



Araştırma Makalesi / Research Article

EL YAZI KARAKTERLERİNİN KAPALI CEBİRSEL EĞRİLERLE
MODELLENMESİ VE SINIFLANDIRILMASI

İhsan PENÇE*, Bayram CETİŞLİ

Süleyman Demirel Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, ISPARTA

Geliş/Received: 06.02.2013 Revised/Düzelme: 26.02.2013 Kabul/Accepted: 27.02.2013

ÖZET

Bu çalışmada el yazısı karakterlerinin kapalı cebirsel eğriler ile modellenip sınıflandırılması amaçlanmıştır. El yazısı rakamların sekizinci dereceden cebirsel denklemleri elde edilerek denklem katsayıları öznel olarak kullanılmıştır. Elde edilen katsayıların değişmez olabilmesi için sadece ölçekleme ve öteleme göre normalizasyonu yapılmıştır. Ayrıca dilsel kuvvetli ve uyarlamalı sinir bulanık sınıflayıcı (USBS) ile öznel seçimi yapılmıştır. Çalışmada Modified National Institute of Standards and Technology (MNIST) el yazısı rakam veri tabanı kullanılarak Bayes ve yapay sinir ağları ile önerilen yöntemin tanıma başarısı ölçülmüştür. Çalışmada elde edilen %92.87 değerindeki tanıma oranı henüz diğer yöntemlerle kıyaslanabilir seviyede değildir. Ancak yöntemin geliştirilerek her karakterin bir denklemle ifade edilebileceği açıktır. Böylece karakterleri görüntü biçiminde saklamak yerine katsayılarla saklayarak daha az bellek kullanımı sağlanabilir. **Anahtar Sözcükler:** El yazısı, karakter tanıma, kapalı cebirsel eğriler, modelleme, Bayes, yapay sinir ağları, uyarlamalı sinir-bulanık sınıflayıcı.

HANDWRITING CHARACTER MODELING WITH IMPLICIT CURVES AND CLASSIFICATION

ABSTRACT

In this study, the classification and modeling of handwriting characters by using the implicit curves were aimed. Also, the coefficients of eighth degree implicit equations were used for classification of handwritten digits. Therefore, a variety of curve fitting methods were tested in the study. To be invariant, the normalization of the obtained coefficients was made according to only scaling and translation. Feature selection was also done with neuro-fuzzy classifier with linguistic hedges. In the study, the recognition rate of the method proposed with the Bayes and neural networks was measured by using entire and the certain part of MNIST database of handwritten digits. Recognition rate of 92.87 % was obtained, which in the study is not yet comparable to other methods. But it is clear that each character can be expressed with an equation by developing the method. By this way, the less memory using can be satisfied by storage of coefficients instead of the storage of image.

Keywords: Handwriting, character recognition, implicit curves, modeling, Bayes, neural network, adaptive neuro-fuzzy classifier.

* Corresponding Author/Sorumlu Yazar: e-mail/e-ileti: ihsan_ce@hotmail.com, tel: (246) 211 15 63

1. GİRİŞ

Karakter tanıma konusu üzerine günümüzde pek çok çalışma olmuştur. Bu çalışmalar genel olarak otomatik karakter tanıma başlığı altında incelenmektedir. Karakter tanıma işlemi pek çok alanda kullanılmakta olup en yaygın örnekleri plaka tanıma, SPAM engelleme ve el yazısı tanıma görülmektedir. Bu çalışmada el yazısı karakterleri üzerinde durulmuştur. El yazısı karakterleri normal karakterlere göre daha zor ayırt edilmekte olup halen daha kesin çözüme ulaştırılmamış bir konu olarak kalmaktadır. Veri seti olarak MNIST [1] veri setinin seçilme sebebiyse fazla sayıda ve çeşitte el yazısı rakamlarını içeriyor olmasıdır. Bu sayede daha gerçekçi ve genel sonuçların alınması amaçlanmaktadır. Çalışma el yazısı karakterlerinin normalizasyonu, kapalı cebirsel eğriler ile modellenmesi, öznelik seçimi ve farklı sınıflayıcılar ile sınıflandırılması üzere üç aşamadan oluşmaktadır. Modelleme işlemi için literatürde birçok yöntem kullanılmıştır.

Bunlar arasında Zernike momentler [2] ve Fourier tanımlayıcılar [3-4] en çok kullanılanlarıdır. Bu çalışmada ise karakterlerin modellenmesi için kapalı cebirsel eğriler kullanılıp elde edilen katsayılar sınıflayıcı için birer öznelik olarak kullanılmıştır.

Diğer çalışmalardan farklı olarak bu çalışmada karakterlerin sadece kenarları değil tamamen kendileri 3-boyutlu olarak modellenmeye çalışılmıştır. Ayrıca günümüzde eğri ve yüzeyleri modellemede sıklıkla kullanılan Spline, Bezier ve NURBS [5] interpolasyonlarının parçalı yapısı yerine bir bütün olarak modelleme üzerinde durulmuştur. Karakterlerin ya da diğer nesnelerin parçalara ayrılmadan modellenmesi, zor bir problem olup eğri ve yüzeylerin matematiksel olarak daha iyi ifade edilmesinde önemli bir yere sahiptir. Bu çalışmada parçalı yapı kullanılmayarak karakterlerin tek bir denklemle ifade edilebilmesi amaçlanmıştır.

Literatürde MNIST karakter veritabanı için elde edilen tanıma oranları incelendiğinde sanal destek vektörlerinin 2. versiyonu (VSV2) ile %99.44 [6], gradiyent öznelikleri üzerinde destek vektör sınıflayıcısı (SVC) ile %99.58 [7], katlamalı yapay sinir ağları ve elastik bozulma ile %99.60 [8], v.b. gibi çeşitli başarıların sağlandığı görülür. Bu çalışmada ise daha çok karakterlerin doğru şekilde modellenmesi ve denklemlerinin oluşturulması üzerine çalışılmıştır. Öznelik olarak sadece elde edilen denklemin katsayıları kullanılmıştır. Literatürdeki diğer yöntemlere ait özneliklerin de çalışmaya dâhil edilmesiyle tanıma oranının artacağı barizdir.

2. NORMALİZASYON İŞLEMİ

Karakterlerin standart boyutta ve biçimde olması, tanımadaki başarı oranında önemli bir yere sahiptir. Bu sebeple veri setindeki rakamlara normalizasyon işlemi uygulanmıştır. Bunun için öncelikle karakterlerin sınırları belirlenip bu sınırlardan kesilmesi işlemi yapılmıştır. Bunun amacı karakterlerin kapladığı alanın dışında kalan siyah bölgenin modelleme işlemi yapılırken sorun oluşturmasıdır. Sınırları yeniden belirlenen karakterlerin normalizasyon işlemine tutulmalarındaki amaç, her bir karakterin ortak sınırlara sahip olmasının istenmesidir. Bu sayede büyük boyutlara sahip rakamların normal boyutlardakine göre orantısız çıkan katsayılarının önüne geçilmektedir. Ayrıca resimdeki yerleri farklı olan aynı sınıftaki rakamlar farklı katsayılar üretebilmektedir. Bu nedenle rakamlar resim içinden çıkarılarak normalizasyona tabi tutulurlar.

Normalizasyon, modellenecek veri kümesini koordinat merkezine taşır, x ve y yönündeki standart sapmaları 1 yapar [9]. Bu sayede rakam hem orijine taşınmış olur hem de düzlemde standart dağılım gösterir. Şekil 1'de orijinal rakamın sınır kümesinin belirlenmesi görülürken Şekil 2'de orijinal rakamın normalizasyon işlemi sonucundaki durumu görülmektedir.

Rakamı ifade eden veri kümesinde toplam N adet nokta bulunduğunu ve $Y_i = [x_i, y_i]^T$ ile ifade edildiğini varsayarsak eğrinin merkezini ifade eden G , Denklem (1)'deki gibi elde edilir [9].

$$G = \frac{1}{N} \sum_{i=1}^N Y_i = \begin{pmatrix} x_c \\ y_c \end{pmatrix} \quad (1)$$

Sınır kümesinin her bir noktasından ağırlık merkezinin çıkartılmasıyla, verilen sınır kümesi koordinat merkezine taşınmış olur. Eğrinin kovaryans matrisi Denklem (2)'deki gibi hesaplanır [10].

$$\Sigma = \frac{1}{N-1} \sum_{i=1}^N (Y_i - G)(Y_i - G)^T \quad (2)$$

Kovaryans matrisi Denklem (3)'teki gibi ayrılabilir.

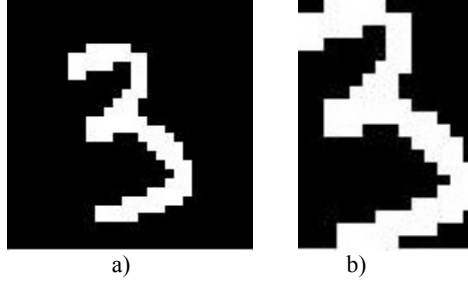
$$\Sigma = \mathbf{W} \mathbf{D} \mathbf{W}^T \quad (3)$$

Burada \mathbf{W} kovaryans matrisinin özvektörlerini, \mathbf{D} ise özdeğerlerini ifade etmektedir.

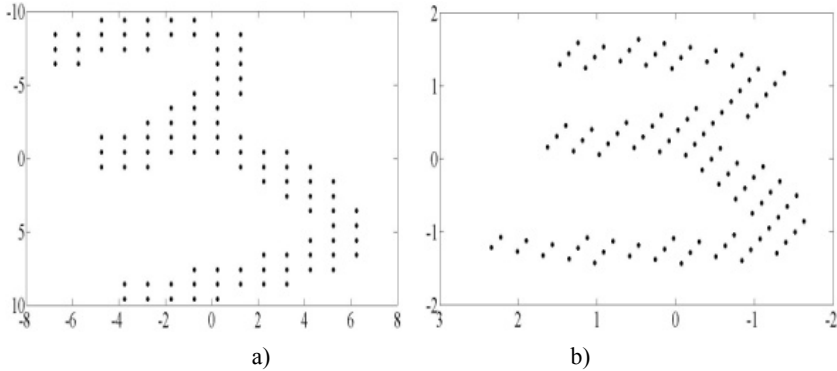
Normalizasyon sonucu oluşan yeni veri kümesi $\hat{Y}_i = [\hat{x}_i, \hat{y}_i]^T$ ile ifade edilirse \hat{Y}_i Denklem (4)'teki gibi hesaplanır.

$$\hat{Y}_i = (\hat{x}_i \hat{y}_i)^T = \mathbf{D}^{-\frac{1}{2}} \mathbf{W}^T (Y_i - G) \quad (4)$$

Rakama ait yeni veri kümesi artık orijindedir.



Şekil 1. a) Orijinal rakam b) Sınırları belirlenmiş rakam



Şekil 2. a) Orijinal rakam b) Normalize edilmiş rakam

3. MODELLEME İŞLEMİ

Modelleme işlemi bir karakterin uygun matematiksel denklemlerle ifade edilmesidir. Bu çalışmada kapalı cebirsel eğriler kullanılarak rakamlar polinom şeklinde ifade edilip katsayıları en

küçük kareler yöntemi ile bulunmuştur. En küçük kareler yöntemi katsayıların bulunmasında en iyi yöntemlerden birisi olmakla beraber oldukça hızlıdır.

Karakterin kapalı cebirsel eğrilerle gösterim şekli olan $f(x,y)$ Denklem (5)'teki gibi hesaplanır [11].

$$f(x,y) = \sum_{0 \leq i,j, 0 \leq i+j \leq d} a_{ij} x^i y^j \quad (5)$$

Burada d , kaçınıcı dereceden kapalı cebirsel eğri ile işlem yapılacağını belirtmektedir. Dördüncü dereceden kapalı cebirsel eğri şu şekilde ifade edilir.

$$f(x,y) = a_{00} + a_{10}x + a_{01}y + a_{20}x^2 + a_{11}xy + a_{02}y^2 + a_{30}x^3 + a_{21}x^2y + a_{12}xy^2 + a_{03}y^3 + a_{40}x^4 + a_{31}x^3y + a_{22}x^2y^2 + a_{13}xy^3 + a_{04}y^4$$

Burada x ve y karakterin koordinat noktalarını gösterip Denklem (6)'da \mathbf{X} ile ifade edilirken, \mathbf{a} katsayılar vektörünü göstermekte olup Denklem (7) ile ifade edilir.

$$\mathbf{X} = \begin{bmatrix} 1 & x & y & x^2 & xy & y^2 & x^3 & x^2y & xy^2 & \dots \\ y^3 & x^4 & x^3y & x^2y^2 & xy^3 & y^4 \end{bmatrix} \quad (6)$$

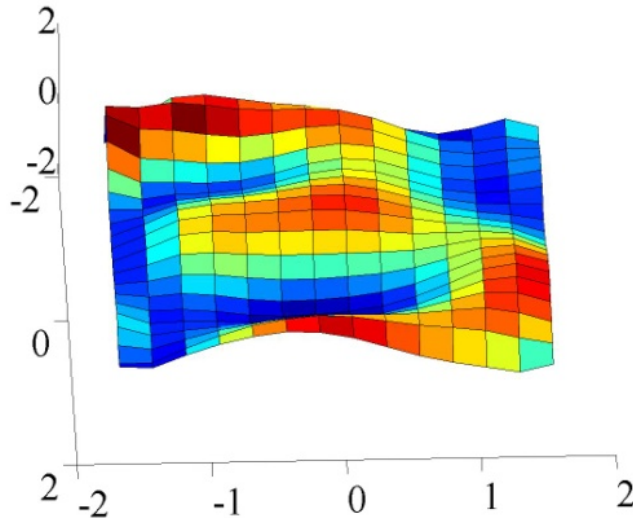
$$\mathbf{a} = \begin{bmatrix} a_{00} & a_{10} & a_{01} & a_{20} & a_{11} & a_{02} & a_{30} & a_{21} & \dots \\ a_{12} & a_{03} & a_{40} & a_{31} & a_{22} & a_{13} & a_{04} \end{bmatrix} \quad (7)$$

Katsayılar vektörü olan \mathbf{a} en küçük kareler yöntemi kullanılarak Denklem (8) ile hesaplanır [12]. En küçük kareler yöntemi karakterin veri kümesinin uydurulacak cebirsel eğriye uzaklığının karesini minimum yapmaya çalışır.

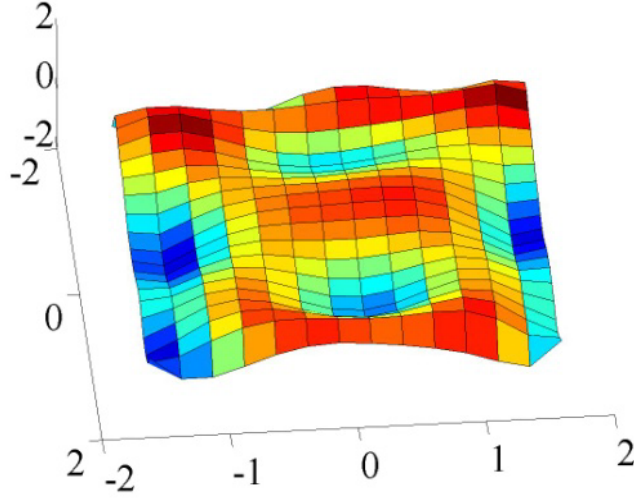
$$\mathbf{a} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y} \quad (8)$$

Burada \mathbf{Y} karakterin veri kümesini ifade etmektedir.

Bu çalışmada kullanılan bütün rakamlar 8. dereceden kapalı cebirsel eğrilerle modellenmiştir. Yapılan çalışmalarda 8. derecenin rakamları modellemek için yeterli olduğu görülmüştür. Üç rakamının kapalı cebirsel eğrilerle oluşturulmuş gösterimi Şekil 3'te görülürken sekiz rakamının gösterimi ise Şekil 4'te görülmektedir.



Şekil 3. Üç rakamının kapalı cebirsel yüzeyi



Şekil 4. Sekiz rakamının kapalı cebirsel yüzeyi

4. ÖZNETELİK SEÇİMİ VE SINIFLANDIRMA

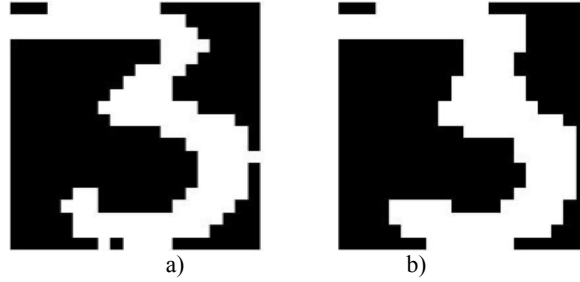
Sınıflandırma günümüzde pek çok uygulamada kullanılmakta olup karakterlerin doğru tespit edilmesinde de önemli bir yere sahiptir. Bu çalışmada Bayes, yapay sinir ağları (YSA) ve USBS kullanılarak sınıflandırma işlemi yapılmıştır [13-16]. USBS sinir ağlarının öğrenme özelliği ile bulanık mantığın insan düşüncesine en yakın eğer-ise çıkarsama mantığını bir araya getiren bir yöntemdir [15]. Bu sınıflayıcının en büyük özelliği bulanık kümelerle uygulanan ve matematiksel olarak üstel kuvvetlere karşılık gelen dilsel kuvvetlerdir. Dilsel kuvvetler, önemli özniteliklerin vurgulanmasını, önemsiz özniteliklerin ise sistemden zayıflatılarak uzaklaştırılmasını sağlamaktadırlar [16].

Çalışmada kullanılan bütün sınıflayıcılar için öznitelik olarak daha önce hesaplanan a katsayılar vektörü kullanılmıştır.

5. DENEYSSEL ÇALIŞMA

Bu çalışmada MNIST [1] veri setinde bulunan 28x28 boyutlarında 60000 adetinin eğitim setini ve 10000 adetinin test setini oluşturduğu el yazısı rakamlar kullanılmıştır. Rakamların hepsi 8. dereceden kapalı cebirsel eğrilerle modellenmiş olup öznitelik olarak oluşturulan polinomun en küçük kareler yöntemi ile bulunan katsayıları kullanılmıştır. Derece olarak 8. derecenin seçilmesindeki amaç bu derecenin altındaki derecelerde el yazısı rakamların anlamlı şekilde modellenememesidir. Ayrıca 8. dereceden daha yüksek derecelerde de modellenmenin daha iyi olduğu fakat tanıma oranının düştüğü görülmüştür. Bu duruma yüksek dereceli polinomların katsayı adetlerinin çok fazla olması sebep olmuştur. İlk bakışta katsayı adetinin fazla olması sınıflayıcı için daha fazla öznitelik anlamına geleceğinden faydalıymış gibi gözükabilir. Fakat 8. derecedeki seçici katsayı büyüklüklerinin ortaya çıkan yeni katsayılarla dağıtılması sonucu sınıflandırmada tanıma oranının düştüğü gözlenmiştir. Bu sebeple 8. derecenin rakamları 3-boyutlu modellemede uygun olduğuna karar verilmiştir.

Modelleme sonucu elde edilen yüzey üzerinde belli bir değerın üzeri alınarak rakamı tekrar elde etmekte mümkündür. Bu durum modellenmenin doğru bir şekilde yapıldığını da gösterir. Örnek olarak üç rakamına ait yüzey üzerinde karakterin çukur veya tümsek bölgelerini belirten z koordinatının $z > 0.5$ için belirttiği bölge Şekil 5'te görülmektedir.



Şekil 5. a) Orijinal rakam b) Modelleme sonucunda geri elde edilen rakam

Bu çalışmada sınıflayıcı olarak YSA ve Bayes kullanılmıştır. Ayrıca USBS'nin seçtiği önemli özniteliklerden faydalanmak için USBS, kullanılan iki sınıflandırma yöntemine dahil edilmiştir. Bu sayede USBS'nin önemli öznitelikleri bulma gücü ile Bayes ve YSA'nın sınıflandırmadaki gücü birlikte kullanılmıştır. Nitekim USBS normalde 45 adet olan öznitelik sayısını 23 adete indirmeyi başarmış olmakla beraber tanıma oranında da iyi sonuçlar vermiştir. Öznitelik sayısının azaltılmasındaki amaç sınıflayıcının eğitim süresini azaltmaktır. Nitekim 45 öznitelik ile YSA kullanılarak eğitim süresi yaklaşık 8 saat olarak ölçülürken, 23 öznitelik ile YSA eğitim süresi yaklaşık 6 saat olarak ölçülmüştür. Farklı sınıflayıcılarda aradaki süre farkının dahada artacağı öngörülebilir.

MNIST veri setinde 60000 eğitim ve 10000 test verisi kullanılarak elde edilen tanıma oranları Çizelge 1 ve Çizelge 2' de görülmektedir.

Çizelge 1. MNIST tanıma oranları

| Sınıflayıcı | Öznitelik S. | Tanıma O.(%) Eğitim S. | Tanıma O.(%) Test S. |
|-------------|--------------|------------------------|----------------------|
| Bayes | 45 | 81.77 | 82.03 |
| YSA | 45 | 95.80 | 92.87 |

Çizelge 2. USBS ile MNIST tanıma oranları

| Sınıflayıcı | Öznitelik S. | Tanıma O.(%) Eğitim S. | Tanıma O.(%) Test S. |
|-------------|--------------|------------------------|----------------------|
| USBS-Bayes | 23 | 84.89 | 85.80 |
| USBS-YSA | 23 | 92.15 | 89.98 |

Çizelgeler incelendiğinde USBS'nin Bayes sınıflayıcısına daha fazla katkıda bulunduğu görülmektedir. Eğitim ve test seti üzerinde en yüksek tanıma oranı ise YSA ile elde edilmiştir.

6. SONUÇLAR VE TARTIŞMA

Bu çalışmada el yazısı rakamların sadece sınır bölgelerini belirten kenarları yerine bir bütün olarak 3-boyutlu şekilde kapalı cebirsel eğriler ile modellenip uygun sınıflayıcı yöntem ile sınıflandırılması üzerinde durulmuştur. Bütün rakamlar 8. dereceden kapalı cebirsel eğrilerle modellenmiş olup öznitelik sayısında çıkan fazlalık USBS'nin yetenekleriyle azaltılmaya çalışılmıştır. Ayrıca oluşturulan 3-boyutlu yüzeylerden orijinal 2-boyutlu rakamların geri elde edilebileceği de görülmüştür. MNIST veritabanı üzerinde alınan %92.87'lik tanıma oranı ve öznitelik sayısı azaltıldığında elde edilen %89.98'lik tanıma oranları ilerisi için umut vaat etmekte

olup yöntemin geliştirilmesiyle birlikte her rakamın birer denklemle ifade edilebileceği açıktır. Böylece karakterleri görüntü biçiminde saklamak yerine katsayılarla saklayarak daha az bellek kullanımı sağlanabilir.

Gelecekte, veri transferinde görüntü yerine transfer edilen katsayılarla veri iletim hızı artırılabilir. Yeni ve daha verimli bir bilgi ya da görüntü sıkıştırma yöntemi, cebirsel eğri yöntemleri ile önerilebilir.

KAYNAKLAR / REFERENCES

- [1] LeCun, Y., Bottou, L., Bengio, Y., et.al., “Gradient-Based Learning Applied to Document Recognition”, Proc. of the IEEE, 86, 11, 2278-2324, November 1998.
- [2] Lisboa, P.J.G. and Perantonis, S.J., “Invariant Digit Recognition by Zernike Moments and Third-Order Neural Networks”, Second Int. Conf. Artificial Neural Networks, 18-20, 1991.
- [3] Jeong, C. and Jeong, D., “Handwritten Digit Recognition Using Fourier Descriptors and Contour Information”, IEEE TENCON, 6, 99, 1999.
- [4] Chung, Y. and Wong, M., “Handwritten Character Recognition by Fourier Descriptors and Neural Network”, IEEE TENCON, Speech and Image Technologies for Computing and Telecommunications, 1997.
- [5] Çınar, A. and Arslan, A., “Bulanık Mantık Tabanlı Yüzey Modelleme ve Üç Boyutta Nesne Kaynaştırma İşlemine Uygulanması”, Gazi Üniversitesi Müh. Mim. Fak. Derg., 17, 4, 2002.
- [6] Decoste, D. and Scholkopf, B., “Training Invariant Support Vector Machines”, Mach. Learn., 46, 1–3, 160–190, 2002.
- [7] Liu, C.L., Nakashima, K., Sako, H., et.al., “Handwritten Digit Recognition Using State-of-The-Art Techniques”, Proceedings of the 8th International Workshop on Frontiers in Handwritten Recognition, 320–325, Ontario, Canada, 2002.
- [8] Simard, P.Y., Steinkraus, D. and Platt, J.C., “Best Practice for Convolutional Neural Networks Applied to Visual Document Analysis”, Proceedings of the 7th Int. Conf. on Document Analysis and Recognition (ICDAR2003), 2, 958–962, Edinburgh, Scotland, 2003.
- [9] Candemir, S., “Kapalı Cebirsel Eğrilerle Karakter Tanıma”, Yüksek Lisans Tezi, Gebze Yüksek Teknoloji Enstitüsü, Mühendislik ve Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı, 2004.
- [10] Unel, M., “Polynomial Decompositions for Shape Modelling”, Object Recognition and Alignment, PhD Thesis, Brown University, 1999.
- [11] Blane, M.M., Lei, Z., Çivi, H., et.al., “The 3L Algorithm for fitting Implicit Polynomial Curves and Surfaces to Data”, IEEE Trans. Pattern Analysis and Machine Intelligence, 22, 3, March 2000.
- [12] Cetişli, B. and Edizkan, R., “Estimation of Adaptive Neuro-Fuzzy Inference System Parameters with the Expectation Maximization Algorithm and Extended Kalman Smoother”, Neural Computing and Applications, 20, 3, 403-415, 2011.
- [13] Trier, O.D., Jain, A.K. and Taxt, T., “Feature Extraction Methods for Character Recognition: A survey”, Revised July 19, 1995.
- [14] Jain, A.K., Mao, J. and Mohiuddin, K., “Artificial Neural Networks: A Tutorial”, Accepted to appear in IEEE Computer Special Issues on Neural Computing, March 1996.
- [15] Cetişli, B., “El Yazısı Karakter Tanıma: Dalgacık Moment Özniteliklerinin Yenilenen ANFIS ile Sınıflandırılması”, Doktora Tezi, Eskişehir Osmangazi Üniversitesi, Fen Bilimleri Enstitüsü, 2005.
- [16] Cetişli, B., “Öznitelik Seçiminde Dilsel Kuvvetli Sinir Bulanık Sınıflayıcı Kullanımı”, Eskişehir Osmangazi Üniversitesi Müh-Mim. Fak. Dergisi, 19, 2, 109-129, 2006.