



**T.C.
DÜZCE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE ARI ALT
TÜRLERİNİN SINIFLANDIRILMASI**

HASAN DEMİR

**YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI**

**DANIŞMAN
DOÇ. DR. PAKİZE ERDOĞMUŞ**

DÜZCE, 2018

T.C.
DÜZCE ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE ARI ALT
TÜRLERİNİN SINIFLANDIRILMASI

Hasan DEMİR tarafından hazırlanan tez çalışması aşağıdaki jüri tarafından Düzce Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Doç. Dr. Pakize ERDOĞMUŞ

Düzce Üniversitesi

Jüri Üyeleri

Prof. Dr. Recep DEMİRCİ
Gazi Üniversitesi

Doç. Dr. Pakize ERDOĞMUŞ
Düzce Üniversitesi

Yrd. Doç. Dr. Meral KEKEÇOĞLU
Düzce Üniversitesi

Tez Savunma Tarihi: 12/01/2018

BEYAN

Bu tez çalışmasının kendi çalışmam olduğunu, tezin planlanmasından yazımına kadar bütün aşamalarda etik dışı davranışımın olmadığını, bu tezdeki bütün bilgileri akademik ve etik kurallar içinde elde ettiğimi, bu tez çalışmasıyla elde edilmeyen bütün bilgi ve yorumlara kaynak gösterdiğimi ve bu kaynakları da kaynaklar listesine aldığımı, yine bu tezin çalışılması ve yazımı sırasında patent ve telif haklarını ihlal edici bir davranışımın olmadığını beyan ederim.

12 Ocak 2018

Hasan DEMİR

TEŞEKKÜR

Yüksek lisans öğrenimimde ve bu tezin hazırlanmasında gösterdiği her türlü destek ve yardımdan dolayı çok değerli hocam Doç. Dr. Pakize Erdoğan'a en içten dileklerimle teşekkür ederim.

Tez çalışmam boyunca değerli katkılarını esirgemeyen Yrd. Doç. Dr. Meral Kekeçoğlu'na da şükranlarımı sunarım.

Bu çalışma boyunca yardımlarını ve desteklerini esirgemeyen, başta eşim, annem ve babam olmak üzere tüm aile ve çalışma arkadaşlarıma sonsuz teşekkürlerimi sunarım.

12 Ocak 2018

Hasan DEMİR

İÇİNDEKİLER

	<u>Sayfa No</u>
ŞEKİL LİSTESİ.....	VII
ÇİZELGE LİSTESİ.....	VIII
KISALTMALAR.....	IX
SİMGELER	X
ÖZET	XI
ABSTRACT	XII
1. GİRİŞ.....	1
2. MAKİNE ÖĞRENMESİ YÖNTEMLERİ	3
2.1. DESTEK VEKTÖR MAKİNELERİ	4
2.2. K-EN YAKIN KOMŞU	6
2.3. K-MEANS.....	7
2.4. YAPAY SİNİR AĞLARI	8
3. ÖZELLİK SEÇİM METODLARI	11
4. ARI ALT TÜRLERİNİN SINIFLANDIRILMASI	13
4.1. LİTERATÜRDEKİ ÇALIŞMALAR	14
4.2. GEOMETRİK MORFOMETRİ	14
4.3. ARI KANADI ÜZERİNDEN ÖZELLİK ALINMASI	16
5. MATERYEL VE YÖNTEMLER.....	19
5.1. ARI KANADI ÖRNEKLERİNİN ALINMASI.....	19
5.2. ARI KANADI ÜZERİNDE KAVŞAK BELİRLEME	20
5.3. ÖNERİLEN KAVŞAK SEÇİM ALGORİTMASI (ÖKSA).....	21
5.4. ÖZELLİK ÇIKARMA	23
6. BULGULAR VE TARTIŞMALAR.....	26
6.1. KANAT İŞARETLEME İLE ALINAN KOORDİNATLAR	26
6.2. NORMALİZASYON İŞLEMİ UYGULANMIŞ VERİLER.....	27
6.3. ÖZELLİK SEÇİM SONUÇLARI	30

6.4. ÖZELLİKLERİN FARKLI BİR BOYUTA TAŞINMASI.....	32
6.5. SINIFLANDIRMA SONUÇLARI VE DOĞRULUK MATRİSİ.....	32
6.5.1. ÖKSA Kullanılmadığı Durumda Sınıflandırma Sonuçları	33
6.5.2. ÖKSA İle 27 Özelliğin Tamamı İçin Başarı Oranları	33
6.5.3. BFM İle Seçilen 22 Özelliğin Sınıflandırılması	35
6.5.4. SFS İle Seçilen Özellikler ve Sınıflandırma.....	36
6.5.5. LDA İle Seçilen Özellikler ve Sınıflandırma Başarısı	37
6.6. GENEL SINIFLANDIRMA BAŞARILARI	39
7. SONUÇ VE ÖNERİLER.....	41
8. KAYNAKLAR.....	43
ÖZGEÇMİŞ	46

ŞEKİL LİSTESİ

	<u>Sayfa No</u>
Şekil 1.1. Arı kanadı üzerindeki bazı morfolojik özellikler.	2
Şekil 2.1. Sınıflandırma.	4
Şekil 2.2. DVM’de hiper düzlem ve ayırıcı vektörler.	5
Şekil 2.3. KNN ile veri sınıflandırma.	7
Şekil 2.4. K-Means ile veri sınıflandırma.	8
Şekil 2.5. Yapay sinir ağlarının genel yapısı.	10
Şekil 3.1. Sarmalayıcı özellik seçim algoritması.	12
Şekil 4.1. Türkiye’deki arı ırk ve ekotipler ile bunların dağılım alanları.	13
Şekil 4.2. Arı kanadı üzerinde alan.	15
Şekil 4.3. Arı bacağı üzerinde uzunluk ölçümü.	15
Şekil 4.4. Balık boyu ve kuyruk açısı ölçümü.	15
Şekil 4.5. Kelebek üzerinde açı ölçümü ve en-boy oranı.	15
Şekil 4.6. Kanat özellikleri.	16
Şekil 5.1. Örneklerin sentilasyon şişelerinde kodlanması.	19
Şekil 5.2. TpsDig32 programı ile kavşak noktası belirleme.	20
Şekil 5.3. Matlab programı ile kavşak noktası belirleme.	20
Şekil 5.4. Kanat üzerinde standart kavşak belirleme işlemi.	22
Şekil 5.5. Kavşak noktalarından ile açı hesaplama.	23

ÇİZELGE LİSTESİ

	<u>Sayfa No</u>
Çizelge 2.1. Veri seti ve sınıf etiketi.....	3
Çizelge 2.2. KNN ile sınıflandırma adımları.....	6
Çizelge 2.3. Toplam fonksiyonları.	9
Çizelge 2.4. Aktivasyon fonksiyonları.	10
Çizelge 3.1. Filtreleme metodu adımları.	11
Çizelge 5.1. Örneklerin alındığı il/ilçeler.	19
Çizelge 5.2. Farklı zamanlarda alınan koordinatlarının değişimi.	21
Çizelge 5.3. ÖKSA ile farklı zamanlarda işaretlenen koordinatlar.	22
Çizelge 5.4. Kullanılan özellikler.	23
Çizelge 5.5. Karışıklık matrisi (Confusion matrix).	25
Çizelge 6.1. ÖKSA kullanılmadan alınan örnek koordinatlar.	26
Çizelge 6.2. ÖKSA ile alınan koordinatlar.	27
Çizelge 6.3. Normalizasyon uygulanmamış özellik örnekleri.	28
Çizelge 6.4. İki ilçeye ait normalizasyon uygulanmış özellikler.	29
Çizelge 6.5. SFS algoritması sonucunda seçilen özellikler.	30
Çizelge 6.6. BFM ile seçilen özellikler.....	31
Çizelge 6.7. LDA ile boyut indirgeme sonuçları.	31
Çizelge 6.8. LDA ile verilerin farklı bir boyuta taşınması.	32
Çizelge 6.9. ÖKSA kullanılmadığı durumda sınıflandırma başarıları	33
Çizelge 6.10. ÖKSA kullanılmadığı durumda örnek bazında doğruluk matrisleri.	33
Çizelge 6.11. ÖKSA kullanıldığı durumda sınıflandırma başarıları	34
Çizelge 6.12. ÖKSA ile 27 özellik kullanıldığında doğruluk matrisi.....	34
Çizelge 6.13. BFM ile sınıflandırma başarıları.....	35
Çizelge 6.14. BFM doğruluk matrisi.	36
Çizelge 6.15. SFS ile sınıflandırma başarıları.	36
Çizelge 6.16. SFS ile yapılan sınıflandırmada doğruluk matrisi.	37
Çizelge 6.17. LDA ile seçilen 9 özellik için başarı yüzdeleri.....	38
Çizelge 6.18. LDA ile seçilen özellikler için doğruluk matrisleri.	38
Çizelge 6.19. Tüm il/ilçeden alınan örneklerin sınıflandırma başarıları.....	39
Çizelge 6.20. Kullanılan sınıflandırma yöntemlerinin sınıflandırma başarıları.	39
Çizelge 6.21. İl/ilçe bazında sınıflandırma başarıları.....	40

KISALTMALAR

BFM	Kaba kuvvet yöntemi
DN	Doğru negatif
DP	Doğru pozitif
DVM	Destek vektör makineleri
KNN	K-nearest neighbors
LDA	Linear discriminant analysis
ÖKSA	Önerilen kavşak seçim algoritması
SFS	İleri sıralı seçim
YN	Yanlış negatif
YP	Yanlış pozitif
YSA	Yapay sinir ağları

SİMGELER

©	Copyright
∈	Elemanı
Σ	Toplam



ÖZET

MAKİNE ÖĞRENMESİ YÖNTEMLERİ İLE ARI TÜRLERİNİN SINIFLANDIRILMASI

Hasan DEMİR

Düzce Üniversitesi

Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Danışman: Doç. Dr. Pakize ERDOĞMUŞ

Ocak 2018, 45 sayfa

Bu araştırmada arı kanatları üzerindeki kavşak noktalarına göre arı türlerinin sınıflandırılması amaçlanmıştır. Beş farklı il/ilçe 'den alınan arı kanat resimleri üzerinde kavşak noktaları belirlenmiştir. Arı kanatları üzerinde kavşak noktalarının belirlenmesi işleminin minimum hata ile yapılması ve standart bir kavşak noktası seçimi için kavşak noktası seçim algoritması önerilmiştir. Kavşak noktaları kullanılarak açı, uzunluk, indeks ve alan bilgileri içeren 27 morfolojik özellik çıkarılmıştır. Bu özellikler arasındaki sayısal farklılıkların giderilmesi ve iş yükünün azaltılması için verilere normalizasyon işlemi uygulanmıştır. Destek Vektör Makineleri (DVM), Yapay Sinir Ağları (YSA), K-Ortalama (K-Means) ve K-en Yakın Komşu (KNN) algoritmaları, sınıflandırma yöntemi olarak kullanılmıştır. Veri boyutunu azaltmak için Kaba Kuvvet Yöntemi (BFM) , İleri Sıralı Seçim (SFS) ve Lineer Discriminant Analiz (LDA) yöntemlerinden yararlanılmıştır. Arı alt türlerinin sınıflandırılmasında en yüksek başarı, K-Means yöntemi ile %50, DVM yöntemi ile % 71, KNN yönteminde ile %55,3 ve YSA ile %82,7 olarak gözlemlenmiştir. Kavşak noktaları için önerilen algoritmanın kullanıldığı ve kullanılmadığı durumlardaki başarı oranları, boyut indirgeme yöntemlerinin kullandığı ve kullanılmadığı durumlardaki başarı oranları incelenmiştir. Görüntüler üzerindeki kavşak noktalarının seçimi, özellik çıkarma ve verilerin sınıflandırılmasında MATLAB® programı kullanılmıştır.

Anahtar sözcükler: Arı kanadı, Makine öğrenmesi, Sınıflandırma.

ABSTRACT

CLASSIFICATION BEE SUBSPACES BY MACHINE LEARNING METHODS

Hasan DEMİR

Düzce University

Graduate School of Natural and Applied Sciences, Department of Computer
Engineering

Master's Thesis

Supervisor: Assoc. Prof. Pakize ERDOĞMUŞ

January 2018, 45 pages

In this research, it is aimed to classify the bee species according to the intersection points on the bee wings. The intersection points on the pictures of bee wings taken from five different country / districts have been determined. The intersection point selection algorithm is proposed for the determination of intersection points on bees' wings with minimum error and for selecting a standard intersection point. Using the intersection points, 27 morphological features including angle, length and area information were extracted. The normalization process has been applied to eliminate numerical differences between these features and to reduce workload. Support Vector Machines (SVM), Artificial Neural Networks (ANN), K-Means (K-Means) and K-Nearest Neighbors (KNN) algorithms were used as the classification method. Brute Force Method (BFM), Forward Sequential Selection(SFS) and Linear Discriminant Analysis (LDA) methods were used to reduce the data size. The highest success rate in the classification of bee subspecies was 50% with K-Means, 71% with SVM, 55,3% with KNN and 82,7% with ANN. The success rates of the proposed and unused algorithms for intersection points and the success rates of cases where size reduction methods are used or not are examined. The selection of the intersection points on the images, feature extraction and the classification of data, MATLAB[®] program is used.

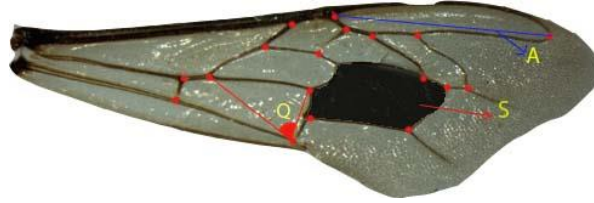
Keywords: Bee wing, Classification, Machine learning.

1. GİRİŞ

Büyük miktarda verilerin işlenmesi ve analiz edilmesi süreci elle yapılmayacak kadar uzundur. Bu nedenle, problemlerde verinin belli kısımlarını kullanarak bir model elde edilmektedir. Makine öğrenmesi ham veriden anlamlı bilgi çıkarımıdır. Veri madenciliği ise bu verilerden işe yarayacak olanların keşfedilmesi sürecidir. Verinin toplanması saklanması ve analiz edilmesi için bilgisayarlara ve daha çok yazılımlara ihtiyaç duyulmaktadır. Veri madenciliği tanımlayıcı ve tahmin edici modeller olarak ikiye ayrılır. Tahmin edici modellerde, sonuçları bilinen verilerden hareket edilerek bir model geliştirilir ve bu model yardımıyla girdi değerlerine karşılık çıktı değerleri tahmin edilmeye çalışılır. Tanımlayıcı modellerde ise karar vermeye rehberlik etmede kullanılabilecek mevcut verilerdeki örüntülerin tanımlanması sağlanmaktadır [1]. Tahmin edici modeller sınıflandırma ve regresyon olarak iki ana başlıkta toplanır. Sınıflandırma; yeni karşılaşılan veri örneklerinin önceden karşılaşılmış olan verilerden elde edilen bilgiler ışığında farklı sınıflara ayrıştırılması işlemidir [2]. Regresyon analizi; eldeki girdi değerlerini kullanarak çıktı için fonksiyon üretme veya eğri uydurma olarak isimlendirilir. En bilinen sınıflandırma yöntemleri k-ortalama, destek vektör makineleri, karar ağaçları, naive bayes, yapay sinir ağları, k-en yakın komşu ve rastgele orman (random forest)'dır. Veri madenciliği, mühendislikten ekonomiye, tarımdan astronomiye kadar hemen her alanda kullanılmaktadır. Biyolojide de canlı türlerine ait bilginin çıkarılması işlemlerinde veri madenciliğinden yararlanılır.

Arıların kretase döneminin sonlarına doğru çiçekli bitkilerin ortaya çıkması ile gelişmeye başladığı bilinmektedir. Çiçekli bitkilerin tozlaşmasına yardımcı en aktif canlı grubunun bal arıları olmasından dolayı çiçekli bitkilerin ve bal arılarının paralel şekilde evrimleşmesi söz konusudur [3]. Tür kendi aralarında üreyebilen canlılara denir. Linnaeus tarafından 1758'de yapılan çalışmada, Apis florea, Apis dorsata, Apis Cerana, Apis mellifera dört arı türünün varlığından söz edilmiştir. Alt tür ise bazı araştırmacılara göre bir türün kalıtsal farklılığını ifade ederken diğer bir kanı, türlerin coğrafi dağılımını ifade etmektedir [4]. Dünyada morfolojik, fizyolojik ve davranış karakterleri açısından tanım ve taksonomik sınıflandırması yapılmış 24 coğrafi arı ırkı belirlenmiştir [5]. Arı ırklarının belirlenmesinde biyolojik ve morfolojik yöntemler kullanılmaktadır.

Morfolojik özelliklerin belirlenmesinde kanat özellikleri, kıl özellikleri, bacak özellikleri, tergite genişliği ve renk kullanılabilmektedir [6]. Kanat özellikleri ile bal arısı ırkının belirlenmesinde daha çok kanat üzerindeki kavşak kesim noktaları belirlenmekte, alan, indeks, boyut, açı gibi özellikler kullanılarak tür tespiti yapılmaktadır. Şekil 1.1’de bu özellikler gösterilmiştir (A=uzunluk, S=Alan, Q= Açı) .



Şekil 1.1. Arı kanadı üzerindeki bazı morfolojik özellikler.

Bal arısı ırklarını belirleme çalışmalarında önceleri klasik morfometri yöntemiyle abdomendeki renklenmeler değerlendirilmiş kanat, bacak, dil ve mum yüzeylerinde birçok morfometrik ölçüm yapılmıştır [7]-[9]. Fakat birden çok karakterin ölçümü zaman aldığından bir çözüm yolu olarak diğer vücut parçalarının yerine kanattaki ölçüm sayısı artırılarak sadece kanatta otomatik ölçüm yapılmasının daha avantajlı olacağı önerilmiştir [10]. Böylece ölçümler hızlı ve daha hassas olacaktır. Dolayısıyla araştırmacılar geometrik morfometrik yöntemlere yönelmişlerdir.

Bal arısı ırklarını sınıflandırma çalışmalarında çeşitli paket programlar (tpsDig, DrawWing, BeeWing, IMP, Morpheus, Morphologika) kullanılmıştır. Bu programlar ile işaretlenen noktaların koordinatları alınmakta ve kaydedilmektedir. Alan belirlemede ise kanat üzerindeki damarlar arasında kalan kısmın belirlenmesi gerekmektedir. Lakin bilgisayar programları, resmin alınması sırasında oluşan deformasyon nedeni ile damar sınırlarının tam belirleyememektedir. Kavşak işaretlemede ise bir noktaya ait aynı zamanda alınan farklı işaretlemeler arasında dahi farklılıklar oluşmakta, bu durum verilerin sınıflandırılmasında zorluğa neden olmaktadır.

Bu çalışmanın amacı arı kanatları üzerindeki kavşak noktalarının seçiminin standartlaştırılması için yeni bir algoritma ve bunun uygulanabilirliğini sağlayan bir program hazırlamak, kavşak noktaları üzerindeki koordinatlara göre özellik çıkartmak ve bu özellikler ile arı ırklarının sınıflandırmasını sağlamaktır. Ayrıca boyut indirgeme algoritmaları ve bu algoritmaların sınıflandırma başarısı üzerine katkısı ele alınmıştır.

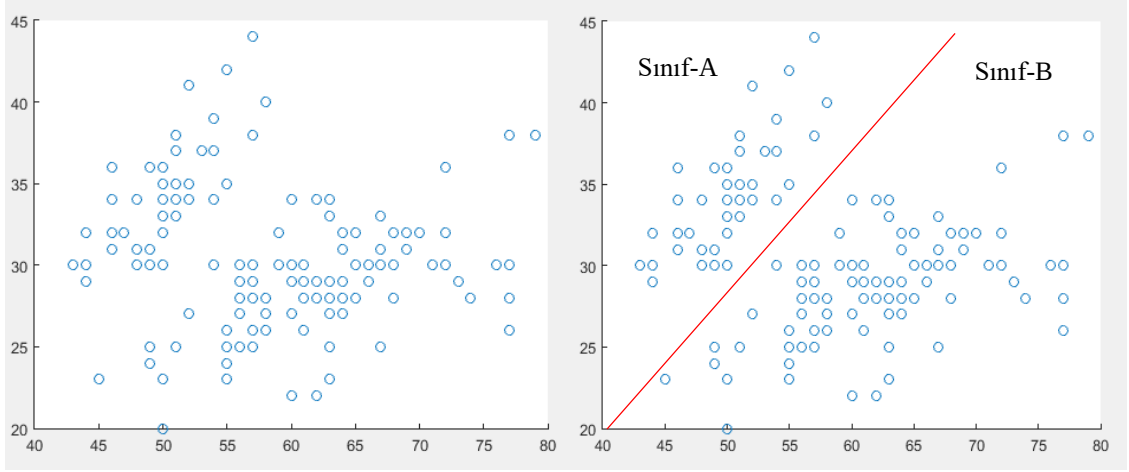
2. MAKİNE ÖĞRENMESİ YÖNTEMLERİ

Makine öğrenmesi bilgisayarların geçmiş bilgilerinden elde edilen tecrübelerden yararlanarak, gelecekteki olayların tahmin edilmesine ve modelleme yapmasına imkân veren bir yapay zekâ alanıdır [11]. Sınıflandırma bir veri setindeki her bir örneğin etiketlenmesi işlemidir. Çizelge 2.1’de veri seti ve sınıflandırma sonucu etiketleme verilmiştir. Bu tabloda verilen sınıf sayısı k ’dır.

Çizelge 2.1. Veri seti ve sınıf etiketi.

Nitelik taşıyan veriler					Sınıf etiketleri
Örnek	<i>Nitelik-1</i>	<i>Nitelik-2</i>		<i>Nitelik-k</i>	<i>Sınıf</i>
	Veri-1.1	Veri-2.1		Veri-k.1	S_1
	Veri-1.2	Veri-2.2		Veri-k.2	S_2
					
	Veri-1.n	Veri-2.n	Veri-....n	Veri-k.n	S_k

Veri sınıflandırma gözetimsiz bir öğrenme algoritmasıdır. Şekil 2.1’de iki eksenli bir doğrudaki sınıflandırma gösterilmektedir. Veriler sınıf etiketi ile etiketlenmesi için niteliklerin ayıricılığı önemlidir.



Şekil 2.1. Sınıflandırma.

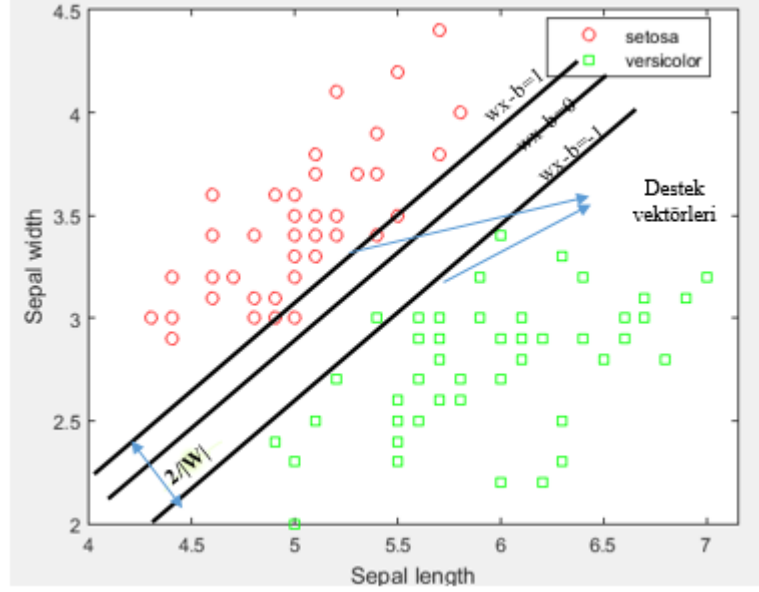
Sınıflandırma algoritmaları ile farklı performans ve başarı yüzdeleri elde etmek mümkündür. Bir grup verinin sınıflandırma başarısı, verilerin ayrılabilirliği ve sınıflandırma algoritmalarının başarısına bağlıdır. Destek vektör makinesi, veri akış madenciliği, naive bayes sınıflandırıcısı, karar ağacı öğrenmesi, k-en yakın komşu ve k-ortalama en çok kullanılan sınıflandırma yöntemleridir.

2.1. DESTEK VEKTÖR MAKİNELERİ

DVM gözetimli bir öğrenme metodudur [12]. Bir başka tanıma göre DVM istatistiksel öğrenme teorisi alanında ortaya çıkmış bir öğrenme metodudur. DVM lineer olmayan örnek uzayını, örneklerin lineer olarak ayrılabilirliği bir yüksek boyuta aktararak, farklı örnekler arasındaki maksimum sınırın bulunması esasına dayanır [13]. DVM’de problemler lineer olarak ayrılabilen ve lineer olarak ayrılamayan olmak üzere ikiye ayrılır. Lineer olarak ayrılan problemlerde özellikler arasından geçen hiper düzlemi bulmak amaçlanmaktadır. Bu hiper düzlem sınıflara ait özelliklerin birbirine en uzak olduğu iki doğrudan oluşur. Doğrusal bir denklem $ax+b$ şeklinde düşünüldüğünde burada amaç sınıfları ayıracak denklem için a ve b değerlerini hesaplamaktır. Günlük hayattaki problemler genelde lineer ayrılamamaktadır. Lineer olarak ayrılabilen problemlerde sınıflar arasında çizilecek doğrunun sınıflara uzaklığı maksimize etmeye çalışılırken, lineer olarak ayrılamayan problemlerde ise hiperdüzlem doğrusal olmayan bir fonksiyon olarak belirlenmektedir.

Bir elemandan oluşan eğitim kümesinin $A=\{a_i, b_i\}$, $i=1,2,\dots,n$ olduğu varsayalım. Burada $b_i \in \{-1,1\}$ etiket değerleri iken a_i ve b_i , n boyutlu bir uzayda özellik vektörüdür.

Doğrusal olarak ayrılabilme durumunda, bu iki değerli veriler bir aşırı düzlem ile ayrılabilir. DVM'nin amacı, bu aşırı düzlemin iki örnek grubuna en uzaklıkta olmasını sağlamaktır [14]. Bu işlem iki grup için çizilecek paralel vektörlerle gerçekleştirilecektir. Bu iki paralel vektöre eşit uzaklıktaki düzlem de aşırı düzlemdir. Şekil 2.2'de iris veri setindeki veriler 100X2 boyutundaki kısmı ve şekilsel olarak hiper düzlem gösterilmiştir.



Şekil 2.2. DVM'de hiper düzlem ve ayırıcı vektörler.

Şekil 2.2'de verilen w katsayısı değiştiğinde vektörün yönü değişecek, b sabiti değiştiğinde ise vektörün eğimi değişmeden düzlem üzerinde kayması sağlanacaktır. $2/W$ iki vektör arasındaki mesafeyi vermektedir.

$$f(x_i) = Wx_i + b = y_i, i = 1, 2, 3, \dots, m (-1 < y_i < 1) \quad (2.1)$$

$$\{x / -1 \leq y \leq 1\} \quad (2.2)$$

Denklem (2.1)'de, i özellik sayısı, y_i ise iki destek vektörü arasında kalan hiper düzleme ait noktaları içeren fonksiyondur. Karar fonksiyonu y_i değerinin -1 ve 1 sınıfına ayırır ve genel olarak Denklem (2.2)'de tanımlanmıştır.

Sonsuz sayıda karar fonksiyonu vardır. Bu karar fonksiyonları ayırıcı hiperdüzlem olarak adlandırılır. Genelleştirme kabiliyeti ayırıcı hiperdüzleme bağlıdır. En geniş sınırlı hiperdüzlem optimal ayırıcı hiperdüzlem olarak tanımlanmaktadır. Optimal ayırıcı hiperdüzlemi seçmek önemlidir; çünkü hatayı minimize eder [15].

2.2. K-EN YAKIN KOMŞU

KNN, sınıflandırma problemini çözen denetimli öğrenme (sınıflandırma için öğrenme kümesi kullanan) algoritmalarından biridir. Sınıflandırma, yeni bir nesnenin özniteliklerini inceleme ve bu nesneyi önceden tanımlanmış bir sınıfa atamaktır [16]. KNN, sınıflandırma yapılacak verilerin öğrenme kümesindeki normal davranış verilerine benzerlikleri hesaplanarak; en yakın olduğu düşünülen k verinin ortalamasıyla, belirlenen eşik değere göre sınıflara atamaları yapılır [17]. KNN yönteminin performansını;

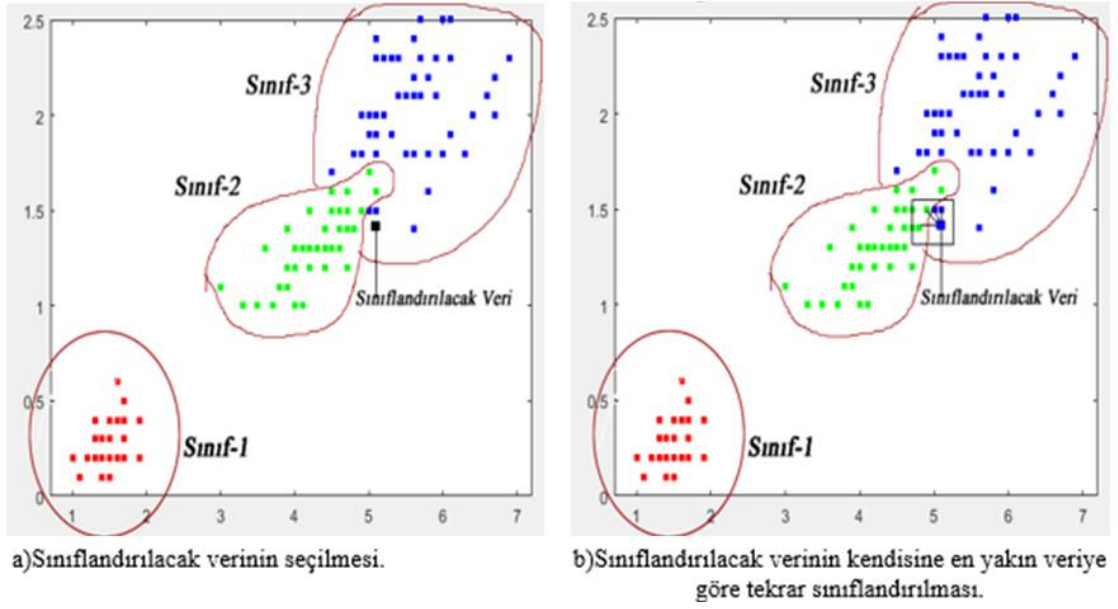
- K en yakın komşu sayısı,
- Eşik değeri,
- Benzerlik ölçütü etkilemektedir [18].

KNN ile sınıflandırmanın en önemli avantajı anlaşılabilir olması ve sınıflandırılacak verinin kendisine komşu olan verilerden yararlanılarak sınıflandırılacağından gürültüden etkilenmemesidir. KNN ile sınıflandırma adımları Çizelge 2.2’de verilmiştir.

Çizelge 2.2. KNN ile sınıflandırma adımları.

Adım-1	k sayısı belirlenir.
Adım-2	Veriye en yakın k adet komşuyu belirle.
Adım-3	Veriyi k adet komşu veri içindeki en yakın komşu verinin sınıf etiketi ile etiketle.
Adım-4	Sınıflandırma bitmedi ise adım-2’ye dön.

Bu algorithmada k değerinin yüksek seçilmesi gürültülü verilerden etkilenme riskini doğurur ve sınıflandırma maliyetini artırır. Şekil 2.3’te üç sınıfa ayrılmış bir veri setinde herhangi bir verinin KNN ile sınıflandırılması gösterilmiştir.



Şekil 2.3. KNN ile veri sınıflandırma.

Şekil 2.3 (a)'da siyah ile belirtilen verinin en yakın iki komşusu sınıf-3 olduğundan, veri sınıf-3 etiketi ile etiketlenecektir. KNN ile iki veri arasındaki mesafeyi hesaplamak için, Denklem (2.3)'te verilen öklid denklemi, Denklem (2.4)'te verilen menhetten denklemi veya Denklem (2.5)'te verilen minkowski denklemlerin birinden yararlanmak mümkündür.

$$d_1 = \sqrt{\sum_{i=1}^k (x_i - y_i)^2} \quad (2.3)$$

$$d_2 = \sqrt{\sum_{i=1}^k |x_i - y_i|} \quad (2.4)$$

$$d_3 = \sqrt{\sum_{i=1}^k |x_i - y_i|^q}^{1/q} \quad (2.5)$$

Denklemlerde verilen k veri sayısı, i ise her bir verinin indis numarasıdır. İki veri arasındaki mesafe ise d ile gösterilmiştir.

2.3. K-MEANS

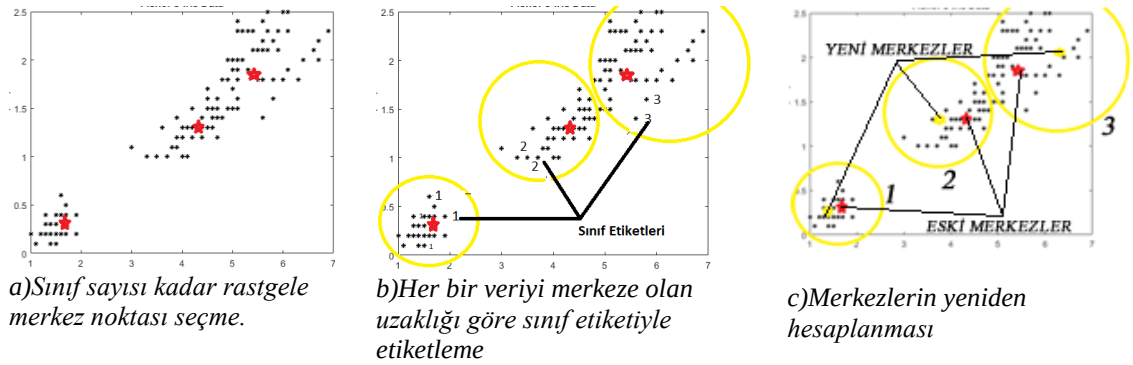
K-Means kümeleme, nesnelerin kullanıcı tanımlı 'k' sayıda kümelerine gruplandığı bölme yöntemidir [19]. K-Means temelde bir kümeleme algoritmasıdır. Tanım olarak, n

elemandan oluşan gözlem kümesini, küme içerikleri ya da merkezlerinin en yakın (uygun) olduğu k adet alt kümeye bölme işlemi şeklinde ifade edilebilir [20].

K-Means kümelemede ilk küme merkezi rasgele seçilir, data içindeki verilerin bu küme merkezine uzaklıkları hesaplanır, yeni küme merkezi belirlenir ve bu işlem kümeleme işlemi için seçilen iterasyon boyunca devam eder. Basit K-Means algoritmasını aşağıdaki gibi ifade edilebilir;

1. Sınıfları için k adet rasgele merkez nokta seç
2. Döngü
 - 2.1. Her verinin sınıf merkezlerine olan uzaklığını hesapla
 - 2.2. Hesaplanan uzaklık değerine göre verileri sınıf etiketi ile etiketle
 - 2.3. Sınıf merkez noktalarının değişimi eşik değerinde ise 4. adıma git
3. Döngü Sonu
4. Bitir

Şekil 2.4'te K-Means ile kümeleme adımları gösterilmiştir. Şekilde ilk olarak (a) sınıf sayısı kadar rasgele küme merkezi seçilmiş, her bir küme merkezi verilerin uzaklığına göre tekrar tekrar hesaplanmıştır.



Şekil 2.4. K-Means ile veri sınıflandırma.

2.4. YAPAY SİNİR AĞLARI

YSA insan beyninin işlemlerini taklit eden ardışık aşamalarındaki yapay zekanın bir türüdür [21]. Yapay sinir ağları gözetimli veya gözetimsiz olarak öğrenebilirler. Gözetimli öğrenen yapay sinir ağlarında bir grup eğitim verisi ile ağ önceden eğitilmekte ve test verisi ile de eğitilmiş ağa gelen veriler karşılığında çıktılar üretilmektedir [22]. Genel

olarak yapay sinir ağırları girdiler, ağırlıklar, toplam fonksiyonu, aktivasyon fonksiyonu ve çıktılardan oluşmaktadır. Burada amaç ağırlık değerlerini en uygun hale getirmektir. Bunun için girdi değerleri sonucu oluşan çıktı değerleri ve gerçek değerler karşılaştırılır, eşik hata değerine ulaşmamışsa tekrar ağırlık değerleri değiştirilerek hata minimum yapılmaya çalışılır. Yapay sinir ağlarında toplam fonksiyonu, ağırlık değerleri ile verileri çarpılarak toplanması sonucunda elde edilir. Çizelge 2.3'te toplam fonksiyonları verilmiştir.

Çizelge 2.3. Toplam fonksiyonları.

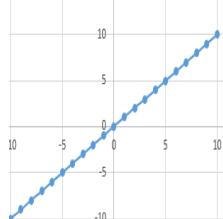
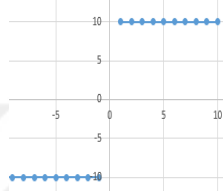
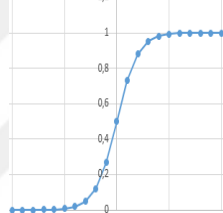
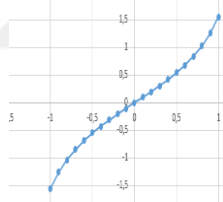
Toplam Fonksiyonu	Net Girdi
Ağırlık Toplam	$Net = \sum_{i=1}^N (X_i * W_i)$
Çarpım	$Net = \prod_{i=1}^N (X_i * W_i)$
Maksimum	$Net = \text{Max}((X_i * W_i))$
Minimum	$Net = \text{Min}((X_i * W_i))$
Çoğunluk	$Net = \sum_{i=1}^N \text{Sgn}(X_i * W_i)$
Kümülatif Toplam	$Net = Net(eski) + \sum_{i=1}^N (X_i * W_i)$

Toplam fonksiyonu ağa giren her bir verinin ağırlıklar ile çarpılarak toplanmasını sağlamaktadır. Çizelge 2.3'te i, her bir verinin indis numarasını, net ise ağırlıklar ve veriler çarpılarak, toplanması ile elde değeri ifade eder. Toplam fonksiyonu için önerilen bir seçim yöntemi yoktur. Toplam fonksiyonu daha çok deneme yanılma yolu ile seçilir. Toplam fonksiyonu seçimi, ağın başarısına fazla etkisi yoktur.

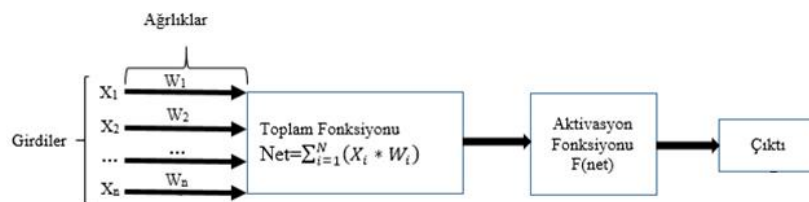
Aktivasyon fonksiyonu, YSA içinde girdilere karşı üretilen çıktıları belirler [23]. Aktivasyon fonksiyonunun seçiminde genellikle verilerin değer aralıkları etkili olmaktadır. Bu nedenle aktivasyon fonksiyonunun yanlış seçimi ağın başarısını düşürür. Yanlış aktivasyon fonksiyonu ile oluşan ağ sınıflandırmada kullanıldığında ise sınıflandırma başarısını düşürür. En fazla kullanılan aktivasyon fonksiyonu sigmoid fonksiyonudur. Sigmoid fonksiyonu için verilerin normalizasyon ile normalleştirilmesi

önerilmektedir. Doğrusal problemlerin çözümünde doğrusal aktivasyon fonksiyonu kullanılmaktadır. Doğrusal problemler için genellikle girdi sayısı düşüktür. Çizelge 2.4'te YSA'da kullanılan aktivasyon fonksiyonları verilmiştir.

Çizelge 2.4. Aktivasyon fonksiyonları.

Doğrusal Aktivasyon Fonksiyonu $F(x)=ax+b$		Doğrusal problem çözümü için kullanılır
Adım Aktivasyon Fonksiyonu $x > \text{eşik değeri} \rightarrow f(x)=1$ $x < \text{eşik değeri} \rightarrow f(x)=-1$		Daha çok iki farklı değer içeren girdiler için kullanılır.
Sigmoid Aktivasyon Fonksiyonu $F(x)=1/(1+e^{-x})$		YSA'da en çok kullanılan fonksiyondur. Girdi değerlerinin her biri için 0 ile 1 arasında değer üretir.
Tanjant Aktivasyon Fonksiyonu $F(x)=\tan(x)$		Tanjant fonksiyonu sigmoid fonksiyonuna benzer. Girdi değerlerinin -1 ile 1 arasında değer alır [24].

Aktivasyon fonksiyonunun ürettiği sonuç ve test verisi sonucunda karşılaştırılarak ağırlıklar tekrar hesaplanır. Hata payı eşik değeri ve eşik değerinin altına düşünceye kadar bu işlem tekrarlanır. Burada ağırlıkların maliyetini en iyi belirleyen ölçüt iterasyon sayısıdır. Eğer iterasyon sayısı fazla olursa başarılı sonuçlar elde edilebilir. Lakin ağırlıkların eğitilme süresi artacaktır. Giriş verileri toplam fonksiyonundan geçtikten sonra tekrar gizli katman denilen katmandan geçirilebilmektedir. Şekil 2.5'te basit bir YSA yapısı görülmektedir.



Şekil 2.5. Yapay sinir ağlarının genel yapısı.

3. ÖZELLİK SEÇİM METODLARI

Örnekleri çok sayıda öznitelik barındıran veri kümelerinin sınıflandırılması uzay ve zaman olarak yüksek maliyetlidir. Çok boyutluluğun laneti (curse of dimensionality) olarak bilinen bu problemi çözmek için öznitelik seçimi ve öznitelik çıkarımı yöntemlerinden oluşan boyut indirgeme yöntemleri geliştirilmiştir [25]. Özellik seçme belli bir veri kümesi içinden daha kullanışlı olanların alınması için kullanılır. Sınıflandırmak üzere elimizde M boyutlu bir veri varsa bu verilerden bazılarının sınıflandırma başarısına etkisi olmayabilir veya çok az etkisi olabilir. Bu durumda bu özelliğin çıkarılması gerekir. Özellik seçiminde genel olarak filtreleme yöntemi, sarmalayıcı yöntem ve yerleşik yöntemler olarak bilinen yöntemler kullanılmaktadır. Filtreleme yöntemleri genellikle bir ön işlem basamağı olarak kullanılır. Özelliklerin seçimi, herhangi bir makine öğrenme algoritmasından bağımsızdır. Bunun yerine, özellikler, sonuç değişkeniyle ilişkilendirilmesi için çeşitli istatistiksel testlerdeki puanlarına dayanarak seçilir. İlişki, buradaki öznel bir terimdir [26]. Filtre yönteminde özellik seçiminde kullanılan algoritmanın performansının optimize edilmesinden ziyade bir ölçüt veya değerlendirme fonksiyonunu maksimum yapan alt küme belirlenir [27]. Filtreleme yöntemi ile özellik seçim adımları Çizelge 3.1’de verilmiştir.

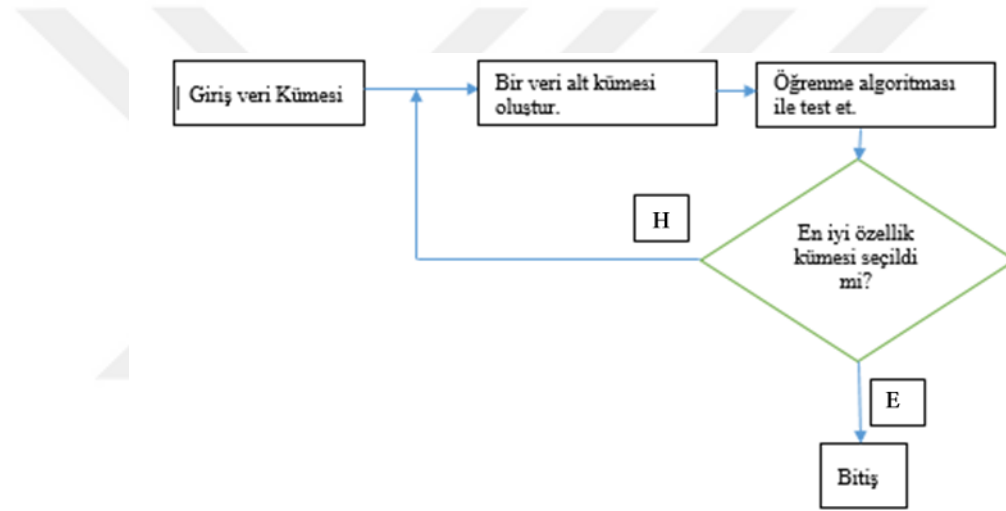
Çizelge 3.1. Filtreleme metodu adımları.

Adım	İşlem
1	Tüm girdi özelliklerini belirle,
2	En iyi özellik kümesini belirle,
3	Öğrenme algoritmalarına seçilen özellik kümesini girdi olarak gönder,
4	Öğrenme algoritmalarının performansını belirler.

Özellik seçiminde LDA, Ki-Kare ve ANOVA bilinen filtreleme metodlarıdır. LDA, veri içerisinde bulunan farklı sınıflara ait grupların doğrusal ayrılabilirliğini maksimize ederek

boyut azaltması yapan bir yöntemdir [28]. LDA, Bir grubun diğerinden iyi bir şekilde ayrılabilmesi için grup içi varyansın minimum, gruplar arası varyansın maksimum olması fikrine dayanır. Boyut azaltmada istatistiksel temele dayanan yöntemler kullanır. Kovaryans matrisi denilen, verilerin birbiri ile ilişkilerini açıklayan matrisi kullanır. ANOVA sınıf içi ve sınıflar arası varyans değerlerine göre boyut azaltmayı sağlayan istatistik temelli boyut indirgeme yöntemidir. Ki-Kare yöntemi verilerin frekans dağılımını kullanan istatistiksel bir yöntemdir.

Sarmalayıcı yöntem algoritmalarında, seçme metodu bir spesifik tahmin edicinin (algoritma) performansını doğrudan optimal yapma amacına yöneliktir. Bu, her bir adımda seçilen öznelilikler için tahmin edicinin genel performansının ölçülmesiyle yapılabilir [29] . Şekil 3.1’de sarmalayıcı metodların genel yapısı görülmektedir.



Şekil 3.1. Sarmalayıcı özellik seçim algoritması.

Bilinen sarmalayıcı özellik seçim algoritmaları, ileri sıralı seçim, geri sıralı seçim, sıralı ileri kayan seçim ve sıralı geri kayan seçimdir. Bunlardan sıralı ileri seçimde sırayla özellikler eklenir ve bu özellikler eklendikten sonra öğrenme algoritmasının başarısı test edilir. Eğer başarı artmışsa bu özellik dahil edilir, diğer durumda özellik dahil edilmez. Geri sıralı seçim algoritmasında ise özelliklerin tamamı eklenir, sırayla özellikler çıkarılır ve çıkarılan özelliğin öğrenme algoritmasına etkisine göre bu özellik dahil edilir veya çıkarılır. Bu tez kapsamında özellik seçim yöntemi olarak ileri sıralı seçim algoritması kullanılmıştır.

4. ARI ALT TÜRLERİNİN SINIFLANDIRILMASI

Türkiye coğrafik konumu, iklim özellikleri ve Asya ve Avrupa arasında bir köprü görevi görmesi nedeniyle zengin bir arı biyoçeşitliliğine sahiptir. Türkiye'nin kuzeydoğusunda A. m. caucasica, güneydoğu Anadolu'da A. m. meda, Suriye sınırında küçük bir alanda A. m. syriaca, Trakya bölgesinde A. m. carnica, Türkiye'nin geri kalan tüm alanlarında A. m. anatoliaca türlerinin bulunduğu bildirilmiştir. Ayrıca bunların yanı sıra göçer arıcılığın yaygın olmadığı kapalı ceplerde de Muğla ve Yığılca arısı gibi bazı yerel ekotiplerin varlığından da söz edilmektedir [30]-[32]. Şekil 4.1'de Türkiye'de arı ırklarının dağılımı görülmektedir [33].



Şekil 4.1. Türkiye'deki arı ırk ve ekotipler ile bunların dağılım alanları.

Türkiye'de gezici arıcılık faaliyetlerinin yürütülmesi arıların melez nesiller üretmesine sebep olmaktadır. Bu durum arı ırk ve ekotiplerinin tanımlanmasını güçleştirmektedir. Bal arısı alt türlerinin belirlenmesinde şimdiye kadar klasik morfometrik yöntem ve mtDNA çalışmaları yapılırken son yıllarda geometrik morfometrik yöntem popüler hale gelmiştir. Bu yöntem ile arıların belli özellikleri değer özellik kümesi olarak seçilir ve bu özelliklere göre sınıflandırma yapılır. Arı kanatları üzerinde damar kavşak noktaları arıların sınıflandırmasında sınıflandırma aracı olarak kullanılmaktadır.

4.1. LİTERATÜRDEKİ ÇALIŞMALAR

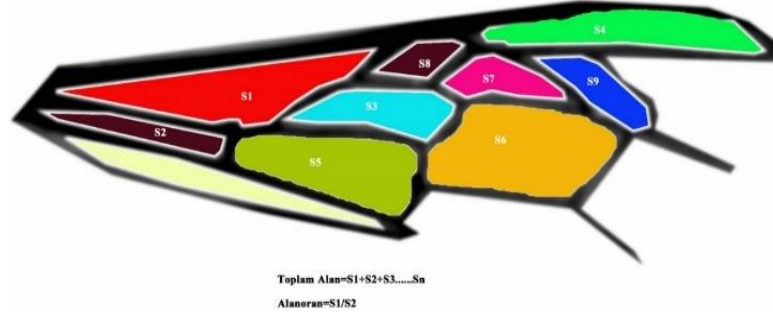
Bir çalışmada Yunanistan’da 11 bal arısı türünden alınan 450 örnek kanat üzerinden 19 damar kavşak noktası ile alt tür tespiti çalışması yapılmıştır [34]. Diğer bir çalışmada arı kanatları üzerindeki 19 kavşak noktası kullanılarak arılardaki parazit istilasının tespiti çalışması yapılmıştır [35]. Arı kanatları üzerindeki alan ve kavşak noktalarının kullanıldığı, sınıflandırma yöntemi olarak k-ortalama ve en yakın komşuluk algoritmaları kullanılarak sınıflandırma yapılabilmektedir [36]. Bal arısı alt türlerinin tespiti için kanat özelliklerinin incelemesi çalışmaları yapılmış, standart ve geometrik olmak üzere iki morfometrik analiz uygulanmıştır [37]. Başka bir çalışmada arı kanadı üzerindeki kavşak noktalarına ait koordinatlar alınmış, bu koordinatlar yardımıyla elde edilen düzlem bir noktanın referans koordinat seçilmesi için döndürülmüş, her bir koordinat noktası da özellik kabul edilerek 97 özellik çıkartılmış, KNN, DVM, NaiveBayes, Logistic, MLP, LDA, C4.5 sınıflandırma algoritmaları kullanılarak sınıflandırma yapılmıştır [38]. Arı kanatları görüntülerinin CCD kamera ve mikroskop yardımıyla alınarak %98 ile 99,8 arasında başarı ile tür tanımlanması yapılmıştır [39]. Bir diğer çalışmada arı kanadı üzerinde 16 uzunluk, 16 açı, 1 kübital indeks, uzunluk ve genişlik özellikleri kullanılarak 3 tür üzerinde sınıflandırma yapılmış ve ortalama %94 başarı sağlanmıştır [40].

4.2. GEOMETRİK MORFOMETRİ

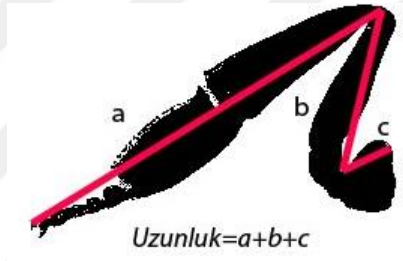
Morfometri, organizmaların sınıflandırılması amacıyla vücudun bazı kısımlarının veya organların morfolojik yapılarının geometri, bilgisayar grafikleri ve biyometrik tekniklerle kombinlenerek sayısal ölçümlerinin yapılması, bu ölçümlerin çok değişkenli istatistiksel metotlarla yorumlanmasını sağlayan bir ölçüm yöntemidir. Klasik ve geometrik olarak iki yönteme ayrılan morfometri, birçok canlı türünün sınıflandırılmasında, cinsiyet tayininde, türler veya ırklar arasındaki evrimsel ilişkinin belirlenmesinde geçmişten günümüze kullanılan bir yöntemdir [41].

Şekil analizi pek çok biyolojik araştırmada kullanılmaktadır. 1980’lerde verilerin toplanma biçimi ve analizinde temel bir değişim başlamıştır. Bu değişim işaret (Landmark) koordinatları ve onların karşılaştırmalı pozisyonları (Relative position) ile ilgili geometrik bilgiye odaklanmıştır. Bu yeni yaklaşım geometrik morfometri olarak adlandırılmıştır [42]. Geometrik morfometri görüntü işleme yöntemleri ile

kullanıldığında uzun zaman alan bazı ölçümler daha kısa zamanda yapılabilir. Arı kanatları üzerindeki kanat uzunluğu, genişliği, indeks ve açı gibi verilerin ölçümü için koordinatlar alınarak ölçümler yapılabilir. Şekil 4.2, Şekil 4.3, Şekil 4.4 ve Şekil 4.5'te bazı örnek geometrik özellikler görülmektedir.



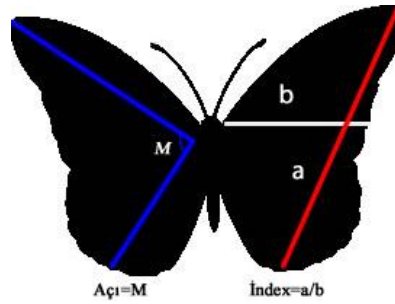
Şekil 4.2. Arı kanadı üzerinde alan.



Şekil 4.3. Arı bacağı üzerinde uzunluk ölçümü.



Şekil 4.4. Balık boyu ve kuyruk açısı ölçümü.

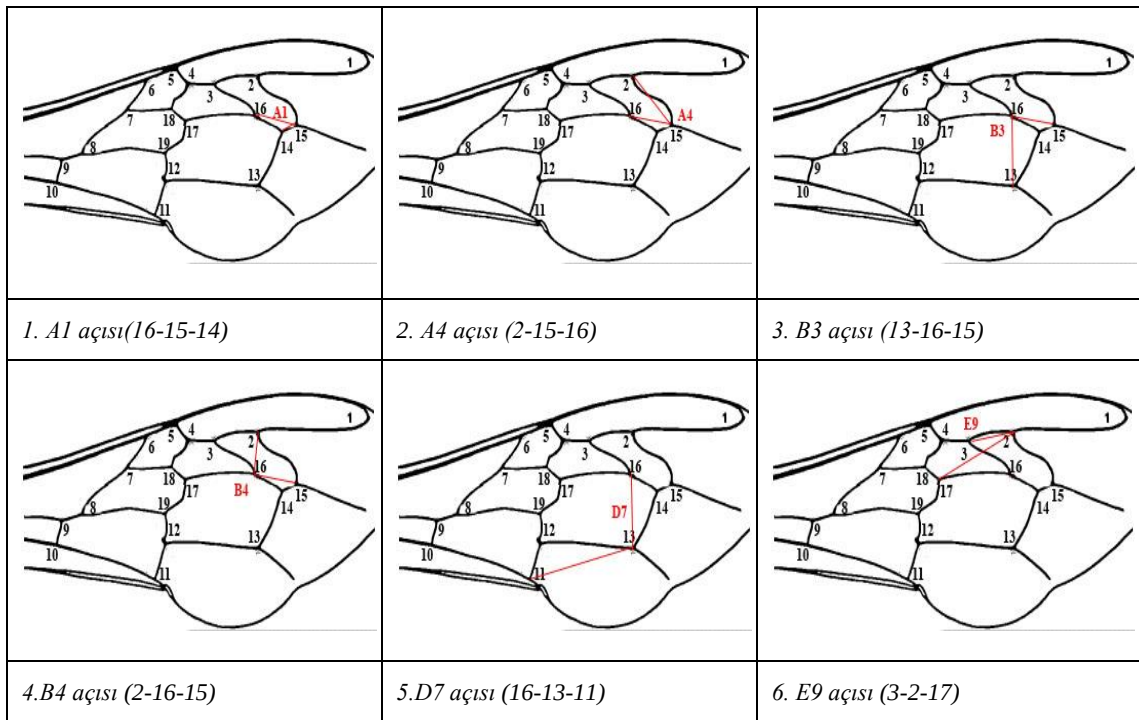


Şekil 4.5. Kelebek üzerinde açı ölçümü ve en-boy oranı.

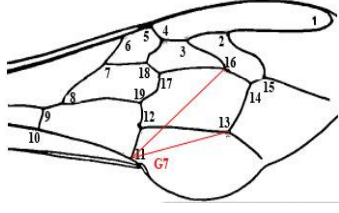
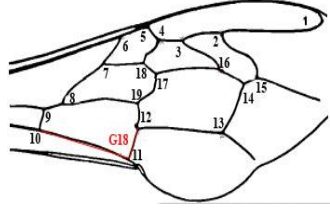
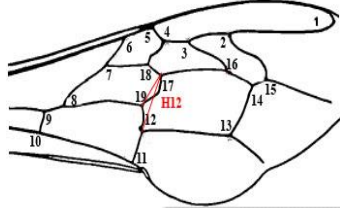
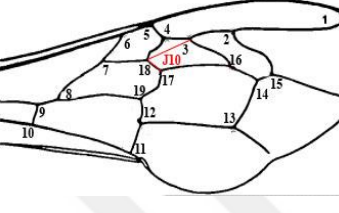
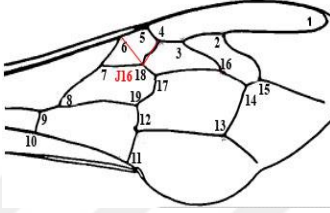
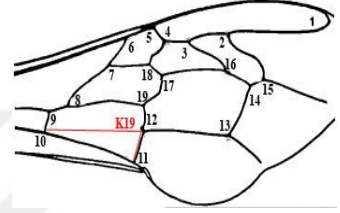
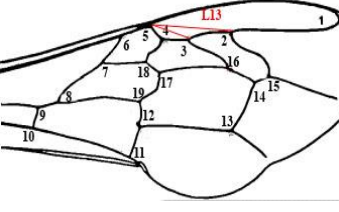
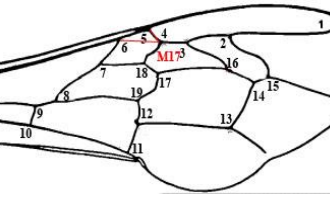
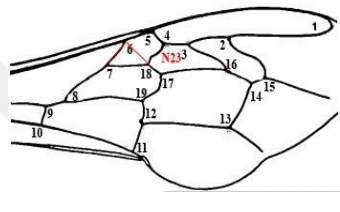
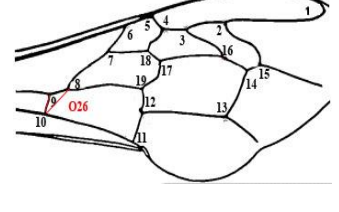
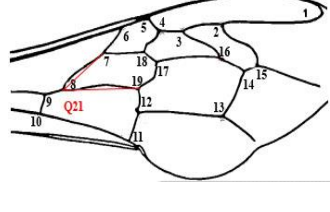
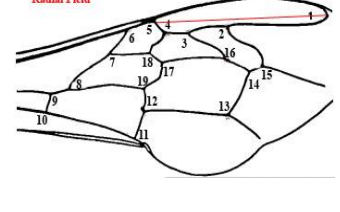
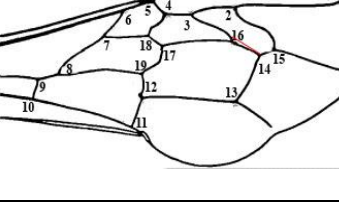
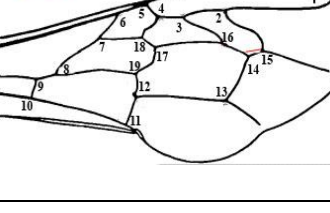
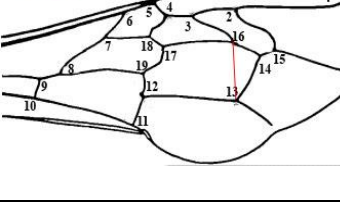
Bu ve benzeri özellikler geometrik morfometrik özelliklerdir ve bu özellikler türlerin sınıflandırılmasında, hastalık tespiti, cinsiyet belirle, yaş hesaplama, ürün kalitesi belirleme gibi işlemlerde veri olarak kullanılabilir. Bazı paket programlar ölçüm yapılacak cismi üç boyutlu uzayda incelemektedir. Bu tür programlarda X, Y, Z koordinatları alınmakta, açı, uzunluk, alan gibi bilgiler bu koordinatlar üzerinden hesaplanmaktadır.

4.3. ARI KANADI ÜZERİNDEN ÖZELLİK ALINMASI

Arı kanadı üzerinden özellik çıkarımı arı için tür, hastalık, cinsiyet gibi bilgilerin elde edilmesini sağlayabilmektedir. Kanat üzerinden özellik çıkarımında renk bilgisi kullanılabilir. Lakin örneklerin alınması esnasında kanat renkleri kullanılan kimyasalın miktarına göre değişeceği, kanat üzerinde renklerin her alanda aynı olmayacağı düşünüldüğünde renklere göre işlem yapmak sağlıklı sonuçlar vermeyebilir. Bunun yerine kanat üzerindeki kavşak noktaları kullanılabilir. Damar kavşakları üzerinde işaretli olan noktalar aslında bir resim üzerindeki koordinatlardır. Bu koordinat noktalarını sınıflandırmada tek başına kullanmak yerine, bu koordinatlardan çıkarılan morfolojik özellikler kullanılmaktadır. Şekil 4.6'daki çalışmada bu koordinat noktaları kullanılarak toplam 30 özellik çıkarılmıştır [43].



Şekil 4.6. Kanat özellikleri.

		
7. G7 açısı (16-11-13)	8. G18 açısı (10-11-12)	9. H12 açısı (19-17-12)
		
10. J10 açısı (3-18-17)	11. J16 açısı (6-18-4)	12. K19 açısı (11-12-10)
		
13. L13 açısı (2-4-3)	14. M17 açısı (5-4-6)	15. N23 açısı (7-6-18)
		
16. O26 açısı (9-10-8)	17. Q21 açısı (7-8-19)	18. Radyal Uzunluk (1-4)
		
19. Uzunluk A (16-14)	20. Uzunluk B (14-15)	21. Uzunluk C (16-13)

Şekil 4.6. (devam). Kanat özellikleri.

Length D	Inner wing length	Inner wing width
22. Uzunluk D (9-19)	23. İç kanat uzunluğu (10-14)	24. İç Kanat genişliği (11-5)
Discoidal shift	Cubital index A/B	Precubital index A/B
25. Discoidal Shift	26. Cubital indeks (16-14)/ (14-15)	27. Precubital Cubital indeks
Dumb-bell index A/B	Radial index A/B	Area
28.Dumb-bell indeks	29.Radial İndeks (A/B)	30. Alan

Şekil 4.6. (devam). Kanat özellikleri.

5. MATARYEL VE YÖNTEMLER

5.1. ARI KANADI ÖRNEKLERİNİN ALINMASI

Arı kanatları örneklerinin alınması, muhafazası ve saklanması Düzce Üniversitesi Biyoloji Bölümü tarafından yapılmış, görüntülerin alınması, Düzce Üniversitesi Fen Edebiyat Fakültesi Biyoloji Bölümü Moleküler Genetik Araştırma Laboratuvarında yürütülmüştür. Örnekler 2014-2016 tarihleri arasında alınmıştır. Kanat görüntüleri için kullanılacak arılar çeşitli il ve ilçelerden alındıktan sonra 2/3'ü %96 etil alkol olan bir çözeltiye konulmuş ve sintilasyon şişelerinde taşınmıştır. Şekil 5.1'de sintilasyon şişesi ve bu şişe üzerine örneğin kodlanması gösterilmiştir.



Şekil 5.1. Örneklerin sintilasyon şişelerinde kodlanması.

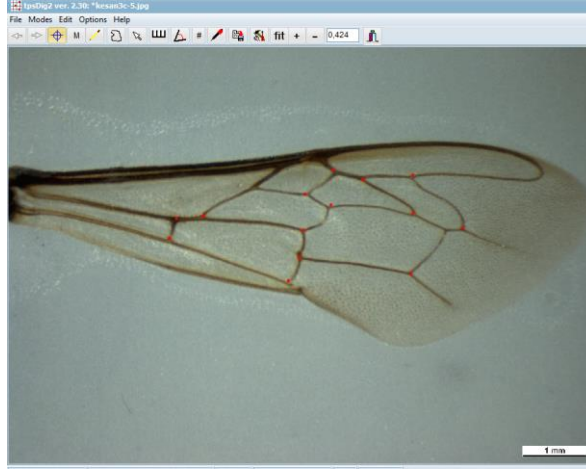
Bu örnekler laboratuvar ortamında +4 derece ısı altında muhafaza edilmiştir. İçerisinde distile su bulunan petri kabına pens yardımıyla dikkatli bir şekilde koparılan sağ ön kanat bırakılmıştır. Kanatlar pens yardımıyla tek tek alınarak 6x26 mm' lik lam üzerine dikkatlice yerleştirilmiş ve aydınlatma bandı ile sabitlenmiştir. Lamin üst kısmına kanadın nereye ait olduğunu belirtmek için kodlama yapılmıştır. Hazırlanan preparatların fotoğrafları BAB STR45 stereozoom mikroskobuna bağlı BAB kamera sistemiyle 1X büyütmede çekilmiştir. Kullanılan arı kanatlarının ait olduğu il ve ilçe ile bu özellikleri elde etmek için kullanılan resim sayısı Çizelge 5.1'de verilmiştir.

Çizelge 5.1. Örneklerin alındığı il/ilçeler.

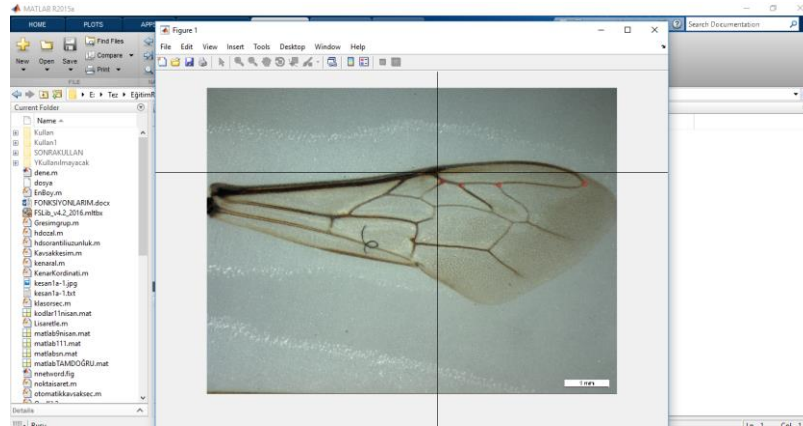
Özellik Alındığı il /ilçe	Örnek Sayısı
1.Bozcaada	30
2.Gökçeada	30
3.Kırklareli	30
4.Saray	30
5.Muğla	30

5.2. ARI KANADI ÜZERİNDE KAVŞAK BELİRLEME

Arı kanatları üzerinde kavşak noktası işaretleme için değişik yazılımlar kullanmak mümkündür. Örneğin Matlab© veya tpsDig232 gibi programlar kavşak noktalarının belirlenmesinde kullanılabilir. Şekil 5.2’de arı kanadı üzerinde tpsDig232 programı ile, Şekil 5.3’te Matlab Programı ile kavşak noktası seçimi gösterilmiştir.



Şekil 5.2. TpsDig32 programı ile kavşak noktası belirleme.



Şekil 5.3. Matlab programı ile kavşak noktası belirleme.

Bu çalışmada örnekler üzerinde kavşak noktası işaretleme için Matlab Programı seçilmiştir. Çünkü resim üzerinde, kavşak işaretleme yapan diğer yazılımların geliştirilebilir olmadığı görüldü. Matlab programının tercih edilmesinin bir diğer nedeni, içinde kavşak işaretleme yapılabilen yazılımın hazırlanmasına olanak tanınmasının yanı sıra, özelliklerin koordinatlara göre çıkarılmasını, görüntü işleme tekniklerini

kullanabileceğimiz birçok bileşeni sunması ve sınıflandırma işlemleri için uygulamalar sunması olmuştur. Ayrıca Matlab, kendi içinde birçok görüntü işleme bileşenini sunmaktadır.

5.3. ÖNERİLEN KAVŞAK SEÇİM ALGORİTMASI (ÖKSA)

Kavşak noktalarının belirlenmesi aşamasında bazı sıkıntılarla karşılaşmıştır. Örneğin bir kişinin aynı kanat görüntüsü üzerinde farklı zamanlarda aldığı koordinatlar farklılık göstermekte, işaretleme elle yapıldığından hata payının yüksek olduğu ve özelliklerin de farklı çıkmasına neden olduğu görülmüştür. Çizelge 5.2’de Saray ilçesinden alınan örnek bir kanat üzerinde farklı zamanlarda yapılan 2 nokta ve 3 farklı işaretlemeye ait koordinatlar verilmiştir.

Çizelge 5.2. Farklı zamanlarda alınan koordinatlarının değişimi.

Nokta	x	y	x	y	x	y	Std.x	Std.y
1	1472	986	1463	984	1469	990	4,242641	4,242641
2	1264	984	1266	982	1264	995	1,414214	9,192388
	1. İşaretleme		2. İşaretleme		3. İşaretleme		Standart Sapma	

Çizelge 5.2’de görüldüğü gibi işaretleme el ile yapıldığında hatanın yüksek çıkma ihtimali vardır. Bu durum alınan özelliklerin de farklı olmasına neden olacak ve sınıflandırma başarısı düşecektir. Bu nedenle aşağıda verilen algoritma ve Matlab’da kavşak işaretleme için program hazırlanmıştır.

Algoritma 1:

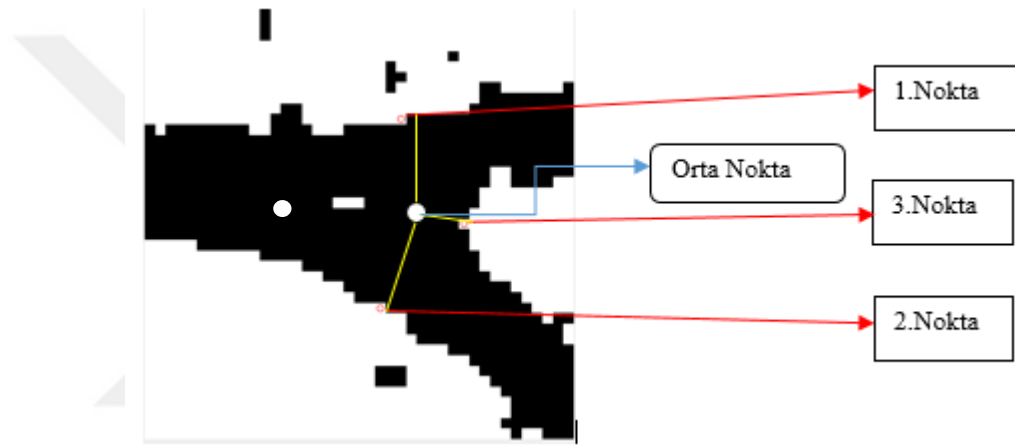
```

1: K=Resmin tıklanan nokta merkez olmak üzere 20X20 piksel boyutunda
alan
2: K=siyahbeyaz(K)
3: A=K resmindeki 3 adet beyaz alan
4: kmesafe=20x20
5: for t=1:3
5.1: Beyazpikselsayisi(t)=size(A(t))
6: 5.dongu sonu
7: for a=1:Beyazpikselsayisi(1)
7.1: for t=1:Beyazpikselsayisi(2)
7.1.1: Mesafe1=Karekök(A(1,a,1)* A(1,a,1) - A(1,a,2)*A(1,a,2))
7.1.2: for c=1:Beyazpikselsayisi(3)
7.1.2.1: Mesafe2=sqrt(A(1,a,1)-A(3,c,1))^2+(A(1,a,2)-A(3,c,2))
7.1.2.2: Mesafe3=sqrt(A(3,c,1)-A(2,t,1))^2+(A(3,c,2)-A(2,t,2))
7.1.2.3: mesafe=Mesafe1+Mesafe2+Mesafe3
7.1.2.4: if(mesafe<kmesafe)
7.1.2.4.1: kmesafe=mesafe;
7.1.2.4.2: yakinnokta=[A(1, :, :) A(2, :, :) A(3, :, :)]
7.1.3: 7.1.2. dongu sonu

```

7.2: 7.1dongu sonu
8: 7.dongu sonu

Bu algoritma kavşak işaretlemeye resmin tıklanan noktasından itibaren 20x20 gibi küçük bir parçasını alır, siyah beyaz yapar, damarların ayırdığı alanları bulur ve bu alanlardaki beyaz noktalar arasında birbirine en yakın beyaz noktaları bulur. Bu beyaz noktaların x ve y koordinatları ortalaması bize kavşak noktasını vermektedir. Burada özellikle damarlarla üç alana ayrılabilen 20x20 resim parçaları bu işleme dâhil edilmekte, üç alana ayrılamayan resim parçalarında kavşak noktaların elle seçim yapılması istenmektedir. Şekil 5.4'te resimden alınan parça ve alanlar üzerinde birbirine en yakın 3 beyaz nokta ve orta noktaları gösterilmiştir.



Şekil 5.4. Kanat üzerinde standart kavşak belirleme işlemi.

Önerilen kavşak seçim algoritması, 20*20 pixel alan tıklanan noktadan itibaren alınmış ve kavşak noktasını kullanıcı farklı bir noktaya da tıklasa orta noktanın en az hatayla bulunması sağlanmıştır. Çizelge 5.3'te Saray ilçesinden alınan örnek bir kanat üzerinde farklı zamanlarda yapılan 2 farklı noktadaki işaretlemeye ait koordinatlar geliştirilen algoritma ile test edilmiş ve sonuçlar aşağıda verilmiştir. Bu algoritma ile işaretleme kavşak kesim noktasına otomatik tutunma etkisi yaratmaktadır. Bu algoritma işaretlemeye oluşabilecek kaymaları önlemesine karşın, döngü sayısının fazla olması algoritmanın yavaş çalışmasına neden olmaktadır.

Çizelge 5.3. ÖKSA ile farklı zamanlarda işaretlenen koordinatlar.

Nokta	X	y	x	y	x	y	Std.x	Std.y
1	1473	568	1473	568	1473	568	0	0
2	1271	568	1271	568	1271	568	0	0
	1. İşaretleme		2. İşaretleme		3. İşaretleme		Standart Sapma	

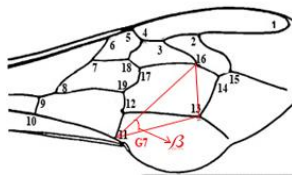
5.4. ÖZELLİK ÇIKARMA

Bu çalışmada, Şekil 4.6’da belirtilen özellikler kullanılarak sınıflandırma yapılmıştır. Bu özelliklerden Çizelge 5.4’te verilenler kullanılmıştır. Bu özelliklerin alan, uzunluk ve açı gibi bilgiler olmasından dolayı Denklem (5.5)’te verilen normalizasyon formülü ile veriler normalleştirilmiştir. Ayrıca Çizelge 5.1’de örneklerin ait olduğu yerler verilmiştir.

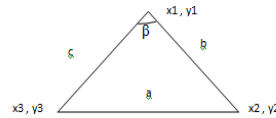
Çizelge 5.4. Kullanılan özellikler.

Özellik Tanım	AÇI																
	1.	2.	3	4.	5.	6.	7.	8.	9.	10.	11.	12.	13.	14.	15.	16.	17.
Özellik Adı	A1	A4	B3	B4	D7	E9	G7	G18	H12	J10	J16	K19	L13	M17	N23	Q26	Q21
Özellik Tanım	Uzunluk																
	18.		19.			20.		21.		22.			23.				
Özellik Adı	RadialFaıld		B			C		D		İnnerWingthLength			İnnerWingthWidth				
Özellik Tanım	İndeks															Alan	
	24.		25.					26.			27.						
Özellik Adı	Cubital Indeks		Precubital Indeks					Dumbbell indeks			Alan						

Arı kanadı üzerindeki damar kavşak noktaları belirledikten sonra üç koordinat noktası bilinen bir üçgenin bir köşesindeki açı hesaplaması için Şekil 5.5’teki gibi kosinüs teoreminden yararlanılmıştır. Burada amaç belirtilen β açısı Denklem (5.1)’deki formülle hesaplanmıştır.



a) β açısı.



b) β açısının üçgen üzerinde gösterimi.

Şekil 5.5. Kavşak noktalarından ile açı hesaplama.

$$\beta = \arccos((b^2 + c^2 - a^2)/2ab) \quad (5.1)$$

$$a = \sqrt{(x_3 - x_2)^2 + (y_3 - y_2)^2} \quad (5.2)$$

$$b = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (5.3)$$

$$c = \sqrt{(x_3 - x_1)^2 + (y_3 - y_1)^2} \quad (5.4)$$

Özellik çıkarmada iki noktasına ait koordinatları verilen bir doğrunun uzunluğu Denklem (5.2)'de öklid denkleminde elde edilir.

Tüm özellikler çıkarıldıktan sonra özelliklerin normalleştirilmesi ve 0 ile 1 arasında değer alması verilerin işlenmesini kolaylaştırır. Bu nedenle Denklem (5.5)'te verilen normalizasyon formülü kullanılmıştır.

$$X_{inorm} = (X_i - X_{min}) / (X_{max} - X_{min}) \quad (5.5)$$

Denklem (5.5)'te X_{inorm} , i.özellik için normalizasyon sonucunu vermektedir. X_{max} özellik kümesindeki en büyük değere sahip veri, X_{min} özellik kümesindeki düşük değere sahip veriyi temsil etmektedir.

Özellik çıkarımından sonra bu özellikler Matlab programında DVM, K-Means, YSA ve KNN kullanılarak sınıflandırılmıştır.

Sınıflandırmada doğruluk oranları ölçmek için karışıklık matrisi (confusion matrix) değerleri incelenmiştir. Çizelge 5.5'te karışıklık matrisi gösterilmektedir. Matris içinde Doğru Pozitif (DP), Doğru Negatif (DN), Yanlış Pozitif (YP), Yanlış Negatif (YN) değerleri Çizelge 5.5'teki gibi hesaplanmıştır [44]. Burada DP değeri sınıflandırma neticesinde doğru sınıflandırılmış verinin, toplam doğru sınıflandırmaya oranıdır. YP değeri, sınıflandırma neticesinde yanlış sınıflandırılmış verinin, toplam yanlış sınıflandırmaya oranıdır. DN değeri ise yanlış sınıflandırılmamış verinin toplam yanlış sınıflandırılmamış veriye oranıdır. YN değeri ise yanlış negatif sınıflandırılmış verilerin toplam negatif sınıflandırmaya oranıdır.

Çizelge 5.5. Karışıklık matrisi (Confusion matrix).

GERÇEK DURUM	POZİTİF DURUM	Doğru Pozitif	Yanlış Negatif (2.Tip Hata)	$DP\ oranı = \frac{\sum DP}{\sum Pozitif\ durum}$	$YN\ Oranı = \frac{\sum YN}{\sum Pozitif\ durum}$
	NEGATİF DURUM	Yanlış Pozitif (1.Tip Hata)	Doğru Negatif	$YP\ Oranı = \frac{\sum YP}{\sum Negatif\ durum}$	$DN\ Oranı = \frac{\sum DN}{\sum Negatif\ durum}$

6. BULGULAR VE TARTIŞMALAR

6.1. KANAT İŞARETLEME İLE ALINAN KOORDİNATLAR

Kavşak noktalarının işaretlenmesi için önce standart kavşak noktası işaretleme programı MATLAB’da geliştirildi. Bu program ile her bir kanat resmi üzerindeki kavşak noktaları işaretlendi. İşaretleme sonucu oluşan 19 koordinat noktası bir .txt uzantılı dosyaya kaydedildi. Çizelge 6.1’de bu işaretlenme sonucunda Muğla ilinden alınan bir kanat resmi üzerinde koordinatlar görülmektedir.

Çizelge 6.1. ÖKSA kullanılmadan alınan örnek koordinatlar.

Nokta	X koordinatı	Y koordinatı	Nokta	X koordinatı	Y koordinatı
1.Nokta	1916	592	11.Nokta	1030	973
2.Nokta	1460	599	12.Nokta	1078	897
3.Nokta	1287	598	13.Nokta	1435	922
4.Nokta	1214	587	14.Nokta	1570	798
5.Nokta	1172	548	15.Nokta	1623	821
6.Nokta	1014	566	16.Nokta	1451	730
7.Nokta	908	651	17.Nokta	1161	711
8.Nokta	736	750	18.Nokta	1099	663
9.Nokta	656	766	19.Nokta	1071	796
10.Nokta	628	833			

Bu kavşak noktaları, işaretleme yapılırken kişiden kişiye farklılık göstermektedir. Çünkü tıklandığımız nokta göz kararı seçilmektedir. Bu durumu standartlaştırmak için geliştirilen algoritma ile Çizelge 6.1’deki bilgilerin ait olduğu örnekten koordinatlar alınmış ve bu koordinat sonucu da Çizelge 6.2’de gösterilmiştir. Bu algoritma ile tıklanan

noktadan 20x20'lik bir parça alınmış, alınan parça siyah-beyaz yapılmış, resmin beyaz alanlarına komşu siyah noktalar bulunmuş, bu siyah noktalardan birbirine en yakın olanların ortalama koordinatı bize orta noktayı vermiştir.

Çizelge 6.2. ÖKSA ile alınan koordinatlar.

Nokta	X koordinatı	Y koordinatı	Nokta	X koordinatı	Y koordinatı
1.Nokta	1922.00	592	11.Nokta	1043.00	988.000
2.Nokta	1469.00	604	12.Nokta	1070.00	894.0000
3.Nokta	1298.00	610	13.Nokta	1442.00	925.0000
4.Nokta	1214.00	595	14.Nokta	1577.00	795.0000
5.Nokta	1190.00	538	15.Nokta	1628.00	781.0000
6.Nokta	1010.00	577	16.Nokta	1448.00	736.0000
7.Nokta	914.00	646	17.Nokta	1169.00	721.0000
8.Nokta	734.00	745	18.Nokta	1109.00	667.0000
9.Nokta	653.00	769	19.Nokta	1091.00	787.0000
10.Nokta	641.00	826			

6.2. NORMALİZASYON İŞLEMİ UYGULANMIŞ VERİLER

Çizelge 5.4'te belirtilen açı, uzunluk ve uzunluk oranları değerlerini elde etmek için Matlab'da program geliştirilmiş, 150 örnek için 27 özellik çıkarılmıştır. Bu özelliklerin normalizasyon işleminden önceki değerleri Çizelge 6.3'te, normalleştirme işleminden sonraki değerleri ise Çizelge 6.4'te verilmiştir.

Çizelge 6.3. Normalizasyon uygulanmamış özellik örnekleri.

	A1	A4	B3	B4	D7	E9	G7	G18	H12	J10	J16	K19	J13	M17	N23	Q26	Q21	RF	B	C	D	IWL	IWW	CI	PI	DI	Alan	Sınıf
	21,60	29,54	83,37	115,27	106,79	20,34	22,82	100,73	16,51	50,66	93,48	78,57	14,21	56,80	91,89	37,79	36,30	734,97	61,07	199,47	441,31	990,85	447,81	1,65	1,51	1,01	292854,00	1
	20,62	30,92	76,24	104,76	97,26	19,38	25,02	98,43	15,77	51,08	97,04	78,44	15,01	51,16	94,09	30,75	36,42	752,65	58,08	205,48	444,38	1019,24	461,38	2,33	1,49	1,19	311398,00	1
	18,00	33,09	82,66	101,87	105,51	19,11	22,51	93,09	14,19	50,80	93,39	76,19	15,45	56,98	90,04	36,54	35,56	729,66	66,12	191,51	432,20	976,88	449,05	1,61	1,56	0,90	287700,00	1
	20,75	35,09	78,99	97,06	102,37	16,31	22,18	92,09	16,79	52,04	98,77	76,85	15,74	52,34	94,73	44,05	35,84	745,58	56,08	189,28	439,03	995,55	444,00	2,09	1,58	0,88	291202,00	1
	22,43	35,16	89,88	103,57	113,63	18,72	21,52	93,16	14,74	50,59	92,72	77,94	17,84	64,11	89,31	40,86	37,07	726,57	59,30	197,12	443,31	992,42	456,36	1,65	1,50	0,82	293652,00	1
	24,59	30,38	80,73	106,46	100,76	19,43	24,07	96,44	15,32	55,56	95,46	76,19	18,17	57,76	94,56	35,34	35,91	736,09	56,57	207,63	452,12	1029,19	460,00	2,37	1,48	1,06	313878,00	1
	27,47	30,26	74,45	96,20	97,90	19,34	25,10	99,29	15,73	50,58	97,37	78,75	16,23	55,23	89,01	33,88	35,72	728,82	55,58	204,41	426,08	985,25	445,86	2,60	1,42	1,10	285767,00	1
	24,27	31,97	75,63	100,30	98,75	17,89	24,64	100,42	14,31	54,64	92,37	80,82	17,19	54,80	86,77	29,09	35,69	717,35	50,36	206,29	423,37	991,16	458,70	2,70	1,41	1,04	294736,00	1
	21,74	31,91	74,90	92,20	98,41	18,25	23,92	97,07	15,73	51,60	94,35	78,69	14,47	53,53	91,45	29,29	37,53	754,86	65,19	205,16	429,19	1024,90	460,79	2,14	1,46	1,05	305422,00	1
	18,31	35,06	86,51	103,09	105,88	18,52	22,20	103,00	17,05	58,52	103,70	77,50	14,55	56,72	96,45	44,19	36,89	747,49	74,03	200,35	439,01	1002,22	439,93	1,29	1,61	0,83	295773,00	1
	23,83	27,87	76,61	113,46	99,19	20,91	23,71	99,87	17,76	55,39	96,89	77,42	14,92	43,88	98,53	42,71	37,28	759,22	51,24	201,72	458,66	1015,64	437,03	2,55	1,51	1,02	304812,00	1
	21,51	34,10	75,97	94,91	97,50	18,35	24,37	94,54	12,31	53,10	99,46	68,47	16,10	60,76	100,03	39,36	37,71	737,91	53,15	199,30	443,76	986,17	441,82	2,41	1,50	0,96	288761,00	1
B	19,84	31,53	85,82	112,14	106,97	20,27	21,78	98,95	18,54	51,98	97,39	81,10	17,80	59,86	96,22	32,36	39,64	756,29	66,37	203,86	446,29	1043,31	467,46	1,56	1,63	0,86	323981,00	1
C	26,66	32,85	78,08	99,98	100,15	18,52	23,50	93,97	15,63	51,37	97,63	77,31	17,50	57,98	93,04	32,96	36,17	754,39	54,59	200,21	434,66	1017,24	441,88	2,43	1,58	0,86	304342,00	1
C	23,25	32,48	83,42	103,84	102,17	19,93	23,16	96,07	18,50	52,98	101,97	82,04	16,58	59,80	95,49	42,99	38,93	755,97	62,51	203,69	451,14	1034,02	459,05	1,98	1,61	0,93	318985,00	1
A	22,18	33,73	78,72	95,07	101,18	17,82	24,83	96,61	18,50	56,84	93,47	83,52	16,15	58,98	89,58	38,01	34,99	705,47	68,07	199,12	433,89	990,00	434,53	1,74	1,47	0,99	287346,00	1
A	22,63	34,64	78,83	94,62	100,54	17,38	24,63	93,96	18,02	50,49	99,40	80,72	17,42	52,09	91,09	38,72	36,53	737,62	56,14	202,67	429,87	987,29	433,94	2,14	1,56	0,84	290602,00	1
D	22,98	26,56	79,67	112,64	105,54	19,62	23,44	91,20	18,44	52,15	99,82	77,82	14,42	53,10	94,56	34,78	35,35	743,96	65,07	207,15	438,65	1016,33	450,63	1,91	1,59	1,05	303468,00	1
A	22,30	34,22	85,60	96,89	105,46	18,97	23,10	100,08	16,47	54,90	99,80	77,78	14,87	49,41	94,99	40,35	37,48	754,86	75,80	204,84	444,50	1018,55	444,63	1,42	1,55	0,87	302834,00	1
	25,26	32,06	79,22	94,08	101,90	18,21	24,70	96,72	18,05	54,93	97,27	78,48	16,31	55,53	95,15	35,24	38,50	721,18	60,53	208,50	440,73	1013,44	452,96	2,24	1,41	1,06	300381,00	1
	23,12	31,37	85,56	116,13	105,77	18,88	22,03	93,54	12,87	49,90	101,55	73,79	17,67	55,76	95,35	35,84	37,62	751,85	54,23	194,28	447,69	1009,04	452,13	1,89	1,62	0,88	306318,00	1
	25,55	30,08	76,15	110,75	99,15	21,22	22,59	94,47	19,03	50,37	99,76	75,50	15,40	54,60	92,23	33,52	35,33	767,61	44,28	197,82	430,51	1011,26	461,29	3,00	1,53	1,03	303115,00	1
	22,31	31,91	76,50	101,42	96,85	18,50	23,86	94,99	15,63	51,39	100,78	73,66	17,22	52,60	91,17	33,60	35,51	737,14	59,14	197,73	431,05	999,00	439,55	2,16	1,55	1,02	289760,00	1
	21,29	31,14	85,16	111,56	104,31	19,90	22,58	89,51	16,90	47,53	104,32	75,02	16,21	53,74	98,86	36,53	36,77	757,03	53,15	190,25	437,12	994,00	429,50	2,13	1,68	0,94	291444,00	1
	23,36	28,22	79,14	112,97	100,68	18,51	23,42	95,24	13,92	54,88	103,69	72,05	14,35	51,85	98,12	37,33	36,56	749,39	59,14	200,26	436,33	995,00	442,91	2,06	1,59	1,10	291036,00	1
	22,29	33,21	82,57	99,67	104,18	15,93	23,08	102,00	17,14	56,43	100,53	81,11	15,40	59,68	96,03	36,42	38,45	736,13	61,29	205,66	435,11	1019,23	459,43	1,96	1,47	1,00	303834,00	1
	21,20	30,90	72,97	93,66	96,08	19,74	25,93	100,80	18,24	55,08	92,45	79,63	15,20	53,23	91,72	40,25	35,33	747,77	59,01	215,33	453,05	1026,02	454,95	2,53	1,42	1,08	309882,00	1
	22,12	32,51	82,55	105,22	104,11	18,58	23,44	102,72	14,73	58,53	92,10	78,47	18,04	55,41	94,08	38,93	37,66	726,89	64,12	202,79	430,43	992,22	439,39	1,81	1,49	0,89	291629,00	1
	24,86	29,38	78,76	108,20	100,10	20,60	25,13	96,16	15,58	54,01	97,58	77,81	16,84	55,92	93,33	42,14	37,84	760,71	55,33	209,75	445,54	1009,31	448,43	2,41	1,51	1,02	305036,00	1
	23,36	32,98	80,75	106,38	104,60	19,57	24,02	94,31	17,29	56,64	95,79	77,74	15,14	51,81	91,67	40,53	36,85	739,65	50,25	212,28	445,29	1002,18	458,23	2,33	1,54	0,88	303570,00	1

	A1	A4	B3	B4	D7	E9	G7	G18	H12	J10	J16	K19	J13	M17	N23	Q26	Q21	RF	B	C	D	IWL	IWW	CI	PI	DI	Alan	Sınıf
G Ö K Ç E A D A	22,15	38,33	87,24	96,47	107,77	17,08	21,81	100,73	18,73	53,02	96,04	79,83	14,96	52,67	97,47	31,57	33,12	751,60	52,47	192,00	442,55	1007,50	434,89	2,08	1,65	0,75	302698,00	2
	19,96	35,80	89,67	105,10	113,52	17,35	20,65	95,05	20,31	51,77	100,98	85,49	13,53	65,44	96,94	36,60	36,09	744,68	57,14	193,13	431,56	1000,18	445,10	1,64	1,76	0,77	297857,00	2
	21,35	39,56	86,22	94,69	105,82	17,75	22,81	99,73	17,95	51,20	97,96	82,41	15,49	58,48	94,06	37,14	35,80	735,53	56,22	203,81	444,17	1011,31	452,25	1,90	1,57	0,76	309808,00	2
	25,98	38,01	83,00	86,94	104,29	16,70	22,95	98,41	20,04	51,39	100,15	84,52	16,19	51,44	94,56	34,75	33,67	760,54	50,64	197,94	452,24	1036,21	445,98	2,61	1,57	0,82	313827,00	2
	21,64	36,50	84,08	97,15	105,08	17,03	22,69	98,76	18,75	50,62	102,24	85,47	16,23	62,30	94,73	39,26	36,69	734,14	50,36	194,33	430,98	997,19	445,28	2,37	1,61	0,83	299385,00	2
	27,58	30,73	81,94	111,60	103,67	18,20	22,24	93,36	15,81	60,63	100,30	77,30	13,97	65,75	103,62	44,15	37,49	747,83	52,77	195,80	435,44	995,13	446,64	2,21	1,56	0,98	290515,00	2
	21,91	32,33	79,60	102,50	103,30	18,35	22,55	95,43	18,71	55,06	96,01	81,77	15,85	61,41	96,63	40,86	36,05	753,72	57,14	191,64	442,95	1002,58	443,36	2,07	1,54	0,89	298087,00	2
	25,97	33,77	76,06	95,67	98,04	18,25	22,90	100,80	16,03	53,76	104,13	82,26	11,50	55,43	98,92	34,95	39,69	723,94	49,65	187,22	409,41	966,29	421,89	2,50	1,57	0,89	275046,00	2
	24,58	35,38	80,26	100,70	101,47	19,22	22,99	94,86	18,66	55,40	98,98	79,41	15,02	59,47	101,31	33,66	36,09	730,31	49,50	188,29	426,73	972,78	429,12	2,27	1,56	0,84	282284,00	2
	23,34	35,23	77,89	87,67	101,49	16,03	24,10	95,10	15,34	59,34	93,64	79,90	12,25	43,67	96,15	38,55	37,92	763,86	61,07	205,20	440,20	1025,02	448,63	2,25	1,39	0,98	308394,00	2
	23,34	35,06	80,76	94,63	103,25	17,73	24,15	97,42	19,28	50,03	94,34	82,11	13,50	55,68	93,19	35,46	38,67	748,10	61,29	210,30	432,95	1014,47	468,33	2,02	1,51	0,95	312703,00	2
	19,79	36,45	83,53	92,10	103,38	17,43	23,10	101,34	18,08	55,59	97,26	79,54	12,34	50,74	97,43	37,27	39,07	759,71	64,28	201,95	425,30	999,45	449,70	1,77	1,53	0,93	296230,00	2
	20,99	34,58	83,05	92,85	105,18	18,49	23,78	90,53	21,33	48,54	100,82	84,01	13,50	52,02	97,86	33,97	35,66	761,65	63,13	199,28	423,80	988,32	436,79	1,76	1,55	0,88	289279,00	2
	27,87	32,03	86,16	104,37	106,90	20,38	23,69	94,76	19,53	49,38	93,34	81,74	12,62	47,50	91,86	36,26	38,21	753,40	63,13	199,28	423,80	988,32	436,79	1,76	1,55	0,88	304790,00	2
	21,01	32,31	80,91	97,69	101,82	18,71	25,14	98,66	19,43	50,05	99,25	81,65	15,50	54,62	95,07	38,21	39,92	745,58	65,07	215,24	439,65	1002,14	462,40	1,97	1,58	1,00	312090,00	2
	24,55	32,38	80,53	100,92	102,15	18,24	23,70	97,17	19,31	58,49	100,73	83,54	15,40	54,29	97,62	33,30	38,67	739,67	55,33	201,40	431,86	1006,10	441,48	1,87	1,48	0,93	295591,00	2
	24,37	33,26	77,94	104,23	102,33	17,52	23,40	91,05	15,96	61,80	97,26	82,54	15,18	55,94	93,81	37,02	37,54	732,45	50,25	191,27	415,48	955,50	448,09	2,30	1,41	0,98	279985,00	2
	28,39	32,51	74,83	94,71	99,21	19,26	23,47	100,36	19,89	54,23	101,23	79,50	16,02	71,43	99,42	37,28	36,16	767,72	45,54	196,57	426,36	986,52	443,50	3,09	1,55	0,90	289799,00	2
	24,13	33,20	77,30	92,06	100,72	15,58	23,51	100,14	16,52	55,09	102,51	81,66	12,52	50,58	96,18	30,92	34,79	731,74	53,00	196,90	431,92	1010,39	435,22	2,49	1,47	1,01	291750,00	2
	23,37	33,51	75,20	96,41	99,48	16,45	23,64	90,71	13,73	65,67	97,42	76,07	16,54	52,20	98,39	38,57	37,46	770,75	54,33	205,16	458,53	1045,41	458,42	2,44	1,45	0,90	319774,00	2
24,05	31,59	80,00	105,12	102,82	19,18	23,84	94,92	13,67	57,90	93,34	78,38	13,55	53,73	92,70	37,67	38,49	751,80	54,33	208,89	446,17	1019,75	460,75	2,33	1,49	0,97	312516,00	2	
21,79	32,69	80,15	103,78	102,24	18,95	23,83	96,69	18,02	58,32	95,12	82,95	14,64	56,15	95,40	34,00	37,35	744,55	63,20	201,68	426,13	990,77	445,28	1,78	1,52	0,92	294845,00	2	
20,44	32,79	83,01	105,65	106,23	18,89	21,41	98,48	15,97	51,66	92,41	77,86	12,18	55,05	91,19	34,46	39,18	741,49	59,14	186,59	420,00	974,79	435,91	1,71	1,58	0,92	279387,00	2	
22,63	31,54	73,57	91,46	96,96	17,71	23,94	95,37	19,05	51,95	97,51	82,59	14,15	46,90	94,11	38,24	37,83	760,23	61,29	196,01	437,25	1013,91	440,52	2,23	1,50	1,04	296056,00	2	
22,82	34,65	76,69	91,30	99,90	17,64	23,99	100,24	17,88	55,08	94,52	83,18	13,69	52,22	92,13	37,43	38,28	754,94	56,14	200,90	434,23	1000,19	451,03	2,33	1,47	0,98	298170,00	2	
33,89	32,69	77,30	99,58	99,66	18,46	24,30	96,51	15,70	56,27	97,25	80,05	12,60	53,71	93,89	34,64	36,03	747,84	46,15	198,98	427,13	977,13	440,10	3,05	1,43	1,08	284788,00	2	
17,29	39,98	80,86	81,63	101,98	15,41	22,69	96,03	17,47	54,87	94,07	85,08	15,24	53,33	93,76	30,74	34,06	790,77	70,18	197,50	455,58	1066,43	450,02	1,69	1,55	0,76	329737,00	2	
22,30	36,09	79,49	89,20	102,58	18,69	24,19	94,54	22,96	52,81	96,06	88,46	16,93	49,21	93,36	34,22	36,14	794,74	64,28	205,29	450,52	1041,77	452,75	1,94	1,53	0,79	325859,00	2	
19,60	34,67	81,66	100,89	104,06	19,18	21,57	102,79	17,07	54,99	97,94	75,57	16,38	55,07	96,38	34,28	37,50	742,89	58,14	188,53	436,72	997,93	451,31	1,89	1,48	0,91	291405,00	2	
21,30	35,65	77,35	93,04	98,47	18,21	23,79	90,96	18,33	56,84	96,38	80,72	17,27	55,90	96,72	37,38	37,65	755,16	60,21	191,44	426,88	985,34	433,24	1,96	1,52	0,87	286189,00	2	

Çizelge 6.4. İki ilçeye ait normalizasyon uygulanmış özellikler.

	A1	A4	B3	B4	D7	E9	G7	G18	H12	J10	J16	K19	J13	M17	N23	Q26	Q21	RF	B	C	D	IWL	IWW	CI	PI	DI	Alan	Sınıf
	0,25	0,26	0,70	0,86	0,70	0,55	0,40	0,88	0,46	0,41	0,21	0,51	0,39	0,43	0,35	0,52	0,41	0,33	0,57	0,59	0,67	0,43	0,61	0,19	0,39	0,60	0,45	1
	0,19	0,35	0,37	0,59	0,29	0,44	0,80	0,78	0,40	0,43	0,40	0,50	0,51	0,24	0,47	0,17	0,43	0,53	0,50	0,74	0,73	0,64	0,84	0,53	0,34	1,00	0,73	1
	0,06	0,50	0,66	0,52	0,65	0,41	0,34	0,53	0,26	0,42	0,21	0,39	0,57	0,43	0,24	0,46	0,32	0,27	0,71	0,38	0,49	0,32	0,63	0,16	0,50	0,35	0,38	1
	0,20	0,63	0,49	0,40	0,51	0,10	0,28	0,49	0,48	0,46	0,49	0,42	0,61	0,28	0,51	0,83	0,35	0,45	0,44	0,33	0,63	0,46	0,54	0,41	0,56	0,31	0,43	1
	0,29	0,64	1,00	0,56	1,00	0,37	0,16	0,54	0,31	0,41	0,17	0,47	0,91	0,67	0,20	0,67	0,51	0,24	0,53	0,53	0,71	0,44	0,75	0,19	0,35	0,16	0,47	1
	0,40	0,33	0,58	0,64	0,44	0,45	0,63	0,69	0,36	0,60	0,32	0,39	0,96	0,46	0,50	0,40	0,36	0,34	0,46	0,80	0,88	0,72	0,82	0,55	0,31	0,71	0,76	1
	0,55	0,31	0,28	0,37	0,32	0,44	0,81	0,82	0,39	0,41	0,42	0,51	0,68	0,38	0,19	0,32	0,34	0,26	0,43	0,71	0,38	0,38	0,57	0,67	0,16	0,80	0,35	1
	0,38	0,42	0,34	0,48	0,36	0,27	0,73	0,87	0,27	0,57	0,15	0,62	0,82	0,36	0,06	0,08	0,33	0,13	0,29	0,76	0,33	0,43	0,79	0,72	0,15	0,65	0,48	1
	0,25	0,42	0,30	0,27	0,34	0,32	0,60	0,71	0,39	0,45	0,26	0,51	0,43	0,32	0,32	0,10	0,57	0,55	0,68	0,73	0,44	0,69	0,83	0,43	0,25	0,68	0,64	1
	0,07	0,63	0,84	0,55	0,66	0,35	0,28	0,98	0,50	0,72	0,76	0,45	0,44	0,43	0,71	0,84	0,49	0,47	0,91	0,61	0,62	0,51	0,47	0,00	0,63	0,18	0,50	1
	0,36	0,15	0,38	0,82	0,37	0,61	0,56	0,84	0,56	0,60	0,39	0,45	0,50	0,01	0,72	0,76	0,54	0,60	0,32	0,64	1,00	0,61	0,42	0,64	0,38	0,62	0,63	1
B	0,24	0,57	0,35	0,34	0,30	0,33	0,68	0,60	0,11	0,51	0,53	0,00	0,66	0,56	0,80	0,60	0,59	0,36	0,37	0,58	0,72	0,39	0,50	0,57	0,35	0,47	0,39	1
C	0,15	0,39	0,81	0,78	0,71	0,54	0,21	0,80	0,63	0,46	0,42	0,63	0,88	0,52	0,59	0,25	0,84	0,57	0,71	0,70	0,76	0,82	0,95	0,14	0,67	0,25	0,91	1
Z	0,51	0,48	0,45	0,47	0,42	0,34	0,52	0,57	0,38	0,44	0,43	0,44	0,86	0,47	0,41	0,28	0,39	0,55	0,40	0,61	0,54	0,63	0,50	0,58	0,55	0,25	0,62	1
A	0,33	0,46	0,70	0,57	0,50	0,50	0,46	0,67	0,63	0,50	0,67	0,68	0,73	0,53	0,55	0,78	0,75	0,57	0,61	0,70	0,86	0,75	0,80	0,35	0,63	0,40	0,84	1
C	0,28	0,54	0,48	0,35	0,46	0,27	0,73	0,69	0,63	0,65	0,21	0,75	0,67	0,50	0,22	0,53	0,24	0,00	0,76	0,58	0,53	0,42	0,38	0,23	0,28	0,54	0,37	1
A	0,30	0,60	0,49	0,33	0,43	0,22	0,73	0,57	0,59	0,40	0,53	0,61	0,85	0,27	0,30	0,56	0,44	0,36	0,44	0,67	0,45	0,40	0,37	0,43	0,50	0,20	0,42	1
D	0,32	0,06	0,53	0,80	0,65	0,47	0,51	0,45	0,62	0,47	0,55	0,47	0,42	0,31	0,50	0,37	0,29	0,43	0,68	0,78	0,62	0,62	0,65	0,31	0,59	0,68	0,61	1
A	0,28	0,58	0,80	0,39	0,65	0,40	0,45	0,85	0,45	0,58	0,55	0,47	0,49	0,19	0,52	0,64	0,56	0,55	0,96	0,72	0,73	0,64	0,55	0,07	0,47	0,27	0,60	1
	0,44	0,43	0,51	0,32	0,49	0,31	0,74	0,70	0,59	0,58	0,41	0,50	0,69	0,39	0,53	0,39	0,69	0,18	0,56	0,82	0,66	0,60	0,69	0,48	0,13	0,71	0,56	1
	0,32	0,38	0,80	0,89	0,66	0,39	0,25	0,55	0,15	0,38	0,64	0,27	0,89	0,40	0,54	0,42	0,58	0,52	0,39	0,45	0,79	0,56	0,68	0,31	0,65	0,30	0,65	1
	0,45	0,30	0,36	0,75	0,37	0,64	0,35	0,60	0,67	0,40	0,55	0,35	0,56	0,36	0,37	0,31	0,28	0,70	0,13	0,54	0,46	0,58	0,84	0,87	0,44	0,63	0,61	1
	0,28	0,42	0,38	0,51	0,27	0,34	0,60	0,62	0,38	0,44	0,60	0,26	0,82	0,29	0,31	0,31	0,31	0,35	0,52	0,54	0,47	0,49	0,46	0,44	0,48	0,61	0,41	1
	0,23	0,37	0,78	0,77	0,60	0,50	0,35	0,37	0,49	0,29	0,79	0,33	0,68	0,33	0,74	0,46	0,47	0,58	0,37	0,35	0,59	0,45	0,29	0,43	0,79	0,44	0,43	1
	0,34	0,17	0,50	0,80	0,44	0,34	0,51	0,63	0,24	0,58	0,76	0,18	0,41	0,27	0,69	0,49	0,44	0,49	0,52	0,61	0,57	0,46	0,52	0,39	0,58	0,80	0,43	1
	0,28	0,51	0,66	0,45	0,59	0,06	0,44	0,94	0,51	0,64	0,59	0,63	0,56	0,52	0,58	0,45	0,69	0,34	0,58	0,75	0,55	0,64	0,81	0,34	0,28	0,56	0,62	1
	0,22	0,35	0,21	0,31	0,24	0,48	0,96	0,88	0,60	0,58	0,16	0,56	0,53	0,31	0,34	0,64	0,29	0,47	0,52	0,99	0,89	0,69	0,73	0,63	0,16	0,74	0,71	1
	0,27	0,46	0,66	0,61	0,59	0,35	0,51	0,97	0,31	0,72	0,14	0,50	0,94	0,38	0,47	0,57	0,59	0,24	0,65	0,67	0,46	0,44	0,46	0,26	0,35	0,32	0,44	1
	0,42	0,25	0,48	0,68	0,41	0,58	0,82	0,67	0,38	0,54	0,43	0,47	0,77	0,40	0,43	0,73	0,61	0,62	0,42	0,85	0,75	0,57	0,62	0,57	0,37	0,61	0,63	1
	0,34	0,49	0,58	0,64	0,61	0,46	0,62	0,59	0,52	0,65	0,34	0,46	0,53	0,27	0,34	0,65	0,48	0,38	0,29	0,92	0,74	0,51	0,79	0,53	0,47	0,29	0,61	1

	A1	A4	B3	B4	D7	E9	G7	G18	H12	J10	J16	K19	J13	M17	N23	Q26	Q21	RF	B	C	D	IWL	IWW	CI	PI	DI	Alan	Sınıf
G Ö K Ç E A D A	0,27	0,85	0,88	0,38	0,75	0,19	0,21	0,88	0,64	0,50	0,35	0,57	0,50	0,29	0,66	0,21	0,00	0,52	0,35	0,39	0,69	0,55	0,38	0,40	0,73	0,00	0,60	2
	0,16	0,68	0,99	0,60	1,00	0,22	0,00	0,62	0,78	0,45	0,61	0,85	0,30	0,71	0,63	0,46	0,38	0,44	0,47	0,42	0,48	0,50	0,56	0,18	1,00	0,04	0,53	2
	0,23	0,93	0,83	0,34	0,66	0,26	0,39	0,83	0,58	0,43	0,45	0,70	0,58	0,48	0,47	0,49	0,35	0,34	0,45	0,70	0,72	0,58	0,68	0,31	0,54	0,01	0,70	2
	0,47	0,83	0,68	0,14	0,60	0,14	0,42	0,78	0,75	0,44	0,57	0,80	0,68	0,25	0,50	0,37	0,07	0,62	0,30	0,55	0,88	0,77	0,57	0,67	0,52	0,16	0,76	2
	0,25	0,73	0,73	0,40	0,63	0,18	0,37	0,79	0,65	0,41	0,68	0,85	0,68	0,61	0,51	0,59	0,46	0,32	0,29	0,45	0,47	0,47	0,56	0,55	0,63	0,19	0,55	2
	0,56	0,34	0,63	0,77	0,57	0,31	0,29	0,55	0,40	0,80	0,58	0,44	0,36	0,72	1,00	0,83	0,56	0,47	0,36	0,49	0,56	0,46	0,59	0,47	0,50	0,52	0,42	2
	0,26	0,45	0,52	0,54	0,55	0,33	0,35	0,64	0,64	0,58	0,35	0,67	0,63	0,58	0,61	0,67	0,38	0,54	0,47	0,39	0,70	0,52	0,53	0,40	0,46	0,33	0,53	2
	0,47	0,54	0,36	0,36	0,32	0,32	0,41	0,88	0,42	0,53	0,78	0,69	0,01	0,38	0,74	0,38	0,85	0,21	0,28	0,27	0,06	0,24	0,16	0,62	0,52	0,33	0,19	2
	0,40	0,65	0,55	0,49	0,47	0,42	0,43	0,61	0,64	0,60	0,51	0,55	0,51	0,52	0,87	0,31	0,38	0,28	0,27	0,30	0,39	0,29	0,28	0,50	0,52	0,20	0,30	2
	0,34	0,64	0,44	0,15	0,47	0,07	0,63	0,63	0,36	0,75	0,22	0,57	0,11	0,00	0,58	0,56	0,62	0,65	0,57	0,73	0,65	0,69	0,62	0,49	0,09	0,53	0,68	2
	0,34	0,63	0,58	0,33	0,55	0,26	0,64	0,73	0,69	0,39	0,26	0,68	0,29	0,39	0,42	0,40	0,72	0,48	0,58	0,87	0,51	0,61	0,96	0,37	0,39	0,46	0,75	2
	0,15	0,73	0,71	0,27	0,56	0,22	0,45	0,91	0,59	0,60	0,41	0,55	0,13	0,23	0,66	0,49	0,77	0,61	0,66	0,65	0,36	0,49	0,64	0,25	0,43	0,42	0,50	2
	0,21	0,60	0,68	0,29	0,63	0,34	0,57	0,42	0,86	0,33	0,60	0,78	0,29	0,27	0,68	0,33	0,33	0,63	0,63	0,58	0,33	0,41	0,41	0,24	0,48	0,30	0,40	2
	0,57	0,43	0,83	0,58	0,71	0,55	0,55	0,61	0,71	0,36	0,20	0,66	0,17	0,13	0,35	0,44	0,66	0,54	0,62	0,78	0,52	0,54	0,67	0,32	0,56	0,56	0,63	2
	0,16	0,45	0,58	0,41	0,49	0,37	0,82	0,79	0,70	0,39	0,52	0,66	0,58	0,36	0,52	0,54	0,88	0,45	0,68	0,99	0,64	0,66	0,86	0,35	0,35	0,56	0,74	2
	0,40	0,45	0,57	0,50	0,51	0,31	0,56	0,72	0,69	0,72	0,60	0,75	0,56	0,35	0,67	0,30	0,72	0,38	0,42	0,64	0,49	0,54	0,50	0,45	0,33	0,40	0,49	2
	0,39	0,51	0,45	0,58	0,51	0,23	0,50	0,44	0,41	0,85	0,41	0,70	0,53	0,40	0,45	0,48	0,57	0,30	0,29	0,38	0,18	0,16	0,61	0,51	0,13	0,53	0,26	2
	0,60	0,46	0,30	0,34	0,38	0,43	0,51	0,86	0,74	0,55	0,63	0,55	0,65	0,91	0,77	0,49	0,39	0,70	0,17	0,51	0,38	0,39	0,53	0,92	0,49	0,34	0,41	2
	0,38	0,51	0,42	0,27	0,44	0,02	0,52	0,85	0,46	0,58	0,69	0,66	0,15	0,23	0,59	0,18	0,22	0,29	0,36	0,52	0,49	0,58	0,39	0,61	0,29	0,60	0,44	2
	0,34	0,53	0,32	0,38	0,39	0,12	0,55	0,43	0,22	1,00	0,42	0,38	0,73	0,28	0,71	0,56	0,56	0,73	0,40	0,73	1,00	0,84	0,79	0,59	0,24	0,33	0,85	2
	0,37	0,40	0,54	0,60	0,53	0,42	0,58	0,62	0,22	0,69	0,21	0,50	0,30	0,33	0,39	0,51	0,69	0,52	0,40	0,83	0,76	0,65	0,83	0,53	0,33	0,51	0,74	2
	0,26	0,47	0,55	0,57	0,51	0,39	0,58	0,70	0,58	0,71	0,30	0,72	0,45	0,41	0,54	0,33	0,55	0,44	0,63	0,64	0,38	0,43	0,56	0,25	0,41	0,39	0,48	2
	0,18	0,48	0,68	0,62	0,68	0,39	0,14	0,78	0,41	0,45	0,16	0,47	0,10	0,37	0,31	0,35	0,78	0,40	0,52	0,26	0,26	0,31	0,40	0,21	0,55	0,39	0,25	2
	0,30	0,39	0,24	0,25	0,28	0,26	0,60	0,64	0,67	0,46	0,43	0,71	0,39	0,11	0,47	0,54	0,61	0,61	0,58	0,50	0,59	0,60	0,48	0,48	0,35	0,66	0,50	2
	0,31	0,60	0,39	0,25	0,41	0,25	0,61	0,86	0,57	0,58	0,27	0,74	0,32	0,28	0,36	0,50	0,67	0,55	0,44	0,62	0,53	0,50	0,66	0,53	0,30	0,51	0,53	2
0,89	0,47	0,42	0,46	0,39	0,34	0,67	0,69	0,39	0,63	0,41	0,58	0,16	0,33	0,46	0,36	0,38	0,47	0,13	0,57	0,40	0,32	0,47	0,89	0,18	0,76	0,33	2	
0,02	0,96	0,58	0,00	0,50	0,00	0,37	0,67	0,34	0,58	0,24	0,83	0,54	0,32	0,62	0,17	0,12	0,96	0,81	0,54	0,94	1,00	0,64	0,20	0,49	0,03	1,00	2	
0,28	0,70	0,52	0,19	0,52	0,36	0,65	0,60	1,00	0,50	0,35	1,00	0,78	0,18	0,44	0,34	0,39	1,00	0,66	0,74	0,84	0,81	0,69	0,33	0,44	0,10	0,94	2	
0,14	0,61	0,62	0,49	0,59	0,42	0,17	0,97	0,51	0,58	0,45	0,36	0,70	0,37	0,60	0,34	0,57	0,42	0,50	0,31	0,58	0,48	0,67	0,31	0,31	0,35	0,43	2	
0,23	0,67	0,42	0,29	0,34	0,31	0,57	0,44	0,61	0,65	0,37	0,61	0,83	0,40	0,62	0,50	0,58	0,56	0,55	0,38	0,39	0,39	0,35	0,34	0,41	0,28	0,35	2	

6.3. ÖZELLİK SEÇİM SONUÇLARI

Özellik seçimi, bir özellik seçim kümesinden belli özelliklerin çıkarılması işlemidir. Arı kanatları için alınan açı, oran, uzunluk ve alan değerleri gibi özelliklerin bazılarının çıkarılması, işlem hızı açısından önem arz etmektedir. Bu nedenle İleri Sıralı Seçim Algoritması (Sequential forward selection) kullanılmıştır. Bu algorithmada verinin %10'luk kısmı test, %90'ı ise eğitim amacı ile kullanılmıştır. Sonuç olarak SFS algoritması ile 1., 2., 6., 8., 10., 12., 13.,15., 16., 17., 20., 21., 24. ve 25. özellikler seçilmiştir. SFS'in çalıştırılması sonucunda Çizelge 6.5'daki sonuçlar elde edilmiştir. Bu algoritma gönderilen verilerin sınıflandırılması için DVM kullanılmış ve DVM ile sınıflandırma sonucuna göre özellikler seçilmiş veya seçilmemiştir. Bunun yanında LDA özellik seçimi için kullanılmış ve LDA'da son 9 özellik seçilerek veri farklı bir boyuta taşındığında da sınıflandırma başarısının arttığı görülmüştür. Ayrıca BFM kullanılarak özelliklerin seçimi için program geliştirilmiş, bu programda 27 özellikten 5 özellik atılarak, 22 özellik seçilmiş ve bu özelliklerden sınıflandırma başarısı en yüksek olan özelliklerin listesi Çizelge 6.6'da verilmiştir. 22 özellik seçilmesinin nedeni programın 22 özellik için, 27'nin 22'li kombinasyonu ($C(27,22)$) hesaplanmış ve bu hesaplama sonunda 80730 ihtimali değerlendirmesi gerektiği görülmüştür. 80730 ihtimal için 80730 iterasyon oluşturulmalı ve 80730 sınıflandırma yapılması, bunun sonucunda en yüksek sınıflandırma başarısına sahip özelliklerin seçilmesi gerekmektedir. Eğer 27 özellik içinden 21 tane seçmek istersek $C(27,21)=296010$ ihtimal oluşmakta, bu da iş yükünü arttırmaktadır. Bu nedenle BFM ile 27 özellik içinden 22 özellik seçilmiştir.

Çizelge 6.5. SFS algoritması sonucunda seçilen özellikler.

Adım	Seçilen Özellik	Türü	Adım	Seçilen Özellik	Türü
<i>Adım:1</i>	12	Açı	<i>Adım:8</i>	17	Açı
<i>Adım:2</i>	6	Açı	<i>Adım:9</i>	13	Açı
<i>Adım:3</i>	15	Açı	<i>Adım:10</i>	1	Açı
<i>Adım:4</i>	10	Açı	<i>Adım:11</i>	24	İndeks
<i>Adım:5</i>	21	Uzunluk	<i>Adım:12</i>	2	Açı
<i>Adım:6</i>	16	Açı	<i>Adım:13</i>	25	İndeks
<i>Adım:7</i>	8	Açı	<i>Adım:14</i>	20	İndeks

Çizelge 6.6. BFM ile seçilen özellikler.

Sıra No	Seçilen Özellik	Türü	Sıra No	Seçilen Özellik	Türü
1	1	Açı	12	12	Açı
2	2	Açı	13	13	Açı
3	3	Açı	14	15	Açı
4	4	Açı	15	17	Açı
5	5	Açı	16	18	Uzunluk
6	6	Açı	17	19	Uzunluk
7	7	Açı	18	20	Uzunluk
8	8	Açı	19	23	Uzunluk
9	9	Açı	20	24	İndeks
10	10	Açı	21	25	İndeks
11	11	Açı	22	26	İndeks

Ayrıca LDA ile boyut azaltmada 9 özellik seçimi yapılmış ve bu verilerin Çizelge 6.7’de verilmiştir.

Çizelge 6.7. LDA ile boyut indirgeme sonuçları.

Sıra No	Seçilen Özellik	Türü	Sıra No	Seçilen Özellik	Türü
1	19	Uzunluk	6	24	İndeks
2	20	Uzunluk	7	25	İndeks
3	21	Uzunluk	8	26	İndeks
4	22	Uzunluk	9	27	Alan
5	23	Uzunluk			

6.4. ÖZELLİKLERİN FARKLI BİR BOYUTA TAŞINMASI

Özellik içi varyansı azaltmak ve özellikler arasındaki varyansı arttırmak için verilerin farklı bir boyuta taşınmıştır. Bu işlem için LDA yöntemi kullanılmıştır. Çizelge 6.7’de verilen 9 özellik farklı boyuta taşınması sonucunda Bozcaada ilçesinden alınan kanat özelliklerinde yapılan boyut indirgeme çalışması sonuçları alınmış ve Çizelge 6.8’deki veriler elde edilmiştir.

Çizelge 6.8. LDA ile verilerin farklı bir boyuta taşınması.

Sıra No	ÖZELLİKLER								
	19	20	21	22	23	24	25	26	27
1	-9,52	-1,4	24,74	26,93	-19,83	-8,98	7,59	-3,64	-10,39
2	-9,62	-1,46	24,91	27,41	-19,93	-8,85	7,6	-3,71	-10,41
3	-9,86	-1,41	24,74	28,03	-19,78	-8,77	7,5	-3,63	-10,34
4	-9,64	-1,38	24,91	27,01	-19,86	-8,86	7,42	-3,67	-10,38
5	-9,81	-1,46	24,77	26,06	-19,81	-9,04	7,53	-3,67	-10,18
6	-9,56	-1,46	24,77	25,57	-19,83	-8,77	7,72	-3,65	-10,17
7	-9,46	-1,52	24,73	24,45	-19,93	-8,91	7,53	-3,64	-10,27
8	-9,39	-1,44	24,78	25,6	-19,82	-8,95	7,59	-3,67	-10,23
9	-9,58	-1,46	24,67	26,68	-19,81	-9,15	7,52	-3,77	-10,34
10	-9,79	-1,36	24,9	28,29	-19,71	-8,69	7,65	-3,67	-10,39
11	-9,34	-1,46	24,7	26,07	-19,78	-8,86	7,64	-3,61	-10,33
12	-9,69	-1,44	24,91	26,26	-19,91	-8,84	7,31	-3,59	-10,26
13	-9,63	-1,4	24,7	27,56	-19,73	-8,99	7,74	-3,75	-10,27
14	-9,37	-1,43	24,83	24,7	-19,92	-9,05	7,48	-3,66	-10,29
15	-9,72	-1,45	24,77	26,35	-19,84	-8,77	7,61	-3,71	-10,35
16	-9,47	-1,37	24,87	26,71	-19,84	-8,92	7,57	-3,66	-10,39
17	-9,41	-1,4	24,8	26,07	-19,83	-8,92	7,36	-3,59	-10,34
18	-9,36	-1,38	24,71	26,72	-19,78	-8,78	7,81	-3,63	-10,37
19	-9,94	-1,41	24,89	26,73	-19,84	-8,84	7,67	-3,71	-10,34
20	-9,51	-1,4	24,8	25,28	-19,84	-8,89	7,54	-3,7	-10,26

Çizelge 6.8’deki veriler, LDA algoritması sonucunda elde edilen ağırlıkların veriler ile çarpılması sonucunda elde edilmektedir.

6.5. SINIFLANDIRMA SONUÇLARI VE DOĞRULUK MATRİSİ

Öncelikle işaretleme önerilen kavşak seçim algoritması kullanılmadan ve önerilen kavşak seçim algoritmasının kullanıldığı durumda başarı yüzdeleri değerlendirilmiştir.

6.5.1. ÖKSA Kullanılmadığı Durumda Sınıflandırma Sonuçları

Kavşak kesim algoritması kullanılmadığı durumda Çizelge 6.9'daki sınıflandırma başarıları elde edilmiştir. Çizelge 6.10'da elde edilen sonuçlar Çizelge 6.12'de elde edilen sonuçlara göre il ve ilçedeki örnekler açısından aşırı farklılık göstermiştir. Yapay sinir ağlarında 24 veri test amaçlı, 126 veri eğitim amaçlı kullanılmıştır. Ayrıca YSA'da giriş nöron sayısı 28'dir. Sonuç olarak YSA ile diğer sınıflandırma yöntemlerinden daha iyi sınıflandırma başarıları elde edilmiştir.

Çizelge 6.9. ÖKSA kullanılmadığı durumda sınıflandırma başarıları.

Sınıflandırma Yöntemi	Başarı Yüzdesi
DVM	%49,33
KNN	%38,6
K-Means	%34
YSA	%60,9

Çizelge 6.10. ÖKSA kullanılmadığı durumda örnek bazında doğruluk matrisleri.

Sınıflandırma Yöntemi	No	TAHMİN EDİLEN SINIFLANDIRMA						Sınıflandırma Başarıları	Toplam Başarı Oranları
		No	1	2	3	4	5		
		Alt tür	Gökçeada	Bozcaada	Kırklareli	Saray	Muğla		
DVM	1	Gökçeada	16	4	5	1	4	%53,3	%49,3
	2	Bozcaada	4	18	5	1	12	%60	
	3	Kırklareli	2	4	15	5	4	%50	
	4	Saray	6	7	7	5	5	%16	
	5	Muğla	2	2	4	2	20	%66	
K-Means	1	Gökçeada	6	7	3	5	9	%20	%34
	2	Bozcaada	4	11	4	7	4	%36	
	3	Kırklareli	7	1	10	2	10	%33	
	4	Saray	2	5	2	11	10	%36	
	5	Muğla	2	3	9	4	12	%40	
KNN	1	Gökçeada	17	7	1	5	0	%57	%38,6
	2	Bozcaada	6	19	2	1	2	%63	
	3	Kırklareli	7	10	3	6	4	%10	
	4	Saray	4	12	3	8	3	%26	
	5	Muğla	7	8	1	3	11	%36	
YSA	1	Gökçeada	3	0	0	0	0	%100	%60,9
	2	Bozcaada	3	4	0	0	1	%50	
	3	Kırklareli	0	1	0	0	0	%0	
	4	Saray	0	0	1	4	0	%80	
	5	Muğla	1	1	1	0	3	%50	

6.5.2. ÖKSA İle 27 Özelliğin Tamamı İçin Başarı Oranları

Kavşak seçim algoritması için hazırlanan program, resmin işaretlenen bölümünden 20X20'lik kısmının kesilmesini sağlamış, bu kısım siyah, beyaz resme dönüştürülmüştür.

Bu siyah beyaz resimde damar kısımları resmi 3 ayrı alana ayırmakta, bu alanlara komşu olan siyah noktaların birbirine olan mesafesine göre orta nokta bulunmaktadır. Kavşak seçim algoritması ile kavşak noktaları standart bir seçim yapılabilir duruma getirilmektedir. Bu nedenle Çizelge 6.11’de görüldüğü gibi sınıflandırma başarıları artmaktadır. 27 özellik için en iyi sınıflandırma YSA ile elde edilmiştir. En kötü sınıflandırmanın Saray ilçesinden alınan örneklerde olduğu görülmüş, en iyi sınıflandırma başarısı ise Kırklareli ilinden alınan örneklerde yakalanmıştır. Çizelge 6.12’de tüm örneklerin başarı yüzdeleri verilmiştir. Bu örnekte YSA için 127 örnek eğitim amaçlı, 23 örnek ise test amaçlı kullanılmıştır.

Çizelge 6.11. ÖKSA kullanıldığı durumda sınıflandırma başarıları.

Sınıflandırma Yöntemi	Başarı Yüzdesi
DVM	%58
KNN	%52
K-Means	%42
YSA	%73,9

Çizelge 6.12. ÖKSA ile 27 özellik kullanıldığında doğruluk matrisi.

Sınıflandırma Yöntemi	No	TAHMİN EDİLEN SINIFLANDIRMA						Sınıflandırma Başarısı	Toplam Başarı Oranları
		No	1	2	3	4	5		
		Alt tür	Gökçeada	Bozcaada	Kırklareli	Saray	Muğla		
DVM	1	Gökçeada	18	2	1	3	6	%60	%58
	2	Bozcaada	4	19	5	2	0	%63,3	
	3	Kırklareli	2	1	19	7	1	%63,3	
	4	Saray	3	4	9	11	3	%36,7	
	5	Muğla	5	1	3	1	20	%66,7	
K-Means	1	Gökçeada	9	5	1	10	5	%30	%42
	2	Bozcaada	7	10	6	6	1	%33	
	3	Kırklareli	1	4	19	4	2	%63,3	
	4	Saray	4	8	6	10	2	%33	
	5	Muğla	3	2	6	4	15	%50	
KNN	1	Gökçeada	16	5	1	1	7	%53,3	%52
	2	Bozcaada	3	22	4	0	1	%73,3	
	3	Kırklareli	3	3	20	2	2	%66,7	
	4	Saray	7	6	12	3	2	%10,0	
	5	Muğla	7	3	3	0	17	%56,7	
YSA	1	Gökçeada	6	0	0	1	0	%85	%73,9
	2	Bozcaada	1	4	0	1	0	%66	
	3	Kırklareli	0	0	2	0	2	%50	
	4	Saray	0	0	0	3	0	%100	
	5	Muğla	1	0	0	0	2	%66	

6.5.3. BFM İle Seçilen 22 Özelliğin Sınıflandırılması

Kaba kuvvet metodu ile 22 özelliğin seçilmek istenmiştir. Bu 22 özellik için $C(n,r)=(n!)/((n-r)!*r!)$ formülünden $C(27,22)$ değeri hesaplandığında 80730 ihtimal olduğu görülmüştür. 27 özellikten 22 özellik seçilebilmesi için 80730 ihtimal vardır. Bu ihtimallerin her biri için, sınıflandırma sonucuna bakılarak en iyi sonucu veren özellikler Çizelge 6.6'da gösterilmiştir. Bu özelliklerle elde edilen sınıflandırma başarısı yüzdeleri Çizelge 6.13'de verilmiştir. Sınıflandırma algoritmalarının doğruluk matrisleri de Çizelge 6.14'te verilmiştir. Burada YSA girişte 28 nöron kullanılmış, aktivasyon fonksiyonu için sigmoid aktivasyon fonksiyonu kullanılmıştır. YSA'nın eğitim aşamasında iterasyon sayısı 50 alınmıştır. DVM ve KNN sınıflandırma yöntemlerinde Saray ilçesinden alınan örneklerin başarısının çok düşük olduğu gözlemlenmiştir. YSA ve K-Means yöntemleri ile yapılan sınıflandırmada da bu ilçeden alınan örneklerin diğerlerinden çok yüksek başarıya sahip değildir. Bu nedenle sınıflandırmanın en kötü yapıldığı örneklerin Saray ilçesine ait olduğu sonucuna varılmaktadır. Bozcaada'dan alınan örneklerin ise sınıflandırma başarısının yüksek olduğu gözlemlenmiştir. Bu ilçeden alınan örnekler sınıflandırmanın kavşak seçim algoritması ile boyut indirgmeden yapılan sınıflandırmada da başarısının yüksek olduğu gözlemlenmiştir.

Çizelge 6.13. BFM ile sınıflandırma başarıları.

Sınıflandırma Yöntemi	Başarı Yüzdesi
DVM	%54
KNN	%52,7
K-Means	%37,3
YSA	%60,9

Çizelge 6.14. BFM doğruluk matrisi.

Sınıflandırma Yöntemi	No	TAHMİN EDİLEN SINIFLANDIRMA						Sınıflandırma Başarısı	Toplam Başarı Oranları
		No	1	2	3	4	5		
		Alt tür	Gökçeada	Bozcaada	Kırkklareli	Saray	Muğla		
DVM	1	Gökçeada	17	4	1	3	5	%56,7	%54
	2	Bozcaada	2	23	3	1	1	%76,7	
	3	Kırkklareli	3	4	19	2	2	%63,3	
	4	Saray	5	7	12	3	3	%10	
	5	Muğla	5	2	3	1	19	%63,3	
K-Means	1	Gökçeada	13	9	2	4	2	%43,3	%37,3
	2	Bozcaada	6	11	3	7	3	%36,6	
	3	Kırkklareli	1	1	11	7	10	%36,6	
	4	Saray	5	3	7	11	4	%36,6	
	5	Muğla	11	7	0	2	10	%33,3	
KNN	1	Gökçeada	17	5	1	1	6	%73,3	%52,7
	2	Bozcaada	2	22	2	2	2	%60	
	3	Kırkklareli	4	2	18	4	2	%16,7	
	4	Saray	5	4	14	5	2	%10,0	
	5	Muğla	6	4	3	0	17	%56,7	
YSA	1	Gökçeada	5	1	0	1	1	%62,5	%60,9
	2	Bozcaada	0	3	0	0	1	%75	
	3	Kırkklareli	0	0	3	1	0	%75	
	4	Saray	0	0	1	1	0	%50	
	5	Muğla	0	2	0	1	2	%40	

6.5.4. SFS İle Seçilen Özellikler ve Sınıflandırma

İleri sıralı seçim ile toplam 14 özellik seçilmiş ve bu özelliklerin sınıflandırma başarıları Çizelge 6.15'te verilmiştir. En iyi sınıflandırma YSA ile elde edilmiştir. YSA giriş nöron sayısı 28 olarak belirlenmiştir. YSA ile sınıflandırmada 127 veri eğitim, 23 veri test amaçlı kullanılmıştır. En kötü sınıflandırma K-means sınıflandırma yönteminde elde edilmiş, Ortalama en kötü sınıflandırma başarıları Saray ilçesinden alınan örneklerde gözlemlenmiştir.

Çizelge 6.15. SFS ile sınıflandırma başarıları.

Sınıflandırma Yöntemi	Başarı Yüzdesi
DVM	%60,7
KNN	%55,3
K-Means	%49
YSA	%82,6

Çizelge 6.16. SFS ile yapılan sınıflandırmada doğruluk matrisi.

Sınıflandırma Yöntemi	No	TAHMİN EDİLEN SINIFLANDIRMA						Sınıflandırma Başarısı	Toplam Başarı Oranları
		No	1	2	3	4	5		
		Alt tür	Gökçeada	Bozcaada	Kırklareli	Saray	Muğla		
DVM	1	Gökçeada	19	2	0	3	6	%63,3	%60,7
	2	Bozcaada	3	17	4	6	0	%56,6	
	3	Kırklareli	2	3	19	6	0	%63,3	
	4	Saray	3	3	8	15	1	%50	
	5	Muğla	4	1	3	1	21	%70	
K-Means	1	Gökçeada	13	0	0	13	4	%43,3	%49
	2	Bozcaada	8	14	3	4	1	%46,6	
	3	Kırklareli	2	5	10	6	7	%33,3	
	4	Saray	2	4	6	15	3	%50	
	5	Muğla	13	0	0	3	14	%46,6	
KNN	1	Gökçeada	22	1	1	2	4	%73,3	%55,3
	2	Bozcaada	2	18	4	3	3	%60	
	3	Kırklareli	5	3	20	1	1	%66,6	
	4	Saray	9	4	11	5	1	%16,6	
	5	Muğla	7	1	3	1	18	%60	
YSA	1	Gökçeada	3	0	0	0	0	%100	%82,6
	2	Bozcaada	0	5	0	0	1	%83,3	
	3	Kırklareli	0	0	4	2	0	%66,6	
	4	Saray	0	0	0	3	0	%100	
	5	Muğla	0	0	0	1	4	%80	

6.5.5. LDA İle Seçilen Özellikler ve Sınıflandırma Başarısı

LDA ile seçilen 9 özellik için sınıflandırma sonuçları Çizelge 6.17’de verilmiştir. Çizelge 6.18’de ise sınıflandırma doğruluk matrisleri verilmiştir. Sınıflandırmada özellikle K-means yöntemi ile sınıflandırma tekrar tekrar yapılmasına rağmen doğru sınıflandırma oranının düşük olduğu gözlemlenmiştir. Bunun nedeni, seçilen 9 özelliğin farklı boyuta taşınması sonucunda oluşan veriler için sınıf merkez noktaların birbirinden yeteri kadar ayrılamamasıdır. Bunun neticesinde veriler sürekli aynı sınıf etiketi ile etiketlenmektedir. Boyut indirgeme sonucunda sınıflandırma başarısının en fazla arttığı yöntem DVM olmuştur. Saray ilçesinden alınan örneklerin yine sınıflandırma başarısının diğer örneklere göre daha düşük olduğu gözlemlenmiştir.

Çizelge 6.17. LDA ile seçilen 9 özellik için başarı yüzdeleri.

Sınıflandırma Yöntemi	Başarı Yüzdesi
DVM	%71
KNN	%55,3
K-Means	%25
YSA	%43,5

YSA’da 28 giriş nöronu kullanılmıştır. İterasyon sayısı 50 alınarak YSA ile sınıflandırma yapılmıştır. Verilerin eğitimi aşamasında aktivasyon fonksiyonu olarak sigmoid fonksiyonu kullanılmıştır. Bozcaada ilçesinden alınan örneklerin daha iyi sınıflandırıldığı gözlemlenmiştir. K-means ile yapılan sınıflandırmada Bozcaada ilçesinden alınan örneklerin sınıflandırma başarısının çok düşük olması bu verilerin K-means ile sınıflandırılmasının doğru olmayacağını göstermektedir.

Çizelge 6.18. LDA ile seçilen özellikler için doğruluk matrisleri.

Sınıflandırma Yöntemi	No	TAHMİN EDİLEN SINIFLANDIRMA						Sınıflandırma Başarısı	Toplam Başarı Oranları
		No	1	2	3	4	5		
		Alt tür	Gökçeada	Bozcaada	Kırklareli	Saray	Muğla		
DVM	1	Gökçeada	23	2	0	1	4	%76,6	%71
	2	Bozcaada	3	22	4	0	1	%73,3	
	3	Kırklareli	2	1	19	7	1	%63,3	
	4	Saray	1	1	8	18	2	%60	
	5	Muğla	2	1	1	1	25	%83,3	
K-Means	1	Gökçeada	2	11	8	3	6	%6,6	%25
	2	Bozcaada	2	11	6	5	6	%36,6	
	3	Kırklareli	3	7	11	3	6	%36,6	
	4	Saray	3	8	9	5	5	%16,6	
	5	Muğla	1	11	6	4	8	%26,6	
KNN	1	Gökçeada	22	1	1	2	4	%73,3	%55
	2	Bozcaada	2	18	4	3	3	%60	
	3	Kırklareli	5	3	20	1	1	%66,6	
	4	Saray	9	4	11	5	1	%16,6	
	5	Muğla	7	1	3	1	18	%60	
YSA	1	Gökçeada	3	2	2	0	0	%42,8	%43,5
	2	Bozcaada	0	0	0	0	2	%0	
	3	Kırklareli	1	0	3	0	1	%60	
	4	Saray	1	0	1	2	0	%50	
	5	Muğla	1	0	1	1	2	%40	

6.6. GENEL SINIFLANDIRMA BAŞARILARI

Tüm sınıflandırma durumlarını ve alınan örneklerin il-ilçe bazında sınıflandırma başarısı Çizelge 6.19’da gösterilmiştir. Çizelge 6.20’de sınıflandırma yöntemlerinin örnekler üzerindeki sınıflandırma başarıları verilmiştir.

Çizelge 6.19. Tüm il/ilçeden alınan örneklerin sınıflandırma başarıları.

Sınıflandırma yöntemi			DVM					KNN				
Yöntem	Özellik sayısı	ÖKSA kullanıldı mı?	Bozcada	Gökçeada	Kırklareli	Saray	Mugla	Gökçeada	Bozcada	Kırklareli	Saray	Mugla
/	27	Hayır	53%	60%	50%	16%	66%	57%	63%	10%	26%	36%
/	27	Evet	60%	63,30%	63,30%	36,70%	66,70%	53,30%	73,30%	66,70%	10%	56,70%
BFM	22	Evet	57%	76,70%	63,30%	10%	63,30%	73,3%	60%	16,7%	10%	56,70%
SFS	14	Evet	63,30%	56,70%	63,30%	50%	70%	73,30%	60%	66,70%	16,70%	60%
LDA	9	Evet	76,70%	73,30%	63,30%	60%	83,30%	73,30%	60%	66,70%	16,70%	60%
Sınıflandırma yöntemi			K-Means					YSA				
Yöntem	Özellik sayısı	ÖKSA kullanıldı mı?	Bozcada	Gökçeada	Kırklareli	Saray	Mugla	Bozcada	Gökçeada	Kırklareli	Saray	Mugla
/	27	Hayır	20%	36%	33%	36%	40%	100%	50%	0%	80%	50%
/	27	Evet	30%	33%	63,30%	33,30%	50%	85%	66%	50%	100%	66%
BFM	22	Evet	43,30%	36,60%	36,60%	36,60%	33,30%	70%	76,60%	70%	76,60%	75%
SFS	14	Evet	43,30%	46,60%	33,30%	50%	46,60%	100%	83%	66%	100%	80%
LDA	9	Evet	6,60%	36,60%	36,60%	16,60%	26,60%	42,8%	0%	60%	50%	40%

Çizelge 6.20. Kullanılan sınıflandırma yöntemlerinin sınıflandırma başarıları.

Yöntem	Özellik sayısı	ÖKSA kullanıldı mı?	SINIFLANDIRMA YÖNTEMİ			
			DVM	KNN	K-Means	YSA
/	27	Hayır	49,3%	38,6%	34%	60,9%
/	27	Evet	58%	52%	42%	73,9%
BFM	22	Evet	54%	52,7%	37,3%	60,9%
SFS	14	Evet	60,7%	55,3%	49%	82,6%
LDA	9	Evet	71%	55,3%	25%	43,5%

Bozcada ilçesinden alınan örneklerin tabloda diğer örneklere göre sınıflandırma başarısının daha yüksek olduğu gözlemlenmiştir. En düşük sınıflandırma başarısı ise Saray ilçesinden alınan örneklerde gözlemlenmiştir. Çizelge 6.21’de tüm sınıflandırma yöntemleri, toplam kullanılan örnek sayısı ve doğru sınıflandırma oranı verilmiştir.

Çizelge 6.21. İl/İlçe bazında sınıflandırma başarısı.

Verinin alındığı il/ilçe	Sınıflandırma için kullanılan toplam örnek sayısı	Doğru sınıflandırılan örnek sayısı	Sınıflandırma başarı oranı
Gökçeada	478	234	%48,9539749
Bozcaada	476	242	%50,84033613
Kırklareli	470	220	%46,80851064
Saray	467	127	%27,19486081
Muğla	474	226	%47,67932489

7. SONUÇ VE ÖNERİLER

Bu çalışma kapsamında 5 farklı il ve ilçeden alınan arı kanadı görüntüleri üzerindeki kavşak noktaları kullanılarak özellikler çıkarılmıştır. Bu özellikler kullanılarak destek vektör makineleri, yapay sinir ağları, k-en yakın komşu ve k-ortalama yöntemleri ile sınıflandırma yapılmıştır. En başarılı sınıflandırma, %82,6 sınıflandırma başarıları ile YSA kullanıldığında elde edilmiştir. Boyut indirgeme algoritmalarından yararlanılarak özellik azaltımına gidilmiştir. Lakin özellik azaltımı farklı öğrenme yöntemlerinde farklı sonuçlar vermiştir. LDA ile 9 özellik seçilip veriler farklı çarpan ile yeni bir eksene taşındığında DVM ile sınıflandırma başarıları artarken K-ortalama ile sınıflandırma başarıları ise düşmüştür. İleri sıralı seçim ile boyut indirgeme yapıldığında 14 özellik kullanılmış, sınıflandırma başarılarında, özellik azaltımı yapılmadığı duruma göre fazla değişim olmasa da gereksiz özellikler çıkartılmıştır. BFM ile 22 özellik seçildiğinde sınıflandırma başarılarında gözle görülür değişim olmamış fakat gereksiz özellikler çıkarılmıştır. Koordinat noktalarının seçimi için hazırlanan algoritma ile koordinat işaretlemesi yapıldığında ise sınıflandırma başarıları önemli ölçüde artmıştır. Bu algoritmanın dezavantajı koordinat seçim işleminde zaman kaybına neden olmasıdır. 150 örnek için koordinat işaretlemesi uzun sürmese de daha fazla veri için koordinat noktası işaretleme uzun zaman almaktadır. Kavşak noktalarının işaretleme yerine otomatik kavşak noktası seçilmesini sağlayacak bir çalışma bu alanda yapılacak çalışmaların kolaylaştırılmasını sağlar. Saray ilçesinden alınan örneklerin sınıflandırma başarıları diğer il/ilçelerden daha az çıkmıştır. Bu ilçeden alınan örneklerin sınıflandırma başarıları DVM ve KNN için %10 ile %15 arası değişmektedir. Bu durum, diğer verilerin sınıflandırılma başarıları yüksek olsa dahi toplam sınıflandırma başarıları düşürmüştür. Bu ilçeden alınan örneklerin alındığı arılar daha melez yapıda olabilirler. En başarılı sınıflandırma Bozcaada ilçesinden alınan örneklerde görülmüştür. Gelecekte yapılacak çalışmalarda örneklerle ait resimlerin daha hassas alınması önerilmektedir. Çünkü resimlerdeki bozukluk koordinat noktalarının kaymasına ve özelliklerin değişmesine neden olmaktadır. Ayrıca kanat üzerinde işaretleme otomatikleştirilmesi için resimler ve arka planın ayırıcılığı artırılmalı, görüntüler beyaz zeminde alınmalıdır. Böylece resmin segmentasyonunu ve kavşak noktalarının seçimi için hazırlanacak programın

verimliliğini arttıracaktır. Yapay sinir ağı ile yapılan sınıflandırmada başarı yüksek çıkmıştır. Lakin verilen tablolarda görüldüğü üzere verilerin %90'lık kısmı eğitim amacı ile kullanılmakta, %10'luk kısmı test amacı ile kullanılmaktadır. Bu durum sınıflandırma başarısı yüksek veriler için rasgelelik oluşmasına neden olabilmektedir. YSA'da test örneklerinde oluşan farklılık her sınıf için 30 örnek olmasında kaynaklanmıştır. Çünkü YSA'da her sınıf için daha fazla örnek kullanmak, sınıflandırma başarısı yüksek bir ağ için gereklidir. DVM ile yapılan sınıflandırma her tekrarda aynı sonucu vermesi bu yöntemin arı kanadı sınıflandırmada daha güvenilir olduğunu göstermektedir. Kavşak noktasına dayalı arı alt türü sınıflandırma için verilerde boyut indirgeme, verilerin LDA gibi yöntemlerle farklı bir boyuta taşıma ve DVM kullanmak daha başarılı sonuçlar verecektir.

8. KAYNAKLAR

- [1] S. Ozekes, “Veri Madenciliği Modelleri ve Uygulama Alanları,” *İstanbul Ticaret Üniversitesi Dergisi*, c. 2, s. 3, ss. 67, 2003.
- [2] E. Çomak, “Destek vektör makinelerinin etkin eğitimi için yeni yaklaşımlar,” Doktora tezi, Elektrik Elektronik Mühendisliği Bölümü, Selçuk Üniversitesi, Konya, Türkiye, 2008.
- [3] D. Keçeci, “Türkiyede farklı illerden toplanan bal arılarında mezofilik bakteri florasının belirlenmesi ve arılar üzerindeki insektisidal etkileri,” Yüksek lisans tezi, Biyoloji Bölümü, Ordu Üniversitesi, Ordu, Türkiye, 2014.
- [4] M. Kekeçoğlu, “Türkiyede bal arılarının mtDNA ve bazı morfolojik özellikler bakımından karşılaştırılmasına yönelik bir araştırma,” Doktora tezi, Zootečni Bölümü, Namık Kemal Üniversitesi, Tekirdağ, Türkiye, 2007.
- [5] F. Ruttner, *Biogeography and taxonomy of honeybees*, 3st ed., vol. 75, New York, USA: Springer-Verlag, 1988, pp. 58–59.
- [6] G.Ö. Özbakır, “Güneydoğu sınırı boyu bal arısı popülasyonunun morfolojik özellikleri,” Doktora tezi, Zootečni Bölümü, Ankara Üniversitesi, Ankara, Türkiye, 2011.
- [7] I. Kandemir, M. Kence and A. Kence, “Genetic and morphometric variation in honeybee (*Apis Mellifera*) populations of Turkey,” *Apidologie*, vol. 31, no. 3, pp. 343–356, 2000.
- [8] D.R. Simith, “Genetic diversity in turkish honey bees,” *Uludağ Arıcılık Dergisi*, c. 2002, s. 3, ss. 10–15, 2002.
- [9] R. Hepburn and S.E. Radloff, “*Apis mellifera capensis*: an essay on the subspecific classification of honeybee,” *Apidologie*, vol. 33, no. 2, pp. 105–127, 2002.
- [10] T.M. Francoy, M. Drauschke, D. Wittmann, S. Müller, V. Steinhage, M.A.F.B. Laure, D.D. Jong and L.Ç. Gonçalves, “Identification of africanized honey bees through wing morphometrics: two fast and efficient procedures,” *Apidologie*, vol. 39, no. 5, pp. 488–495, 2008.
- [11] Ö. Ünsal, “Meslek alan seçimlerinin makine öğrenmesi algoritması kullanılarak belirlenmesi,” Yüksek lisans tezi, Bilgisayar Eğitimi Bölümü, Gazi Üniversitesi, Ankara, Türkiye, 2011.
- [12] E. Carrizosa, A.N. Gomez and D.R. Moralez, “Clustering categories support vector machines,” *Omega (UK)*, vol. 66, no. 1, pp. 28–37, 2017.
- [13] D.A. Demirci, “Destek vektör makineleri ile karakter tanıma,” Yüksek lisans tezi, Bilgisayar Mühendisliği, Yıldız Teknik Üniversitesi, İstanbul, Türkiye, 2007.
- [14] O. Eray, “Destek vektör makineleri ile ses tanıma uygulaması,” Doktora tezi, Elektrik-Elektronik Mühendisliği, Pamukkale Üniversitesi, Denizli, Türkiye, 2008.
- [15] M. Erdem, “Mesleki bel ağrısı riskinin destek vektör makineleri ile

- sınıflandırılması,” Yüksek lisans tezi, Endüstri Mühendisliği, Gazi Üniversitesi, Ankara, Türkiye, 2008.
- [16] S.K. Çalışkan, “K’knn:kümeleme ve en yakın komşu yöntemleri ile ağlarda nüfus tespiti,” Yüksek lisans tezi, Bilgisayar Mühendisliği, Gebze Yüksek Teknoloji Enstitüsü, Gebze, Türkiye, 2008.
- [17] S.K. Çalışkan, İ. Soğukpınar, “K-means ve k-en yakın komşu yöntemleri ile ağlarda nüfus tespiti,” 2. Ağ ve Bilgi Güvenliği Ulusal Sempozyumu, Girne, Kuzey Kıbrıs Türk Cumhuriyeti. 2008.
- [18] F. Köktürk, “K-en yakın komşuluk, yapas sinir ağları ve karar ağaçları yöntemlerinin karşılaştırılması,” Doktora tezi, Biyo istatistik Anabilim Dalı, Bülent Ecevit Üniversitesi, Zonguldak, Türkiye, 2012.
- [19] S. Kapil, M. Chawla and M.D. Ansari, " On k-means data clustering algorithm with genetical agorithm," presented at 4th International Conference On Parallel, Distributed and Grid Computing (PDGC), Solan, India. 2016.
- [20] Ö.F. Saraç, “Identification of facial expressions using artificial neural network (ann),” Yüksek lisans tezi, Bilgisayar Mühendisliği, Kocaeli Üniversitesi, Kocaeli, Türkiye, 2014.
- [21] M. Altınay, “Yapay sinir ağları ve k-means kullanarak sınır değerlerine göre yazılım efor tahmini,” Yüksek lisans tezi, Bilgisayar Mühendisliği, Atılım Üniversitesi, Ankara, Türkiye, 2012.
- [22] M. Alma, J. J. Martinez, I. D. Landau, and G. Buche, “Design and tuning of reduced order h-infinity feedforward compensators for active vibration,” *IEEE Transactions on Control Systems Technology*, vol. 20, no. 2, pp. 554–561, 2012.
- [23] Ş. Emir, “Yapay sinir ağları ve destek vektör makineleri yöntemlerinin sınıflandırma performanslarının karşılaştırılması: borsa endeks yönünün tahmini jzerine bir uygulama,” Yüksek lisans tezi, İşletme Ana Bilim Dalı, İstanbul Üniversitesi, İstanbul, Türkiye, 2013.
- [24] Mühendislik Fakültesi. (2017, 14 Ekim). *Yapay Sinir Ağları*. Erişim:<http://www.ibrahimcayiroglu.com/Dokumanlar/IleriAlgoritmaAnalizi/IleriAlgoritmaAnalizi-5.Hafta-YapaySinirAglari.pdf>
- [25] G. Biricik, "Comparing the impacts of dimension reduction methods that use class labels on text classification," *2012 20th Signal Processing and Communications Applications Conference (SIU)*, Muğla, 2012, pp. 1-4.
- [26] Anonim, (09 Kasım 2017). [Online]. Erişim: [https:// www.analyticsvidhya.com /blog /2016/12/introduction-to-feature-selection-methods-with-an-example-or-how-to-select-the-right-variables/](https://www.analyticsvidhya.com/blog/2016/12/introduction-to-feature-selection-methods-with-an-example-or-how-to-select-the-right-variables/)
- [27] İ. Taskesen, T. Kavzoğlu, "Nesne tabanlı sınıflandırmada filtreleme tabanlı özellik seçim algoritmalarının kullanımı ve sınıflandırma doğruluğunun etkilerinin incelenmesi," *6.Uzaktan Algılama ve Coğrafi Bilgi Sempozyumu*, Adana, 2016, ss. 445-453.
- [28] Anonim, (10 Kasım 2017). [Online]. Erişim: [http:// bmb.cu.edu.tr /uorhan/ DersNotu/ Ders11.pdf](http://bmb.cu.edu.tr /uorhan/ DersNotu/ Ders11.pdf)
- [29] T. Güneş, E. Polat, “Yüz ifade analizinde öznitelik seçimi ve çoklu svm sınıflandırıcılarına etkisi,” *Gazi Üniversitesi Mühendislik ve Mimarlık Fakültesi*,

c. 24, s. 1, ss. 7–14, 2009.

- [30] M. Kekeçoğlu, M.I. Soysal, “Genetik diversity of bee ecotypes in turkey and evidence for geographical differences,” *Romanian Biotechnological Latter*, vol. 15, no. 5, pp. 5646–5653, 2010.
- [31] M. Kekeçoğlu, M.I. Soysal, M. Bouga ve P. Harizanis, “Morphometrics as a tool for the study of genetic variability of honey,” *Tekirdağ Ziraat Fakültesi Dergisi*, vol. 4, no. 1, pp. 7–15, 2007.
- [32] A. Özkan and İ. Kandemir, “Discrimination of western honey bee populations in turkey using geometric morphometric methods,” *Bee Word Journal*, vol. 87, no. 2, pp. 24–26, 2015.
- [33] Anonim, (27 Ekim 2017). [Online]. Erişim: [http:// www.marmarisbalevi.com.tr /tr / aricilik/ turkiyede-aricilik](http://www.marmarisbalevi.com.tr/tr/aricilik/turkiyede-aricilik)
- [34] A. Güler, “A morphometric model for determining the effect of commercial queen bee usage on the native honeybee (*Apis mellifera* L.) population in a Turkish province,” *Apidologie*, vol. 41, no.2010, pp. 622–635, 2010.
- [35] İ. Çakmak, İ. Kandemir, A. Özkan ve S.S. Çakmak, “Bal arılarında kanatların geometrik morfometrik metodu ile analiz edilerek farklı seviyedeki varroa (*varroa destructor*) parazitin bulaşıklık seviyesinin belirlenmesi konusunda ön çalışma,” *Uludağ Arıcılık Dergisi*, c. 2011, s. 4, ss. 118–123, 2011.
- [36] C.J. Hall, “an automated approach to bee identification from wing venatio,” M.S. thesis, Department of Electrical Engineering, Wisconsin University, Madison, USA, 2011.
- [37] H.A. Sahara, “Morphological characterization and wing description of *vespa orientalis* queens,” *Journal of Apiculturel*, vol. 28, no. 2, pp. 76-80, 2013.
- [38] F.L.D. Silva, “Automated bee species identification through wing images,” M.S. thesis, Department of Biology, Sao Paulo University, Sao Paulo, Brezilya, 2015.
- [39] Agriculture and Nature - Ministry of Environment. (2017, 24 Sep.). *Automated identification by image analysis of wings*. Agriculture and Nature - Ministry of Environment. Available: http://www.webbee.org.br/bpi/pdfs/livro_03_schroder.pdf
- [40] F.J. Szymule and W. Skowronek, “Use of various morphological traits measured my microscope or computer methods in the honeybee taxonomy,” *Journal of Apiculturel*, vol. 54, no. 2, pp. 88-96, 2010.
- [41] M. Kambur, “Türkiyede bal arısı biyoçeşitliliğinin geometrik morfometrik yöntemler ile belirlenmesi,” Yüksek lisans tezi, Biyoloji Ana Bilim Dalı, Düzce Üniversitesi, Düzce, Türkiye, 2017.
- [42] A. Kaya, “Güneydoğu anadolu bölgesindeki gerbillinae (mammalia: rodentia) alt familyası türlerinin karyolojik ve morfolojik özellikleri,” Doktora tezi, Biyoloji Ana Bilim Dalı, Dicle Üniversitesi, Diyarbakır, Türkiye, 2010.
- [43] Bee Research Institute Dol, Ltd. (2017, 21 November). *Morphometrical analysis and taxonometical classification of honey bees by wing characters*. Available: http://beedol.cz/dawino/DAWINO_prosp_EN.pdf
- [44] Anonim, (23 Nisan 2017). [Online]. Erişim: https://en.wikipedia.org/wiki/Confusion_matrix

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : Hasan DEMİR
Doğum Tarihi ve Yeri : 01.01.1986-Şalpazarı
Yabancı Dili : İngilizce
E-posta : hasandemir61@gmail.com

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Y. Lisans	Bilgisayar Mühendisliği	Düzce Üniversitesi	2018
Lisans	Elektronik ve Bilgisayar Öğretmenliği	Abant İzzet Baysal Üniversitesi	2008
Lise		Şalpazarı Lisesi	2002