

TÜRKÇE ELYAZISI TANIMA SİSTEMLERİNDE ÖZNETELİK ÇIKARMA VE SINIFLANDIRMA YÖNTEMLERİNİN KARŞILAŞTIRMASI

Abdulkerim ÇAPAR, Kadim TAŞDEMİR, Özlem KILIÇ, Muhittin GÖKMEN
İstanbul Teknik Üniversitesi
Çoğulortam Merkezi
Maslak, İstanbul

capar@itu.edu.tr, tasdemis@be.itu.edu.tr, kilicoz@itu.edu.tr, gokmen@cs.itu.edu.tr

Özet: Bu çalışmanın amacı Türkçe elle yazılmış büyük harfleri tanımadır. Öncelikle, İstanbul Teknik Üniversitesi öğrencilerinden alınan örneklerle Türkçe El yazısı Veri tabanı oluşturulmuştur. Bu veri tabanında yaklaşık 20000 büyük harf ve 7000 rakam örneği bulunmaktadır. Bu çalışmada Türkçe karakterler için daha başarılı tanıma sistemi bulmak üzere, çeşitli öznetelik bulma ve sınıflandırma teknikleri gerçekleştirilmiş ve denenmiştir. Bu veri tabanındaki görüntülerin Karhunen-Loève Dönüşümü, Açısız Radyal Dönüşüm, Zernike Momentleri, Şekilsel Öznetelikleri bulunarak elde edilmiş öznetelikler, Yapay Sinir Ağları, K-En Yakın Komşuluk, En Yakın Ortalama, Bayes, Parzen ve Boyut Bağımlı Eksi-Log-Olabilirlik (Size Dependent Negative Log-Likelihood) metotlarıyla sınıflandırılmışlardır.

1. Giriş

Optik Karakter Tanıma (OKT) örüntü tanıma alanında en çok umut veren alanlardan biridir. Fakat, bu uygulamalar sadece yüksek kalitede basılmış metinleri tanımak için kullanılmıştır. El yazısı karakterlerinin bilgisayarlar tarafından tanınması optik karakterlere göre daha az ilgi çekmiştir [1]. El yazısı karakter tanıma başarımını arttırmak için iki önemli hedefe ulaşılması gerekmektedir: anlamlı ayrıntıları kaçırmayan bir öznetelik bulma yöntemi tasarımı ve genelleme gücü yüksek ama ayırt etme hatası düşük sınıflandırıcı saptanması [2]. Türk alfabesinde bulunan özel bazı karakterlerden dolayı (Ö, Ü, İ, Ğ, Ş, Ç gibi), Türkçe karakterlerin tanınması, Latin harflerinin tanınmasından biraz daha zor bir problemdir.

Bu konuda bazı ticari çalışmalar olduğu bilinmektedir, fakat bu çalışma Türkçe el yazısı tanıma alanında, anlaşılır ve yeniden kullanılabilir karakter veri tabanına sahip olmasıyla, akademik anlamdaki birkaç çalışmadan biridir. Bu çalışmadaki hedef Türkçe karakterlerin tanınması için özel tanıma kuralları çıkarılmasıdır.

2. Elyazısı Tanıma (ET) Sistemi

El yazısı, insanla bilgisayar arasında iletişimde en genel ve en doğal araçtır. ET sistemlerinin tarihsel gelişimi OKT (Optik Karakter Tanıma) sistemlerine dayanmaktadır [3]. OKT problemi ise bilgisayarın tarihi kadar eskidir. Şu anki araştırmalar elle yazılmış rakam ve karakter tanıma odaklanmıştır. Ne yazık ki, el yazısındaki yüksek değişimlerden dolayı OKT' deki başarıya AKT' de ulaşamamıştır. Suen çeşitli sınıflandırma yöntemleri kullanarak elle yazılmış tanıma sistemleri kuran araştırmacıların çalışmalarını bir araya getirmiştir [4]. Suen, Yüksek başarıma ulaşmaktaki anahtarın güçlü öznetelikler bulmak olduğuna dikkat çekmiştir.

Bir ET sistemi üç ana parçadan oluşur: Ön işlemler, öznetelik bulma ve sınıflandırma.

A. Ön İşlemler

Ön işlemler ileriki adımlar için gerekli bilgileri sağlamaktadır. Başka bir deyişle gerçek dünyadaki karakter ile tanıma motoru arasındaki ilişkiyi kurabilmek için ön işlemlerin yapılması gereklidir.

Ön işlem adımları sayısallaştırma, ikili görüntü elde etme, karakter bölütleme, düzgeleme ve inceltmeden oluşmaktadır.

B. Öznetelik Bulma

Öznetelik bulma sorunu işlenmemiş veriden sınıflar arası değişintiyi yüksek tutan ve sınıf içi değişintiyi azaltan nitelikteki bilgilere ulaşmaktır.

Bu çalışmada kullanılan öznetelik bulma yöntemleri, Karhunen-Loève dönüşümü, Açısız Radyal dönüşüm, Zernike momentleri ve şekilsel özneteliklerin bulunması olarak sıralanabilir.

1) Karhunen-Loève Dönüşümü (KLD): KLD, örüntü tanıma problemlerinde sıkça kullanılmakta ve güvenilir sonuçlar vermektedir. KLD, görüntüleri, eğitim kümesindeki karakter görüntülerince yaratılmış olan ortak değişinti matrisinin, özvektörlerine izdüşüren lineer bir dönüşümdür [5].

2) Zernike Moments: Zernike [6] birim çember içinde karmaşık tam dikgen küme çokterimliler tanımlamıştır ve bu çokterimlileri Kartezyen konaçlarda $V_{nm}(x, y)$ olarak göstermiştir. Kutupsal konaçlarda ise Zernike çokterimlileri ;

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho) \exp(jm\theta) \quad (1)$$

olarak ifade edilir.

n	sıfır veya pozitif tam sayı
m	pozitif ve negatif tamsayı öyle ki $n- m $ her zaman çift ve $ m \leq n$
ρ	(x, y) beneğinin orijine olan Euclid uzaklığı
θ	(x, y) vektörünün x eksenine ile saat yönünde yaptığı açı

Zernike momentlerinin hesabı için önce görüntü belirli bir β oranıyla yeni bir boyuta düzgelendir [7].

3) Açısal Radyal Dönüşüm: Açısal Radyal Dönüşüm birim çember içinde kutupsal konaçlarda tanımlı karmaşık bir dönüşümdür,

$$F_{nm} = \langle V_{nm}(\rho, \theta), f(\rho, \theta) \rangle = \int_0^{2\pi} \int_0^1 V_{nm}^*(\rho, \theta) f(\rho, \theta) \rho d\rho d\theta \quad (2)$$

Öyle ki, F_{nm} n. ve m. dereceden ARD katsayısı, $f(\rho, \theta)$ kutupsal konaçlardaki görüntü işlevi, ve $V_{nm}(\rho, \theta)$ ise ARD temel işlevidir [8]. ARD temel işlevi açısal ve radyal kısımlarına ayrılabilir,

$$V_{nm}(\rho, \theta) = A_m(\theta) R_n(\rho) \quad (3)$$

ARD'nin açısal ve radyal kısımları aşağıdaki eşitliklerle tanımlanmıştır:

$$A_m(\theta) = \frac{1}{2\pi} \exp(jm\theta), \quad (4)$$

$$R_n(\rho) = \begin{cases} 1 & n = 0 \\ 2 \cos(\pi m \rho) & n \neq 0 \end{cases} \quad (5)$$

4) Karakterlerin Şekilsel Öznitelikleri: KLD insan algılama mekanizmasına benzer hiçbir bir model kabul etmez, yalnızca istatistiksel olarak el yazısı karakterlerinin nasıl oluştuğuyla ilgilenir [9]. Aynı karakterin farklı çeşitlerde yazılmış olanlarını tanımak için, genelleşmiş bazı ek öznitelikler kullanılmalıdır.

Karakterin son noktaları ve bağlantı noktaları, özellikle Türkçe'deki 'Ç' ve 'Ş' harfleri için, karakterle ilgili sağlıklı bilgiyi taşıyabilecek şekilsel özniteliklerdir. Çoğu karakter tanıma sistemi karaktere ait beneklerle ilgilenmektedir, fakat karakteri çevreleyen arka planın da boyut, şekil ve pozisyonu ile ilgili bilgileri de şekilsel öznitelikler arasında sıralayabiliriz. Ayrıca bu arka planlar üzerinde Hu'nun ikinci ve üçüncü dereceden merkezi moment değişmezleri de hesaplanmıştır [9]. Karakterdeki bağlantılı bileşenlerin sayısı da, özellikle 'İ', 'Ü' ve 'Ö' gibi bazı Türkçe karakterler için şekilsel öznitelik olarak kullanılabilir.

Sonuç olarak karakterlerin şekilsel özellikleri kullanılarak 39 adet öznitelik elde edilmiştir.

C. Sınıflandırma

K-en yakın komşu, en yakın ortalama, Bayes, Parzen, boyut bağımlı NLL (Size Dependent Negative Log-Likelihood) ve yapay sinir ağları oluşturulan öznitelik vektörlerini sınıflandırmada kullanılmıştır.

1) K-En Yakın Komşu: Tanıma, tanınacak öznitelik vektörüne en yakın öznitelik vektörlerinin bulunmasına bağlıdır. En yakın öznitelik vektörlerini bulurken çeşitli metrikleri kullanılabilir. Bu bildiride, Euclid metriği kullanılmıştır. Tanıma yapılacak öznitelik vektörüne en yakın k komşu bulunur. Daha sonra bu k komşu en fazla hangi sınıfa ait ise, o sınıf tanıma sonucu olarak atanır [2].

2) En Yakın Ortalama Sınıflandırıcı: Sadece birinci dereceden istatistikler kullanılır. İlk önce, sınıfların ortalamaları elde edilir. Daha sonra, çeşitli uzaklık metrikleri kullanılarak tanınacak örneğe en yakın ortalamaya sahip sınıf tanıma sonucu olarak atanır. Bu çalışmada Euclid metriği kullanılmıştır [2].

3) Bayes Sınıflandırıcısı: Bayes sınıflandırıcısında, ayrıtaç işlevi

$$g_c(x) = -\frac{1}{2}(x - \mu_c)^T \Gamma_c^{-1}(x - \mu_c) - \frac{1}{2} \ln(\Gamma_c^{-1}) + \ln P_c \quad (6)$$

şeklindedir. Bu denklemde μ_c , c sınıfındaki örneklerin ortalama değeri; Γ_c c sınıfının ortak değışinti matrisidir.

P_c değeri ise c sınıfının önsel dağılım değeridir ve bütün sınıflar için eşit alınmıştır. Bu sınıflandırıcıda, en yüksek ayrıtaç işlevi değerine sahip sınıf tanıma sonucu olarak değerlendirilir [9].

4) Parzen Sınıflandırıcısı: Parzen sınıflandırıcısı tasarlamak için Hummels ve Fukunaga [10]'nın çalışmasındaki yöntem kullanılmıştır. Bu yöntem (6) de denklemi verilen Gauss çekirdeğini kullanarak her sınıf için koşullu olasılık değerlerini hesaplar. Gauss çekirdeğindeki d değeri Mahalonobis uzaklık metriğidir.

$$k(x) = \frac{1}{\sqrt{(2\pi)^n |\Gamma_c| h^n}} e^{-\frac{d^2(x)}{2h^2}} \quad (7)$$

Bu çekirdek kullanılarak hesaplanan koşullu olasılık değeri,

$$p(x) = p(x | w) = \frac{1}{N_c} \sum_{i=1}^{N_c} k(x - x_i^{(c)}) \quad (8)$$

denklemini ifade edilir. Hummels ve Fukunaga yanlılık değeri için bir eşik değeri tanımlamışlardır. Farklı yanlılık şekilleri için dört adet eşik bulma yöntemi kullanılmıştır. Bu bildiride kullanılan yöntem, dağılımları Gauss

olarak varsayan yöntemdir. Bu yöntem için yanlışlık değeri,

$$t = \frac{1}{2} \left(\frac{h^2}{h^2 + 1} \right) \ln(\Gamma_c) + \frac{1}{h^2 + 1} \ln P_c \quad (9)$$

denklemleri ile gösterilir. Elde edilen yanlışlık değerini de ayrıca işlevine eklersek, işlevin yeni gösterimi

$$g_c(x) = -\ln p_c(x) - t \quad (10)$$

olur. Bu ifade P_c değeri (8) nolu denklemden koşullu olasılık değeridir. Bayes de olduğu gibi, en yüksek ayırtaç işlevi değerine sahip sınıf, tanıma sonucudur.

5) Boyut Bağımlı Eksi-Log-Olabilirlik (Size Dependent Negative-Log-Likelihood): Bu sınıflandırıcı Hwang ve Weng [11] tarafından ortaya atılmıştır. Olabilirlik kapsamında kullanılan ortak değişim uzaklığı, Mahalonobis uzaklığı, ve Euclid uzaklığı [11] de birleştirilmiştir. Her üç metrik için gerekli örnek sayısı birbirinden farklıdır. Bu üç metriğin de en verimli şekilde kullanılabilmesi için örnek sayısına uygun olarak bu metrikleri uygun katsayılar kullanarak birleştirmek gereklidir [11]. Birleştirmek için gerekli katsayılar (11) nolu denklemden verilmiştir. Bu denklemden n sayısı o sınıfa ait örnek sayısını temsil ederken q sayısı sınıf sayısını ifade eder.

$$b_e = n-1; \quad b_m = \frac{2(n-q)}{q}; \quad b_g = \frac{2(n-q)}{q^2} \quad (11)$$

Tanıma aşamasında kullanılan birleştirilmiş matris

$$W_i = w_e \rho^2 I + w_m S_w + w_c \Gamma_i \quad (12)$$

denklemleri ile gösterilir. Bu denklemden ρ standart sapma değerini gösterirken I birim matristir. S_w matrisi, i sınıfının sınıf içi dağılım matrisi; Γ matrisi ise, ortak değişim matrisidir. w_e, w_m, w_c ile gösterilen değerler, b_e, b_m, b_g nin düzeltilmiş halidir. Sonuç olarak elde edilen ayırtaç işlevi

$$g_c(x) = \frac{1}{2} (x-c)^T W_i^{-1} (x-c) + \frac{q-1}{2} \ln(2\pi) + \frac{1}{2} \ln(|W_i|) \quad (13)$$

denklemleri ile ifade edilir. Tanıma sonucu en büyük ayırtaç işlevi değerine sahip sınıftır.

6) Yapay Sinir Ağları: Bu çalışmada, sinir ağlarının en önemlisi olan çok katmanlı ileri beslemeli ağ (multilayer feedforward networks) üzerinde çalışılmıştır. Genel olarak bu ağlar, bir dizi algı birimlerinden oluşan giriş katmanından , hesaplama düğümlerinden oluşan bir veya daha fazla sayıda gizli katmandan, ve hesaplama düğümlerinden oluşan bir tane çıkış katmanından oluşur [12].

Bu çalışmada, çok katmanlı ileri beslemeli sinir ağlarını eğitim aşamasında gradyan-descent yöntemlerini kullanan hata geri yayılım algoritmaları kullanılmıştır.

3. Deneysel Sonuçları

Bu bölümde öznetelik çıkarım yöntemleri ile sınıflandırıcıların değişik şekillerde birleştirilmesi sonucu elde edilen sonuçlar anlatılacaktır. Yapılan çalışmalarda 20000 büyük Türkçe harften oluşan İTÜ Türkçe karakter veri tabanı kullanılmıştır. Veri tabanındaki karakterlerden 12644 tanesi eğitim aşaması için, 6322 tanesi ise test aşaması için kullanılmıştır.

TABLO I

İSTATİSTİKSEL VE ŞABLON TABANLI SINIFLANDIRICILAR İLE ÖZNETELİK SETLERİNİN AYRI AYRI SINIFLANDIRILMASIYLA ELDE EDİLEN YÜZDE BAŞARIMLAR

Sınıflandırıcı	Öznetelik				
	KLT (64)	KLT (128)	ZERNIKE (47)	ART (35)	ART (63)
k-NN k=1	92,0	91,7	79,8	77,5	80,1
k-NN k=3	91,9	91,6	80,7	78,5	80,9
k-NN k=5	91,9	91,5	80,9	79,0	81,1
NMN	75,3	75,4	58,7	56,0	57,7
Bayes Quad.	92,2	90,6	82,0	82,9	83,5
Parzen	92,3	90,9	81,5	83,2	84,1
SDNLL	93,0	92,2	83,0	83,1	84,2

TABLO II

İSTATİSTİKSEL VE ŞABLON TABANLI SINIFLANDIRICILAR İLE BİRLEŞTİRİLMİŞ ÖZNETELİK SETLERİNİN SINIFLANDIRILMASIYLA ELDE EDİLEN YÜZDE BAŞARIMLAR

Sınıflandırıcı	Öznetelik		
	KLT (64)+ ZERNIKE (47)	KLT (64)+ ART (35)	KLT (64)+ ART (63)
k-NN k=1	83,8	90,1	90,2
k-NN k=3	84,3	90,0	90,1
k-NN k=5	84,7	90,0	90,1
NMN	63,6	75,4	75,4
Bayes Quad.	91,9	92,1	92,1
Parzen	92,5	92,8	92,7
SDNLL	93,4	93,6	93,4

Sınıflandırıcı	Öznitelik					
	KLT (64)	KLT (128)	ART (35)	KLT (64) + ART (35)	KLT (64)+ ART (63)	KLT (64) + Geometric (39)
ANN	85,9	88,3	69,1	89,2	90,4	91,9

TABLO III
Yapay sinir ağları ile öznitelik setlerinin sınıflandırılması sonucu elde edilen başarımlar

Tablo I de istatistiksel ve şablon tabanlı sınıflandırıcıların tekil öznitelik vektörleri ile kullanıldığı zaman elde edilen yüzde başarımlar görülmektedir. Yapılan çalışmalar sonucunda, SDNLL sınıflandırıcısı ART (35) dışındaki tüm öznitelik yöntemleri için en iyi tanıma sonucunu vermiştir. ART (35) için ise en iyi sonuç, Parzen sınıflandırıcısı ile elde edilmiştir. Sonuç olarak bu özniteliklerle elde edilen en yüksek başarımlar % 93 tür.

KLT, harflerin görünüş karakteristiğini içermediği için harflerin şekil özelliklerini içeren öznitelik vektörleri eklenmelidir. Böylece başarımlar artırılabilir. Bunun için, yapılan çalışmada, KLT öznitelikleri ile Zernike öznitelikleri, ART öznitelikleri, ve şekilsel öznitelikler ayrı ayrı birleştirilmiştir. Table II de birleştirilen öznitelik vektörleri ile kullanılan sınıflandırıcıların başarımları gösterilmiştir. k-NN, NMN ve Bayes sınıflandırıcıların KLT özniteliği ile elde edilen başarımlar düşük olmasına rağmen, SDNLL ve Parzen sınıflandırıcısında bir başarımlar artışı En yüksek başarımlar Zernike ve ART özniteliklerinin eklenmesiyle artmıştır. Tanıma yüzdesi KLT ile ART için % 93.6'e; KLT ile Zernike için % 93.4'e çıkmıştır.

Yapay sinir ağları ile elde edilen sonuçlar Tablo III de gösterilmiştir. Birleştirilmiş özniteliklerin performansı KLT ve ART özniteliklerinin tek başına kullanılmasından daha iyi sonuç vermiştir. ANN için en yüksek başarımlar KLT ile eşitlenmiş geometrik öznitelikler kullanıldığında elde edilmiştir.

k- en yakın komşu sınıflandırıcısı için en iyi başarımlar, KLT özniteliklerinin 64 boyutlusu ile elde edilmiştir. En yakın ortalama sınıflandırıcısı için en iyi performans 128 boyutlu KLT ile sağlanmıştır. Bayes sınıflandırıcısı için en iyi sonuç yine KLT öznitelikleri ile sağlanmıştır. Parzen ve SDNLL sınıflandırıcıları için classifier and SDNLL, KLT ile ART in birleştirilmesiyle oluşturulan öznitelik vektörü en iyi sonucu vermiştir.

3. Sonuç

Bu bildiriye, el yazısı ile yazılmış Türkçe büyük karakterlerin tanınması üzerine çalışılmıştır. Bu çalışma için 20000 karakterden oluşan ilk Türk el yazı karakter veri tabanı oluşturulmuştur.

Yapılan çalışmada farklı öznitelik çıkarım yöntemleri kullanılmıştır. Kullanılan öznitelik yöntemleri KLT, Zernike, ART öznitelik çıkarım yöntemleridir. Bununla beraber, Türkçe karakter özelliklerine göre oluşturulmuş geometrik öznitelik vektörü kullanılmıştır. Bu özniteliklerin tek tek kullanılmasının yanında, birleştirilmiş şekilleri de öz nitelik çıkarım yöntemi olarak kullanılmıştır. Birleştirme KLT ile diğer yöntemlerden birisi ile yapılmıştır.

k-en yakın komşu sınıflandırıcısı, en yakın ortalama sınıflandırıcısı, Bayes sınıflandırıcısı, Parzen sınıflandırıcısı, SDNLL sınıflandırıcısı, ve ANN sınıflandırıcıları tanıma için kullanılmıştır. En yüksek başarımlar SDNLL ile elde edilmiştir. Birleştirilmiş öznitelikler tekil olarak kullanılanlardan daha iyi sonuçlar vermiştir. Yakalanan en iyi başarımlar oranı % 93.6' dir.

4. Kaynaklar

- [1] J. Hu, "HMM Based On-Line Handwriting Recognition", IEEE Trans. On Pattern Analysis and Machine Intelligence, v. 18, s. 1039-1045, 1996.
- [2]. K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2. baskı., Boston: Academic Press, 1990
- [3] B. Verma, M. Blumenstein, ve S. Kulkarni, "Recent Achievements in Off-Line Handwriting Recognition Systems", School of Information Technology Griffith University -- Gold Coast Campus PMB 50, 1997.
- [4] C.Y. Suen, et al., "Building a New Generation of Handwriting Recognition System", Pattern Recognition Letters, v. 14, s. 303-315, 1993.
- [5] P. J. Grother, "Karhunen-Loève Feature Extraction For Neural Handwritten Character Recognition", Image Recognition Group, National Institute of Standards and Technology, 1709, s. 155-166, 1992.
- [6] F. Zernike, *Physica*, v. 1. s. 689, 1934.
- [7] A. Khonzad, Y. H. Hong, "Invariant Image Recognition by Zernike Moments", v. 12, No. 5, Mayıs 1990.
- [8] B.D. Ripley, *Pattern Recognition and Neural Networks*, Cambridge University Press, 1996.
- [9] O. D. Trier, A. K. Jain ve T. Taxt, "Feature Extraction Methods For Character Recognition: A survey", Revised Temmuz 19, 1995.
- [10]. K. Fukunaga, and R.R. Hummels, "Bayes error estimation using Parzen and k-NN procedures", IEEE Trans. Pattern Analysis and Machine Intelligence, v.9, no 5, s. 634-643, Eylül 1987
- [11]. W.S. Hwang, J. Weng, "Hierarchical Discriminant Regression", IEEE Trans. Pattern Analysis and Machine Intelligence, v.22, no 11, s. 1-17, Kasım. 2000
- [12] A. K. Jain, J. Mao ve K. Mohiuddin, "Artificial Neural Networks: A Tutorial", Accepted to appear in IEEE Computer Special Issues on Neural Computing, Mart 1996.