HORA 2019
www.horacongress.com

International Congress on Human-Computer Interaction, Optimization and Robotic Applications
July 5-7, 2019, Urgup, Nevşehir, Turkey

SETSCI Conference Proceedings
www.set-science.com
8: 184-187, 2019

Kaydırma Biyometriği ile Kişilerin Doğrulaması için Çekirdek Fisher Ayırtacına Dayalı Yeni Bir Yaklaşım

Orhan Sivaz[1*] ve Murat Aykut[1]

[1]*Bilgisayar Müh. Böl. / Müh. Fakültesi, Karadeniz Teknik Üniversitesi, TÜRKİYE*
[*]*(osivaz@ktu.edu.tr)*

Özet **–** Son yıllarda bilgi güvenliğini sağlamada biyometrik doğrulama sistemlerinin kullanımı yaygınlaşmış ve farklı biyometrikler ortaya atılmıştır. Çalışmamızda dokunmatik ekran üzerinde yapılan hareketlerden kaydırma hareketine bağlı olarak kişilerin ayırt edilmesi ele alınmıştır. Bu kapsamda, öncelikle kişilerin dokunmatik ekran üzerinde kaydırma hareketinden elde edilen giriş özellik sayısı 28'den 37'ye çıkartılmıştır. Bu özellikler doğrusal olmayan Çekirdek Fisher Ayırtacı yöntemi ile daha ayırt edilebilir özelliklere dönüştürülmüş ve Destek Vektör Makineleri ile sınıflandırılmıştır. Bu sınıflandırmaya paralel olarak daha önce imza tanıma sistemlerinde kullanılan 100 özellikten kaydırma biyometriği ile ilgili 61 özellik Sıralı İleri Kayan Arama (SFFS) yöntemi ile 5'e düşürülüp GMM yöntemi ile de istatistiksel sınıflandırma yapılmıştır. Her iki sınıflandırmadan elde edilen skor değerleri ortalama alınarak birleştirilerek nihai karar verilmiştir. Deneyler için 190 kişiden iki farklı oturumda yatay ve dikey konumlarda sola, sağa, yukarı ve aşağı kaydırma hareketleri ile oluşturulan ortak kullanıma açık Serwadda Veritabanı kullanılmış, EER üzerinden başarı değerlendirmesi yapılmıştır. Sonuçlar incelendiğinde giriş özellik sayısının 37'ye çıkartılması ve ön işlem olarak Çekirdek Fisher Ayırtacının kullanılmasının sonuca pozitif yönde etki ettiği görülmüştür. Ayrıca, literatürdeki yöntemlerle kıyaslandığında başarıda azımsanmayacak bir iyileşmenin sağlandığı görülmüştür.

Anahtar Kelimeler – kaydırma biyometriği, ÇFA, DVM, GMM.

A Novel Approach based on Kernel Fisher Discriminants for User Authentication with Swipe Biometrics

Abstract **–** Recently, the use of biometric verification systems to secure the information has become popular and various biometrics have been proposed. In our study, the recognition of the individuals via swipe gesture on a touch screen is discussed. In this context, firstly, the number of input features obtained from a swipe gesture of an individual on a touch screen has been increased from 28 to 37. Then, these features are transformed into more distinctive features by nonlinear Kernel Fisher Discriminants method and classified with Support Vector Machines. Parallel to this classification, 61 features related to swipe biometry from 100 features previously used in signature recognition systems are reduced to 5 by Sequential Forward Floating Search (SFFS) method and statistical classification is made by GMM method. The final decision was made by fusing the average score values obtained from both classifiers. For the experiments, publicly available Serwadda Database, which was constituted by the left, right, up and down swipe gesture of 190 individuals, acquired horizontally and vertically in two different sessions, was used and performance evaluation has been performed through EER. When the results are analyzed, it can be seen that increasing the number of input features to 37 and using the Kernel Fisher Discriminants as a preprocessing step has a positive effect on the result. Furthermore, a significant improvement in performance has been observed when compared to the methods in the literature.

Keywords – swipe biometric, KFD, SVM, GMM.

I. GİRİŞ

Günümüzde kullanıcıları dokunmatik cihazlarda doğrulamak için kullanılan çoğu yöntem bir giriş noktası tanımlar. Tipik olarak kullanıcıdan bir şifre girilmesi beklenir ve eğer şifre doğruysa sisteme erişim izni verilir. Bu giriş tabanlı yöntemler çok popüler olsa da kullanılabilirlik ve güvenlik açısından bazı eksikleri vardır. Şifreler ve PIN kodları kısa ve hatırlanması kolay olduğundan kırılmaları da kolaydır. Gizli dokunma modelleri gibi diğer kimlik doğrulama alternatifleri yaygın olmalarına rağmen sınırlamalara sahiptirler. Örneğin aynı modele sık sık girdikten sonra cihazın ekranında kalan kalıntıları takip etme gibi saldırılara karşı hassastırlar.

Yukarıda belirtilen dezavantajlara ek olarak geleneksel güvenlik sistemleri tarafından sunulan temel sınırlama, kullanıcıların yalnızca oturumun başında bir kez doğrulanması gerçeğidir. Kullanıcı sadece oturumun başında bir kez doğrulanır ve oturum sonlanana kadar önemli bilgiler tehlikeye girebilir. Bu durum sürekli biyometri olarak adlandırılan, kullanıcının periyodik olarak kimliğini doğrulayan ve böylece giriş noktasının ötesinde cihazda güvenliği garanti eden, bir araştırma alanına yol açmıştır.

DOI: 10.36287/setsci.4.5.036

Kullanılan yaklaşım, kullanıcının cihazla olan etkileşimini izler ve sistem meşru kullanıcının yaptığı kaydırma hareketlerine bakarak, cihazı kullanıp kullanmadığını kontrol eder. Tek parmak dokunuşlu kaydırma hareketi, kullanıcının bir parmağı dokunmatik ekrana yerleştirdiği ve tipik olarak kaydırma amacıyla yatay ve dikey olarak hızlıca hareket ettirdiği dokunma hareketi olarak kabul edilir.

Yapılan çalışmalar incelendiğinde;
Fierrez vd. [1], 190 kullanıcılı veri tabanından yerel ve global özellikler çıkarmış, En Küçük Kareler-Destek Vektör Makineleri (EKK-DVM) ve Gauss Karışım Modeli'nin füzyonu olan bir sistem önermişlerdir.

Serwadda vd. [2], 190 kullanıcı içeren bir veri tabanı kullanarak 28 özellik çıkarmış ve bu özellikleri içeren sistemi Lojistik Regresyon, Destek Vektör Makineleri ve Rastgele Orman (Random Forest) algoritmalarını kullanarak eğitmişlerdir.

Xu vd. [8], 32 kullanıcılı bir veri tabanından 37 özellik çıkarmış ve bu özelliklere Destek Vektör Makineleri algoritmasını uygulamışlardır.

Antal vd. [3], 71 kullanıcı içeren bir veri tabanından 15 özellik çıkarmış ve bu özelliklere Destek Vektör Makineleri, Rastgele Orman (Random Forest) ve K-En Yakın Komşu algoritmalarını uygulamışlardır.

Manhub vd. [9], 48 kullanıcı bir veri tabanı kullanarak 24 özellik çıkarmış ve bu özellikleri K-En Yakın Komşu, Destek Vektör Makineleri ve Rastgele Orman (Random Forest) gibi algoritmalarla eğitmişlerdir.

Bu çalışmada [1]'den farklı olarak yerel özelliklerin sayısı 28'den 37'ye çıkarılmış ve özellik çıkarımı aşamasında bu 37 özelliğe sınıf içi varyansı minimize edip sınıflar arası varyansı maksimize eden Çekirdek Fisher Ayırtaçları (Kernel Fısher Discriminant Analysis-KFDA) yöntemi uygulanmıştır. Sınıflandırma aşamasında ise [1]'e benzer şekilde En Küçük Kareler-Destek Vektör Makineleri (Least Squares Support Vector Machines – LS-SVM) ve Gauss Karışım Modelinin (Gaussian Mixture Model-GMM) skor füzyonu olan ve başarıyı önemli derecede arttıran bir sistem kullanılmıştır. Bu çalışmada sadece oturumlar arası senaryo incelenmiştir. Oturumlar arası senaryo kullanıcının uzun süre sonra dokunmatik ekranla etkileşime girdiği senaryodur. Bu senaryoda ilk oturumdaki veriler eğitim için kullanılırken, ikinci oturumdakiler ise test için kullanılır.

II. MATERYAL VE YÖNTEM

A. Ön İşlemler

Ön işlem adımında yatay ve dikey yönden gelen veriler ayrı ayrı ele alınır. Veri tabanındaki kaydırma hareketleri yatay ve dikey olarak gruplandırılır. Daha sonra hareketler yönlerine göre yukarı, aşağı, sol ve sağ olmak üzere ayrılır. Yönü belirlenen her bir veri seti için iki özellik vektörü hesaplanır. Bu çalışmada ilk olarak [1]'de kullanılan 28 boyutlu özellik vektörüne 9 özellik daha eklenmiştir. Bu özellikler şunlardır:

- Kaydırma hareketinin uzunluğunun, başlangıç ve bitiş noktaları arasındaki mesafeye oranı [7].
- Başlangıç ve bitiş noktaları arasındaki doğrunun kaydırma hareketine olan uzaklıklarının vektörü. Bu uzaklık vektörünün ortalaması, standart sapması, birinci, ikinci ve üçüncü çeyrek bilgisi.
- Başlangıç ve bitiş noktaları arasındaki doğrunun kaydırma hareketine olan maksimum uzaklığı ve maksimum uzaklığın koordinatları [3].

Diğer özellik vektörü ise çevrimiçi imza doğrulaması için [4]'de sunulan özellik vektöründen uyarlanmıştır. Bu özellik vektörünün boyutu 100 olmasına rağmen sadece 61 tanesi kaydırma biyometriği ile ilgilidir. Sequential Forward Floating Search (SFFS) algoritması bu 61 özelliğe uygulanarak en iyi alt kümeyi ifade eden 5 özellik seçilmiştir. Sonuç olarak biri 37 boyutlu diğeri 5 boyutlu iki özellik vektörü elde edilmiştir. Her iki özellik vektörü için (0-1) aralığına normalleştirme tanh normalizasyonu ile yapılmıştır.

B. Özellik Çıkartma

1.Çekirdek Fisher Ayırtaçları

Doğrusal Fisher Ayırtacı, sınıflar arası varyansı en üst düzeye çıkararak ve sınıf içi varyansı en aza indirerek en iyi ayırma hiperdüzlemini bulmayı amaçlamaktadır. Çekirdek Fisher Ayırtaçları (ÇFA) adı verilen doğrusal olmayan Fisher Ayırtaçları, bu yapıya çekirdek hilesi eklenerek doğrusal olmayan biçime kavuşturulmuş şeklidir [5]. Alternatif olarak, Mika vd. [6] çok sayıda örnek için çıkabilecek problemlerin üstesinden gelmek amacıyla dışbükey karesel programlamayı (optimizasyon problemi) kullanmayı önermiştir:

$$\min_{a,b,\xi} ||\xi||^2 + CP(\alpha) \quad (1)$$

(2) ve (3) nolu şartlar altında (1) nolu denklem minimize edilir.

$$\mathbf{K}\alpha + \mathbf{1}b = \mathbf{y} + \xi \quad (2)$$

$$\mathbf{1}_i^T \xi = 0, i = 1,2 \quad (3)$$

$\alpha, \xi \in \mathbb{R}^l \; ve \; b, C \in \mathbb{R}, C \geq 0$. Burada **K** çekirdek matrisi, b bias terimi, C ve P düzenleştirme fonksiyonu ve düzenleştirme parametresidir. $||\xi||^2$ terimi hatanın varyansını minimize eder. İlk koşul her bir örneğin çıktısını sınıf etiketine çeker; ikinci koşul ise her sınıf için ortalama çıktının etiket değerine eşit olmasını garanti eder. Daha sonra x test örneği için ÇFA özellikleri (4) nolu denklemle hesaplanır.

$$(\mathbf{w}\cdot\boldsymbol{\phi}(\mathbf{x})) = \sum_{i=1}^{l} \alpha_i k(\boldsymbol{x_i}, \boldsymbol{x}) \quad (4)$$

Burada $k()$ çekirdek fonksiyonu ve l örnek sayısıdır. Çalışmamızda çekirdek fonksiyonu olarak Radial Basis Function (RBF) kullanılmış, parametreler deneysel olarak belirlenmiştir.

C. Sınıflandırma

Sınıflandırma aşamasında biri ayırıcı bir diğeri istatistiksel olmak üzere iki farklı sınıflandırıcı kullanılmıştır. Ayırıcı kısım RBF çekirdek ile Destek Vektör Makinerinden oluşurken (DVM) [11], istatistiksel kısım ise Gauss Karışım Modellerinden oluşmaktadır.

1.En Küçük Kareler Destek Vektör Makinaları

DVM [11], genelleştirme hatasının üst sınırını en aza indiren yapısal risk minimizasyonuna dayalı ayırıcı öğrenme yöntemidir. DVM'nin ana mantığı maksimum marjine sahip ayırma hiperdüzleminin oluşturulmasıdır. Çekirdek fonksiyonunun kullanılması ile açıkça özellik uzayına eşlemeden ayırma hiperdüzlemi doğrusal olmayan bir yapıya kavuşturulmuş olur [12]. Bu çalışmada kullanılan EKK-DVM daha hızlı eğitim gerçekleştiren bir DVM çeşitidir ve genel yapısı şu şekildedir:

$$y = \text{sgn}\left(\sum_{i=1}^{N} y_i \alpha_i K(x, x_i) + b\right) \quad (5)$$

Burada x, d boyutlu giriş örneği, y sınıf etiketi, x_i de i. eğitim örneği vektörüdür. α vektörü (7) ve (8) koşulları altında (6) numaralı ikinci dereceden programlama probleminin çözümü ile elde edilir.

$$\min_{\alpha} \alpha^T \Omega \alpha + CP(\alpha), \quad C = \frac{1}{\gamma},$$
$$P(\alpha) = \|\alpha\|^2, \ \Omega_{kl} = y_k y_l K(x_k, x_l), \ k, l = 1, \dots, N \quad (6)$$

$$y^T \alpha = 0 \quad (7)$$
$$yb + y_k y_l \, K(x_k, x_l)\alpha + \frac{aI}{\gamma} = 1 \quad (8)$$

Burada C, marjinin geniş olması ile marjin hatalarının küçüklüğü arasındaki ödünleşimi ifade eden önceden tanımlanmış bir parametredir [13].

2.Gauss Karışım Modeli

Eğitim verisini, birbirinden bağımsız birden fazla Gauss dağılımının bileşimiyle tanımlayan bir modeldir. Bu modelde eğitim örneklerinin, Gauss dağılımına sahip birden fazla bağımsız kaynaktan üretildiği varsayılıp, bu kaynaklara ait Gauss parametrelerinin, karışımın olasılık yoğunluk işlevini maksimize edecek şekilde optimizasyonu gerçekleştirilir. Böylece, veri setinin tek bir dağılımdan üretildiğinin varsayılıp, gerekli dağılım parametrelerinin kestirimiyle modelleme yapan sistemlerin yetersiz kaldığı durumlarda başarılı sonuçlar verebilmektedir [10].

3.Füzyon Sistem

Füzyon sistemde her iki yöntemin skorları, öncelikle tanh normalizasyonu ile normalize edilir. Daha sonra skorların ortalaması alınarak nihai skor oluşturulur. Bu hibrit sistem sayesinde ayırıcı veya istatistiksel sistemden oluşabilecek negatif etki diğeriyle dengelenmektedir.

III. BULGULAR VE TARTIŞMA

A. Veri Seti

Bu çalışmada kullanılan Serwadda veri seti 197 kullanıcının yaptığı kaydırma hareketlerinden oluşmaktadır. Dokunmatik ekranı dikey kullanan kullanıcıların sayısı 138 iken, yatay kullananların sayısı 59'dur. Tablo I'de eğitim aşamasında kullanılacak örnek sayısına göre kullanıcı sayısı verilmiştir. Bazı kullanıcılara ait örnek sayısı az olduğundan tabloda belirtilen minimum eğitim örneği sayısını sağlayan kişi sayısı da değişiklik göstermektedir.

B. Deneyler

Çalışmamızda doğrulama performansını ölçmek için literatürde en yaygın kullanılan EER (Equal Error Rate) ölçütünden yararlanılmıştır. Veri setinden elde edilen 37 boyutlu ilk özellik vektörüne uygulanan boyut indirgeme algoritmalarından ÇFA için gauss genişliği t=25 alınmıştır. EKK-DVM için düzenlileştirme ve çekirdek parametreleri deneysel olarak belirlenmiştir. 5 boyutlu özellik vektörüne uygulanan GKM için uygunluk terimi r=30 seçilmiştir.

TABLO I
KULLANILAN ÖRNEK SAYISI VE DOKUNMATİK EKRANIN KULLANIM YÖNÜNE GÖRE KULLANICI SAYISI

	Dikey			Yatay		
Yön/Örnek Sayısı	**40**	**60**	**80**	**40**	**60**	**80**
Yukarı	124	74	36	54	44	25
Aşağı	132	124	99	55	53	51
Sol	104	73	38	39	22	9
Sağ	118	97	67	45	36	22

Tablo II incelendiğinde, sınıflandırma için kullanılacak örnek sayısı 40 olarak belirlendiğinde dikey kullanımda EER %20'lere, yatay kullanımda ise %11'lere kadar düşmektedir. Örnek sayısı arttıkça hem dikey hem de yatay kullanımda sınıflandırma performansı artmakta, örnek sayısı 80 olduğunda ise dikey kullanımda EER %14'lere, yatay kullanımda ise %4'lere kadar düşmektedir. Eklenen 9 özellik dikey kullanımda yukarı, aşağı ve sola yapılan kaydırmalarda sınıflandırma performansını %1'e yakın artırırken, sağa yapılan kaydırmalarda sınıflandırma performansı aynı kalmaktadır. Yatay kullanımda ise ekran boyutları düşük olduğu için performans çok az derecede kısmen artmaktadır.

TABLO II
EER AÇISINDAN DVM SINIFLANDIRMA PERFORMANSI

Örnek Sayısı	**Yön**	**Önerilen Yöntem (ÇFA+EKK-DVM)**		**Fierrez ve diğ.[11]**			
		Dikey	**Yatay**	**EKK-DVM**		**GKM**	
				Dikey	**Yatay**	**Dikey**	**Yatay**
40	**Yukarı**	**21.4**	15.2	23.8	**14.7**	22.7	17.9
	Aşağı	**20.5**	16.9	22.9	**16.1**	20.6	18.9
	Sol	**20.0**	**11.0**	21.6	13.2	18.2	19.5
	Sağ	**21.2**	14.9	22.9	**11.8**	20.8	19.4
60	**Yukarı**	**19.1**	**14.0**	-	-	21.5	18.0
	Aşağı	**18.4**	**15.1**	-	-	20.0	18.6
	Sol	**16.6**	**7.7**	-	-	17.5	16.7
	Sağ	**19.3**	**11.1**	-	-	20.5	17.6
80	**Yukarı**	**18.7**	**11.2**	-	-	21.0	17.5
	Aşağı	**16.0**	**13.6**	-	-	20.4	18.9
	Sol	**14.2**	**4.1**	-	-	16.1	11.6
	Sağ	**18.0**	**8.5**	-	-	20.0	22.3

Tablo III veri setinden elde edilen özellik vektörlerine uygulanan EKK-DVM ve GKM yöntemlerinin ortalama skor füzyonunu göstermektedir. Örnek sayısı 80 olarak seçildiğinde dikey kullanımda EER %10'lara, yatay kullanımda ise %4'lere kadar gerilemektedir.

TABLO III
EER AÇISINDAN FÜZYON SINIFLANDIRMA PERFORMANSI

		Önerilen Yöntem (ÇFA+DVM+GKM)		**Fierrez ve diğ.[11]**	
		Dikey	**Yatay**	**Dikey**	**Yatay**
40	**Yukarı**	**16.2**	11.0	17.4	**10.7**
	Aşağı	**14.7**	12.7	15.9	**12.2**
	Sol	**13.6**	**10.0**	13.9	12.1
	Sağ	**15.1**	12.5	16.1	**12.0**
60	**Yukarı**	**15.3**	**11.3**	-	-
	Aşağı	**14.1**	**11.9**	-	-
	Sol	**12.7**	**7.2**	-	-
	Sağ	**15.3**	**10.0**	-	-
80	**Yukarı**	**14.2**	**10.6**	-	-
	Aşağı	**12.5**	**11.2**	-	-

	Sol	10.2	4.0	-	-
	Sağ	14.9	10.3	-	-

IV. SONUÇLAR

Bu çalışmada dokunmatik ekran ile olağan etkileşimi kullanan kaydırma biyometrileri üzerinde çalışılmıştır. Bu kaydırma biyometrilerinden biri 37 diğeri 5 boyutlu olmak üzere yerel ve global özellik vektörleri oluşturulmuş ve bunlardan 37 boyutlu özellik vektörüne boyut indirgeme algoritmalarından olan ÇFA uygulanmıştır. Boyut indirgemeden sonra ise EKK-DVM ile sınıflandırma yapılmıştır. 5 boyutlu diğer özellik vektörüne ise GKM algoritması uygulanmıştır. Biri ayırıcı bir diğeri istatistiksel olan iki sistemin skorları normalize edilip, ortalamaları alınarak füzyon bir sınıflandırma oluşturulmuştur. Gerek giriş boyutundaki artırımın gerekse ÇFA yönteminin uygulanmasının sonuçlara pozitif yansıdığı deneysel sonuçlarla gözlemlenmiştir. Ayrıca yatay kullanımda elde edilen performansın, dikey kullanımdan daha iyi olduğu görülmüştür..

Gelecekteki çalışmalar için kaydırma hareketlerinin istikrarsızlığı probleminde daha iyi performans gösterebilecek yöntemler üzerinde durulabilir.

KAYNAKLAR

[1] J. Fierrez, A. Pozo, M. M. Diaz, J. Gallbally and A. Morales, "Benchmarking Touchscreen Biometrics for Mobile Authentication, "*IEEE Trans. Inf. Forensics Security*, vol. 13, no. 11, pp. 2720-2733,Nov. 2018

[2] A. Serwadda, V. V. Phoha, and Z. Wang, "Which verifiers work?:A benchmark evaluation of touch-based authentication algorithms," in *Proc. IEEE BTAS*, 2013, pp. 1–8.

[3] M. Antal, Z. Bokor, and L. Z. Szabó, "Information revealed from scrolling interactions on mobile devices," *Pattern Recognit. Lett*.,vol. 56, pp. 7–13, Apr. 2015.

[4] M. Martinez-Diaz, J. Fierrez, R. P. Krish, and J. Galbally, "Mobile signature verification: Feature robustness and performance comparison," *IET Biometrics*, vol. 3, no. 4, pp. 267–277, 2014.

[5] S. Mika, G. Ratsch, J. Weston, B. Scholkopf and K.R. Mullers, "Fisher discriminant analysis with kernels," in *Proc. Neural Networks for Signal Processing IX: Proceedings of the 1999 IEEE Signal Processing Society Workshop*, 1999, pp. 41-48

[6] S. Mika, G. Ratsch and K.R. Müller, "A mathematical Programming Approach To the Kernel Fisher Algorithm," in *Proc. NIPS*, 2000, pp. 591-597

[7] M. Frank, R. Biedert, E. Ma, I. Martinovic, and D. Song, "Touchalytics: On the applicability of touchscreen input as a behavioral biometric for continuous authentication," *IEEE Trans. Inf. Forensics Security*, vol. 8, no. 1, pp. 136–148, Jan. 2013.

[8] H. Xu, Y. Zhou, and M. R. Lyu, "Towards continuous and passive authentication via touch biometrics: An experimental study on smartphones," in *Proc. SOUPS*, 2014, pp. 187–198.

[9] U. Mahbub, S. Sarkar, V. M. Patel, and R. Chellappa, "Active user authentication for smartphones: A challenge data set and benchmark results," in *Proc. IEEE BTAS*, 2016, pp. 1–8.

[10] K. Herkiloğlu, "Gauss arışım modelleri kullanılarak ses imzalarının sınıflandırılması," M.Sc.Thesis, İTÜ, İstanbul, Turkey, Aug. 2005.

[11] V. Vapnik, *The Nature of Statistical Learning Theory*, 1st ed., Springer-Verlag, 1995.

[12] S. Abe, *Support Vector Machines for Pattern Classification*, 1st ed., Springer, 2005.

[13] M. Aykut, E. Gedikli and M. Ekinci, "Avuç İzine Dayalı Kimlik Doğrulama Sistemi," in *Proc. SIU*, 2010, pp. 117-120