

Doğrusal Ayırt Etme Ölçütlerinin Teorik İncelenmesi

Theoretical Analysis of Linear Discriminant Analysis Criteria

Hakan Çevikalp

Elektrik ve Elektronik Mühendisliği Bölümü
Eskişehir Osmangazi Üniversitesi
hakan.cevikalp@gmail.com

Özetçe

Fisher's Linear Discriminant Analysis (FLDA) yöntemi sınıflar arası ayrıştırabilirliği büyütürken, sınıflar içi saçılımı küçültmeyi amaçlayan doğrusal bir öznelik özütleme yöntemidir. Bu sebeple yöntem, Fisher's Linear Discriminant Analysis ölçütünü eniyileştirme amacını güder. Bu bildiride öncelikle FLDA ölçütünün eksiklikleri tartışılmış ve bu eksiklikleri gideren yeni ölçütler verilmiştir. Son olarak bu yeni ölçütlerin yeterliliği teorik olarak ve deneysel çalışmalarla ispatlanmıştır.

Abstract

The Fisher's Linear Discriminant Analysis (FLDA) is a successful linear feature extraction method which aims to maximize the between-class separability and to minimize the within-class variability. In order to accomplish its goal, FLDA maximizes the Fisher's Linear Discriminant Analysis criterion given in the paper. In this paper we first address the limitations of the classical FLDA criterion and then discuss new criterion functions which were introduced to overcome those limitations. Finally, we demonstrate the capability of the new criteria both theoretically and empirically.

1. Giriş

Fisher's Linear Discriminant Analysis (FLDA) yöntemi görüntü tanıma, tıbbi uygulama alanları ve çoğul ortamlı bilgi alımı gibi daha birçok alanda oldukça sık kullanılan doğrusal bir öznelik özütleme yöntemidir [1]. Bu nedenle diğer öznelik özütleme yöntemlerinde olduğu gibi, bu yöntem kullanılarak örnek uzayı genellikle daha küçük boyutlu bir uzaya aktarılır. Burada amaç verilerin sınıflandırılması için gerekli olan ayırt edici öznelikleri seçip, ayırt edici olmayan öznelikleri elemektir. FLDA yöntemi özneliklerin özütlenmesi esnasında bir taraftan sınıflar arası ayrıştırabilirliği büyütürken, diğer taraftan sınıflar içi saçılımı küçültme amacını güder. Bu amaçla verilerin özütlenmesi için kullanılacak dönüşüm matrisini bulmak için aşağıda verilen FLDA ölçütü kullanılır,

$$J_{FLDA}(W) = \max \frac{|W^T S_B W|}{|W^T S_W W|}. \quad (1)$$

Bu eşitlikte W yeni özneliklerin elde edilmesi için kullanılacak dönüşüm matrisini, S_W sınıflar içi saçılım matrisini, S_B sınıflar arası saçılım matrisini ve $|\cdot|$ işleci ise matris belirtenini (determinant) ifade etmektedir. FLDA yöntemi direkt olarak bir sınıflandırıcı olarak kullanıldığı gibi, daha karmaşık doğrusal olmayan sınıflandırıcılar için bir önışlemci olarak da kullanılabilir. Bu yöntem kullanılarak elde edilen özneliklere uygulanan sınıflandırıcılarla elde edilen

tanıma oranları genellikle aynı sınıflandırıcıların doğrudan orjinal veri uzayına uygulanmasıyla elde edilen tanıma oranlarından daha iyi sonuçlar vermektedir. FLDA yöntemi farklı dağılımlara sahip sınıflara uygulanırsa da, aynı kovaryans matrisine sahip olan Gauss dağılımlı sınıflara uygulandığında optimal Bayes sınıflandırıcı olacaktır.

FLDA yöntemi birçok alanda başarı ile uygulanmasına rağmen, FLDA ölçütünün yetersiz olduğu bazı durumlar vardır. İkinci bölümde FLDA ölçütünün karakteristik özellikleri incelenip, bu ölçütün yetersiz olduğu durumlar teorik olarak anlatılacaktır. Üçüncü bölümde FLDA ölçütünün eksiklerini gideren yeni ölçütler incelenecektir. Dördüncü bölümde deneysel çalışmalar açıklanacak ve son bölümde elde edilen sonuçlar değerlendirilecektir.

2. FLDA Ölçütünün Karakteristik Özellikleri

Eğitim setinde her biri N_i örnek içeren C tane sınıf olduğunu varsayalım. Her bir sınıftaki m 'nci örneği d boyutlu örnek uzayında x_m^i , $i=1, \dots, C$, ile gösterelim. Bu durumda eğitim setinde toplam $M = \sum_{i=1}^C N_i$ örnek olacaktır. Sınıflar içi saçılım matrisi S_W , sınıflar arası saçılım matrisi S_B ve toplam saçılım matrisi S_T aşağıdaki eşitlikler kullanılarak bulunabilir:

$$S_W = \sum_{i=1}^C \sum_{m=1}^{N_i} (x_m^i - \mu_i)(x_m^i - \mu_i)^T, \quad (2)$$

$$S_B = \sum_{i=1}^C N_i (\mu_i - \mu)(\mu_i - \mu)^T, \quad (3)$$

$$S_T = \sum_{i=1}^C \sum_{m=1}^{N_i} (x_m^i - \mu)(x_m^i - \mu)^T = S_W + S_B. \quad (4)$$

Bu eşitliklerde μ_i , i 'nci sınıfa ait ortalama değer vektörünü ve μ , eğitim setindeki tüm örneklerin ortalama değer vektörünü ifade eder. Saçılım matrisleri simetrik ve yarı kesin artı (positive semi-definite) matrisler olduğundan, bu matrislere ait özdeğerler yalnızca sıfır yada pozitif değerler alırlar.

FLDA ölçütü, saçılım matrislerinin skalar olarak ölçülmesi için matris belirtenlerini kullanır. Bir saçılım matrisinin belirtenini o matrisin özdeğerlerinin çarpımı olacağından, FLDA ölçütü bir anlamda çok boyutlu uzaydaki elipsel saçılımların hacmini ölçmektedir.

Dönüşüm matrisi W 'nin sütunlarını oluşturan izdüşüm vektörleri w_k , $S_W^{-1} S_B$ matrisinin sıfırdan farklı özdeğerlerine karşılık gelen özvektörler olarak seçilirse FLDA ölçütü en büyük değerine ulaşır. İzdüşüm vektörlerinin normları FLDA ölçütünün büyüklüğünü etkilemez. Bu sebeple özütleme için seçilen özvektörlerin normları genellikle 1 olacak şekilde ayarlanır. $S_W^{-1} S_B$ matrisi genellikle simetrik bir matris değildir. Bu sebeple özdeğerlerin ve özvektörlerin hesaplanması

esnasında kararlılık (stability) sorunlarıyla karşılaşılabilir. Bu sorunu çözmek için “simultaneous diagonalization” [2] yöntemine başvurulur. Bu işlem sonucunda elde edilen izdüşüm vektörleri S_W üzerinden ortogonal olacaktır. Bu durum matematiksel olarak aşağıdaki gibi ifade edilebilir,

$$w_j^T S_W w_j = \delta_{ij}. \quad (5)$$

Bu eşitlikte δ_{ij} , Kronecker delta'yı ifade etmektedir. Bu nedenle, FLDA yöntemiyle elde edilen izdüşüm vektörleri genellikle ortonormal değildir.

FLDA ölçütünün dezavantajlarından biri, bu ölçüt kullanılarak elde edilen izdüşüm vektör sayısının en fazla $C-1$ olmasıdır. Bu durum sınıflar arası saçılım matrisinin kertesinin en fazla $C-1$ olmasından kaynaklanır. Bu durumda $S_W^{-1} S_B$ matrisinin sıfırdan farklı pozitif özdeğerlerinin sayısı $C-1$ sayısını geçemez. Bazı uygulamalarda bu sayı, veri örneklerinin sınıflandırılması için yeterli olmayabilir. Bu sorunu çözmek için Foley ve Sammon, iki sınıfa ait örneklerin ayırt edilmesi için Foley ve Sammon Discriminant Analysis yöntemini önermişlerdir [3]. Bu yöntem özyineli bir yöntem

olup her bir özyinelemede $\frac{w_k^T S_B w_k}{w_k^T S_W w_k}$ oranını enbüyüten

vektör izdüşüm vektörü olarak seçilir. Bu durumda elde edilecek izdüşüm vektör sayısı d' 'ye kadar çıkabilir. Bu yöntem sınıf sayısının ikiden fazla olduğu durumlar için Okada ve Tomita tarafından geliştirilmiş ve Orthonormal Discriminant Vector (ODV) yöntemi olarak adlandırılmıştır [4]. Bu yöntemle elde edilen izdüşüm vektörleri birbirine ortonormaldir. Bu yaklaşımların en büyük dezavantajı algoritmaların çok fazla işlem gerektirmesi ve bu sebeple sınıflandırma işleminin oldukça yavaş olmasıdır.

FLDA ölçütünün en büyük eksikliği, bu ölçütün eğitim setindeki örnek sayısının örnek uzayının boyutundan küçük olduğu durumlarda uygulanamamasıdır. Bu sorun örüntü tanıma alanında “small sample size problem” olarak isimlendirilmiştir [2]. S_W matrisinin kertesini en fazla $M-C$ olduğundan, eğitim setindeki veri sayısı M , örnek uzayı boyutu d' 'den küçük olduğu durumlarda S_W matrisi tekil bir matris olacaktır. Bu durumda S_W matrisinin tersi alınmaz. Aynı sorun ODV yöntemi için de geçerlidir. Bu problemi çözmek için önerilen yöntemler iki grup altında toplanabilir. Birinci gruptaki yöntemler S_W matrisinin tersini alma sorununu lineer cebir teknikleri uygulayarak çözmeye çalışırlar. Örnek olarak Tian ve arkadaşları S_W^{-1} matrisi yerine onun pseudo tersini kullanmışlardır [5]. Bazı yazarlar ise S_W matrisini tekil olmayan bir matris haline getirmek için, bu matrise elemanları küçük olan köşegen bir matris eklemişlerdir [6]. Fakat bu yöntemler örnek uzay boyutunun çok büyük olduğu durumlarda kullanılamaz. İkinci gruba ait yöntemlerde ise örnek uzayının boyutu azaltılır. Bu amaçla Swets and Weng, Principal Component Analysis (PCA) uygulamak suretiyle örnek uzayını S_W matrisinin tekil olmadığı daha küçük boyutlu bir uzaya aktarmışlardır [7]. Daha sonra FLDA yöntemi bu yeni örnek uzayına aktarılan örneklerle uygulanmıştır. Fakat PCA yöntemiyle boyut indirgenmesi sırasında verilerin sınıflandırılması için önemli ayırt edici bilgileri barındıran öz nitelikler elenebilir. Bu sebeple bu önerilen yöntem optimal bir yöntem olmayacaktır [8].

Özetlemek gerekirse FLDA yöntemi yada ODV yöntemi genel olarak S_W matrisinin tekil olmadığı durumlarda başarı ile kullanılabilir. Fakat her iki yöntemde, eğitim setindeki örnek sayısı örnek uzayı boyutundan küçük olduğu durumlarda kullanılamaz. Aşağıdaki bölümde bu sorunu çözen yeni ölçütler verilip, bu ölçütler teorik olarak incelenecektir.

3. Yeni Doğrusal Öznitelik Özütleme Ölçütlerinin Teorik İncelenmesi

Eğitim setindeki örnek sayısının örnek uzayı boyutundan küçük olduğu durumlarda FLDA ölçütünün yerine kullanılacak alternatif ölçütleri incelemeye önce, bu ölçütlerin önerilmesinin altında yatan motivasyonu daha iyi anlayabilmek için, toplam saçılım matrisinin sıfır altuzayının verilerin sınıflandırılması için gerekli hiç bir ayırt edici özellik taşımadığını ispatlayacağız.

Teorem: Toplam saçılım matrisinin sıfır altuzayı, $N(S_T)$, veri örneklerinin sınıflandırılması için gerekli ayırt edici bilgi içermez.

İspat: Yukarıda verilen teoremi, $N(S_T)$ 'den gelen izdüşüm vektörlerinin FLDA ölçütünün eniyilenmesi üzerinde hiç bir etkisi olmadığını göstererek ispatlayacağız. S_T matrisinin sıfır altuzayından gelen herhangi bir vektör q seçelim, $q \in N(S_T)$.

Bu durumda $q^T S_T q = 0$ olacaktır. Saçılım matrisleri S_W ve S_B yarı kesin artı matrisler olduğundan aşağıdaki eşitlikler geçerlidir:

$$q^T S_T q = q^T (S_W + S_B) q = 0 \Leftrightarrow q^T S_W q = 0 \text{ \& } q^T S_B q = 0.$$

Aynı şekilde z vektörünün S_T matrisinin erim altuzayından gelen herhangi bir vektör olduğunu varsayalım, $z \in R(S_T)$ ve $z^T S_T z \neq 0$. S_T matrisinin erim ve sıfır uzaylarının toplamı örnek uzayını doğurduğundan $(N(S_T) \oplus R(S_T) = R^d)$, öz nitelik özütlemeye kullanılacak izdüşüm vektörü w , q ve z vektörlerinin toplamı olarak ifade edilebilir. Başka bir deyişle, $w = z + q$. Bu durumda FLDA ölçütünü w vektörüne göre yada z vektörüne göre enbüyütmeye çalışmak aynı sonuçları verir,

$$J(w) = \max_{z+q} \frac{(z+q)^T S_B (z+q)}{(z+q)^T S_W (z+q)} = \max_z \frac{z^T S_B z}{z^T S_W z} = J(z). \quad (6)$$

Bu da $N(S_T)$ uzayının sınıflandırma için yararlı hiç bir ayırt edici özellik içermediğini ispatlar. \square

Teorem 1 aynı zamanda Çevikalp ve arkadaşları tarafından eğitim setindeki tüm örneklerin $N(S_T)$ üzerindeki izdüşümlerinin aynı olduğu gösterilerek ispatlanmıştır [9].

Eğitim setindeki örnek sayısının örnek uzayı boyutundan küçük olduğu durumlarda kullanılmak üzere Liu ve arkadaşları [10] aşağıda verilen geliştirilmiş FLDA ölçütünü önermişlerdir,

$$J(W) = \max \frac{|W^T S_B W|}{|W^T S_T W|}. \quad (7)$$

Bu ölçütün klasik FLDA ölçütünden farkı payda kısmında sınıflar içi saçılım matrisi S_W yerine toplam saçılım matrisi S_T 'nin kullanılmasıdır. Geliştirilmiş FLDA ölçütü en büyük değerine $S_T^{-1} S_B$ matrisinin sıfırdan farklı özdeğerlerine karşılık gelen özvektörlerin özütlemeye için kullanılan izdüşüm

vektörleri olarak seçilmesiyle ulaşır. Bu ölçüt, örnek sayısının örnek uzayı boyutundan büyük olduğu durumlarda klasik FLDA ölçütüyle aynı sonuçları üretir. Örnek sayısının yetersiz olduğu durumlarda ise, S_T matrisi tekil bir matris olacağından bu matrisin tersi alınmaz. Bu sebeple orjinal örnek uzayı PCA uygulanarak S_T matrisinin erim uzayına aktarılır. Bu işlem sadece S_T matrisinin sıfır altuzayını elediğinden, Teorem 1 kullanılarak hiçbir ayırt edici bilgi kaybı olmadığı ispatlanabilir. Daha sonra bu yeni uzaya aktarılan verilerin saçılım matrisleri kullanılarak geliştirilmiş FLDA ölçütü eniyilenmeye çalışılır. Bu işlem sonucunda geliştirilmiş FLDA ölçütü en büyük değeri olan 1 sayısına ulaşır. Bu yöntem PCA+Generalized Discriminant Analysis (PCA+GDA) olarak adlandırılmıştır. FLDA yönteminde olduğu gibi PCA+GDA yöntemi kullanılarak en fazla $C-1$ izdüşüm vektörü elde edilebilir. Bu işlemler sonucunda elde edilen izdüşüm vektörleri *optimal ayırt edici altuzayı* olarak isimlendirdiğimiz S_T matrisinin erim uzayı ile S_W matrisinin sıfır altuzayının kesişim uzayından $(N(S_W) \cap R(S_T))$ gelir [9]. Matematiksel olarak bu durum aşağıdaki gibi ifade edilebilir:

$$W^T S_W W = 0, \quad W^T S_T W \neq 0. \quad (8)$$

Elde edilen izdüşüm vektörleri aynı zamanda S_T matrisi üzerinden ortogondur, $w_i^T S_T w_j = \delta_{ij}$.

Geliştirilmiş FLDA ölçütünün en büyük eksikliği seçilen izdüşüm vektörlerinin ayırt edici gücünün numerik olarak iyi ifade edilememesidir. Çünkü bu ölçüt $W^T S_W W = 0$ ve $W^T S_T W \neq 0$ şartlarını sağlayan herhangi bir dönüşüm matrisi için en büyük değerine ulaşır. Fakat bu şartı sağlayan bir çok matris bulunabilir. Bununla beraber bu matrislerin bir çoğu test setindeki örneklerin doğru olarak sınıflandırılmasında yetersiz kalır. Bu durumda aşağıda verilen ölçüt kullanılırsa seçilen izdüşüm vektörlerinin ayırt edici gücü daha iyi ölçülebilir,

$$J(W) = \max_{\|W\|_{S_W W}=0} |W^T S_B W| = \max_{\|W\|_{S_W W}=0} |W^T S_T W|. \quad (9)$$

Bu ölçütü *sıfır alt uzay ölçütü* olarak isimlendirdik. Fakat bu ölçüt, izdüşüm vektörlerinin normlarına bağlı olduğundan, izdüşüm vektörlerinin normlarının 1 olarak alınma şartı unutulmamalıdır. Bu yeni ölçüt kullanılarak sadece bir tane izdüşüm vektör matrisi bulunur ve bu matris aynı zamanda geliştirilmiş FLDA ölçütünün en büyük değeri olan 1 sayısına ulaşır. Sıfır Alt Uzay ölçütünün en iyilenmesi için eğitim setindeki tüm örneklerin S_W matrisinin sıfır altuzayındaki izdüşümleri bulunur. Daha sonra bu uzaya aktarılan örnekler PCA uygulanarak optimal izdüşüm vektörleri elde edilir. Bu yöntem Ayırt Edici Ortak Vektör (AEOV) yöntemi olarak adlandırılmıştır [8]. Diğer yöntemlerde olduğu gibi AEOV yöntemi kullanılarak elde edilen izdüşüm vektörlerinin sayısı en fazla $C-1$ olabilir. Seçilen izdüşüm vektörleri ortonormaldir ve bu ortonormal vektörler optimal ayırt edici altuzayını doğurur.

PCA+GDA ve AEOV yöntemleri ile bulunan izdüşüm vektörleri optimal ayırt edici altuzayını doğururlar. İzdüşüm vektörleri S_W matrisinin sıfır altuzayından geldikleri için, her sınıftaki tüm örneklerin bu vektörler üzerindeki izdüşümleri aynı olacaktır. Bu iki yöntem arasındaki tek fark AEOV yöntemi kullanılarak elde edilen izdüşüm vektörlerinin ortonormal olması ve PCA+GDA yöntemiyle elde edilen

izdüşüm vektörlerinin ise S_T matrisi üzerinden ortogonal olmasıdır. Bu sebeple bu yöntemler farklı tanıma sonuçları üretebilir. W matrisinin AEOV yöntemi kullanılarak elde edilen dönüşüm matris olduğunu ve \tilde{W} matrisinin ise PCA+GDA yöntemiyle elde edilen dönüşüm matrisi olduğunu varsayalım. Bu durumda dönüşüm matrisleri arasında aşağıdaki ilişki vardır:

$$W W^T = \tilde{W} (\tilde{W}^T \tilde{W})^{-1} \tilde{W}^T. \quad (10)$$

Yukarıdaki eşitliğin her iki tarafında optimal ayırt edici uzayının ortogonal izdüşüm işlecini (orthogonal projection operator) verir.

4. Deneysel Çalışmalar

Bu bildiriye incelenen ölçütlerin kullanımını, örnek uzayı boyutunun eğitim setindeki örnek sayısına olan oranı belirler. Bu sebeple ölçütlerin performansını test etmek için, örnek uzayı boyutunun veri tabanındaki örnek sayısına olan oranı oldukça farklı iki tip veri tabanı kullanılmıştır. İlk kullanılan veri tabanı elle yazılmış sayılara ait veri tabanıdır [11]. Bu veri tabanındaki örnek sayısı örnek uzayı boyutundan oldukça büyüktür. İkinci veri tabanı olarak ise AR yüz veri tabanı [12] kullanılmıştır. Bu veri tabanındaki örneklerin sayısı örnek uzayı boyutundan oldukça küçüktür.

4.1 Sayı Veri Tabanıyla Yapılan Deneyler

Bu veri tabanı her biri 200 örnek içeren 10 sınıfa ait toplam 2000 örnekten oluşur. Her bir örnek ikili imge halindedir. Bu imgeler kullanılarak 76 boyutlu Fourier katsayıları elde edilmiş ve örnekler 76 boyutlu vektörlerle belirtilmişlerdir.

Her bir sınıftan rastgele 100 örnek seçilerek eğitim seti oluşturulmuş ve geriye kalan 100 örnek ise test seti için kullanılmıştır. Eğitim setindeki örnek sayısı (1000), örnek uzayı boyutundan (76) büyük olduğu için FLDA ölçütünün kullanılmasında hiç bir sorun yoktur. Bu bölümde bu veri tabanı kullanılarak FLDA ölçütünü kullanan FLDA yöntemiyle, ortonormal izdüşüm vektörlerini kullanan ODV yöntemi karşılaştırılmıştır. Kullanılan izdüşüm vektör sayısı her iki yöntem içinde $C-1=9$ 'dur. Tanıma sırasında 1-en yakın komşu (nearest neighbor - NN) ve en yakın ortalama (nearest mean - NM) algoritmaları kullanılmıştır. Sonuçlar Tablo 1 ve Tablo 2'de verilmiştir.

TABLO 1
Eğitim Seti Tanıma Oranları

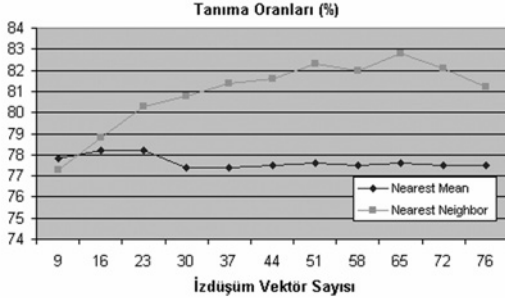
Yöntemler	FLDA Ölçütünün Sayısal Değeri	Tanıma Oranları	
		NN	NM
FLDA	27.4	%83.1	%84.6
ODV	1.32e-7	%78.5	%79.6

TABLO 2
Test Seti Tanıma Oranları

Yöntemler	FLDA Ölçütünün Sayısal Değeri	Tanıma Oranları	
		NN	NM
FLDA	27.4	%80.5	%80.7
ODV	1.32e-7	%77.3	%77.8

Tablo 1 ve 2'de görüldüğü gibi, FLDA ölçütü en büyük değerine FLDA yöntemi kullanılarak ulaşır. Tanıma oranları da FLDA ölçütünün sayısal değeri ile doğru orantılı olup, FLDA yöntemine ait tanıma oranları ODV yöntemiyle elde edilen tanıma oranlarından çok daha iyidir. Daha öncedende belirtildiği üzere ODV yöntemi kullanılarak FLDA yöntemiyle elde edilen izdüşüm sayısından daha fazla sayıda

izdüşüm vektörü elde edilebilir. İzdüşüm vektör sayısının tanıma oranlarındaki etkisini gözlemek için FLDA yöntemi ile bulunabilen maksimum izdüşüm vektör sayısı olan 9'dan başlayarak, izdüşüm vektör sayısını kademeli olarak 76'ya kadar çıkardık. Eğitim setine ait tanıma oranları Şekil 1'de verilmiştir. Şekil 1'de görüldüğü üzere izdüşüm vektör sayısının artırılması tanıma oranlarını iyileştirmiştir.



Şekil 1: Değişik sayıdaki izdüşüm vektörleri için test seti tanıma oranları.

4.2 AR Yüz Veri Tabanıyla Yapılan Deneyler

AR yüz veri tabanı 126 kişiye ait imgelerden oluşur. Veri tabanında her bir kişiye ait aydınlatma, yüz ifadesi ve kapatılma farklılıkları içeren 26 imge vardır. Deneyler için bu veri tabanından rastgele 30'u erkek 20'si bayan 50 kişi seçilmiş ve daha sonra bu kişilere ait imgeler gözler aynı konuma gelecek şekilde hizalandırılmıştır. İmgelerin sadece yüz içeren kısımları alınıp artalanlar kesilmiş ve son olarak bu imgeler boyutları 99x134 olacak şekilde örneklendirilmiştir. Her sınıftan kapatılma içeren 12 imge bu deneysel çalışmada kullanılmamış olduğundan, geriye kalan 14 imgeden yararlanılmıştır. Her sınıftan rastgele 7 imge eğitim seti için kullanılmış ve geriye kalan 7 imge ise test seti için kullanılmıştır.

Örnek uzayı boyutu (99x134=13266) eğitim setindeki örnek sayısından (350) oldukça büyüktür. Bu nedenle FLDA ölçütü bu durumda kullanılamaz. Daha önceden de belirtildiği üzere bu durumda uygulanabilecek iki yol vardır. İlk yol PCA+GDA yöntemini uygulamaktır. Bu yöntemde eğitim setindeki tüm örnekler toplam saçılım matrisinin erim uzayına aktarılmış ve daha sonra bu uzayda geliştirilmiş FLDA ölçütü en büyütülmeye çalışılmıştır. İkinci yol ise Ayırt Edici Ortak Vektör yöntemini uygulamaktır. Bu yöntemde ilk olarak eğitim setindeki tüm örneklerin sınıflar içi saçılım matrisinin sıfır altuzayındaki izdüşümleri bulunmuş ve daha sonra bu izdüşüm uzayındaki yeni örneklerle PCA uygulanarak dönüşüm matrisini oluşturan izdüşüm vektörleri elde edilmiştir. Her iki yöntem sonucunda da her sınıftaki örnekler o sınıfı temsil eden ortak bir öznitelik vektörü oluşturduğundan, tanıma esnasında test örneğine ait öznitelik vektörü ile ortak vektörler arasındaki Öklid uzaklığına bakılmış ve test örneği en yakın uzaklığı veren sınıfa atanmıştır. Eğitim setine ait tanıma oranları her iki yöntem içinde %100 olduğundan sadece test setine ait tanıma oranları bulunmuştur. Eşitlik 9'da verilen Sıfır Alt Uzayı Ölçütünün sayısal değerleri ile tanıma oranları Tablo 3'de verilmiştir.

Tablo 3'de görüldüğü üzere AEOV yöntemi PCA+GDA yönteminden daha iyi tanıma oranları üretmiştir. Test seti tanıma oranları sıfır altuzayı ölçütünün sayısal değerleri ile orantılı çıkmıştır. Buda sıfır altuzayı ölçütünün, eğitim setindeki örnek sayısının örnek uzayı boyutundan küçük

olduğu durumlar için ideal bir ölçüt olduğunu deneysel olarak ispatlamaktadır.

TABLO 3
Test Seti Tanıma Oranları

Yöntemler	Sıfır Alt Uzayı Ölçütünün Sayısal Değeri	Tanıma Oranları
AEOV	8.02e+290	%99.14
PCA+GDA	3.54e+286	%98.29

5. Sonuç

Bu bildiriye FLDA ölçütünün eksiklikleri incelenmiş ve bu ölçütün yetersiz kaldığı durumlarda kullanılmak üzere önerilen yeni ölçütler anlatılmıştır. Bu yeni ölçütlerin yeterliliği teorik olarak tartışılmış ve bu teorik bilgiler deneysel çalışmalarla desteklenmiştir. Yapılan deneysel çalışmalar teorik bilgileri destekleyen sonuçlar üretmişlerdir. Bu bilgiler ışığında örnek uzayı boyutunun eğitim setindeki örnek sayısından küçük olduğu durumlarda FLDA yönteminin ODV yönteminden daha etkili bir yöntem olduğu ve örnek uzayı boyutunun eğitim setindeki örnek sayısından büyük olduğu durumlarda ise AEOV yönteminin PCA+GDA yönteminden daha etkili bir yöntem olduğu gözlenmiştir.

6. Kaynakça

- [1] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *In Annals of Eugenics*, vol. 7, pp 179-188, 1936.
- [2] K. Fukunaga, *Introduction to Statistical Pattern Recognition*. New York: Academic Press, 1990.
- [3] D. H. Foley, J. W. Sammon, "An optimal set of discriminant vectors," *IEEE Transactions on Computers*, c-24-3 (1975) 281-289.
- [4] T. Okada, S. Tomita, "An optimal orthonormal system for discriminant analysis," *Pattern Recognition*, 18, 139-144, 1985.
- [5] Q. Tian, M. Barbero, Z. H. Gu, and S. H. Lee, "Image classification by the Foley-Sammon transform," *Opt. Eng.*, vol. 25, no. 7, pp. 834-840, 1986.
- [6] Z. Q. Hong, J. Y. Yang, "Optimal discriminant plane for a small number of samples and design method of classifier on the plane," *Pattern Recognition*, 24, 317-324, 1991.
- [7] D. L. Swets, J. Weng, "Using discriminant eigenfeatures for image retrieval," *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 18-8 (1996) 831-836.
- [8] H. Cevikalp, M. Neamtu, M. Wilkes, and A. Barkana, "Discriminative common vectors for face recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, pp. 4-13, January 2005.
- [9] H. Cevikalp, M. Neamtu, and M. Wilkes, "Discriminative common vectors with kernels," *IEEE Transactions on Neural Networks*, accepted.
- [10] K. Liu, Y-Q Cheng, and J-Y Yang, "A generalized optimal set of discriminant vectors," *Pattern Recognition*, 25-7, 731-739, 1992.
- [11] M. van Breukelen, R.P.W. Duin, D.M.J. Tax, and J.E. den Hartog, "Handwritten digit recognition by combined classifiers," *Kybernetika*, 34-4 pp. 381-386, 1998.
- [12] A. M. Martinez and R. Benavente, "The AR face database," CVC Tech. Report #24, 1998.