullanılmaya başlanmıştır. Yapay Arı Kolonisi algoritması, bu amaçla kullanılan optimizasyon yöntemleri içinde oldukça başarılı sonuçlar veren bir yaklaşımdır. Doğadaki arıların davranışlarından esinlenilerek geliştirilen bu yöntemde; bir kolonide işçi, gözcü ve kâşif arılar olmak üzere üç grup arı bulunmaktadır. Algoritmada, her bir besin olası çözümü ifade eder ve kovan için en uygun besin kaynağı (optimum çözüm) bulunmaya çalışılır. ABC Algoritmasının temel adımları aşağıda verilmektedir (Karaboga, 2011):

1. Başlangıç besin kaynağı bölgelerinin üretilmesi
- TEKRARLA
2. İşçi arıların besin kaynak bölgelerine gönderilmesi
 3. Olasılıksal seleksiyonda kullanılacak olasılık değerlerinin hesaplanması
 4. Gözcü arıların olasılık değerlerine göre besin kaynak bölgelerini seçmeleri
 5. Terk edilecek kaynakların terk edilmesi ve kâşif arıların üretimi
- ŞU ANA KADAR (Maks Iterasyon)

Yağ tabakası tespiti için YSA kullanılan çalışmada, ağırlıklar ve eşik(bias) değerlerinin hesaplanması için ABC algoritması kullanılmış ve Levenberg-Marquardt yaklaşımına göre oldukça başarılı sonuçlar elde edilmiştir. Şekil 2' de, ABC algoritması ile Yapay Sinir Ağlarının eğitimi görselleştirilmiştir.



Şekil 2. YSA eğitiminde yapay arı kolonisi algoritmasının kullanımı.

Şekil 2’de w_n ve b_i sırasıyla algoritma tarafından hesaplanan ağırlık ve eşik(bias) değerlerini, x_1, x_2, \dots, x_n giriş; y_i çıkış; f_i ise transfer fonksiyonunu belirtmektedir.

4. UYGULAMA

Destek Vektör Makineleri ve Rastgele Orman yöntemi, RADARSAT-1 uydusuna ait görüntüden elde edilen 9 özellik bilgisine uygulanmış ve sonuçlar bildirim ortak yazarları tarafından daha önceki çalışmalardan elde edilmiş olan sonuçlarla karşılaştırılmıştır. Yazarlar makalelerinde yağ tabakalarının belirlenmesinde en çok kullanılan yöntemlerden olan yapay sinir ağlarının eğitiminde klasik optimizasyon yöntemlerinden Levenberg-Marquardt’a ek olarak başarılı bir yapay zeka optimizasyon yöntemi olan Yapay Arı Koloni Algoritması’nı kullanmışlardır (Ozkan, vd., 2011). Uygulamanın sonucunda ise ABC ile elde edilen sonuçların, Levenberg-Marquardt’a göre daha robust ve başarılı olduğu görülmüştür. Önceki çalışmalardan elde edilen sonuçlar ve Destek Vektör Makineleri ile Rastgele Orman Yönetimi kullanılarak elde edilen sınıflandırma doğrulukları Çizelge 3’te sunulmuştur.

Çizelge 3. Kullanılan yöntemlere ait sınıflandırma doğrulukları.

	Destek Vektör Makineleri				Rastgele Orman Yöntemi			YSA	
	Linear	Polinomial ($d=2$)	RBF ($\gamma=5, C=100$)	Kuadratik	Ağaç Sayısı			LM	ABC
					10	50	100		
Genel Doğruluk(%)	100	100	100	100	100	100	100	89.29	91.07

* LM: Levenberg-Marquardt, ABC: Artificial Bee Colony

5. SONUÇLAR

Günümüzde yağ tabakası kirliliği denizleri ve deniz canlılarını etkileyen en önemli tehditlerden biri haline gelmiştir. Bu tehdidin üstesinden gelinebilmesi ancak yağ tabakasının çok hızlı ve hassas olarak saptanmasına bağlıdır. Yapay Açıklıklı Radar tüm hava ve ışık koşullarında yağın belirlenmesine olanak sağladığı için en çok kullanılan veri kaynaklarından biridir. Literatürde yağ tabakası tespitinde yapay açıklıklı radar verilerini kullanarak başarılı sonuçlar alınan birçok uygulama mevcuttur. Yağ tabakasının saptanmasındaki en büyük zorluk ise, yağ ile yağ benzeri alanların birbirinden ayırt edilmesidir. Bu amaçla ilk olarak segmentasyon işlemiyle görüntüdeki koyu renkli objeler çıkartılarak uygun özellikler hesaplanır ve son olarak seçilen bir sınıflandırıcı ile yağ ve yağ benzeri objeler ayırt edilir.

Sınıflandırma işleminin doğruluğunu arttıran en önemli etkenlerden biri tercih edilecek sınıflandırma yöntemidir. Yağın tespitinde kullanılan en popüler yöntemlerden biri Yapay Sinir Ağları (YSA)dır. Yapay zeka optimizasyon yöntemlerinin gelişimi ve YSA eğitiminde (eşik (bias) ve ağırlıkların hesaplanmasında) kullanılmaya başlanması ile bazı çalışmalarda bu yöntemin doğruluğu arttırılmaya çalışılmıştır.

Bu çalışmada yağ ile yağ benzeri bölgelerin sınıflandırılması için Destek Vektör Makineleri ve Rastgele Orman yöntemleri kullanılmış ve bildiri ortak yazarlarına ait başka bir çalışmadaki Yapay Arı Kolonisi, Levenberg-Marquardt optimizasyon yöntemleri ile eğitilmiş YSA sonuçlarıyla kıyaslanmıştır. Söz konusu çalışmadaki aynı veri grubu ile yapılan analizler sonucunda, Destek Vektör Makineleri ve Rastgele Orman yöntemlerinin tüm eğitim ve test verilerini %100 genel doğrulukla hatasız olarak sınıflandırdığı gözlemlenmiştir. Destek Vektör makineleri ve Rastgele Orman yöntemlerinin birbirlerine göre başarı durumlarını irdeleyebilmek için ileriki çalışmalarda farklı uydu verilerinin kullanılması planlanmaktadır. Son olarak bu çalışma ile özellikle yağ tabakalarının saptanmasında, diğer bir ifade ile yağ tabakasının diğer koyu renkli objelerden ayırt edilmesinde bu iki yöntemin kullanımının son derece başarılı sonuçlar üreteceği gösterilmiştir.

KAYNAKLAR

- Alpers, W., Wismann, V., Theis, R., Huhnerfuss, H., Bartsch, N., Moreira, J., ve Lyden, J.**, 1991, The Damping of Ocean Surface Waves by Monomolecular Sea Slicks Measured by Airborne Multi Frequency Radars During the SAXON-FPN Experiment, *Proceedings of the International Geoscience and Remote Sensing Symposium*, Helsinki, Finland,, pp.1987-1990.
- Atasever, Ü., H.**, 2011, Uydu Görüntülerinin Sınıflandırılmasında Hızlandırma (Boosting), Destek Vektör Makineleri, Rastgele Orman (Random Forest) ve Regresyon Ağaçları Yöntemlerinin Kullanılması, *Yüksek Lisans Tezi*, Erciyes Üniversitesi, Kayseri.
- Bava, J., Tristan, O., ve Yasniouski, J.**, 2005, Earth Observation Applications Through Systems and Instruments of High Performance Computer, ASI/CONEA training course, September- December, Matera, Italya.
- Breiman, L.**, 2001, Random Forests, *Machine Learning*, 45(1), pp.5-32.
- Brekke C., ve Solberg A. H. S.**, 2005, Feature Extraction for Oil Spill Detection Based on SAR Images. *SCIA 2005*, LNCS, 3540, pp.75-84.
- Brekke, C., ve Solberg, A., H., S.**, 2005, Oil Spill Detection by Satellite Remote Sensing, *Remote Sensing of Environment*, 95(1), 1-13.
- Cortes, C. ve Vapnik, V.**, 1995, Support-vector Network, *Machine Learning*, Kluwer Pub.,Boston, Vol.20, pp.273–297.
- Del Frate F., Petrocchi A., Lichtenegger J., ve Calabresi G.**, 2000, Neural Networks for Oil Spill Detection Using ERS-SAR Data, *IEEE Transactions on Geoscience and Remote Sensing*, Vol.38, pp.2282-2287.
- Demirci, D.**, 2007, Destek Vektör Makineleri İle Karakter Tanıma, *Yüksek Lisans Tezi*, Yıldız Teknik Üniversitesi, İstanbul.
- Hovland, H., A., Johannessen J. A., ve Digranes G.**,1994, Slick Detection in SAR Images, *Proceedings of the International Geoscience and Remote Sensing Symposium*, Pasadena CA, pp. 2038-2040.
- İnternet-1**, <http://www.eia.gov/countries/regions-topics.cfm?fips=WOTC>, 18.06.2012
- İnternet-2**, <http://tr.wikipedia.org/wiki/L%C3%BCbnan>, 15.06.2012
- İnternet-3**,
http://www.hypertextbookshop.com/dataminingbook/working_version/contents/chapters/chapter001/section003/blue/page002.html, 20.06.2012
- Karaboga, D.**, 2011, Yapay Zeka Optimizasyon Algoritmaları, *Nobel Basım Dağıtım*, Türkiye.
- Karathanassi, V., Topouzelis, K., Pavlakis, P., ve Rokos, D.**, 2006, An Object-oriented Methodology to Detect Oil Spills. *International Journal of Remote Sensing*, Vol.27, pp.5235-5251.
- Ozkan, C., C. Ozturk, Sunar, F., ve Karaboga, D.**, 2011, The Artificial Bee Colony Algorithm in Training Artificial Neural Network for Oil Spill Detection, *Neural Network World*, 21(6), 473-492.
- Özkan, Y.**, 2008, Veri Madenciliği Yöntemleri, s.187-188, *Papatya Yayıncılık*, İstanbul
- Pal, M.**, 2005, Random Forest Classifier for Remote Sensing Classification, *International Journal Of Remote Sensing*, Vol.26, pp.217-222,
- Sabins, F., F.**,1997, Remote Sensing, Principles and Interpretation (W. H. Freeman and Company, New York).
- Solberg, R., ve Theophilopoulos, N. A.**, 1997, Envisys – A Solution for Automatic Oil Spill Detection in the Mediterranean, *Proceedings of 4th Thematic Conference on Remote Sensing for Marine and Coastal Environments*, Environmental Research Institute of Michigan, Ann Arbor, Michigan, pp.3-12.
- Topouzelis, K., Stathakis, D., ve Karathanassi, V.**, 2009, Investigation of Genetic Algorithms Contribution to Feature Selection for Oil Spill Detection, *International Journal of Remote Sensing*, Vol.3, pp.611-625.