

# Destek Vektör Makinaları Kullanılarak Yaprak Tanıma ile Üzüm Çeşitlerinin Sınıflandırılması

## Classification of The Grape Varieties based on Leaf Recognition by Using SVM Classifier

Muammer Türkoğlu  
Bilgisayar Mühendisliği  
Bingöl Üniversitesi  
Bingöl/TÜRKİYE  
mturkoglu@bingol.edu.tr

Davut Hanbay  
Bilgisayar Mühendisliği  
İnönü Üniversitesi  
Malatya/TÜRKİYE  
dhanbay@inonu.edu.tr

### Özetçe

Bu çalışmada, üzüm ağacı türlerini sınıflandırmak için yaprak görüntülerinden elde edilen özellikler çoklu-sınıf destek vektör makinaları kullanılarak sınıflandırılmıştır. Özellik çıkarma aşamasında üzüm yapraklarından 9 farklı özellik çıkarılmıştır. Görüntü işleme aşaması gri tonlu arama, medyan filtreleme, kontrast, eşikleme ve morfolojik-mantıksal işlemleri kapsamaktadır. Sınıflandırma aşamasında, çoklu-sınıf destek vektör makinaları(ÇSDVM) yardımı ile elde edilen özellikler kullanılarak sınıflandırma işlemi gerçekleştirilmiştir. Test aşamasında, farklı yaprak görüntüleri uygulanarak modelin başarımı hesaplanmıştır. Bu çalışmada, MATLAB yazılımı kullanılmıştır. Test sonucunda toplam başarı oranı %90.7 olarak hesaplanmıştır.

**Anahtar kelimeler:** Destek Vektör Makinaları, Görüntü İşleme, Üzüm Türleri, Yaprak Tanıma

### Abstract

In this paper, to classify the grape tree species, the extracted features from leaf images are classified using a multi-class support vector machines. Feature extraction stage, the grape leaves are calculated by using 9 different features. Image processing stage involves gray tone dial, median filtering, contrast, thresh holding and morphological-logical processes. In the classification stage, the obtained properties with the help of multi-class support vector machines (MCSVM) is performed classification process. In the testing phase, by applying the different leaf images is calculated the performance of model. In this study, MATLAB software was used. At the end of the test was determined the total success rate of 90.7%.

**Keywords:** Image Processing, Grape Varieties, Leaf Recognition, Support Vector Machines

### I. GİRİŞ

Günümüzde bilgisayar teknolojilerinin gelişmesi ile birlikte görsellik de önem kazanmıştır. Tıp, güvenlik, üretim, bilim ve uzay, askeri ve ziraat gibi birçok alanda görüntü işleme teknolojisi sayesinde yenilikler ve kolaylıklar sağlamıştır.

Görüntü işleme teknolojisi ile verilerin analizi ve değerlendirilmesi kısa zamanda gerçekleştirilerek yüzde yüze yakın sonuçlar elde edilebilmektedir. Bu sayede önceden tek tek insanlar tarafından incelenen veriler, günümüzde daha kolay, hatasız ve zaman kaybı olmadan en kısa sürede gerçekleştirmeye olanak sağlamıştır [1-2].

Bitkiler besin sektörü, endüstri ve medikal alanda yaygın şekilde kullanılmaktadır. Kullanım yoğunlukları yanı sıra, bitki türleri üzerindeki çalışmalar özellikle doğal ve çevre koruma için çok önemlidir. Bitki türleri doğal dengenin bozukluğu ve çevre kirliliği yüzünden nesli tükenmekte ve yakın gelecekte bazıları yok olma tehlikesindedirler. Bu yüzden, biyolojik çeşitlilik veri tabanını oluşturmak, dünyadaki bitkileri korumaya yardımcı olabilmek için çok önemlidir [2].

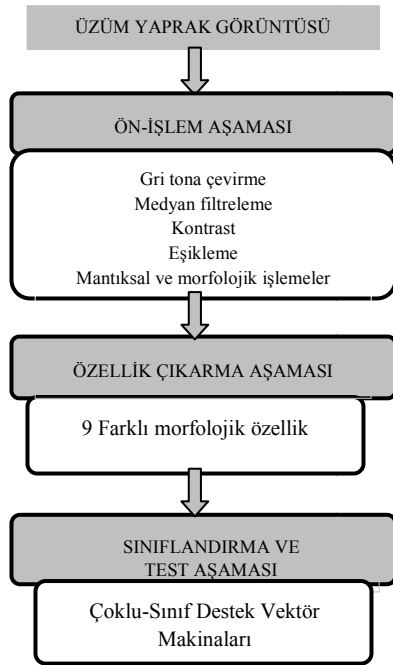
Dünya genelinde yaklaşık 10.000 üzüm çeşidi vardır. Bunların yaklaşık 3.000 tanesi yaygın bir şekilde kullanılır. Üzüm çeşitlerinin belirlenmesi genetik kaynakları korumada, yeni türleri belirlemede ve kaynak istatistikleri için büyük öneme sahiptir [1].

Bitki sınıflandırması ya geleneksel metotlar ile ya da DNA analizi ile yapılır. Otomatik bitki tanıma sistemleri bitkileri tanımak ve sınıflandırmak için yaprağı ve yaprağın biçimsel özelliklerini kullanmaktadırlar [2].

Bitki tanıma ve sınıflandırma uygulamalarında birçok yöntem bulunmaktadır. Yapılan çalışmalarda genelde morfolojik özellikler kullanılmıştır. (Uluturk, 2009) yapmış oldukları çalışmada 7 temel morfolojik özellik ve buna ek olarak yarım yaprak görüntülerini kullanılarak 10 özellik elde etmişlerdir. Veri olarak 32 farklı bitki türünden oluşan 1120 yaprak resmi sınıflandırma aşamasında PNN(Probabilistic Neural Network) kullanılarak sınıflandırılmış ve başarı oranı

%92,5 olarak hesaplanmıştır [2]. Bir diğer çalışmada özellik çıkarma aşamasında 5 temel değişkene dayalı olarak 12 farklı özellik çıkarılmıştır. Veri olarak 32 farklı bitki türlerinden oluşan 1800 yaprak resmi kullanılmıştır. Sınıflandırıcı olarak PNN, PCA, SVM-BDT ve Fourier moment teknikleri kullanılmış ve modellerin performansları birbirleri ile kıyaslanmıştır. Başarı oranı olarak ortalama %90 doğruluk tespit edilmiştir[3]. Başka bir çalışma da ise özellik çıkarma aşamasında Hu's değişmez moment kullanılmıştır. Veri olarak 20 çeşit üzüm türüne ait olan 354 yaprak resmi kullanılmıştır. Sınıflandırıcı olarak Öklid uzaklığı(Euclidean distance) temel alınmıştır. Başarı oranı %87 olarak tespit edilmiştir [1].

Gerçekleştirilen bu çalışmada, kolay ve doğru bir şekilde yaprağın özelliklerini çıkartan etkili bir algoritma kullanılmıştır. Veri olarak 18 adet üzüm türünden 270 adet yaprak resmi kullanılmıştır [4]. Özellik çıkarma aşamasında işlenen her bir resimden 9 adet farklı özellik çıkartılmıştır. Bu elde edilen özellikleri sınıflandırmak için DVM sınıflandırıcısı kullanılmıştır ve doğruluk yüzdesi %90.7 olarak tespit edilmiştir.



Şekil 1. Önerilen sisteminin genel aşamaları

## II. GÖRÜNTÜ İŞLEME

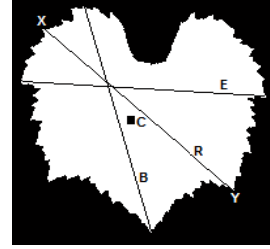
Görüntü işleme teknikleri resme ait özelliklerin belirlenmesi için kullanılır. Bu makale de kullanılan görüntü işleme adımları şunlardır;

- Gri tona çevirme,
- Medyan filtreleme,
- Görüntüyü iyileştirmek ve belirginleştirme,
- Eşikleme
- Aşınma, Dolu ve Mantıksal işlemler

## III. ÖZELLİK ÇIKARMA

Yapılan çalışmada, literatürde yaygın olarak kullanılan 6 temel geometrik özelliğe dayalı olarak 9 farklı morfolojik özellik çıkartılmıştır.

### a) Temel Geometrik Özellikler



Şekil 2. Geometrik özelliklerin semboller ile gösterilmesi(1)

- 1) **Boyu:** Yaprak resminin düşey olarak iki uç arasındaki mesafedir. B sembolü ile ifade edilir.
- 2) **Eni:** Yaprak resminin düşey olarak iki uç arasındaki mesafedir. E sembolü ile ifade edilir.
- 3) **Alan:** Yaprığın bölgedeki piksel sayısını temsil eder. İkili biçiminde yaprak görüntünün arka planı siyah ve yaprak beyazdır. Yani görüntüde, beyaz piksel sayısı yaprak alanını temsil eder. A sembolü ile ifade edilir.
- 4) **Çevre:** Çevre yaprağın etrafını gösteren kapalı eğrinin uzunluğudur. Yani yaprağın çevresi, yaprağın kenarlarını oluşturan pikseller sayılarak hesaplanır. P sembolü ile ifade edilir.
- 5) **Ağırlık merkezi:** Yaprığın matrisindeki her bir boyuta ait elemanların o boyuttaki konumları toplamının alana bölünmesiyle elde edilir. C sembolü ile ifade edilir.
- 6) **Çap:** Yaprığın birbirine en uzak iki noktası arasındaki mesafe olarak tanımlanır.

### b) 9 Morfolojik Özellik

1. Yaprığın boyunun, yaprağın enine oranıdır.

$$\text{Birinci özellik} = \left( \frac{B}{E} \right)$$

2. Yaprığın çevresinin, boyu ile eninin toplamına oranıdır.

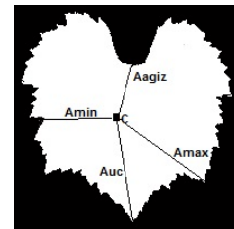
$$\text{İkinci özellik} = \frac{P}{E+B}$$

3. Yaprığın alanının, boyu ile eninin çarpımına oranıdır.

$$\text{Üçüncü özellik} = \frac{A}{E*B}$$

4. Yaprığın dışında kalan kısmın alanının(S), yaprağın alanına oranıdır.

$$\text{Dördüncü özellik} = \frac{S}{A}$$



Şekil 3. Geometrik özelliklerin semboller ile gösterilmesi(2)

5. Yaprığı çevreleyen sınır eğrilerinin ağırlık merkezine olan maksimum uzaklığının, ağırlık merkezine minimum uzaklığına oranıdır.

$$\text{Beşinci özellik} = \frac{A_{\max}}{A_{\min}}$$

6. Ağırlık merkezinin yaprağı çevreleyen sınır eğrilerine olan ortalama uzaklığı

7. Ağırlık merkezinin yaprağın ağız kısmına olan uzaklığının, yaprağın uç kısmına olan uzaklığına oranıdır.

$$\text{Yedinci özellik} = \frac{A_{\text{ağız}}}{A_{\text{uç}}}$$

8. Yaprığın çapının uzunluğunun, yaprağı çevreleyen sınır eğrilerinin ağırlık merkezine olan maksimum ve minimum uzaklıkları toplamına oranıdır.

$$\text{Sekizinci özellik} = \frac{R}{A_{\max} + A_{\min}}$$

9. Yaprığın birbirine olan en uzak iki noktanın (X ile Y), ağırlık merkezine olan uzaklıkları oranıdır.

$$\text{Dokuzuncu Özellik} = \frac{\sqrt{(X(1)-C(1))^2 + (X(2)-C(2))^2}}{\sqrt{(Y(1)-C(1))^2 + (Y(2)-C(2))^2}}$$

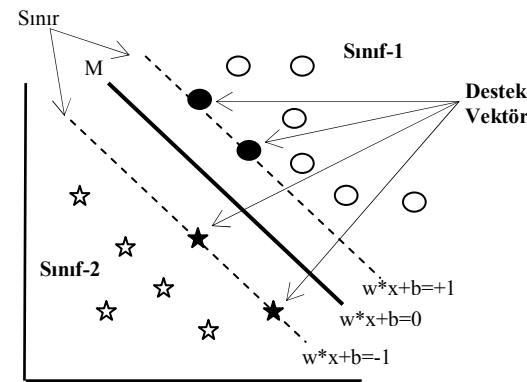
#### IV. SINIFLANDIRMA

Destek vektör makineleri (DVM), Vapnik tarafından geliştirilen ve istatistiksel öğrenme teorisine dayanan bir yöntemdir [5,6]. DVM 'nin diğer sınıflandırıcılardan farkı yanlış sınıflandırma ihtimalini en aza indirgeyecek çözümler bulmaya çalışır [7]. DVM yöntemi son yıllarda birçok alanda kullanılmaya başlanmıştır. Resim ve nesne tanıma, ses tanımlama, parmak izi tanıma, el yazımı tanınması ve daha birçok alanlar da kullanılmıştır.

Farklı 2 örnek verilmiş sınıftan oluşan veri vektörü n-boyutlu uzayda  $\{x_i\} \in R^n$  ve  $(x_i, y_i)$ ,  $i = 1, 2, 3, \dots, N$  olmak üzere buna karşılık gelen sınıf etiketi  $y_i \in \{+1, -1\}$  olsun.

Bir hiperdüzlem  $w^t x + b = 0$  tarafından ayrıştırılabilir. Buradaki w n-boyutlu bir vektör, b ise eşik değeridir.

$$y_i(w^t x + b) \geq 1, \quad i = 1, 2, 3, \dots, N$$



Şekil 4. DVM' nin hiperdüzlemlen ayrılması

Maksimum aralık  $2/\|w\|$  alınır. Böylece  $\|w\|^2$  minimize ederek maksimum marjın veren hiperdüzlemi bulur. Bu problem Lagrange optimizasyon yöntemi kullanarak ikili sorun çözülebilir.

$$\text{Maksimize } Q(a) = \sum_{i=1}^N a_i - \frac{1}{2} \sum_{i,j=1}^N a_i a_j y_i y_j K(x_i, x_j)$$

$$\sum_{i=1}^N y_i a_i = 0$$

$$0 \leq a_i \leq C, \quad i = 1, 2, 3, \dots, N$$

$K(x_i, x_j)$ , kernel fonksiyonudur. C ise hata parametresidir. Buradaki  $a = (a_1, a_2, \dots, a_i)$  Lagrange çarpanlarıdır. Eğitim setindeki her bir örnek  $a_i$ 'ye karşılık gelmektedir.

Optimizasyon teorisinin Kuhn-Tucker teorisine göre sıfır olmayan Lagrange çarpanlarına sahip örnekler destek vektör olarak adlandırılır. Eğer  $a_i$  en iyi çözüm ise, sınıflandırıcı karar fonksiyonu aşağıdaki biçimde ifade edilebilir.

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^N y_i a_i K(x, x_i) + b^* \right\}$$

Bu çalışmada birden fazla sınıf olduğu için sınıflandırma problemi bulunmaktadır. Bunun çözümü içinde çok sınıflı bir DVM sınıflandırıcısına ihtiyaç doğar. Çok sınıflı sınıflandırma, ikili sınıflandırıcıların birleştirilmesiyle oluşur [6].

Çalışmamızda çok sınıflı DVM için bire-karşı-bir (BKB) yöntemi kullanılmıştır. k sınıflı bir sınıflandırma problem için sınıfların tüm kombinasyon çiftlerini göz önünde bulundurarak k(k-1)/2 ikili sınıflandırıcılar oluşturulur. Bu algoritmanın basit olmasının yanı sıra eğitim ve sınıflandırma işlemleri hızlı bir şekilde uygulanır [8].

#### V. METODOLOJİ

Bu çalışmada, 18 adet üzüm türünden 270 adet renkli yaprak resmi kullanılmıştır. İlk olarak bu elde edilen resimleri görüntü işleme yöntemleri kullanılarak arka plandan ayırma işlemleri yapılmıştır.



Şekil 5. a) Gerçek resim  
b) Bölütleme uygulandıktan sonraki resim

Elde edilen bu görüntülerden özellikler çıkarabilmek için 9 farklı morfolojik özellik kullanılmıştır. Aşağıdaki 9 farklı özellik üzüm türlerinden olan "Abondant" bitkisine aittir.

$$[1.0336 \ 2.3109 \ 0.6622 \ 1.3129 \ 17.1266 \ 60291 \ 0.9946 \ 1.3417 \ 1.0361]$$

Bu özellikleri sınıflandırabilmek için de çoklu-sınıf destek vektör makineleri kullanılmıştır. Eldeki verilerin yaklaşık yüzde 80'i eğitim için, kalan yüzde 20' u ise test aşaması için kullanılmıştır.

Tablo1

ÇSDVM göre sistemin sınıflandırma doğruluğu

Eğitim Örnekleri	Doğru Tanıma Sayısı	Yanlış Tanıma Sayısı	Başarı oranı
54	49	5	%90.7

## VI. TARTIŞMA VE SONUÇ

Bugüne kadar pek çok nesne tanıma çalışmaları yapılmıştır ve yapılmaya devam etmektedir. Bununla birlikte birçok özellik çıkarma ve sınıflandırma yöntemleri kullanılmıştır. Bu çalışmada ise, geliştirilen bu sistem ile farklı türdeki üzüm yapraklarından elde edilen özellikleri sınıflandırabilmek için ÇSDVM sınıflandırıcısı kullanılmıştır. Bu sınıflandırma aşamasında yüksek bir başarı oranı sağlamıştır. Elde edilen doğru sınıflandırma oranı %90.7'dir. Yaklaşık %9.3 oranındaki hatanın nedeni verilerin az ve net olmamasından kaynaklanmaktadır.

Bu çalışmada elde edilen sonuçlar, bitki yapraklarını doğru sınıflandırmak için önerilen bu sistemin en önemli parçası görüntü işleme aşaması olduğu görülmüştür.

## KAYNAKÇA

1. Zhang J., Yanne P. And Li H., "Identification of Grape Varieties via Digital Leaf Image Processing by Computer", Bulletin de l'OIV 84 (959-961), 5-14, 2011.
2. Uluturk C. and Ugur A., "Recognition of leaves based on morphological features derived from two half-regions", IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA), pp.1-4, 2012.
3. S. G. Wu, F. A. Bao, E. Y. Xu, Y. Wang, Y. Chang, Q. Xiang, "A leaf recognition algorithm for plant classification using probabilistic neural network," IEEE 7<sup>th</sup> International Symposium on Signal Processing and Information Technology, pp. 11-16, 2007.
4. Vitis International Variety Catalogue(2007) 19 Aralık 2014 tarihinde <http://www.vivc.de/> adresinden erişilmiştir.
5. Ekici S., Yıldırım S., Poyraz M., "Mwsafe koruma için örüntü tanıma uygulaması " Gazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi, Cilt24, No 1, pp.51-61, 2009.
6. Jiang, Z.G., Fu, H.G., Li, L.J., "Support vector machine for mechanical faults classification", Journal of Zhejiang University Science, 6A, 5, pp. 433-439, 2005.
7. V. N. Vapnik, "The nature of statistical learning theory," New York: Springer Verlag, 1995.
8. J. Li, C. Zhang, and Z. Li, "Battlefield Target Identification Based on Improved Grid-Search SVM Classifier," 2009 Int. Conf. Comput. Intell. Softw. Eng., pp. 1-4, Dec. 2009.