

## Leaf Recognition and Classification Based On Geometric and Statistical Features

**Muammer Türkoğlu**

*Bingöl Üniversitesi, Bilgisayar Mühendisliği, Bingöl*

**Ömer Faruk Alçın**

*Bingöl Üniversitesi, Elektrik-Elektronik Mühendisliği, Bingöl*

**Mehmet Üstündağ**

*Bingöl Üniversitesi, Elektrik-Elektronik Mühendisliği, Bingöl*

**Davut Hanbay**

*İnönü Üniversitesi, Bilgisayar Mühendisliği, Malatya*

### ABSTRACT

In this work, Centroid-Contour Distances (CCD), a method which can rapidly and easily recognize patterns and also simultaneously increase system performance, is used to develop a leaf recognition system. By using different angles, we obtained different numbers of features. Adding morphological and statistical features further to these calculated features, the system is made robust. Obtained features were classified by using Extreme Learning Machines (ELM). To increase reliability of the experimental studies, 10-fold cross-validation method was used. As a result, we obtained 94.33% calculation accuracy at 20° angle.

**Keywords:** Leaf Recognition, Morphological Features, CCD, ELM

## Geometrik ve İstatistiksel Özelliklere Dayalı Yaprak Sınıflandırma Sistemi

### ÖZET

Bu çalışmada, örüntü tanıma sistemlerinde hızlı ve kolay hesap yapabilen ve aynı zamanda sistem performansını arttıran, Merkez Kenar Uzaklıkları (MKU) metodu kullanılarak yaprak tanıma sistemi geliştirilmiştir. Yapılan çalışmada farklı açılar kullanılarak farklı sayıda öznitelikler elde edilmiştir. Hesaplanan özniteliklere morfolojik özellikler ve istatistiksel özellikler de eklenerek sistem dayanımlı hale getirilmiştir. Elde edilen öznitelikler Aşırı Öğrenme Makineleri (AÖM) kullanılarak sınıflandırılmıştır. Sistemin performansını belirlemek için 10-kat çapraz geçerlilik testi kullanılmıştır. Yapılan çalışmalarda en yüksek 20° için %94.33 sınıflandırma doğruluğu elde edilmiştir.

**Anahtar Kelimeler:** Yaprak Tanıma, Morfolojik Öznitelikler, MKU, AÖM.

## 1 GİRİŞ

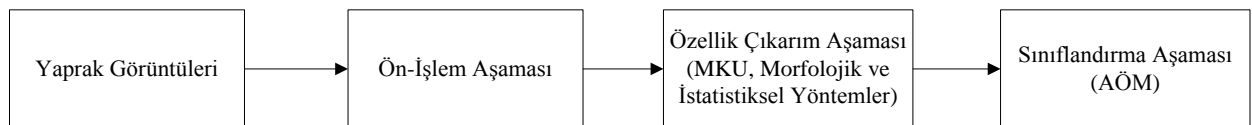
Bir bitkinin hangi türe ait olduğunun belirlenmesi zordur. Ayrıca çevre koşullarının değişmesi ile birlikte bitki türlerinin nesillerinin tükenmeye başlaması bu işlemi daha da zorlaştırmaktadır. Bundan dolayı bitkiler üzerine yapılan araştırmalar bitki türlerinin korunmasında önemli olmaktadır [1, 2]. Çoğu zaman sadece uzmanların yapabileceği geleneksel bitki türlerini belirleme işlemlerinin zaman alıcı olduğu bilinmektedir. Örüntü tanıma yöntemleri ile bitkileri sınıflandırmak daha kolay ve hızlı yapılabilmektedir. Bu yaklaşım sayesinde yapraklar arasında en uygun ayırt edici öznelikler elde edilerek bitki türlerini sınıflandırma işlemleri başarılı bir şekilde gerçekleştirilebilir [1].

Günümüzde bitki yapraklarından özneliklerin hesaplanabilmesi için çok sayıda yöntem kullanılmaktadır. [3]'de 32 farklı bitki türüne ait 1800 görüntüden 5 temel geometrik özneliklere dayalı 12 öznelik parametresi Olasılıksal Sinir Ağları (OSA) kullanılarak %90 doğruluk ile sınıflandırılmıştır. Du ve diğ. geometrik özelliklere dayalı 8 temel morfolojik özneliklere ek olarak değişmez moment uygulanarak toplam da 15 öznelik çıkarmıştır. Her sınıf için 20 örnek alınarak 20 farklı bitki türü incelenmiştir. Sınıflandırıcı olarak Move Median Centers (MMC) yöntemi kullanılarak başarı oranları ve sınıflandırma zamanları diğer çalışmalar ile kıyaslanmıştır [4]. Iwata ve diğ. şekil öznelikleri, değişmez momentler ve renk öznelikleri kullanarak toplam da 23 öznelik çıkarmıştır. Her sınıf için 20 örnek alınarak 96 farklı bitki türü incelenmiştir. Sınıflandırma aşamasında Rastgele Orman (RO) yöntemi kullanılmışlardır [5]. Ayrıca çok sayıda MKU tabanlı çalışmalar yapılmıştır. [6]'de özellik çıkarım aşamasında merkez uzaklık ve en az eylemsizlik eksen metotları kullanılmıştır. Literatür de yaygın olarak kullanılan flavia [7] ve swedish [8] yaprak verileri kullanılmıştır. Sınıflandırıcı olarak OSA yöntemini kullanmışlardır. Sistemlerinin performansını Flavia verisi için %82.05 Swedish verisi için %80.1 olarak hesaplanmıştır. Fern ve diğ. MKU yöntemi ile önerilen Merkez Kenar Eğimi (MKE) yöntemi uygulanarak kıyaslanmıştır. Her bir sınıf için 50 resim olmak üzere 5 sınıf incelenmiştir. Sınıflandırıcı olarak İleri beslemeli geri yayılım ağı kullanılmışlardır. Ortalama başarı oranları MKU yönteminde %74.4, MKE yönteminde ise %96.6 dır [9]. Ayrıca MKU yöntemi kullanılarak yaprak türlerini tanıma ve sınıflandırılmasının gerçekleştirildiği başka çalışmalar da mevcuttur [10, 11, 12].

Bu çalışmada, literatür de en fazla kullanılan geometrik tabanlı metotlar denenmiştir. Bu metotlar MKU, geometrik öznelikler ve istatistiksel özneliklerden oluşmaktadır. Veri olarak bir bitkiye ait 9 adet yaprak türünden 600 adet yaprak resmi kullanılmıştır [13]. Farklı açılar ile performansları tespit edildiğinden dolayı her bir resimden farklı sayıda öznelik parametreleri bulunmaktadır. Bu elde edilen öznelikleri sınıflandırmak için AÖM sınıflandırıcısı kullanılmıştır ve en yüksek doğruluk yüzde %94.333 olarak tespit edilmiştir. Önerilen sistemin 2. bölümünde çalışmanın görüntü işleme, özellik çıkarımı ve sınıflandırma aşamalarından ve çalışmada kullanılan veritabanından bahsedilmiştir. 3. bölümde ise benzetim çalışmaları anlatılmıştır. Son olarak da sonuçlar 4. bölümde yer almıştır.

## 2 MATERYAL VE YÖNTEM

Önerilen yöntemin işlem adımları genel haliyle Şekil 1' de gösterilmiştir.



Şekil 1. Yapılan çalışmanın genel akış şeması.

Şekil 1'de verilen işlem adımları devam eden alt başlıklarında detaylandırılmıştır.

## 2.1 Veriler

Bu çalışmada kullanılan Acer (Japon akçaağacı) bitkisine ait veriler Middle European Woods veri tabanı içerisinde alınmıştır [13]. Bu bitki türüne ait 9 cins bulunmaktadır. Ayrıca her cins için 54 ile 97 arası yaprak örneğine ait 600 yaprak görüntüsü içerir. Örnek resimler değişik çözünürlükte ve 3 kanalıdır.

## 2.2 Ön-İşlem Aşaması

Bu aşamada, Şekil 2'de görüldüğü gibi yaprak görüntüsünün arka plandan ayırma işlemleri yapılır. Ayırma işlemi aşağıdaki adımlar aracılığıyla elde edilmiştir;

- RGB formatındaki renkli yaprak resimleri gri tonlu resme dönüştürülmüştür.
- Uygun eşikleme değeri seçilerek yaprak resmini arka plandan ayırma işlemi yapılmıştır.
- Muhtemel piksel taşmalarını gidermek için aşındırma ve dolun işlemleri uygulanmıştır.
- En son aşama olarak yaprağın sınır uçları belirlenerek kırpma işlemi yapılmıştır.



a)

b)

Şekil 2. Örnek bir yaprağa ait a) Gerçek resim, b) Bölütlenmiş resim

## 2.3 Özellik Çıkarım Aşaması

Yapılan çalışmada, örüntü tanıma işlemlerinde yaygın olarak kullanılan morfolojik özellikler, istatistiksel özellikler ve MKU özellik çıkarma metotları kullanılmıştır. Bunlar;

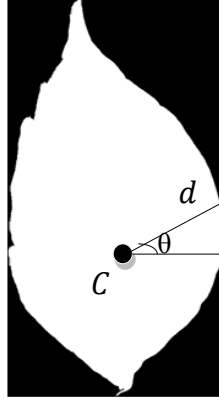
### 2.3.1 Merkez Kenar Uzaklıkları (MKU)

Merkez kenar uzunluğu, şeklin ağırlık merkezinin sınır eğrilerine olan uzaklığını ifade eder (Şekil 3). Şeklin sınır eğrileri

$$(x(i), y(i)), \quad i = 0, 1, 2, \dots, M - 1, \quad (1)$$

gibi tanımlansın. Denklem (1)'de  $M$  sınır eğri sayısını,  $(x, y)$  ise sınır eğrilerini göstermektedir. Bu nesnenin ağırlık merkezi  $C(g(x), g(y))$  olarak tanımlansın. Bir sınır eğrisinin ağırlık merkezine olan Öklid uzaklık fonksiyonu Denklem (2)'deki gibi yazılır.

$$d = \sqrt{(x(i) - g(x))^2 + (y(i) - g(y))^2} \quad (2)$$



Şekil 3. Örnek bir yaprak için MKU Yaklaşımı

Tablo 1’de gösterildiği gibi, ağırlık merkezindeki x eksenini sıfır derece kabul edersek her uzaklık değerinden sonra  $\theta$  derece artırımla toplamda  $N$  ( $N = 360/\theta$ ) adet  $d$  değeri hesaplanmış olur. Elde edilen  $d$  değerleri görüntünün öznitelikleri olarak kabul edilir.

Tablo 1. MKU yönteminin açıya göre öznitelik sayısı

Açı (°)	5	8	9	10	12	15	18	20	24	30	40	45
Öznitelik sayısı	72	45	40	36	30	24	20	18	15	12	9	8

### 2.3.2 Morfolojik Özellikler

Morfolojik özellikler, yaprağın geometrik şekline bağlı öznitelikler elde edilmesini sağlar. Kullanılan morfolojik özellikler Tablo 2’de özetlenmiştir.

Tablo 2. Yapılan Çalışmada Kullanılan Morfolojik Özelliklerin Tanımı

<b>Alan</b>	Bölgedeki yaprak görüntüsünün piksel sayısını
<b>Çevre</b>	Yaprağın sınır eğrilerinin uzunluklarının toplamı
<b>Çap</b>	Yaprağın sınır eğrilerinin birbirine en uzak iki nokta arasındaki uzaklık
<b>Boyu</b>	Yaprağın düşey olarak birbirine en uzak iki noktası
<b>Eni</b>	Yaprağın yatay olarak birbirine en uzak iki noktası
<b>Ağırlık Merkezi</b>	Yaprak görüntüsünün ağırlık merkezinin, yaprağı çevreleyen sınır eğrilerine
<b>Uzaklıkları Toplamı</b>	uzaklıkları toplamı

### 2.3.3 İstatistiksel Özellikler

Ön-işlem aşamasından sonra elde edilen iki boyutlu imgeler  $M \times N$  boyutlu  $A_{ij}$  matrisi olsun. Denklem (3)’de  $\mu_{ij}$  ile gösterilen ortalamayı, Denklem (4)’de ise  $V_{ij}$  ile gösterilen varyansı ifade eder.

$$\mu_{ij} = \frac{1}{MN} \sum_i^N \sum_j^M A_{ij} \quad (3)$$

$$V_{ij} = \frac{1}{MN} \sum_i^N \sum_j^M (A_{ij} - \mu_{ij})^2 \quad (4)$$

### 2.3.4 Sınıflandırma

AÖM Huang ve arkadaşları tarafından TGKİB sinir ağı için önerilen bir öğrenme algoritmasıdır [14-16]. Geleneksel ağların aksine AÖM’de giriş ağırlıkları ve gizli katman eşikleri gibi parametreler rastgele başlatılır ve sabit kalır, çıkış ağırlıkları ise en küçük kareler yöntemi ile hesaplanır. AÖM hızlı öğrenme süreci ve daha iyi bir genelleme başarımı gibi üstünlüklere sahiptir. Sıfır hataya yakınsadığı kabul edilen  $N$  nöronlu,  $g(x)$  aktivasyon fonksiyonuna sahip bir TGKİB ağıın matematiksel modeli;

$$\sum_{i=1}^M \beta_i g(w_i \cdot x_j + b_i) = y_j, \quad j = 1, \dots, N \quad (5)$$

şeklinde yazılabilir. Denklem (5)’de  $[x_i, y_i]$  giriş-çıkış ikilisini,  $w_i$  giriş ağırlıklarını,  $b_i$  gizli katman eşikliğini ve  $\beta_i$  çıkış ağırlıklarını ifade etmektedir. Denklem (5) matris formunda Denklem (6)’de belirtildiği gibi yazılabilir.

$$Y = H\beta$$

$$Y = \begin{bmatrix} y_1^T \\ \vdots \\ y_N^T \end{bmatrix}_{N \times 1}, \quad H = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \dots & g(w_N \cdot x_1 + b_N) \\ \vdots & \dots & \vdots \\ g(w_1 \cdot x_M + b_1) & \dots & g(w_N \cdot x_M + b_N) \end{bmatrix}_{N \times M}, \quad \beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_M^T \end{bmatrix}_{M \times 1} \quad (6)$$

Burada  $Y$  ağıın çıkışını,  $H$  gizli katman çıkış matrisini ve  $\beta$  ise çıkış ağırlıklarını temsil etmektedir. Denklem (6)’in en küçük kareler çözümü ile  $\beta$  çıkış ağırlıkları hesaplanır;

$$\beta = H^+ Y \quad (7)$$

Burada  $H^+$  gizli katman çıkış matrisi  $H$ ’ın Moore Penrose genelleştirilmiş tersini ifade etmektedir.

## 3 BENZETİM SONUÇLARI

Bu çalışmada, Acer türüne ait 9 cins bitkiye ait 600 adet renkli yaprak resmi sınıflandırılmıştır. Ön-İşlem aşamasında görüntüyü arka plandan ayırma ve belirginleştirme işlemleri yapılmıştır. AÖM sınıflandırıcısının güvenilirliğini arttırmak için 10-kat çapraz geçerlilik yöntemi kullanılmıştır. Yapılan çalışmada ilk adımda; morfolojik ve istatistiksel yöntemler ile her bir resim için toplam da 8 öznitelik elde edilmiştir. Bu öznitelikler ile %65.33 doğruluk performansı hesaplanmıştır. Daha sonra MKU yöntemi ile 5°, 8°, 9°, 10°, 12°, 15°, 18°, 20°, 24°, 30°, 40° ve 45° derecelik açılar için Tablo 1’de belirtildiği gibi farklı sayıda öznitelikler çıkarılmıştır. Bu öznitelikler kullanılarak yapılan sınıflandırma sonuçları Tablo 3’te verilmiştir.

**Tablo 3.** MKU yöntemine göre sınıflandırma sonuçları

Açı (°)	5	8	9	10	12	15	18	20	24	30	40	45
Öznitelik sayısı	72	45	40	36	30	24	20	18	15	12	9	8
Doğruluk (%)	86.83	88.5	88.5	88.5	87.33	87.66	86.83	87.66	83.16	80.16	74.83	60.33
k (%)	3	7	0	9	3	7	3	7	7	7	3	3

Tablo 3’te en düşük başarıım 45° için elde edilirken en yüksek başarıım ise 10°’de elde edilmiştir. Ayrıca 20°’den itibaren doğrulukta sürekli düşüş gözlenmiştir. Daha sonra özellik çıkarım aşamasında MKU, morfolojik ve istatistiksel yöntemler bir arada kullanılarak elde edilen öznitelikler ile devam edilmiştir. Bu öznitelikler kullanılarak elde edilen sınıflandırma sonuçları Tablo 4’de verilmiştir.

**Tablo 4.** MKU, morfolojik ve istatistiksel yöntemlere göre sınıflandırma sonuçları

Açı (°)	5	8	9	10	12	15	18	20	24	30	40	45
Öznitelik sayısı	80	53	48	44	38	32	28	26	23	20	17	16
Doğruluk (%)	92.66	93.16	93.33	93.16	92.16	93.16	94.33	92.66	91.5	89.16		
	7	7	3	4	7	7	7	3	7	2	0	7

Tablo 4'te görüldüğü en iyi performans 20° MKU, morfolojik ve istatistiksel yöntemler için elde edilmiştir. Sadece MKU yöntemine göre (Tablo 3) en iyi performans oranı 10°'de elde edilirken MKU, morfolojik ve istatistiksel yöntemleri ile en yüksek performans 20°'de elde edilmiştir. Ayrıca 20°'den itibaren yine sınıflandırma performansında sürekli azalma söz konusudur. MKU, morfolojik ve istatistiksel yöntemler bir arada kullanılması ile sınıflandırıcı performansında yaklaşık %6 artış sağlanmıştır.

#### 4 SONUÇ

Bu çalışmada örüntü tanımda yaygın olarak kullanılan şekilsel ve istatistiksel yöntemlere dayalı bir yaprak türü belirleme sistemi geliştirilmiştir. İlk önce morfolojik ve istatistiksel özelliklere dayalı sınıflandırma işlemi yapılmıştır. Daha sonra MKU farklı açılar ve buna bağlı olarak farklı sayılarda öznitelikler için performans analizi gerçekleştirilmiştir. Önerilen sistemin performansını artırmak için bu iki yaklaşım bir arada kullanılarak farklı açılarda performans karşılaştırılması yapılmıştır. Yapılan çalışmaların sonuçlarına göre en yüksek başarımlar 20° MKU, morfolojik ve istatistiksel yöntemlere dayalı öznitelikler için elde edilmiştir.

#### REFERANSLAR

- [1] Chavoshi, S. H., and Milani, M. M. R. A., "Bulanık Mantığı Kullanarak Bitki Tanıma Sistemi", Akademik Bilişim 2013, 23-25 Ocak 2013, Akdeniz Üniversitesi ANTALYA.
- [2] Kır, B., Öz, C., Gülbağ, A., "Leaf recognition using K-NN classification algorithm", Signal Processing and Communications Applications Conference (SIU), 2012 20th, , 1-4, (2012).
- [3] Stephen Gang Wu, Forrest Sheng Bao, Eric You Xu, Yu-Xuan Wang, Yi-Fan Chang, Qiao-Liang Xiang, "A Leaf Recognition Algorithm for Plant Classification Using Probabilistic Neural Network", International Symposium on Signal Processing and Information Technology, IEEE, pp 11-16, 2007.
- [4] Du, J. X., Wang, X. F., and Zhang, G. J., "Leaf shape based plant species recognition", Applied Mathematics and Computation 185, pp 883-893, 2007.
- [5] Iwata, T., and Saitoh, T., "Tree recognition based on leaf images." SICE Annual Conference (SICE), 2013 Proceedings of. IEEE, 2013.
- [6] Mahdikhanlou, K., and Ebrahimnezhad, H., "Plant leaf classification using centroid distance and axis of least inertia method." Electrical Engineering (ICEE), 2014 22nd Iranian Conference on. IEEE, 2014.
- [7] Laga, S., Kurtek, A., Srivastava, M., Golzarian and S. Miklavcic, "A riemannian elastic metric for shape-based plant leaf classification," In Digital Image Computing: Techniques and Applications, 2012.

- [8] O.J.O. Söderkvist, "Computer vision classification of leaves from Swedish trees," Master's thesis, Linköping University, SE-581 83 Linköping, Sweden, September 2001.
- [9] Fern, B. M., Sulong, G. B., and Rahim, M. S. M., "Leaf Recognition Based on Leaf Tip and Leaf Base Using Centroid Contour Gradient" *Advanced Science Letters* 20.1 (2014): 209-212.
- [10] Chaki, Jyotismita, and Ranjan Parekh. "Plant leaf recognition using shape based features and neural network classifiers." *International Journal of Advanced Computer Science and Applications (IJACSA)* 2.10 (2011).
- [11] Wang, Zhiyong, Zheru Chi, and Dagan Feng. "Shape based leaf image retrieval." *Vision, Image and Signal Processing, IEE Proceedings-*. Vol. 150. No. 1. IET, 2003.
- [12] Hong, Anxiung, et al. "Region-of-Interest based flower images retrieval." *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP'03). 2003 IEEE International Conference on*. Vol. 3. IEEE, 2003.
- [13] Novotný, P., Suk, T., "Leaf Recognition Of Woody Species In Central Europe", *Biosystems Engineering* 115(4) 444-452, (2013).
- [14] Huang, G. B., Zhu, Q. Y., ve Siew, C. K. (2006). "Extreme learning machine: theory and applications", *Neurocomputing*, 70(1), 489-501.
- [15] Alçın, Ö. F., Şengür, A. ve İnce, M. C. (2015), İleri-geri takip algoritması tabanlı seyrek aşırı öğrenme makinesi. *Journal of the Faculty of Engineering & Architecture of Gazi University*, 30(1), 126-132.
- [16] Huang, G. B., Zhu, Q. Y., ve Siew, C. K. (2004), Extreme learning machine: a new learning scheme of feedforward neural networks, *IEEE International Joint Conference on Neural Networks*, Budapest, 25-29 July, 2, 985-990.