

RADYOGRAFİ GÖRÜNTÜLERİ VE SINIFLANDIRMA ALGORİTMALARI KULLANILARAK OMUZ PROTEZLERİNİN ÜRETİCİLERİNİN BELİRLENMESİ

Ebru Efeoğlu¹, Gürkan Tuna^{2*}

¹ İstanbul Gedik Üniversitesi, İktisadi İdari ve Sosyal Bilimler Fakültesi, Yönetim Bilişim Sistemleri, İstanbul

² Trakya Üniversitesi, Edirne Teknik Bilimler MYO, Edirne

Öz

Omuz protezlerinin zamanla farklı nedenlerden dolayı bakımının yapılması ya da değiştirilmesi gerekebilir. Bu bakım işlemleri yine ameliyatlara yapılır. Farklı türlerde ve farklı üreticiler tarafından üretilmiş omuz protezleri bulunmaktadır ve her birinin çıkarılmasında ve bakımının yapılmasında farklı ekipmanlar kullanılması gereklidir. Protez türü ile ilgili yeterli bilginin sağlanamadığı durumlarda bazı sorunlar ile karşılaşılabilir. Radyografi görüntülerinin görsel muayenesi ve karşılaştırılmasının uzmanlar tarafından yapılması hem yorucudur hem de süreci uzatır. Ameliyattan önce doğru donanım ve prosedürlerin seçilmesi için ameliyatı gerçekleştirecek olan cerraha bilinmeyen protezleri tanımda yardımcı olacak, hızlı ve yüksek doğruluk oranına sahip bir çözüme ihtiyaç duyulmaktadır. Bu çalışmada 3 farklı üreticiye ait omuz protezlerinin radyografi görüntülerinden tanınması için 12 farklı sınıflandırma algoritması kullanılmış ve bu algoritmaların performansları karşılaştırılmıştır. K-En Yakın Komşu algoritmasının diğer algoritmalara göre daha iyi performans sergilediği görülmüştür. Radyografi görüntülerinden protez tanımda bu algoritmanın kullanılmasının doğru tercih olacağı ve diğer protez türlerini tanımda da kullanılabileceği düşünülmektedir.

Anahtar Kelimeler: Omuz protezleri, Radyografi görüntüleri, Sınıflandırma algoritmaları, K-en yakın komşu algoritması, Performans analizi

DETERMINATION OF MANUFACTURERS OF SHOULDER PROSTHESES USING RADIOGRAPHY IMAGES AND CLASSIFICATION ALGORITHMS

Abstract

Shoulder prostheses may need to be maintained or replaced over time for different reasons. These maintenance procedures are also performed by surgeries. There are different types of shoulder prostheses produced by different manufacturers, and different equipment is required to remove and care for each. In cases where sufficient information about the prosthesis type cannot be provided,

Sorumlu Yazar: Gürkan TUNA, gurkan@trakya.edu.tr

some problems may be encountered. Visual examination and comparison of radiographic images by experts is both tiring and prolonged. In order to select the correct equipment and procedures before surgery, a fast and highly accurate solution is needed to assist the surgeon who will perform the operation in identifying unknown prostheses. In this study, 12 different classification algorithms were used to identify shoulder prostheses from 3 different manufacturers from radiographic images and the performances of these algorithms were compared. It has been observed that K-Nearest Neighbor algorithm performs better than other algorithms. It is thought that this algorithm will be the right choice for prosthesis recognition from radiography images and can be used to identify other prosthesis types.

Keywords: Shoulder prostheses, Radiography images, Classification algorithms, K-nearest neighbor algorithm, Performance analysis

1. GİRİŞ

Özellikle yaşlılarda kemik kırılabilirliği artmaktadır. Eklemlerdeki kırılma veya deformasyon durumlarında en iyi tedavi yöntemlerinden biri protez takma ameliyatlarıdır [1, 2]. Fakat bu protezlerde zamanla meydana gelebilecek hasarlar hastada acıya neden olmaktadır. Bunun tedavisi için yapılması gereken protezlerin bakımının yapılması veya değişimidir. Çeşitli protezler bulunmaktadır ve protezlerin revizyon oranı artmaktadır [3, 4]. Hasarlı omuz ekleminin protez ile değiştirildiği ameliyat türüne Total Omuz Artroplastisi (TOA) adı verilir. Revizyon cerrahisinin başarısı, mevcut protezler hakkındaki tip ve boyut gibi bilgilerin tespitine bağlıdır.

Radyografi görüntülerinden tespit ve sınıflandırma yapabilmek için farklı algoritmalar denenmiştir. Örneğin radyografi görüntülerinden diz osteoartritini tespit etmek için görüntü işleme ve sınıflandırma tekniklerine dayalı bir algoritma önerilmiştir [5]. Şablon eşleştirme [6] ve görüntü işleme [7] teknikleri kullanılarak radyografi görüntülerinde diz protezi modellerinin tanımlaması yapılmıştır. Fakat kontrast ayarı, gürültü vb. nedenlerden dolayı şablon eşleştirme yönteminde olumlu sonuçlar alınamamıştır. Öte yandan, makine öğrenmesine dayalı yöntemler daha pratik olup, yüksek doğruluğa sahip sonuçlar vermektedir. Literatürde K-En Yakın Komşu (K-EYK) [8], Destek Vektör Makinesi (DVM) [9], Naive Bayes (NB) [10], Lojistik Model Ağacı (LMA) [11], Ekstra Ağaç (EA) [12] ve Hoefding Ağacı (HA) [13] da dahil olmak üzere birçok makine öğrenmesi yöntemi bulunmaktadır ve bu yöntemler radyografi görüntülerinden sınıflandırma yaparken kullanılabilir.

Makine öğrenme algoritmaları insan kemik röntgen görüntülerinden kırık tespitinde [14], omurga

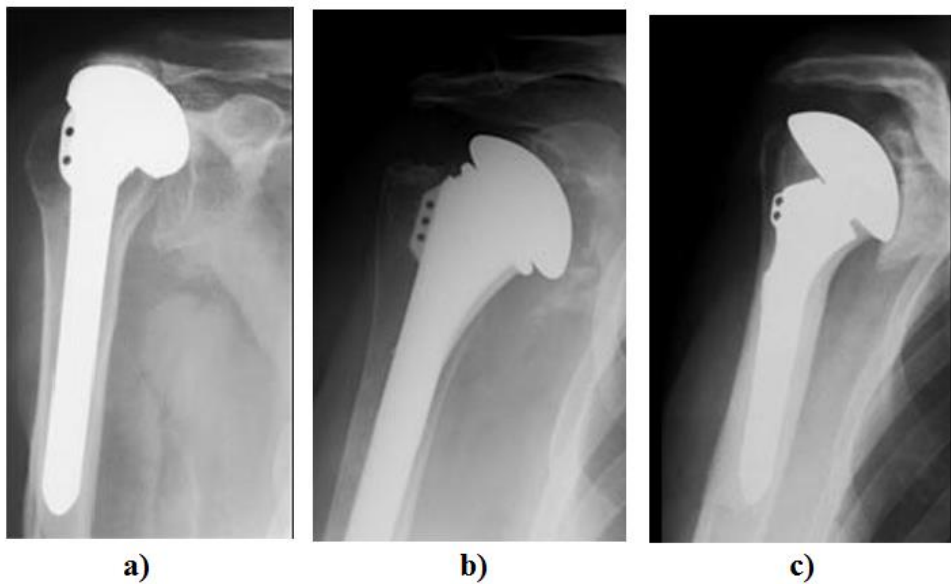
röntgen görüntülerinden kemik yoğunluğunu tahmin etmede [15], kalça protezi tanımada [16], farklı diş implant markalarının doğruluğunu sınıflandırmada [17], diz osteoartritinde [18] kullanılmıştır. Kotti, insan yürüyüşünün karmaşıklığı ve diz osteoartriti incelediği bir çalışmada NB algoritmasını kullanmıştır [19]. Geçmişte omuz protezlerinin sınıflandırılmasında Hough dönüşümü [20] yöntemi yaygın olarak kullanılmıştır. Son yıllarda, omuz protezlerinin sınıflandırılması için bölümlenme tabanlı bir algoritma kullanılmıştır [21]. TOA modellerini saptamada ve tanımlamada Derin Evrişimli Sinir Ağlarının performansını geliştirmek ve değerlendirmek için de çalışmalar yapılmıştır [22, 23]. Bu makalede farklı üreticilere ait omuz protezlerinin radyografi görüntülerinden tanınması için en popüler makine öğrenme yöntemleri kullanılmış ve performans değerlendirmesi yapılmıştır.

2. MATERYAL VE METOT

Bu bölümde, çalışmada kullanılan veri seti, sınıflandırma algoritmaları ve performans metrikleri hakkında bilgiler verilmektedir.

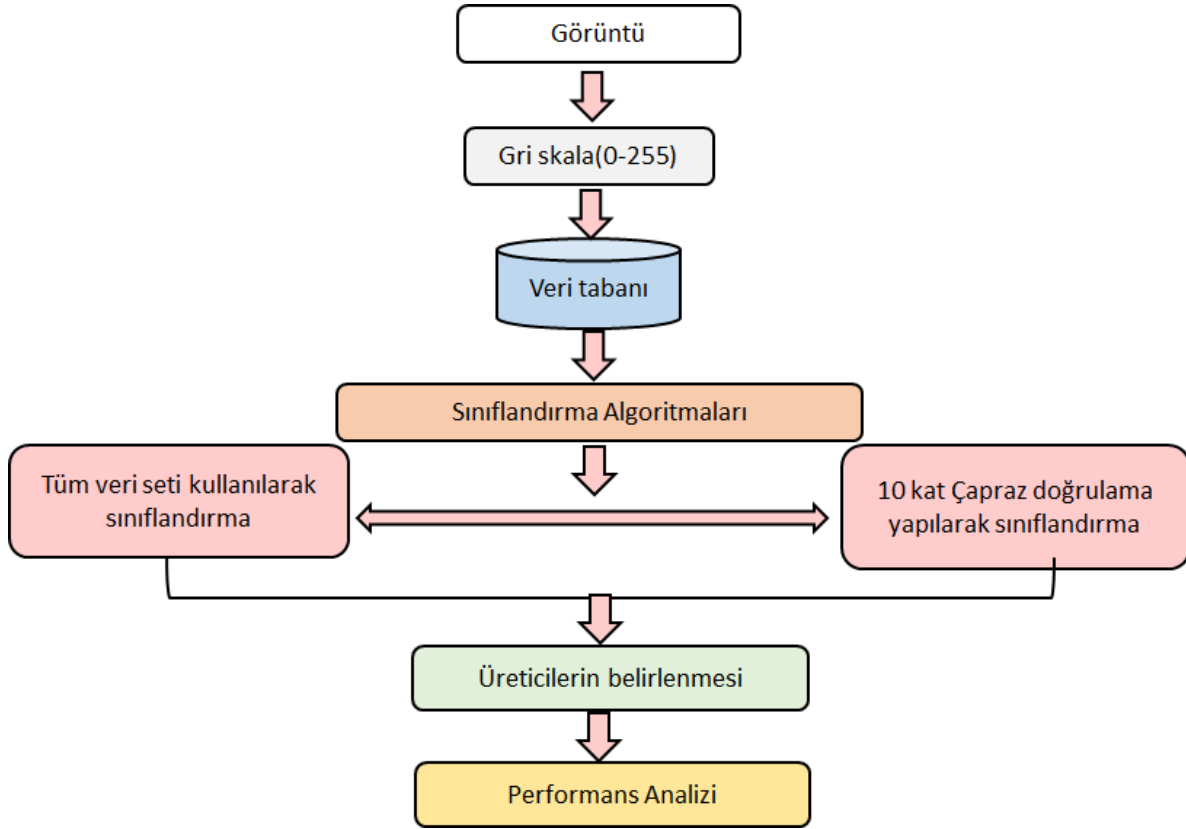
2.1. Veri Seti

Bu çalışmada kullanılan veri seti, farklı üreticiler tarafından üretilmiş omuz protezlerine ait 349 adet radyografi görüntüsünden oluşmaktadır. Veri setinde 67 adet Cofield, 239 adet Depuy ve 43 adet Zimmer tarafından üretilmiş omuz protezlerine ait görüntüler bulunmaktadır. Şekil 1'de veri setinden temsili örnek radyografi görüntüleri gösterilmektedir.



Şekil 1. Veri seti örnekleri a) Cofield b) Depuy c)Zimmer

Bu görüntüler, en yüksek değer 255 ve en düşük değer ise 0 olacak şekilde gri tonlamalı görüntüye dönüştürülmüş ve bu değerler sınıflandırma algoritmalarında kullanılmıştır. Yöntemin akış diyagramı Şekil 2’de verilmiştir.



Şekil 2. Sınıflandırma çalışmasının akış diyagramı

2.2. Sınıflandırma Algoritmaları

NB: Bu yöntemde, bir test veri kümesinden bir verinin doğru sınıfını bulma olasılığı, o sonucu etkileyen tüm faktörlerin olasılıklarının çarpılması ile hesaplanır [24]. Bu hesaplamalar sonucunda, test verilerinin sınıfı belirlenir [24].

DVM: DVM, istatistiksel öğrenme teorisi alanındaki bir yöntem olup, bir düzlemde bulunan örnekler arasına bir sınır çizerek örnekleri iki gruba ayırır [9]. İlgili sınırın çizileceği yer, iki grubun üyelerine en uzak olan yer olup, iki gruba da yakın ve aynı zamanda birbirine paralel olan iki çizgi çizilir ve bu çizgiler birbirine yaklaştırılarak sınır çizgisi elde edilir. DVM, veriyi birbirinden ayırmak için en uygun fonksiyonun tahmin edilmesine dayanır.

Sıralı Minimal Optimizasyon Algoritması (SMO): DVM ve Karesel Programlama (KP) problemini ekstra matris alanı ve nümerik KP optimizasyon adımları kullanmadan hızlıca çözebilen bir algoritmadır [25]. SMO yakınsamayı koruyarak Osuna teoremi yoluyla tüm KP problemlerini alt problemlere ayrıştırmaktadır. SMO her adımda en küçük optimizasyon problemini çözmeyi seçmektedir. Lagrange çarpanları lineer eşitlik kurallarına uymak zorunda olduğundan standart DVM KP problemi için en küçük optimizasyon problemi iki Lagrange çarpanı içermektedir. Her adımda SMO iki Lagrange çarpanını birlikte optimize ederek bu çarpanlar için en iyi değeri bulmaktadır ve DVM’i yeni en iyi değerleri yansıtarak güncellemektedir.

K-EYK Algoritması: Öznitelik uzayındaki en yakın eğitim örneklerine dayanarak nesnelere sınıflandırmayı amaçlar. K-EYK algoritmasında test edilecek örnek, eğitim kümesindeki her bir örnek ile tek tek işleme alınır. Test edilecek örneğin sınıfını tespit etmek için eğitim kümesindeki ilgili örneğe en yakın k adet örnek seçilir. Seçilmiş olan örneklerden oluşan küme içerisinde hangi sınıfa ait en çok örnek varsa test edilen örnek bu sınıfa aittir. Örnekler arası uzaklıklar farklı uzaklık ölçütleriyle belirlenebilmekte olup, sıklıkla Öklid (*Euclidean*) uzaklığı tercih edilir.

K Algoritması*: Algoritma, test veri setinde bulunan, özneteliği bilinmeyen bir örneğin, veri tabanı içerisinde daha önceden sınıflandırılmış ama ortaya çıkmamış gizlide kalmış olan eğitim veri setindeki örnekler ile karşılaştırılması esasına dayanır [26].

Doğrusal Ayrıcılık Analizi (DAA): Matematiksel olarak DAA sınıf içi ve sınıf içi değişkenlik oranını maksimuma çıkarır. DAA, sınıflar arasında ayırım yapmak için özellik alanındaki bir hiper düzlemi temsil eden verilen sınıflar arasında doğrusal bir karar bölgesi çizer. Sınıf sayısının ikiden büyük olması durumunda, birden fazla doğrusal düzlem çizilecektir [27].

REPTree Algoritması: İlk olarak Quinlan tarafından önerilen bu algoritmada [28] regresyon ağacı mantığı kullanılmaktadır. Bu mantıkla farklı yinelemelerde birçok ağaç oluşturulmakta ve daha sonra oluşturulan bu karar ağaçları arasından en iyi ağaç seçilmektedir. Kısaca, REPTree varyansdan kaynaklanan hatayı en aza indirme ilkesine ve entropi ile bilgi kazanımına dayanmaktadır [29].

Rastgele Model Ağaçlar (RMA) Algoritması: Bu algoritma özellikle büyük miktarda verilerin sınıflandırılmasında gösterdiği iyi performans nedeniyle tercih edilmektedir. Algoritma tek bir karar ağacı üretmek yerine her biri farklı eğitim kümelerinde eğitilmiş olan çok sayıda karar ağacını birleştirmektedir [30].

CART Algoritması: Bu algoritmada karar ağaçları, her bir karar düğümünün çeşitli ayrılma kriterleri kullanılarak iki farklı dala ayrılmasıyla oluşturulmaktadır [31].

LMA Algoritması: Bu algoritma, ağacın her düğümünde bir lojistik regresyon modeli üretmek için LogitBoost algoritması kullanır. Daha sonra düğümler C4.5 kriterleri kullanılarak bölünür [32].

EA Algoritması: Algoritma, birden fazla ilişkisiz karar ağacının sonuçlarını bir araya getiren bir toplu öğrenme tekniğidir. RMA algoritmasına benzerlik göstermektedir. Her karar ağacı, orijinal eğitim örneğinden oluşturulur. Daha sonra, her test düğümünde, Her ağaç, her bir karar ağacının verileri Gini İndeksine göre bölünür. En iyi özelliği seçmesi gereken özellik kümesinden rastgele bir k özellik örneği ile sağlanır. Bu rastgele özellik örneği, birden çok ilişkisiz karar ağacının oluşturulmasına yol açar.

HA Algoritması: Bu algoritma ilk olarak Hulten ve ark. tarafından ileri sürülmüştür [33]. HA algoritması, karar ağacının her bir düğümünde, düğümün nasıl parçalanacağına ilişkin kararı verebilmek için Hoeffding sınırı olarak bilinen istatistiksel bir değeri kullanır. Her bir örneği en çok bir kez okur ve uygun bir zaman aralığında işler [33].

2.3. Performans metrikleri

Algoritmaların sınıflandırma performansını karşılaştırmak için kullanılan bazı performans ölçütleri vardır. Bu ölçümler, belirli bir sınıflandırma probleminde hangi sınıflandırma algoritmasının daha iyi performans sergilediğini gösterir. Sınıflandırma yapıldığında dört farklı durum söz konusudur. Bu durumlar aşağıda belirtildiği şekilde açıklanabilir.

- *DP (Doğru pozitif)* gerçek durum pozitifken test sonucunun pozitif olması,
- *YP (Yanlış pozitif)* gerçek durum negatifken test sonucunun pozitif olması,
- *DN (Doğru negatif)* gerçek durum negatifken test sonucunun negatif olması,
- *YN (Yanlış negatif)* gerçek durum pozitifken test sonucunun negatif olması durumunu ifade etmektedir.

Bu durumlar birçok sınıflandırma ölçütünde kullanılarak ölçüt değerleri hesaplanır ve bu ölçütlerle sınıflandırma performans değerlendirme yapılır [34]. Sıklıkla kullanılan performans ölçütleri bu çalışmada kullanılmıştır.

Doğruluk: Doğru sınıflandırılmış numunelerin toplam örnekleme oranını veren bir ölçüttür.

Sınıflandırma algoritmasının doğruluk oranı denklem (1) kullanılarak hesaplanır.

$$\text{Doğruluk} = \frac{DP+DN}{DP+DN+YP+YN} \quad (1)$$

Sınıflandırma algoritmasının tüm tahminleri sayıca fazla olan sınıftan yapması durumunda, algoritma neredeyse hiçbir şey öğrenmediği halde doğruluk oranı yüksek çıkabilir. Bu nedenle sadece doğruluk ölçütüne bakılarak performans değerlendirmesi yapmak yeterli değildir [35].

Kesinlik: Tahmin edilen sınıftaki örneklerin gerçekten kaç o sınıfta sorusuna verilen cevaptır. Düşük kesinlik değeri, birçok yanlış pozitifin bir göstergesi olarak da düşünülebilir [36]. Denklem (2) kullanılarak hesaplanır.

$$\text{Kesinlik} = \frac{DP}{DP+YP} \quad (2)$$

Duyarlılık: Tespit edilmek istenen sınıfı doğru tespit etme oranıdır. Doğru pozitif değerinin doğru pozitif ile yanlış negatif toplamına oranıdır (3).

$$\text{Duyarlılık} = \frac{DP}{DP+YN} \quad (3)$$

Yüksek kesinlik ve duyarlılık değerlerine sahip bir sınıflandırma algoritması iyi bir sınıflandırma yapmış demektir. Bu ölçütler 0 ile 1 arasında değer alırlar ve ölçütlerin 1'e yakın bir değer alması sınıflandırmanın başarısının yüksek olduğunu göstermektedir.

F1 skoru: Kesinlik ve duyarlılık arasında bir denge aranması gerektiğinde veya eşit olmayan bir sınıf dağılımı olduğunda kullanımı faydalı olan bir ölçüttür. Bu ölçüt kesinlik ve duyarlılık ölçütlerinin harmonik ortalaması alınarak hesaplanır (4).

$$F1 \text{ skoru} = 2 \times (\text{Duyarlılık} \times \text{Kesinlik}) / (\text{Duyarlılık} + \text{Kesinlik}) \quad (4)$$

Kappa: Gerçek durum ve tahmin edilen durumun ne kadar uyumlu olduğunu gösterir. Kappa istatistik değeri -1 ile +1 arasında değişen değerler alır. Kappa değeri 1 olması mükemmel uyumun var olduğu anlamına gelir. Bu nedenle Kappa değeri 1'e ne kadar yakınsa algoritma o kadar iyi bir sınıflandırma yapmıştır. Denklem (5) kullanılarak hesaplanır.

$$\text{Kappa} = \frac{P(a)-P(e)}{1-P(e)} \quad (5)$$

Burada, P(a) algoritmanın doğruluğunu P(e), aynı veri kümesinde rastgele tahminler yapan algoritmanın beklenen doğruluğunun ağırlıklı ortalamasını ifade eder.

Kök Ortalama Karese Hata (KOKH): Algoritmaların ne kadar hata ile sınıflandırma yaptığını gösterir. a tahmini değerleri, b ise gerçek değerleri temsil ettiğinde KOKH değeri denklem (6) kullanılarak hesaplanır.

$$KOKH = \sqrt{\frac{(a_1 - b_1)^2 + \dots + (a_n - b_n)^2}{n}} \quad (6)$$

ROC Eğrisi Altında Kalan Alan (AUC): ROC eğrisinde, yatay eksen yanlış pozitif oranı (YPO)'nı, düşey eksen ise doğru pozitif oranı (DPO)'nı gösterir. Bu eğrinin altında kalan alan (AUC) sınıflandırma ölçütü olarak kullanılır. Alanın mümkün olduğunca 1'e yakın bir değer çıkması istenir. Sırasıyla YPO ve DPO oranı denklem (7) ve denklem (8) kullanılarak hesaplanır.

$$YPO = \frac{YP}{YP+DN} \quad (7)$$

$$DPO = \frac{DP}{DP+YN} \quad (8)$$

3. BULGULAR VE TARTIŞMA

Veri setinde bulunan 3 farklı üreticiye ait 349 tane omuz protezinin radyografi görüntüleri bulunmaktadır. Bu görüntülerden omuz protezinin hangi üreticiye ait olduğunun tespiti 12 farklı algoritma kullanılarak yapılmıştır. Algoritmaların hem çapraz doğrulama yapılmadan eğitim setinin kullanılması sonucunda yapılan sınıflandırmadaki performansları hem de çapraz doğrulama sonrası performansları karşılaştırılmıştır. algoritmaların yanlış ve doğru tahmin ettikleri protez sayılarını gösteren hata matrisleri Şekil 3'te verilmiştir. Bu matrislerde Cofield sınıfı (C), Depuy sınıfı (D) ve Zimmer sınıfı (Z) ile gösterilmiştir. Matrislerin köşegen değerleri yeşil renkli kutularda verilmiştir. Bunlar algoritmaların doğru tahmin ettikleri protez sayılarını ifade etmektedir. Çapraz doğrulama yapılmadan önce tüm eğitim setinin kullanıldığı sınıflandırmada, Şekil 3'te gösterildiği gibi NB, DAA, DVM, SMO [37], K-EYK, K*, EA, HA, LMA, RMA, REPTree [38] ve CART [39] algoritmalarının her sınıfa ait yanlış tahmin edilen protezleri bulunmaktadır.

		NB						DAA						DVM			
		Tahmin edilen						Tahmin edilen						Tahmin edilen			
Gerçek		C	D	Z	Gerçek		C	D	Z	Gerçek		C	D	Z			
		C	31	30		6		C	29		34	4		C	60	7	0
		D	28	191		20		D	12		216	11		D	0	239	0
		Z	4	19		20		Z	3		20	20		Z	0	5	38
		SMO						K-EYK						K*			
		Tahmin edilen						Tahmin edilen						Tahmin edilen			
Gerçek		C	D	Z	Gerçek		C	D	Z	Gerçek		C	D	Z			
		C	1	66		0		C	67		0	0		C	67	0	0
		D	0	239		0		D	0		239	0		D	0	239	0
		Z	0	43		0		Z	0		0	43		Z	0	0	43
		EA						HA						LMA			
		Tahmin edilen						Tahmin edilen						Tahmin edilen			
Gerçek		C	D	Z	Gerçek		C	D	Z	Gerçek		C	D	Z			
		C	67	0		0		C	31		31	5		C	19	47	1
		D	0	239		0		D	33		187	19		D	5	230	4
		Z	0	0		43		Z	3		17	23		Z	2	34	7
		RMA						REPTree						CART			
		Tahmin edilen						Tahmin edilen						Tahmin edilen			
Gerçek		C	D	Z	Gerçek		C	D	Z	Gerçek		C	D	Z			
		C	67	0		0		C	21		44	2		C	29	35	3
		D	0	239		0		D	5		230	4		D	3	234	2
		Z	0	0		43		Z	2		26	15		Z	2	26	15

Şekil 3. 10 kat çapraz doğrulama öncesinde hata matrisleri

Tablo 1. 10 kat çapraz doğrulama öncesinde performans metrikleri

Algoritma	Sınıf	Kesinlik	Duyarlılık	F1 skoru	AUC	Kappa
NB	Cofield	0,492	0,463	0,477	0,815	0,3588
	Depuy	0,796	0,799	0,797	0,768	
	Zimmer	0,435	0,465	0,449	0,826	
DAA	Cofield	0,659	0,433	0,523	0,829	0,445
	Depuy	0,800	0,904	0,849	0,801	
	Zimmer	0,571	0,465	0,513	0,840	
DVM	Cofield	1,000	0,896	0,945	0,948	0,9254
	Depuy	0,952	1,000	0,976	0,945	
	Zimmer	1,000	0,884	0,938	0,942	
SMO	Cofield	1,000	0,015	0,029	0,649	0,0135
	Depuy	0,687	1,000	0,814	0,505	
	Zimmer	---	0	---	0,700	
K-EYK	Cofield	1	1	1	1	1
	Depuy	1	1	1	1	
	Zimmer	1	1	1	1	
K*	Cofield	1	1	1	1	1
	Depuy	1	1	1	1	
	Zimmer	1	1	1	1	
EA	Cofield	1	1	1	1	1
	Depuy	1	1	1	1	
	Zimmer	1	1	1	1	
HA	Cofield	0,463	0,463	0,463	0,813	0,3625
	Depuy	0,796	0,782	0,789	0,764	
	Zimmer	0,489	0,535	0,511	0,825	
LMA	Cofield	0,731	0,284	0,409	0,760	0,2821
	Depuy	0,740	0,962	0,836	0,766	
	Zimmer	0,583	0,163	0,255	0,820	
RMA	Cofield	1	1	1	1	1
	Depuy	1	1	1	1	
	Zimmer	1	1	1	1	
REPTree	Cofield	0,750	0,313	0,442	0,740	0,3879
	Depuy	0,767	0,962	0,853	0,725	
	Zimmer	0,714	0,349	0,469	0,728	
CART	Cofield	0,853	0,433	0,574	0,735	0,4855
	Depuy	0,793	0,979	0,876	0,735	
	Zimmer	0,750	0,349	0,476	0,742	

Tablo 1’de görülebileceği gibi kesinlik ve duyarlılık değerleri değişkenlik göstermektedir. K-EYK, K*, EA ve RMA algoritmaları tüm sınıflara ait protezlerin tamamını doğru tahmin ettikleri için kesinlik ve duyarlılık değerleri 1 tam puan almıştır. DVM algoritmasında Depuy olarak tahmin edilenlerin tamamı gerçekte Depuy sınıfına ait olduğu için duyarlılık 1 değerini almıştır. DVM algoritması gerçekte Cofield sınıfına ait olan 7 tane protezi, Zimmer sınıfına ait 5 tane protezi ise doğru tespit edemediğinden sırasıyla Cofield sınıfı için duyarlılık değeri 0,89 ve Zimmer sınıfı için duyarlılık değeri 0,88 olmuştur. DVM algoritmasının Cofield ve Zimmer sınıfında olduğunu tahmin ettiklerinin hepsi gerçekte Cofield ve Zimmer sınıfında olduğundan Cofield ve Zimmer sınıfları için kesinlik değerleri 1 olarak hesaplanmıştır. SMO algoritması ise, Cofield sınıfına ait hiçbir protezi doğru tespit edemediği için duyarlılık değeri 0’dır ve kesinlik değeri ise hesaplanamamıştır. Aynı zamanda en düşük AUC ve Kappa değeri de SMO algoritmasına aittir. Bu durumda en kötü performansı sergilen algoritmanın SMO algoritması olduğu ifade edilebilir. En iyi performansı sunan algoritmalar ise kesinlik, duyarlılık, F1 skoru, AUC ve Kappa değeri 1 tam puan alan K-EYK, K*, EA ve RMA algoritmalarıdır.

Sınıflandırıcıların 10 kat çapraz doğrulamalı performans değerlendirmeleri yapıldığında, Şekil 4’teki hata matrislerinden DVM ve SMO algoritmalarının Zimmer sınıfını tanımada yine başarısız olduğu ve Zimmer sınıfına ait hiçbir protezi tanımadıkları görülmektedir. Çapraz doğrulama yapılmadan önce bütün protezleri doğru tahmin eden K-EYK, K*, EA ve RMA algoritmaları aynı başarıyı çapraz doğrulamada gösterememişlerdir. K-EYK algoritması yine tüm algoritmalar arasından en fazla doğru sınıflandırma yapan algoritma olmuştur. K-EYK algoritmasının doğru sınıflandırdığı protez sayısı 260, yanlış sınıflandırdığı protez sayısı ise 89’dur. Her sınıfa ait doğru ve yanlış sayıları göz önünde bulundurulursa, K-EYK algoritması 67 tane Cofield sınıfına ait protezden 27 tanesini doğru tahmin etmiş, 34 tanesinin Depuy ve 6 tanesinin ise Zimmer sınıfında olduğunu tahmin etmiştir. 239 tane Depuy sınıfına ait protezden 217 tanesini doğru tahmin etmiş, 14 tanesinin Cofield ve 8 tanesinin ise Zimmer sınıfında olduğunu tahmin etmiştir. Zimmer sınıfındaki protezleri tespit ederken 43 protezden 16 tanesini doğru sınıflandırmış fakat 4 tane protezi Cofield ve 23 tanesini ise Depuy olarak sınıflandırmıştır. Bu nedenle, K-EYK algoritması en yüksek kesinlik ve duyarlılık değerlerini Depuy sınıfında elde etmiştir. K-EYK algoritmasının diğer performans metriklerine bakıldığında ortalama metrik değerlerinin diğer algoritmalarından daha yüksek olduğu görülmektedir. K-EYK ayrıca en yüksek Kappa değerine erişmiş algoritmadır. DVM ve SMO algoritmaları özellikle Zimmer sınıfının tespitinde başarılı olamamışlardır. Bu

algoritmaların Zimmer sınıfına ait doğru tespit ettikleri protez sayıları 0'dır. Dolayısıyla, Zimmer sınıfının duyarlılık değerleri 0 ve kesinlik değerleri ise hesaplanamamıştır. 10 kat çapraz doğrulama sonrasında elde edilen performans metriklerinin tümü Tablo 2'de listelenmektedir.

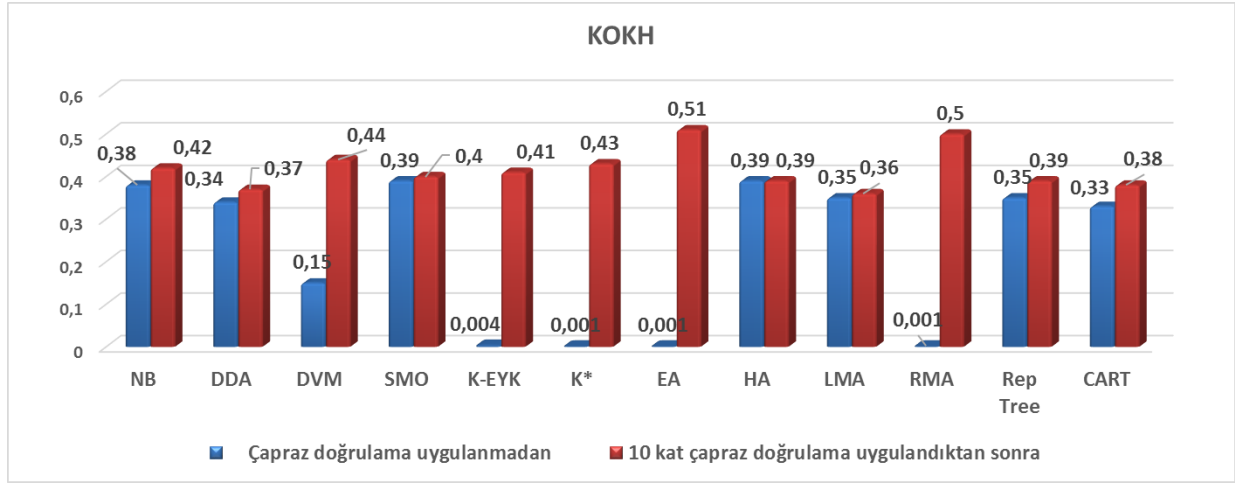
		NB						DAA						DVM			
		Tahmin edilen						Tahmin edilen						Tahmin edilen			
			C	D	Z				C	D	Z				C	D	Z
Gerçek	C	25	37	5	Gerçek	C	23	38	6	Gerçek	C	5	62	0			
	D	33	184	22		D	17	208	14		D	0	239	0			
	Z	5	21	17		Z	5	20	18		Z	0	43	0			
		SMO						K-EYK						K*			
		Tahmin edilen						Tahmin edilen						Tahmin edilen			
			C	D	Z				C	D	Z				C	D	Z
Gerçek	C	2	65	0	Gerçek	C	27	34	6	Gerçek	C	23	39	5			
	D	0	239	0		D	14	217	8		D	24	204	11			
	Z	0	43	0		Z	4	23	16		Z	6	29	8			
		EA						HA						LMA			
		Tahmin edilen						Tahmin edilen						Tahmin edilen			
			C	D	Z				C	D	Z				C	D	Z
Gerçek	C	22	38	7	Gerçek	C	0	67	0	Gerçek	C	22	42	3			
	D	35	176	28		D	0	239	0		D	11	221	7			
	Z	6	25	12		Z	0	43	0		Z	6	27	10			
		RMA						REPTree						CART			
		Tahmin edilen						Tahmin edilen						Tahmin edilen			
			C	D	Z				C	D	Z				C	D	Z
Gerçek	C	30	30	7	Gerçek	C	21	44	2	Gerçek	C	22	41	4			
	D	35	178	26		D	16	213	10		D	14	222	3			
	Z	3	30	10		Z	3	32	8		Z	1	34	8			

Şekil 4. 10 kat çapraz doğrulama sonrasında hata matrisleri

Tablo 2. 10 kat çapraz doğrulama sonrasında performans metrikleri

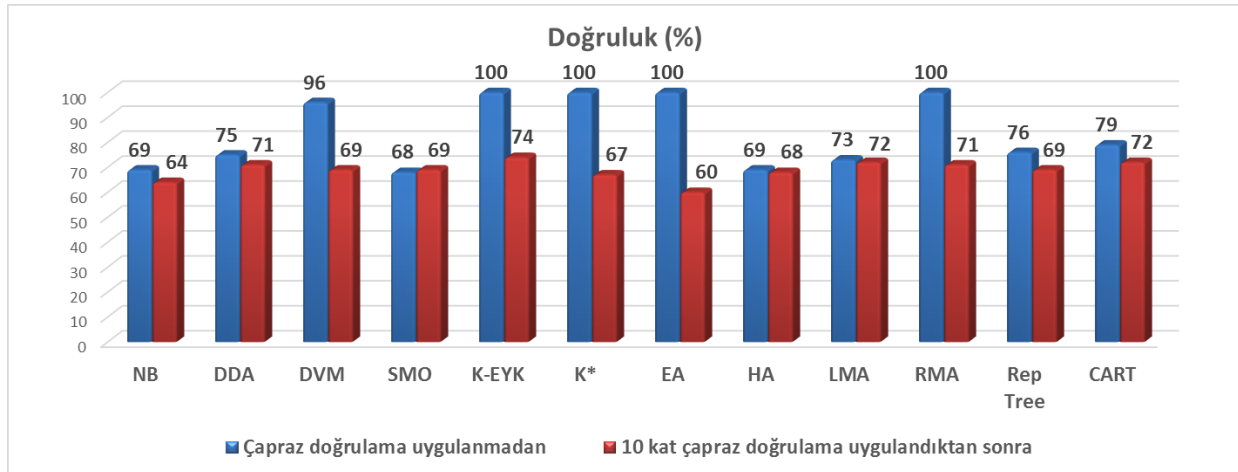
Algoritma	Sınıf	Kesinlik	Duyarlılık	F1 skoru	AUC	Kappa
NB	Cofield	0,397	0,373	0,385	0,752	0,258
	Depuy	0,760	0,770	0,765	0,712	
	Zimmer	0,386	0,395	0,391	0,304	
DAA	Cofield	0,511	0,343	0,411	0,739	0,3486
	Depuy	0,782	0,870	0,824	0,737	
	Zimmer	0,474	0,419	0,444	0,749	
DVM	Cofield	1,000	0,075	0,139	0,537	0,0664
	Depuy	0,695	1,000	0,820	0,523	
	Zimmer	---	0	---	0,50	
SMO	Cofield	1	0,030	0,058	0,597	0,0269
	Depuy	0,689	1,000	0,816	0,509	
	Zimmer	----	0	----	0,675	
K-EYK	Cofield	0,600	0,403	0,482	0,672	0,4028
	Depuy	0,792	0,908	0,846	0,687	
	Zimmer	0,533	0,372	0,438	0,637	
K*	Cofield	0,434	0,343	0,383	0,730	0,238
	Depuy	0,750	0,854	0,798	0,727	
	Zimmer	0,333	0,186	0,239	0,763	
EA	Cofield	0,349	0,328	0,338	0,591	0,1699
	Depuy	0,736	0,736	0,736	0,582	
	Zimmer	0,255	0,279	0,267	0,582	
HA	Cofield	---	0	---	0,482	0
	Depuy	0,685	1	0,813	0,498	
	Zimmer	---	0	----	0,471	
LMA	Cofield	0,564	0,328	0,415	0,743	0,3165
	Depuy	0,762	0,925	0,836	0,748	
	Zimmer	0,500	0,233	0,317	0,792	
RMA	Cofield	0,441	0,448	0,444	0,657	0,2187
	Depuy	0,748	0,745	0,746	0,600	
	Zimmer	0,233	0,233	0,233	0,562	
REPTree	Cofield	0,525	0,313	0,393	0,646	0,2408
	Depuy	0,737	0,891	0,807	0,636	
	Zimmer	0,400	0,186	0,254	0,653	
CART	Cofield	0,595	0,328	0,423	0,648	0,2902
	Depuy	0,747	0,929	0,828	0,660	
	Zimmer	0,533	0,186	0,276	0,700	

Çalışmada kullanılan algoritmaların sınıflandırmayı ne kadar hata ile yaptıklarını göstermek için kullanılan KOKH değerlerini gösteren grafik Şekil 5'te verilmiştir. Çapraz doğrulamadan önce en az hata değerine sahip olan algoritmalar K-EYK, K*, EA ve RMA algoritmalarıdır. En yüksek hata değeri 0,39'dur ve bu değer HA ve SMO algoritmalarına aittir. Çapraz doğrulamadan sonra hata değerleri artmıştır ve en yüksek hata değerini veren algoritma EA algoritması olmuştur. Öte yandan, en düşük hata değeri LMA algoritmasının olmuştur.



Şekil 5. KOKH değerleri

Çalışmada kullanılan algoritmaların doğruluk oranlarını gösteren grafik Şekil 6'da verilmiştir. Bu grafik dikkate alındığında çapraz doğrulamadan önce ve sonra en yüksek doğruluk değerini veren algoritmanın K-EYK olduğu görülebilir.



Şekil 6. Doğruluk değerleri

K-EYK algoritmasının çapraz doğrulamada doğru tespit ettiği protez sayısındaki düşüşün, veri setinde bulunan görüntülerin farklı boyutta olmasından ve çoğu görüntünün çözünürlüğünün ve kontrastının düşük olmasından kaynaklandığı düşünülmektedir. Ayrıca, sınıflar arasındaki dengesizliğin de sınıflandırıcının performansını düşüren önemli bir etken olduğu düşünülmektedir. Bu sorunlar, sayıca az olan veri sınıfına ait yüksek çözünürlüğe sahip görüntüler eklenerek giderilebilir ve algoritmanın başarısı artırılabilir. Sonuç olarak radyografi görüntülerinden protez tanımada K-EYK algoritmasının kullanımının yararlı olacağı anlaşılmıştır.

4. SONUÇ

Omuz protezlerinin değiştirilmesi veya bakımlarının yapılması öncesinde radyografi görüntülerinin incelenmesinin gerekmektedir. Ameliyat öncesinde, uygun donanım ve prosedürlerin seçilmesi için ameliyatı gerçekleştirecek olan cerraha değişim veya bakım gerektiren protezleri tanımada yardımcı olabilecek, hızlı ve yüksek doğruluk oranına sahip bir çözüme ihtiyaç duyulmaktadır. Bu çalışmada, 3 farklı üreticiye ait omuz protezlerinin radyografi görüntülerinden tanınabilmesi için 12 farklı sınıflandırma algoritması kullanılmış ve bu algoritmaların performansları karşılaştırılmıştır. Radyografi görüntülerinden protezlerin tanınmasında K-EYK algoritmasının diğer algoritmalara göre daha iyi bir performans sergilediği görülmüş olup, diğer protez türlerini tanımada da kullanılabileceği öngörülmektedir.

KAYNAKLAR

- [1] Grau, L. et al. Operative trends in the treatment of hip fractures and the role of arthroplasty, Geriatric orthopaedic surgery & rehabilitation Vol.9, 2151459318760634, 2018
- [2] Gasbarra, E. et al. Total hip arthroplasty revision in elderly patients, Aging clinical and experimental research 25, Vol.1, 61-63, 2013
- [3] Lee, J.-M. The current concepts of total hip arthroplasty, Hip & pelvis 28 Vol.4, 191, 2016
- [4] Schwartz, B.E. et al. Revision total hip arthroplasty in the United States: national trends and in-hospital outcomes. International orthopaedics 40, Vol.9, 1793-1802, 2016.
- [5] Saleem, M. et al. X-ray image analysis for automated knee osteoarthritis detection, Signal, Image and Video Processing 14, Vol.6, 1079-1087, 2020.
- [6] Bredow, J. et al. Software-based matching of X-ray images and 3d models of knee prostheses, Technology and Health Care 22, Vol.6, 895-900, 2014.
- [7] Wu, J. and Mahfouz, M.R. Robust x-ray image segmentation by spectral clustering and active shape model, Journal of Medical Imaging 3 Vol.3, 034005, 2016.
- [8] Cover, T. and Hart, P. Nearest neighbor pattern classification, IEEE transactions on

information theory 13, Vol.1, 21-27, 1967.

- [9] Cortes, C. and Vapnik, V. Support-vector networks. *Machine learning* 20, Vol.3, 273-297, 1995.
- [10] John, G.H. and Langley, P. Estimating continuous distributions in Bayesian classifiers, arXiv preprint arXiv:1302.4964, 2013.
- [11] Landwehr, N. et al. Logistic model trees, *Machine learning* 59 (1-2), 161-205, 2005.
- [12] Geurts, P. et al. Extremely randomized trees, *Machine learning* 63, Vol.1, 3-42, 2006.
- [13] Hulten, G. et al., Mining time-changing data streams, *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 97-106, 2001.
- [14] Umadevi, N. and Geethalakshmi, S., Multiple classification system for fracture detection in human bone x-ray images, *Third International Conference on Computing, Communication and Networking Technologies (ICCCNT'12)*, IEEE, 2012, pp. 1-8, 2012.
- [15] Lee, S. et al. The exploration of feature extraction and machine learning for predicting bone density from simple spine X-ray images in a Korean population, *Skeletal radiology* 49 Vol.4, 613-618, 2020.
- [16] Kang, Y.-J. et al. Machine learning–based identification of hip arthroplasty designs. *Journal of orthopaedic translation* 21, 13-17, 2020.
- [17] Sukegawa, S. et al. Deep neural networks for dental implant system classification, *Biomolecules* 10 Vol.7, 984, 2020.
- [18] Kokkotis, C. et al. Machine learning in knee osteoarthritis: A review, *Osteoarthritis and Cartilage Open*, 100069, 2020.
- [19] Kotti, M. et al. The complexity of human walking: a knee osteoarthritis study, *PloS one* 9 Vol.9, e107325, 2014.
- [20] Hough, P.V., Method and means for recognizing complex patterns, *Google Patents*, 1962.
- [21] Stark, M.B.C.G. Automatic detection and segmentation of shoulder implants in x-ray images, 2018.
- [22] Urban, G. et al. Classifying shoulder implants in X-ray images using deep learning, *Computational and structural biotechnology journal* 18, 967-972, 2020.
- [23] Yi, P.H. et al. Automated detection and classification of shoulder arthroplasty models using deep learning, *Skeletal radiology* 49, 1623-1632, 2020.
- [24] Yang, G. et al. Tree Species Classification by Employing Multiple Features Acquired from Integrated Sensors. *Journal of Sensors* 2019.
- [25] Keerthi, S.S. et al. (2001) Improvements to Platt's SMO algorithm for SVM classifier design. *Neural computation* 13 Vol.3, 637-649.
- [26] Aha, D.W. et al. Instance-based learning algorithms. *Machine learning* 6 (1), 37-66, 1991.
- [27] McLachlan, G.J. *Discriminant analysis and statistical pattern recognition*, John Wiley & Sons, 2004.



- [28] Quinlan, J.R. Simplifying decision trees. *International journal of man-machine studies* 27 Vol.3, 221-234, 1987.
- [29] Srinivasan, D.B. and Mekala, P. Mining social networking data for classification using reptime. *International Journal of Advance Research in Computer Science and Management Studies* 2 Vol.10., 2014.
- [30] Pfahringer, B. *Random model trees: an effective and scalable regression method*, 2010.
- [31] Breiman, L. et al. *Classification and regression trees*, CRC press, 1984.
- [32] Landwehr, N. et al. *Logistic model trees*. *Machine learning* 59 (1-2), 161-205, 2005.
- [33] Hulten, G. et al., *Mining time-changing data streams*, *Proceedings of the seventh ACM SIGKDD international conference on Knowledge discovery and data mining*, 2001, pp. 97-106.
- [34] Runkler, T.A. *Data Visualization*. In *Data Analytics*, pp. 37-59, Springer, 2020.
- [35] Joshi, R. Accuracy, precision, recall & f1 score: Interpretation of performance measures, Retrieved April 1 (2018), 2016.
- [36] Brownlee, J. *Classification accuracy is not enough: More performance measures you can use*, *Machine Learning Mastery* 21, 2014.
- [37] Platt, J. *Sequential minimal optimization: A fast algorithm for training support vector machines*, 1998.
- [38] Al Snousy, M.B. et al. *Suite of decision tree-based classification algorithms on cancer gene expression data*, *Egyptian Informatics Journal* 12, Vol.2, 73-82, 2011.
- [39] Bhargava, N. et al., *An approach for classification using simple CART algorithm in WEKA*, *2017 11th International Conference on Intelligent Systems and Control (ISCO)*, IEEE, 2017, pp. 212-216.