

YALOVA ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

**KOLONİ MORFOLOJİ TAHMİNİNDE YENİ YAPAY ÖĞRENME
MODELLERİ GELİŞTİRİLMESİ**

YÜKSEK LİSANS TEZİ

Volkan ALTUNTAŞ

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

HAZİRAN 2016

YALOVA ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

**KOLONİ MORFOLOJİ TAHMİNİNDE YENİ YAPAY ÖĞRENME
MODELLERİ GELİŞTİRİLMESİ**

YÜKSEK LİSANS TEZİ

**Volkan ALTUNTAŞ
(145105014)**

Bilgisayar Mühendisliği Anabilim Dalı

Bilgisayar Mühendisliği Programı

Tez Danışmanı: Doç. Dr. Murat GÖK

HAZİRAN 2016

YALOVA Üniversitesi Fen Bilimleri Enstitüsü'nün 145105014 numaralı Yüksek Lisans **Volkan ALTUNTAŞ**, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “ **KOLONİ MORFOLOJİ TAHMİNİNDE YENİ YAPAY ÖĞRENME MODELLERİ GELİŞTİRİLMESİ** ” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Doç. Dr. Murat GÖK**
Yalova Üniversitesi



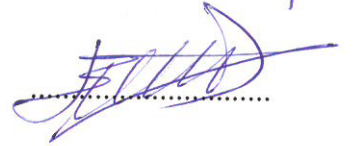
Jüri Üyeleri : **Doç. Dr. Murat GÖK**
Yalova Üniversitesi



Yrd. Doç. Dr. Ali DURSUN
Yalova Üniversitesi



Yrd. Doç. Dr. Emre DANDIL
Bilecik Şeyh Edebali Üniversitesi



Teslim Tarihi :
Savunma Tarihi :



Aileme,



ÖNSÖZ

Tez danışmanım sayın Doç. Dr. Murat Gök hocama, tez çalışma sürecimde yaptığı değerli katkılardan dolayı eşim Gıda Yüksek Mühendisi Seda Altuntaş'a ve çalışma arkadaşlarıma teşekkür ederim. Her zaman yanımda olan aileme sonsuz minnetlerimi sunarım.

Haziran 2016

Volkan Altuntaş



İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	vii
İÇİNDEKİLER	ix
KISALTMALAR	xi
ÇİZELGE LİSTESİ.....	xiii
ŞEKİL LİSTESİ.....	xv
ÖZET.....	xvii
SUMMARY	xix
1. GİRİŞ	1
1.1 Tezin Amacı	2
1.2 Literatür Araştırması	2
1.3 Tezin İçeriği	5
2. GIDA MİKROBİYOLOJİSİ	7
2.1 Mikroorganizmalar.....	7
2.1.1 Küf	7
2.1.2 Maya.....	8
2.1.3 Bakteri	8
2.2 Gıdalara Etkileri	9
2.3 Geleneksel Yöntemler ile Mikrobiyal Tespit	10
3. YAPAY ÖĞRENME SİSTEMLERİ.....	15
3.1 Veri Toplama.....	15
3.2 Ön İşleme	16
3.2.1 Segmentasyon	16
3.2.1.1 Eşikleme tabanlı yöntemler.....	17
3.2.1.2 Kenar tabanlı yöntemler	17
3.2.1.3 Bölge tabanlı yöntemler	18
3.3 Öznitelik Çıkarımı	20
3.3.1 HOG	20
3.3.2 Fraktal geometri	21
3.3.3 SURF.....	22
3.3.4 HU'nun değişmez momentleri	22
3.3.5 GLCM	23
3.3.6 Histogram.....	23
3.4 Boyut Azaltımı	24
3.4.1 Öznitelik dönüşümü	25
3.4.1.1 Temel bileşenler analizi	25
3.4.1.2 Doğrusal ayrıştırma analizi	26
3.4.2 Öznitelik seçimi	27
3.4.2.1 Ardışık öznitelik seçimi	28
3.4.2.2 Relief algoritması.....	28
3.5 Sınıflandırma ve Değerlendirme	29

3.5.1 Naive Bayes sınıflandırıcısı	29
3.5.2 C 4.5 sınıflandırıcısı	31
3.5.3 K-EYK sınıflandırıcısı	32
3.5.4 YSA sınıflandırıcısı.....	33
3.5.5 Performans değerlendirme	35
4. GELİŞTİRİLEN SİSTEM.....	39
4.1 Görüntülerin Elde Edilmesi	39
4.2 Nesnelerin Ayrıştırılması.....	40
4.3 Nesne Özniteliklerinin Çıkarılması	43
4.3.1 Şekil tabanlı yöntem 1	43
4.3.2 Şekil tabanlı yöntem 2.....	44
4.3.3 Renk tabanlı yöntem 1	44
4.3.4 Renk tabanlı yöntem 2	46
4.3.5 Öznitelik seçim ve öznitelik dönüşüm yöntemleri.....	47
4.4 Performans Değerlendirme.....	48
5. SONUÇLAR VE DEĞERLENDİRME.....	55
5.1 Sonuç	55
5.2 Tavsiyeler ve Gelecek Çalışmalar	55
KAYNAKLAR.....	57
EKLER.....	63
ÖZGEÇMİŞ.....	65

KISALTMALAR

ELS	: Elastic Light Scatter
Log	: Laplacian of Gaussian
ANN	: Artificial Neural Network
BP	: Backpropagation
KOB	: Koloni Oluşturan Birim
CFU	: Colony Forming Unit
K-NN	: K Nearest Neighbors
K-EYK	: K En Yakın Komşu
YSA	: Yapay Sinir Ağı
SURF	: Speeded Up Robust Features
SIFT	: Scale Invariant Feature Transform
HOG	: Histogram of Oriented Gradients
HU	: Hu's Invariant Moments
GLCM	: Gray Level Co-Occurrence Matrix
PCA	: Principal Component Analysis
TBA	: Temel Bileşenler Analizi
SVD	: Singular Value Decomposition
LDA	: Linear Discriminant Analysis
DAA	: Doğrusal Ayırıştırma Analizi
SBS	: Sequential Backward Selection
SGS	: Sıralı Geriye Seçim
SFS	: Sequential Forward Selection
SİS	: Sıralı İleri Seçim



ÇİZELGE LİSTESİ

	<u>Sayfa</u>
Çizelge 2.1 : Bazı küf cinsleri.....	8
Çizelge 2.2 : Bazı maya cinsleri.	8
Çizelge 2.3 : Bazı bakteri cinsleri.....	9
Çizelge 3.1 : Çok sınıflı bir sınıflandırıcı için performans ölçümleri.....	36
Çizelge 4.1 : Öznitelik seçimi sonrası değişken sayıları.	48
Çizelge 4.2 : Tür ve adet tahmini için başarımlar sonuçları.	49
Çizelge 4.3 : Tür tahmini için başarımlar sonuçları.....	51
Çizelge 4.4 : Adet tahmini için başarımlar sonuçları.....	52



ŞEKİL LİSTESİ

Sayfa

Şekil 2.1 : Ölçüm pipeti.	10
Şekil 2.2 : Petri kabı.	11
Şekil 2.3 : İnkübatör.	11
Şekil 2.4 : Petri sonucu görüntüsü.	12
Şekil 2.5 : Maya kolonisi görüntüsü.	12
Şekil 2.6 : Küf kolonisi görüntüsü.	12
Şekil 2.7 : Bakteri kolonisi görüntüsü.	13
Şekil 2.8 : Seyreltme işlemi.	13
Şekil 3.1 : HOG genel yapısı.	21
Şekil 3.1 : Ayçiçeği tarlasından SURF ile tespit edilen ilgi noktaları.	22
Şekil 3.1 : Gri seviye eş oluşum matrisi.	23
Şekil 3.2 : Örnek bir karar ağacı.	31
Şekil 3.3 : Doğal nöronun yapısı.	33
Şekil 3.4 : Yapay nöronun yapısı.	34
Şekil 3.5 : Yapay sinir ağının yapısı.	34
Şekil 4.1 : Sistem mimarisi.	39
Şekil 4.2 : Fotoğraf örneği.	40
Şekil 4.3 : Otsu algoritması maya-zemin eşikleme problemi görüntüsü.	41
Şekil 4.4 : Çok dokulu küf nesnesi Gabor segmentasyon görüntüsü.	41
Şekil 4.5 : Çoklu dokulu küf nesnesi renk tabanlı segmentasyon görüntüsü.	42
Şekil 4.6 : Watershed segmentasyon.	42
Şekil 4.7 : Ayrıştırılmış küf kolonileri.	42
Şekil 4.8 : Petri Uzmanı ekran görüntüsü.	43
Şekil 4.9 : Şekil tabanlı yöntem 1 ile öznitelik oluşturma.	44
Şekil 4.10 : Şekil tabanlı yöntem 2 ile öznitelik oluşturma.	44
Şekil 4.11 : Farklı konum ve renk uzaylarındaki renk değişim eğrileri.	45
Şekil 4.12 : Renk tabanlı yöntem 1 ile öznitelik oluşturma.	45
Şekil 4.13 : Renk tabanlı yöntem 2 ile öznitelik oluşturma.	47
Şekil 4.14 : Performans arabirimi ekran görüntüsü.	53
Şekil A.1 : Petri görüntüleri.	64



KOLONİ MORFOLOJİ TAHMİNİNDE YENİ YAPAY ÖĞRENME MODELLERİ GELİŞTİRİLMESİ

ÖZET

Besiyerleri mikroorganizmaların geliştirilmesi için formülize edilmiş ortamlardır. Bunlar, mikroorganizmaların geliştirilmesi, izolasyon, identifikasyon, sayım, duyarlılık testleri, sterilite testleri, klinik örneklerin incelenmesi, gıda, su ve çevre kontrolleri, biyolojik ürünlerin elde edilmesi, antibiyotik ve vitamin analizleri, endüstriyel analizler vb. gibi çok farklı amaçlara yönelik olabilir. Mikroorganizmaların besiyeri üzerinde çoğalarak oluşturduğu ve gözle görülebilen yapılara koloni denir. Agar üzerinde oluşan koloniler, mikroorganizma ve üreme ortamına bağlı olarak farklı morfolojik özelliklerde görüntüler oluşturmaktadır. Biyoteknoloji ve patoloji gibi alanlardaki birçok uygulamada ihtiyaç duyulan koloni sayımı, büyük sayıdaki kolonilerin elle sayılmasında sıkıcı, zaman alıcı ve insan hatası meyilli bir işlemdir. Laborantın yorgunluğu ile ilişkili olarak değişkenlik gösterebilen karar ve yeteneklerine bağımlıdır, laborantın çalışma hızı, belirli zaman aralığında değerlendirilebilecek petri sayısını sınırlandırmaktadır. Bu tezde, süt endüstrisinde ürünlerin mikrobiyolojik analiz sonuçlarının belirlenmesi için kullanılan küf, maya ve bakteri koloni görüntülerinin, görüntü işleme ve makine öğrenmesi teknikleri ile bilgisayar destekli sınıflandırılması üzerine çalışılmıştır.

Görüntü işleme, imgeden bazı yararlı bilgileri çıkarmak amacıyla, görüntüyü dijital forma dönüştürmek ve üzerinde bazı işlemleri gerçekleştirmek için bir yöntemdir. Bu işlem girdisi video karesi veya fotoğraf gibi imge olan ve çıktısı yine imge veya imgeye ait nitelikler olan bir sinyal işleme türüdür. Yapay öğrenme sistemleri, objelerin, sinyallerin, imgelerin bilgisayarlar tarafından otomatik olarak tanınması, belirli parametrelere istinaden otomatik karar verme süreçlerinin bilgisayarlar tarafından yapılması gibi ihtiyaçlardan ortaya çıkmıştır.

Küf, maya ve bakterilerin insan tarafından sınıflandırılmasında iki temel bilgi bulunmaktadır: renk bilgisi ve şekli bilgisi. İnsan gözü değerlendirme yaparken, koloni görüntüsündeki farklı renk sayısı, renk değişim ve geçişleri, koloni görüntüsündeki farklı geometrik şekiller ve kenar şekli değişimleri verilecek kararı etkilemektedir. İnsan gözünün algılamasında etkili olduğunu düşündüğümüz bu özelliklerden yola çıkarak görüntü işleme teknikleri ile ŞTY1 ve ŞTY2 olarak isimlendirilen iki adet şekil tabanlı ve RTY1 ve RTY2 olarak isimlendirilen 2 adet renk tabanlı öznitelik çıkarım yöntemi geliştirilmiştir.

Problemin çözümüne yönelik uygun sınıflandırma sisteminin tayini için veri setlerine TBA, DAA öznitelik dönüşüm, SGS öznitelik seçim algoritmaları uygulanmış ve Naive Bayes, k-EYK, C4.5 ve YSA algoritmaları ile sınıflandırılmıştır. Elde edilen deneysel sonuçlara göre, RTY2 öznitelik çıkarım, TBA öznitelik dönüşüm ve YSA sınıflandırmasından oluşan sistem ile başarımları, diğer sistemlere göre daha yüksek elde edilmiştir.



DEVELOPMENT OF NEW ARTIFICIAL LEARNING MODELS FOR COLONY MORPHOLOGY PREDICTION

SUMMARY

The Medium is environment that have been formulated for the growth of microorganisms. These can be for the different purposes such as growth, isolation, identification, counting, sensitivity tests of microorganisms, sterility testing, analysis of clinical samples, food, water, environmental controls, the acquisition of biological products, antibiotics and vitamins analysis, industrial analysis and so on. Structures which are formed by reproduced microorganisms and can be seen by eye are called colony. Colonies formed on the agar, creating images of different morphological characteristics depending on the microorganism and growth media. In the large number of colonies counted by hand, colony counting which is needed for many applications in areas such as biotechnology and pathology is boring, time consuming and it a process prone to human error. It depending on the decisions and abilities which can vary in relation to the laboratory worker fatigue, laborant's work rate limit the number of petri dishes can be evaluated at specific time interval. In this thesis, computer aided classification with image processing and machine learning methods on mold, yeast and bacteria colonie images which is used to determine the results of microbiological analysis of products in the dairy industry.

Image processing is a method to convert the image to digital form and perform some operations on in order to extract some useful information from the image. This is a kind of signal processing whose input is an image such as video frame or photo and output is an image or features of the image. Artificial learning systems have emerged from needs such as automatic recognition of objects, signals, images by computers and automated decision-making process based on certain parameters by computers.

There are two basic information on the classification of molds, yeasts and bacteria by the people: color information and shape information. When human eye making an assessment, the number of different colors in the colony image, color changes and transitions, different geometric shapes and corner shape changes in the colony image affects on the decision to be made. This hypothesis lead us to develeop two shape based and two color based feature extraction technique named ŞTY1,ŞTY2,RTY1 and RTY2 based on the characteristics that we think they effective to detection by human eye.

For the determination of the appropriate classification system for solving the problem we applied PCA, LDA feature transformations, SBS feature selection algorithms to dataset and they have been classified by Naive Bayes, k-NN, C4.5 and ANN algorithms. According to empirical results obtained, the system that composed RTY2 feature extraction, PCA feature transformation and ANN classifier, have achieved better performance compared to other systems.



1. GİRİŞ

Mikroorganizmalar, atmosferin birçok tabakasında, yer kürenin oldukça derinliklerinde, akarsularda, okyanusların en derin noktalarında, kayaların içlerinde, ekosistemin her bölgesinde yaşayabilmekte ve yaşamlarımızın her alanında varlıklarını sürdürmektedirler.

Katı besiyeri kullanımı mikroorganizmaların tespit edilmesinde kullanılan yöntemlerden biridir. Küf, maya ve bakteriler, buldukları üreme ortamına (besiyeri) ve mikroorganizma cinsine göre farklı morfolojik özelliklerde koloni görüntüleri oluştururlar. Oluşan bu görüntülerden mikroorganizmanın cinsine karar verilmektedir.

Görüntü işleme ve makine öğrenmesi ile imge sınıflandırılması birçok problemin çözümünde önemli rol oynamaktadır. Görüntü işleme, imgeden bazı yararlı bilgileri çıkarmak amacıyla, görüntüyü dijital forma dönüştürmek ve üzerinde bazı işlemleri gerçekleştirmek için bir yöntemdir. Bu işlem girdisi video karesi veya fotoğraf gibi imge olan ve çıktısı yine imge veya imgeye ait nitelikler olan bir sinyal işleme türüdür. Yapay öğrenme sistemleri, objelerin, sinyallerin, imgelerin bilgisayarlar tarafından otomatik olarak tanınması ve belirli parametrelere istinaden otomatik karar verme süreçlerinin bilgisayarlar tarafından yapılması gibi ihtiyaçlardan ortaya çıkmıştır.

Bu çalışmada mevcut öznitelik çıkarım teknikleri araştırılmış, daha yüksek başarımla elde edilmesini sağlayan dört adet yeni öznitelik çıkarım tekniği ve problemin çözümünü sağlayan yapay öğrenme modeli geliştirilmiştir.

1.1 Tezin Amacı

Bu tez çalışmasında petri görüntülerinden yapay öğrenme sistemleri ile kolonilerin morfolojik yapılarının tespit edilmesi ve tür tahminlerinin yapılması amaçlanmıştır. Yapılacak çalışma için paylaşılan bir veri tabanı bulunmaması sebebi ile Marmara bölgesindeki süt işletmelerinden numune toplanarak uygun aydınlatma, arka plan ve görüntüleme cihazları için çalışmalar yapılması ile görüntülerin oluşturulması planlanmaktadır. Öznitelik çıkarım için kullanılacak olan koloni ve koloni olmayan bölgelere ait görüntülerin ayrıştırılması için toplanan petri görüntülerinde yer alan bölgelerin segmentasyonunu en iyi performans ile gerçekleştirecek segmentasyon teknikleri denenerek belirlenecektir. Oluşturulan görüntü veri tabanından literatürde mevcut olan öznitelik çıkarım teknikleri ve problemin çözümü için geliştirilecek olan öznitelik çıkarım tekniklerimiz ile öznitelik veri tabanı oluşturulacaktır. Öznitelik veri tabanımızda en iyi performans ile işlem yapabilecek sınıflandırma algoritmaları tespit edilecek, geliştirdiğimiz öznitelikler ve mevcut özniteliklere ait performans kıyaslamaları yapılarak problemin yapay öğrenme sistemleri ile insan müdahalesi olmadan otomatikleştirecek sistemin belirlenmesi amaçlanmaktadır.

1.2 Literatür Araştırması

Hücre kolonileri gibi dairesel nesnelere sayılması biyologlar için önemli bir bilgi kaynağıdır. Bu işlem zaman alıcı ve öznel olmasına rağmen büyük bir çoğunlukla el ile yapılmaktadır. Quentin Geissmann (2012), dairesel kolonilerin petri fotoğrafından sayılması için görüntü işleme tabanlı açık kaynak kodlu OpenCFU aracını geliştirmiştir. C++ dili ile geliştirilen program kullanılarak “doğrudan eşikleme” ve “morfolojik segmentasyon” teknikleri ile dairesel koloniler tespit edilmekte ve sayılmaktadır. Uygun koloni görüntülerinde başarılı sonuçlar alınabilmektedir.

Bakterilerin konvansiyonel teşhis ve tanımlaması, numunelerin genetik ve biyokimyasal analizler için laboratuvarlara sevk edilmesini gerektirmektedir. Bu işlem günler sürebilir, zamanında müdahalenin yaşam kurtarabileceği ve bununla ilişkili maliyetleri azaltabileceği durumlarda önemli gecikmelere sebep olur. İşlem süresinin kısaltılması için Euiwon Bae ve diğ. (2012), dairesel kolonilerin petri fotoğrafından sayılması ve ELS ile oluşan görüntülerden bakteri tanımlamasını yapabilen taşınabilir cihaz geliştirmişlerdir. Geliştirilen cihaz kolay taşınabilir olma özelliği ile ülke genelinde biyo-güvenlik, besin güvenliği, hastane enfeksiyonlarının izleme ve önlenmesi için kullanılması amaçlanmaktadır.

Uzun ve zahmetli bir işlem olan hücre kolonilerinin sayılması laborantın yorgunluğu ile ilişkili olarak değişkenlik gösterebilen karar ve yeteneklerine bağlıdır. Laborantın çalışma hızı, belirli zaman aralığında değerlendirilebilecek petri sayısını sınırlandırmakta ve işlemin bir sistem tarafından otomatik olarak yapılma ihtiyacını ortaya koymaktadır. U. Bottigli ve diğ. (2007), dairesel kolonilerin petri fotoğrafından sayılması için sistem geliştirmiştir. Koloni bölgelerinin tespit edilmesi için “region-growing” algoritması, bölgelerin niteliklerinin tespiti için “Sanger neural net” ve sınıflandırma işlemi için “İleri Beslemeli sinir Ağı”, “K-En Yakın Komşusu”, “Doğrusal Ayırma Fonksiyonu” kullanılmıştır. Birbirine çok yakın veya çakışmış olan koloni görüntülerinden gerçek koloni sayısının tespit edilememesi problemi kullanılan makine öğrenmesi teknikleri ile çözülmesi amaçlanmıştır. Deney sonuçları hedef mikroorganizmalar için başarılı sonuçlar alındığını göstermektedir.

Silvio D. Brugger ve diğ. (2012) tarafından, zahmetli ve hataya meyilli işlem olan kan agar petrisinden elle bakteri koloni sayım işlemi için bir sistem geliştirilmiştir. Görüntü filtreleme ve segmentasyon teknikleri kullanılan bu çalışmada, dairenin geometrik özellikleri (majör minör eksen uzunluğu oranı ve nesne alanı oranı) kullanılarak mükemmel daireye yakın kolonilerin tespit edilmesi (-+ %20 tolerans konulmuştur) ve bayes sınıflandırıcısı kullanılarak, birbirine çok yakın olduğu için zincir görüntüsüne sahip olan kolonilerin sayısının belirlenmesi hedeflenmiştir. Sonuç olarak *S. pneumoniae*, *P. aeruginosa*, *M. catarrhalis* bakteri kolonileri için %90 düzeyinde başarı ile sayım yapılması sağlanmıştır. Aydınlatmanın görüntü üzerine olan etkileri de araştırılmış ve sayım yapılacak agar için en iyi aydınlatmanın mavi ışık kaynaklı karanlık alan aydınlatma tekniği olduğu tespit edilmiştir. Projede donanım olarak IDS uEye UI-1640-C kamera, C1614-M lens, karanlık alan aydınlatma ünitesi kullanılmış ve program MATLAB de yazılmıştır.

Yapay görme, çeşitli türde bilimsel ve endüstriyel uygulamalar için imgelerin gerçek zamanlı yakalanması, işlenmesi ve analiz edilmesi ile ilişkili olarak ortaya çıkan bir alandır. Biyoteknoloji ve patoloji gibi alanlardaki birçok uygulamada ihtiyaç duyulan bakteri sayımı, büyük sayıdaki kolonilerin elle sayılmasında sıkıcı, zaman alıcı ve insan hatası meyilli bir işlemdir. Yapay görme kavramı ile otomatikleştirilebilecek olan bu işlem için Er. Monita Goyal (2012), görüntü filtreleme, eşikleme ve Heywood dairesellik faktörü tekniklerini kullanarak dairesel özellikte ve belirli boyutlarda oluşan kolonilerin sayımını yapan bir sistem geliştirmiştir. Projede NI IEEE-1394 DCAM kamera ve NI vision Assistant 7.1 görüntü işleme programı kullanılmıştır. Hedef koloni sistemi için başarı sağlanmıştır.

Wei-Bang Chen and Chengcui Zhang (2008), önerdikleri koloni sayma sisteminde kontrast eşitleme ve otsu segmentasyon algoritması ile sayım yapılacak petri alanının otomatik tespitini, görüntüleri kromatik ve akromatik olarak gruplandırarak farklı filtreleme parametreleri ve otsu segmentasyon algoritması ile kolonilerin tespit edilmesini, watershed algoritması ile birbirine yakın kolonilerin ayrıştırılmasını sağlamışlardır. Program MATLAB ile yazılmış, 8.0, 6.0, 5.0, 3.2, 2.0 megapixel beş farklı kamera, Mutans Streptococci nin izole edildiği mavi renk Mitis-Salivarius agar ve *Escherichia Coli* nin izole edildiği şeffaf LB agar kullanılarak yapılan denemelerde %80 nin üstünde başarı oranı elde edilmiştir.

S. Sieuwerts ve diğ. (2008), görüntü eşikleme, filtreleme ve watershed algoritması kullanarak koloni sayma sistemi önermişlerdir. Eşikleme ve filtreleme işleminin kullanıcı tarafından yapıldığı sistemde Canon EOS 350D yüksek çözünürlüklü dijital kamera, Sigma 105 mm macro lens kullanılarak siyah zemin üzerinde yandan aydınlatma ile görüntüler elde edilmiştir. Yazılım ImageJ görüntü işleme programında hazırlanmıştır.

Gurpreet Kaur ve Poonam Sethi (2012), görüntü işleme teknikleri ile basit ve uygun maliyetli bir koloni sayma yöntemi önermişlerdir. Filtreleme, eşikleme, segmentasyon, sınır belirleme algoritmalarının kullanıldığı çalışmada görüntülerin kolay işlenebilmesi için triphenyle tetrazolium chloride kullanılarak koloniler ile besiyeri arasındaki renk ve kontrast farkının artırılması (koloniler koyu kırmızı olmuştur) sağlanmıştır. Düşük kontrast, yüksek kontrast, gürültülü, gürültüsüz, çakışan ve çakışmayan koloniler içeren görüntüler ile yapılan denemelerde %90 düzeyinde başarı oranı elde edilmiştir.

1.3 Tezin İçeriği

Tezin ilk bölümünde tezin amacı ve literatür özeti verilmiştir. İkinci bölümde gıda mikrobiyolojisi hakkında bilgi verilmiştir. Mikroorganizmalar ve çeşitleri tanıtılmış, gıdalara olan etkileri ve tespit yöntemleri hakkında bilgi paylaşılmıştır. Üçüncü bölümde yapay öğrenme sistemleri hakkında bilgi verilmiş ve bir yapay öğrenme sistemi geliştirmek için kullanılan, veri toplama, ön işleme, öznitelik çıkarımı, boyut değişimi, sınıflandırma ve performans değerlendirme aşamaları hakkında bilgi paylaşılmıştır. Dördüncü bölümde tez kapsamında geliştirilen sistem ve geliştirilen öznitelik çıkarım teknikleri detaylı olarak tanıtılmış, sistemin performans sonuçları paylaşılmıştır. Beşinci ve son bölümde çalışmaya ilişkin sonuçlar, ileride yapılması düşünülen çalışmalar ve tavsiyeler ile ilgili bilgiler verilmiştir.



2. GIDA MİKROBİYOLOJİSİ

2.1 Mikroorganizmalar

Mikroorganizma yunanca “küçük” ve “organizma” kelimelerinin birleşmesi ile oluşan, tek veya çok hücreli mikroskobik canlı organizmaları temsil eden bir kavramdır. Mikroorganizmaların incelendiği mikrobiyoloji 1674 yılında Antonie Van Leeuwenhoek tarafından kendi tasarımı olan mikroskopuyla mikroorganizmaları keşfi ile başlamıştır. Mikroorganizmalar çıplak gözle göremeyeceğimiz kadar küçük canlılardır, atmosferin birçok tabakasında, yer kürenin derinliklerinde, akarsularda, okyanusların derin noktalarında, kayaların içlerinde, ekosistemin her bölgesinde yaşayabilmekte ve yaşamlarımızın her alanında varlıklarını sürdürmektedirler. Ekolojik dengede rol oynayan mikroorganizmalar temel olarak bakteriler, mantarlar, algler, virüsler ve protozoalardan oluşur. (Volkan Altuntaş ve diğ., 2016, Url-1, 2016).

2.1.1 Küf

Ökaryotik organizmaların büyük bir üyesi olan mantarlar âleminde yer almaktadırlar. Küfler hif olarak adlandırılan, uzun dallanma liflerinin büyük, çok hücreli kümelerini oluştururlar. Bitkisel ve üreme olmak üzere iki çeşit hif bulunur. Boyutu şekli ve yapısı ile mantarlar âleminde sınıflandırma ve tanımlamada kullanılan sporlar üreme hifi üzerinde oluşurlar. Tüp benzeri hifler makroskobik küf kolonisinin kabarık görüntüsünün sebebidir. Küf hücreleri art arda dizilerek hif yapılarını, hiflerde çeşitli şekillerde dallanma yaparak hif topluluğunu bir araya getirirler. Oluşan bu hif topluluğuna miselyum adı verilir. Küfler havaya dağılıp, beslenmeye elverişli bir yer bulduklarında, oraya yerleşerek çoğalıp gelişirler. Küfler besin maddelerini kendileri yapmazlar, saprofit ya da parazit olarak beslenirler. Bazılarından, antibiyotikler elde edilir, bazı türleri aflatoksin adı verilen zehirli maddeler üretir. Çok sayıda küf cinsi bulunmaktadır ve bunlardan gıda mikrobiyolojisi için önemli olan bazıları Çizelge 2.1 de listelenmektedir (Morphology and General Properties of Fungi, t.y., Genel Mikrobiyoloji, 2006, Url-4, 2016, Url-5, 2016).

Çizelge 2.1 : Bazı küf cinsleri.

Alternaria	Penicillium	Geotricum
Cladosporium	Aureobasidium	Trichothecium
Mucor	Fusarium	Byssochlamys
Aspergillus	Rhizopus	Monilia
Colletotrichum	Botrytis	Wallemia

2.1.2 Maya

Küfler gibi, ökaryotik organizmaların büyük bir üyesi olan mantarlar âleminde yer almaktadırlar. Mayalar nadiren lif oluşturan, tek hücreli (5-8µm) canlılardır. Bazı maya cinsleri tomurcuklanma ile bazılarında bölünme ile çoğalırlar. Mayalara daha çok şeker içeren ortamlarda, çiçek balözlerinde veya meyvelerin yüzeylerinde rastlanır. Koşullar uygunsa şeker içeren maddelere kısa sürede etki ederek bileşimlerinde önemli değişikliklere sebep olurlar. Çok sayıda maya cinsi bulunmaktadır ve bunlardan gıda mikrobiyolojisi için önemli olan bazıları Çizelge 2.2 de listelenmiştir (Genel Mikrobiyoloji, 2006, Morphology and General Properties of Fungi, t.y.).

Çizelge 2.2 : Bazı maya cinsleri.

Brettanomyces	Torulaspora	Rhodotorula
Issatchenkia	Cryptococcus	Zygosaccharomyces
Schizosaccharomyces	Pichia	Hanseniaspora
Candida	Trichosporan	Saccharomyces
Kluyveromyces	Debaryomyces	

2.1.3 Bakteri

Bakteriler klorofil içermeyen, tek hücreli, prokaryotik mikroorganizmalardır. Su, hava, toprak, bitki, hayvan, insan vb. her yerde bulunabilirler. Çok hızlı üreme gösterirler, besinsel gereksinimleri basittir, çevresel etmenlere uyum yetenekleri yüksektir, eşeyli ve eşeysiz olarak çoğalabilirler. Morfolojik yapıları kok, basil ve spiral şeklindedir. Kok formda bulunan bakteriler 0,5-1 µm çapında, basil formda bulunan bakteriler 1-5µm uzunluğunda, 0,5-1µm çapında, spiral formda bulunan mikroorganizmaların uzunlukları ise 1 µm den 100 µm ye ulaşabilir. Çok sayıda bakteri cinsi bulunmaktadır ve bunlardan gıda mikrobiyolojisi için önemli olan bazıları Çizelge 2.3 de listelenmiştir (Genel Mikrobiyoloji, 2006, Morphology and Classification of Bacteria, t.y.).

Çizelge 2.3 : Bazı bakteri cinsleri.

Acetobacter	Proteus	Hafnia
Erwinia	Alcaligenes	Psychrobacter
Pediococcus	Flavobacterium	Bacteroides
Aeromonas	Pseudomonas	Kocuria
Escherichia	Bacillus	Salmonella

2.2 Gıdalara Etkileri

Koliform grubu bakterilerin ısıtılma işlemi görmüş, tüketime hazır gıdalarda bulunması yeterli hijyenik koşullarda üretilmediğinin göstergesidir. Pastörizasyon normlarına oldukça duyarlı olan koliform grubu bakterilerin pastörize ürünlerde yer almaları etkili bir pastörizasyon işleminin yapılmadığını ya da pastörizasyon sonrası ekipman yüzeylerine insan ve çevreden bulaşmanın olduğunu gösterir. Bu gıdaların tüketilmesi sonucu diyare, enfeksiyon hastalıkları vb. meydana gelebilir. Bozulmaya yol açan maya ve küfler gıdalarda acı tat ve kötü koku, gaz, istenmeyen gözenekli yapı oluşumuna neden olmaktadır. Bazı küf türleri bunların yanı sıra bulaştıkları gıda maddesinde gelişerek salgıladıkları toksik metabolitler, mikotoksinler nedeniyle, gıda maddesinin tüketilmesi durumunda ölüme sonuçlanabilen zehirlenmelere yol açmaktadır. Protein, karbonhidrat ve mineral bakımından zengin besin içeriğine sahip olan süt ürünleri, patojen ve patojen olmayan mikroorganizmalar için oldukça uygun bir gelişme-çoğalma ortamıdır. Bu nedenle gıdalarda olması istenmeyen mikroorganizma türleri, ürünlerde tespit edilmeye çalışılır, tüm çeşitlerin aranması yerine temel bilgi sağlayan bazı indikatör mikroorganizmalar kullanılmaktadır. İndikatör mikroorganizmalar gıdaya özgü belirlenir. Süt endüstrisinde koliform, küf ve maya indikatör mikroorganizma olarak aranır. İndikatör mikroorganizmalar, gıdaların raf ömrünü etkileyen mikrobiyel kalitenin belirlenmesinde ve gıda kaynaklı patojenlerin bulaşma olasılığının tespitinde önemli rol oynarlar. (Ahmed E. Yousef ve diğ., 2003, Volkan Altuntaş ve diğ., 2016).

2.3 Geleneksel Yöntemler ile Mikrobiyal Tespit

İndikatör mikroorganizmaların tespiti için, katı besiyeri kullanılan yöntemler, en muhtemel sayı yöntemi, membran filtrasyon, titre belirlenmesi, metabolizmaya dayalı hızlı analizler, mikroskopik sayımlar, canlandırma, var/yok testleri, biyolojik stabilite testi, yüzey ve hava kontrolü, vb. birçok geleneksel mikrobiyal tespit yöntemi bulunmaktadır. Yöntemler arasında maliyet, zaman, uygulanabilirlik gibi avantajlar veya dezavantajlar bulunmakla birlikte, gıda endüstrisinde ve biyoloji gibi farklı birçok alanda indikatör mikroorganizma tespiti için katı besi yeri kullanılan yöntemler yaygın olarak kullanılmaktadır. Katı besiyeri kullanılarak yapılan sayımın prensibi "her canlı hücrenin belirli bir inkübasyon sonunda 1 adet koloni oluşturması" dır. Bununla birlikte, canlı olduğu halde hasar görmüş ve gelişip koloni oluşturamayacak canlı hücreler de gıda maddesinde bulunabilir. Bu nedenle sayım sonuçları sadece koloni oluşturabilenlerin sayıldığını göstermek üzere, koloni oluşturan birim, KOB(colony forming unit, cfu) olarak verilmektedir. Yöntem aşağıdaki adımlardan oluşmaktadır;

- Ürün ve besiyeri Şekil 2.1 ve Şekil 2.2 deki ekipmanlar kullanılarak petri kabında karıştırılır.
- Şekil 2.3 de gösterildiği gibi 25 ± 2 °C de, 3 gün inkübasyonda bırakılır.
- Süreç sonunda oluşan Şekil 2.4 deki gibi görüntüler değerlendirilir.



Şekil 2.1 : Ölçüm pipeti.



Şekil 2.2 : Petri kabı.



Şekil 2.3 : İnkübatör.

Besiyeri üzerinde oluşan koloniler, mikroorganizma ve üreme ortamına bağlı olarak farklı morfolojik özelliklerde görüntüler oluşturmaktadır. Oluşan koloni görüntüleri morfolojik özelliklerine göre kategorik olarak sayılır. Kategori değerlendirme işleminde aşağıdaki özelliklere dikkat edilmektedir;

- Maya
 - Şekil 2.5 de görüldüğü gibi kolonilerin kesin kenarları vardır.
 - Koloniler farklı renk ve büyüklüklerde olabilir.
 - Koloniler tümsekli olarak (3 boyutlu) görünebilir.
 - Genelde koloninin ortasında odak noktası (koyu renkli merkez nokta) yoktur.

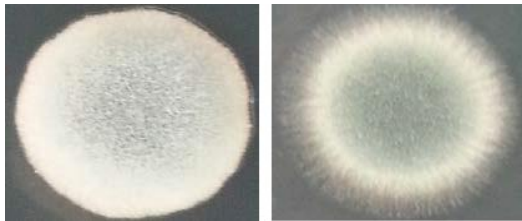


Şekil 2.4 : Petri sonucu görüntüsü.



Şekil 2.5 : Maya kolonisi görüntüsü.

- Küf
 - Şekil 2.6 da görüldüğü gibi kolonilerin kenarları dağınıktır.
 - Kolonilerin rengi değişkendir (küfler kendi pigmentlerini üretebilir).
 - Koloniler farkı büyüklüklerde olabilir.
 - Genelde koloninin ortasında odak noktası bulunur.



Şekil 2.6 : Küf kolonisi görüntüsü.

- Bakteri
 - Şekil 2.7 de görüldüğü gibi kolonilerin kesin kenarları vardır.
 - Koloniler farklı renk ve büyüklüklerde olabilir.
 - Genelde çekirdek şekline sahiptirler ve koloninin ortasında odak noktası bulunur (koyu renkli merkez nokta).

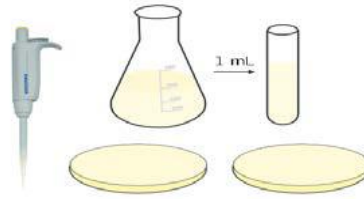


Şekil 2.7 : Bakteri kolonisi görüntüsü.

Seyreltme uygulanmamış ise bulunan sayı cfu/ml olarak ifade edilir. Seyreltme uygulanmış ise bulunan sayı seyreltme faktörü ile çarpılarak cfu/ml olarak ifade edilir. Seyreltme faktörünün amacı, petri üzerinde sayılamayacak oranda koloni üremesini engellemektir.

- Seyreltme işlemi için yöntem

- Şekil 2.8 de yer alan seyreltilmemiş karışım = 1 ml gıda + 15 ml besiyeri
- 10-1seyreltim için, 10 ml gıda + 90 ml ringer solüsyonu ile hazırlanmış karışımdan 1 ml + 15 ml besiyerinden oluşur, bu durumda bulunan sayı 10 ile çarpılır.



Şekil 2.8 : Seyreltme işlemi.

Bulunan küf, maya, bakteri, koliform grubu bakteri sayım değerleri işletmelerin kendi belirledikleri limitlere göre ürün kalitesi ve ürün sevkini belirlemede kullanılır. 1200 ton/gün kapasiteye sahip bir süt işletmesinde bu değerlendirmeler için günde 700 adet petri görüntüsü incelenmektedir. Bunlardan ortalama 580 adedi (%83) küf, maya, bakteri, koliform grubu bakteri petri sayımlarından oluşmaktadır. (Seda Altuntaş ve diğ., 2015, Volkan Altuntaş ve diğ., 2016).



3. YAPAY ÖĞRENME SİSTEMLERİ

3.1 Veri Toplama

Ses, ışık, basınç vb. biçimlerde bulunan fiziksel bilgilerin, sensörler, sayısallaştırma cihazları, tarayıcılar vb. vasıtasıyla ölçülerek bilgisayarlar tarafından işlem yapılabilecek biçimlerde veriye dönüştürülme işlemidir. Başarılı bir yapay öğrenme sisteminin tasarlanabilmesi için, uygun bir sınıflandırma algoritması kullanmanın yanında, yeterli büyüklükte ve problemi temsil edebilen eğitim ve test veri setlerine sahip olunması da önemli bir gereksinimdir. Yeterlilik, karar sınırlarının, öznelitektörleri ve sınıf etiketleri arasındaki fonksiyonel haritalama olarak öğrenilmesi için yeterli miktarda verinin bulunmasını sağlayabilmektir. Hangi miktarda verinin yeterli olduğunu belirleyen genel bir kural bulunmamaktadır. Temsilcilik, oluşabilecek tüm anlamlı veri kombinasyonlarının eğitim ve test veri setlerinde örneklenmiş olmasını sağlayabilmektir. Eğitim ve test veri setinde yer alan bilgilerin tüm problem uzayını kapsadığının belirlenmesi, karşılanması zor bir durumdur. Belirtilen özelliklerin sağlanmasında seçilecek cihazların türleri ve ölçüm mekanizmaları dikkat edilmesi gereken bir konudur (Rohi Polikar, 2006, Seema Asht ve diğ., 2012, Selim Aksoy, 2015).

3.2 Ön İşleme

Gürültü giderme, filtreleme, normalizasyon, istisnai bilgilerin temizlenmesi, önemli bölgelerin belirlenmesi ve izole edilmesi işlemleridir. Çeşitli etmenlerden dolayı toplanan veride gürültü bulunabilir, veride bulunan gürültü hakkında bilgi sahibiysek, uygun filtreleme teknikleri kullanılarak veri gürültüden ayrılabilir. Farklı ölçeklere sahip verilerin, sonuçları farklı ağırlıklarda etkilemesinin önlenmesi için verinin ortalama ve varyans değerleri kullanılarak veri normalize edilebilir. Veri dağılımına göre çeşitli eşik seviyeleri ile istisnai değerlerin belirlenmesi ve çıkartılması, veri setinde uygulanacak farklı algoritmalar ile önemli bölgelerin belirlenmesi ve ayrıştırılması işlemleride verinin daha doğru ve hızlı işlenmesine yardım olacak ön işlem basamaklarındandır (Robi Polikar, 2006, Al Bovik, 2009, Mehmet Erdal Balaban ve diğ., 2015).

3.2.1 Segmentasyon

Segmentasyon imgenin farklı nesne veya nesne parçalarına karşılık gelen bölge veya kategorilere ayrılmasıdır. İmgedeki her piksel bir bölge ile ilişkilendirilir. İyi bir segmentasyon aşağıdaki özelliklere sahiptir:

- Aynı bölgedeki pikseller benzer değerlere sahip olarak birleşik bir bölge oluştururlar.
- Farklı bölgelerdeki komşu pikseller farklı değerlere sahip olurlar.
- Tüm bölgelerin birleşimi imgeyi oluşturur.
- Herhangi iki farklı bölgenin kesişimi boş kümedir.

Segmentasyon görüntü analizinde kritik bir adımdır, birçok pikselden oluşan imgedeki nesnelere çalışmak için her piksel incelenir. Segmentasyondaki başarı görüntü işleminin diğer aşamalarını da olumlu etkiler fakat otomatik segmentasyon algoritmaları kullanıldığında özel durumlara özgü başarılar elde edilir ve harici durumlarda problemlerin üstesinden gelebilmek için elle müdahale gereklidir. Eşikleme tabanlı yöntemler, kenar tabanlı yöntemler ve bölge tabanlı yöntemler olmak üzere segmentasyon için farklı yaklaşımlar bulunmaktadır (Mubarak Shah, 1997, Bryan S. Morse, 2000, Fritz Albregtsen, 2010).

3.2.1.1 Eşikleme tabanlı yöntemler

En basit ve yaygın kullanılan segmentasyon yöntemidir. Belirlenen bir eşik değeri t ve (i,j) pozisyonunda bulunan pikselin gri seviye değeri f_{ij} olmak üzere, eğer (3.1) şartı sağlanıyor ise piksel bölge 1 e dâhil edilir, diğer koşullarda bölge 2 ye dâhil edilir.

$$f_{ij} \leq t \quad (3.1)$$

İlgilenilen bölgenin ayrılması için t eşik değeri birçok durumda farklı değerler denenerek araştırmacı tarafından belirlenir. Aynı bölümde yer alan pikseller ($0 - t$ ya da $t+1 - 255$) her zaman tek bir bağlı nesne oluşturmazlar bu yüzden eşikleme işlemi devamında bölge birleştirme ya da bölge parçalama işlemleri uygulanır. İmgeyi ikiden fazla bölgeye ayırmak için birden fazla eşik değeri kullanılabilir, imgenin gri seviye histogram grafiği vb. bilgiler kullanılarak t eşik değeri otomatik olarak hesaplanabilir (Al Bovik, 2009, Mark S. Nixon, 2002, Mubarak Shah, 1997).

Eşikleme tabanlı yöntemlerden olan Otsu algoritması, otomatik kümelene tabanlı görüntü eşikleme veya bir ikili görüntü için görüntünün gri seviye indirgenmesini gerçekleştirmek için kullanılır. Algoritmada imgenin, çift durumlu histogramıyla ilişkili iki piksel gurubu içerdiği varsayılır (ön plan pikselleri ve arka plan pikselleri) ve iki sınıfı birbirinden ayıracak, sınıf içi varyansın minimum, sınıflar arası varyansın maksimum olduğu optimum eşik değer hesaplanır. Orijinal metottan farklı olarak çok seviyeli eşik kullanan versiyon Çoklu Otsu olarak adlandırılır (Otsu, 1979, Liao ve diğ., 2001, Sezgin ve diğ., 2004).

3.2.1.2 Kenar tabanlı yöntemler

İmgeye kenar filtresi uygulanır, bu şekilde imgede tespit edilen süreksizlikler (ani değişiklikler) kenar olarak adlandırılır ve filtrenin çıkışına göre işlem yapılan piksel, kenar veya kenar-değil bölümlerine ayrılır. Kenar filtresi olarak genellikle türev işlemi kullanılır, birinci derece (3.2) ve ikinci derece (3.3) türev işlemleri çoğunlukla tercih edilir. Canny, LoG, Prewitt, Roberts, Sobel en yaygın kullanılan kenar bulma algoritmalarından bazılarıdır.

$$\frac{\partial f}{\partial x} = f(x) - f(x-1) \quad (3.2)$$

$$\frac{\partial^2 f}{\partial x^2} = f(x+1) + f(x-1) - 2f(x) \quad (3.3)$$

Tespit edilen kenarlar kullanılarak nesnelerin dış hatlarının oluşturulması için kenar bağlama, model uydurma (Hough dönüşümü), seviye setleri vb. işlemler uygulanır (Mubarak Shah, 1997, Bryan S. Morse, 2000, Fritz Albrechtsen, 2010).

3.2.1.3 Bölge tabanlı yöntemler

Bu yöntemler iterative olarak komşu ve benzer değerlere sahip pikselleri gruplamaya, aynı bölümdeki piksellerden bağlantılı nesnelere oluşturulmaya, farklı değerlere sahip pikselleri de bölümlenmeye çalışırlar. Kenar tabanlı yöntemlerde nesne sınırlarının bulunması, birleştirilmesi ve doldurulması ile nesnelere oluşturulmaya çalışılırken bölge tabanlı yöntemlerde nesnenin ortasından başlanarak dışarıya doğru büyüme ile nesne sınırları belirlenmeye çalışılır. Bölgeler kenarlardan daha fazla piksel içerebileceği için bölgenin belirlenmesi işlemine daha fazla bilgi sunabilirler. Bölgelerin oluşturulması için region merging, region splitting, split-and-merge gibi algoritmalar kullanılmaktadır (Mubarak Shah, 1997, Bryan S. Morse, 2000, Fritz Albrechtsen, 2010).

Bölge tabanlı yöntemlerden olan Gabor filtreleri ile doku segmentasyonu, imgeyi doku özellikleri temelinde bölgelere ayırma işlemidir. Dokuların yorumlanması için insan görme istemindeki çok kanallı işlemlerin keşfi, çalışmaların, farklı doku bölgelerinin tespiti için Gabor filtreleri temelinde çok kanalı yaklaşımlara yönelmesini sağlamıştır. Deneyimler göstermektedir ki Gabor filtreleri memeli görme sistemindeki basit hücrelerin makul bir modelidir, bu nedenle insanların dokuları kolayca ayırt edebilmesinin iyi bir modeli olarak düşünülmektedir. Gabor filtreleri ile doku segmentasyonu temel üç aşamadan oluşmaktadır;

1. Tasarlanan Gabor filtre dizisinin gri seviye kaynak imgeye uygulanması.
2. Filtrelenmiş imgeden öznitelik çıkarımı.
3. Kümeleme.

İlk olarak Gabor büyüklüklerinin (magnitude) çıkarılması için gri seviyeye dönüştürülmüş imge, tasarlanan Gabor filtre dizisi ile filtrelenir. İmgenin filtreye yanıtı olan Gabor büyüklüklerinin sınıflandırmada kullanılabilmesi için Gauss yumuşatma, öznitelik kümesine bölgesel bilgi eklenmesi, kümeleme algoritmaları için özniteliklerin yeniden şekillendirilmesi ve normalizasyon gibi işlemlere tabi tutularak Gabor öznitelikleri elde edilir. Son adım olarak pikseller kümeleme ile doku bölgelerini temsil eden bölgeler halinde kümelenir (Anil K. Jain ve diğ., 1990, Naotoshi Seo, 2006, M. Sivalingamaiah ve diğ., 2012).

K-ortalamlar ve renk tabanlı algoritmada, imgeler RGB renk uzayından CIELAB renk uzayına dönüştürülür, Bu sayede parlaklık değişiklikleri ihmal edildiğinde oluşan 3 renkten görsel farklılıkların ölçülebilmesi sağlanır. CIELAB renk uzayında a* ve b* katmalarında bulunan renk bilgisi k-ortalamlar algoritması ile kümelenecek bölütleme gerçekleştirilir (Ming-Ni Wu ve diğ., 2007, Anil Z Chitade ve diğ., 2010).

Watershed algoritması ile segmentasyonda, yerel minimumlarından delinmiş bir yüzeyin suya batırıldığı hayal edilirse, en düşük yükseklikli minimumdan başlayarak su adım adım farklı havzalara dolacaktır. İki farklı havzadan gelen suların birleştiği noktalara barajlar inşa edilirse, bu batırma işleminin sonunda her bir minimum, bu minimumla ilişkili havzayı sınırlayacak şekilde tamamen barajlarla çevrilir. Su seviyesi yüzeyin en yüksek tepesine ulaştığı zaman bu işleme son verilir. Sonuç olarak, yüzeyi farklı bölgelere ya da havzalara ayıran barajlar havza çizgileri ya da sadece havzalar olarak adlandırılır. Aşırı bölütleme, watershed dönüşümünün en büyük dezavantajıdır. Aşırı bölütleme problemini ortadan kaldırmanın bir yolu işaretçi kullanmaktır. Önce nesnelere işaretlemek için gerekli özellikler tanımlanır ve bunlar nesne işaretçileri olarak adlandırılır. Herhangi bir nesneye ait olmadığı kesin olarak bilinen bir imge parçası da işaretlenir ve bunlar da arka plan işaretçileri olarak adlandırılır. Geri kalan bölümler bütün uygulamalar için aynıdır. İşaretçiler arasında kalan ilgilenilen bölümün sınırları muhafaza edilecek şekilde, gradyan imgesi değiştirilir. Daha sonra değiştirilmiş gradyan imgesi üzerinde watershed dönüşümü kullanılarak son sınır belirlemesi gerçekleştirilir (Fernand Meyer, 1990, Pierre Soille ve diğ., 1990, Mevlüt Topaloğlu ve diğ., 2006).

3.3 Öznitelik Çıkarımı

Özellik çıkarma ve sınıflandırma arasındaki kavramsal bağlar kişisel bakış açısına göre değişebilmektedir. İdeal bir öznitelik çıkarıcı sınıflandırma işini önemsiz hale getiren iyi bir temsil verimi sağlar. Aynı şekilde başarılı bir sınıflandırıcının da gelişmiş bir öznitelik çıkarıcıya ihtiyacı bulunmayacaktır. Bu farklıklar bizi teorik gerekçeler yerine pratik gerekçeler üzerine eğilmeye yöneltir. Genel olarak öznitelik çıkarımı sınıflandırmaya göre daha fazla problem ve çalışma alanı bağımlıdır bu yüzden çalışma alanıyla ilgili bilgi gereksinimi duyulur. Balıkları sıralamak için kullanılan iyi bir öznitelik çıkarıcı, parmak izi tanıma veya kan hücrelerinin mikroskop görüntülerinden ayırt edilmesinde yüksek olasılıkla çalışmayacaktır. Öznitelik çıkarmadaki temel amaç toplanan veriden sınıflandırma işlemi için ayırt edici veya bilgi verici, tercihen daha az boyutlu özelliklerin bulunmasıdır. Boyut azalmasının yapay öğrenmedeki önemi göz ardı edilemez. Az ama bilgi verici öznitelik seti sınıflandırma algoritmasının karmaşıklığını, algoritmanın çalışması için gerekli zaman ve hafıza gereksinimlerini belirgin bir şekilde azaltmanın yanında aşırı öğrenme olasılığına düşürür. Öznitelik çıkarımı genellikle toplanan veri üzerinde uygulanan matematiksel dönüşümler ile yapılır (Richard O. Duda, 2000, Robi Polikar, 2006).

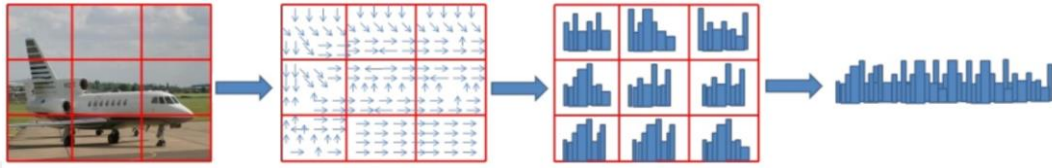
3.3.1 HOG

Son yıllarda imge tanıma konsepti içinde kenar oryantasyon histogramlarının kullanımı büyük bir ivme kazanmıştır. Ancak yoğun ve lokal gradyan oryantasyon histogramlarının kullanımı ilk defa Shashua ve Dalal tarafından önerilmiştir. Bu yöntemdeki amaç imgeyi bir grup lokal histogramlar olarak tanımlamaktır. Bu histogramlar, imgenin lokal bir bölgesindeki gradyanların oryantasyonlarının oluşma sayılarını içerir. HOG tanımlayıcıları gerçekleşmesindeki işlem basamakları şu şekilde sıralanabilir;

1. RGB imge gri seviyeli imgeye dönüştürülür ve gri seviyeli imgenin yatay ve dikey gradyan değerleri elde edilir.
2. Yatay ve dikey gradyan değerleri kullanılarak gradyan yönelimi ve gradyan büyüklüğü değerleri elde edilir.

3. Gradyan yönelimleri imgesi boyutunda hücrelere bölünür ve her bir hücre içerisindeki gradyan yönelimleri 45'şer derecelik 8 farklı durumla etiketlenir.
4. Belirlenen her bir blok için gradyan büyüklüğü ve gradyan yönelimleri kullanılarak lokal histogramlar oluşturulur.
5. Oluşturulan lokal histogramlar normalize edilir.

(Fuat Karakaya ve diğ., 2009, Nur Banu Albayrak ve diğ., 2012, Can Erhan ve diğ., 2013, Murat Yoldaş ve diğ., 2013).



Şekil 3.1 : HOG genel yapısı.

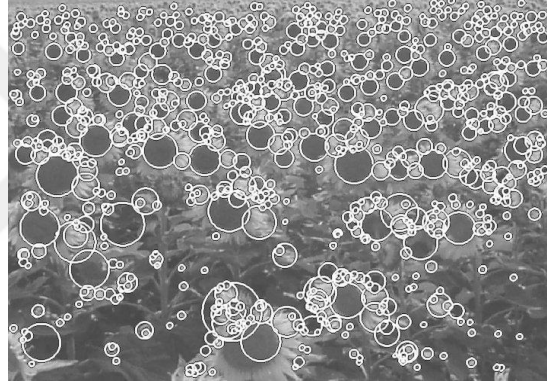
3.3.2 Fraktal geometri

Kendine benzeyen ve düzensiz olan geometrik öğeler fraktal olarak adlandırılır. Fraktal boyut fraktalların geometrik yapısı hakkında bilgi içerdiği için önemli bir özelliktir. Bir nesnenin topolojik boyutu (d) nesne hangi dönüşüme uğrarsa uğrasın değişmez. Fraktal dünyada, fraktal boyutun bir tamsayı olması gerekmez. Bir nesnenin fraktal boyutu (D) normalde onun topolojik boyutundan daha büyüktür ($D \geq d$). Fraktal boyut, Eukleidesçi şekiller için topolojik boyutlarına eşit, fraktallar için topolojik boyutlarından büyüktür. Bir şekil kendisine benzeyen N kadar kopyadan oluşuyor ve her bir kopya özgün şekle göre, $1/M$ büyüklüğünde ölçeklenmiş ise, bu şeklin kendine benzeme boyutu aşağıdaki formül ile hesaplanır (Benoit B. Mandelbrot, 1983, Alceu Ferraz Costa ve diğ., 2012, P. Shanmugavadivu ve diğ., 2012).

$$D = \frac{\log(N)}{\log(M)} \quad (3.4)$$

3.3.3 SURF

SURF, nesne tanıma, sınıflandırma, 3D oluşturma gibi işleri yerine getirmek için kullanılan lokal bir öznitelik çıkarım tekniğidir. SIFT den esinlenilerek geliştirilen SURF ün standart versiyonu SIFT e göre daha hızlıdır ve görüntü dönüşümlerinden daha az etkilenmektedir. SURF algoritmasındaki hız artışı, her görüntü dizisinden elde edilen farklı ölçeklerde süzgeçlenmiş görüntü piramidi yerine farklı ölçeklerde oluşturulan süzgeçin sabit boyuttaki görüntü dizilerine uygulanması ve böylece her görüntü dizisi için ayrı işlem yapılmaması ile gerçekleştirilmiştir. SURF algoritmasında temel olarak Hessian matrisi ve basit Laplace hesaplaması ile tespit edilen ilgi alanlarının, Haar-Wavelet algoritmasına tepkisi bulunur ve tanımlama olarak kullanılır (Herbert Bay ve diğ., 2006, Luo Juan ve diğ., 2009, Gürkan Kocadere, 2010, Melih Evren Buruş ve diğ., 2014).



Şekil 3.1 : Ayçiçeği tarlasından SURF ile tespit edilen ilgi noktaları.

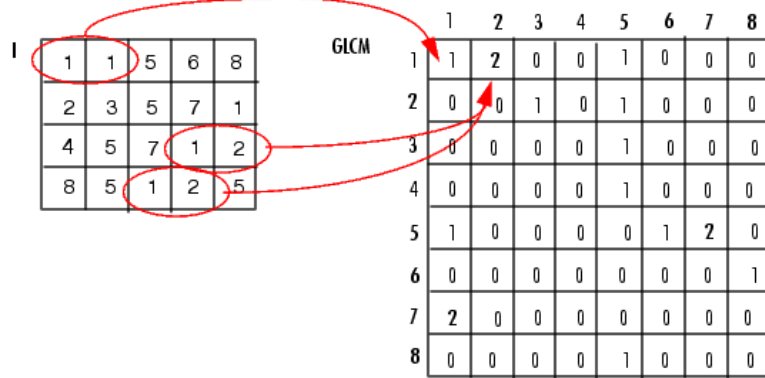
3.3.4 HU'nun değişmez momentleri

1960'ların başında, matematiksel moment teorisinden Eşitlik 3.5 de yer alan yedi değişmez moment Hu tarafından geliştirilmiştir. Bu yedi moment, ölçeklendirme, döndürme ve dönüşüm işlemleri altında değişmemekte ve karakter tanıma gibi iki boyutlu görüntü tanıma problemlerinin çözümünde gözle görülür başarı sağlamaktadır. Bu özelliği ile karakter tanıma, nesne tanıma, radar görüntülerinden uçak ve gemi tanıma gibi uygulamalarda kullanılmıştır (Ming-Kuei Hu, 1962, Franz L. Alt, 1968).

$$\begin{aligned}
\varphi_1 &= \eta_{20} + \eta_{02} \\
\varphi_2 &= (\eta_{20} - \eta_{02})^2 + 4\eta_{11} \\
\varphi_3 &= (\eta_{30} - 3\eta_{12})^2 + (3\eta_{21} - \eta_{03})^2 \\
\varphi_4 &= (\eta_{30} + \eta_{12})^2 + (\eta_{03} + \eta_{21})^2 \\
\varphi_5 &= (\eta_{30} - 3\eta_{12})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
&+ (3\eta_{21} - \eta_{03})(\eta_{21} + \eta_{03})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] \\
\varphi_6 &= (\eta_{20} - \eta_{02})[(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2] + 4\eta_{11}(\eta_{30} + \eta_{12})(\eta_{21} + \eta_{03}) \\
\varphi_7 &= (3\eta_{21} - \eta_{03})(\eta_{30} + \eta_{12})[(\eta_{30} + \eta_{12})^2 - 3(\eta_{21} + \eta_{03})^2] \\
&- (\eta_{30} - 3\eta_{12})(\eta_{03} + \eta_{12})[3(\eta_{30} + \eta_{12})^2 - (\eta_{21} + \eta_{03})^2]
\end{aligned} \tag{3.5}$$

3.3.5 GLCM

GLCM, gri tonlu imgenin özniteliğini çıkarmaya yarayan, Haralick tarafından geliştirilmiş bir öznitelik çıkarma yöntemidir. GLCM de, iki komşu piksel arasındaki ilişkiyi tanımlar. Bu piksellerden birincisi referans pikseli, diğeri de komşu piksel olarak kabul edilir ve komşu pikseller arasındaki frekans matrisi oluşturulur. Matris oluşturulurken imgeye ait gri seviye değerleri ölçeklenir, pikseller arasındaki mesafe ve açı dikkate alınır. GLCM, N boyutlu bir kare matris olup, matrisin her bir elemanı d mesafesindeki i ve j piksel değerli çiftin oluşum sayısını belirtir (P. Mohanaiah ve diğ., 2013).



Şekil 3.1 : Gri seviye eş oluşum matrisi.

3.3.6 Histogram

Bir imgenin gri seviye histogramı imgedeki piksellerin gri seviye ölçeğindeki dağılımlarını gösterir. Her piksel renk yoğunluğuna tekabül eden gri seviye kümeye dâhil edilir. Kümelere ait eleman sayıları ilgili gri seviyenin histogram değerine karşılık gelir. Özniteliğin oluşturulması için histogram değerlerinden ortalama, varyans, çarpıklık, basıklık, enerji, entropi değerleri hesaplanarak bir dizide birleştirilir.

3.4 Boyut Azaltımı

Veri toplama ve depolama yeteneklerindeki gelişmeler birçok bilim dalında aşırı bilgi yüklemesine yol açmıştır. Biyoloji, kimya, mühendislik, astronomi gibi deneysel yaşam bilimleri gibi alanlarda çalışan araştırmacılar, deneylerden gelen daha büyük gözlem sonuçları ile günlük olarak karşı karşıya kalmaktadır. Laboratuvar aletleri daha karmaşık hale gelip bir deney için yüzlerce ya da binlerce ölçümü rapor ederken, veri analizi yöntemleri zorlu bir görev olan yüksek boyutlu veriyle çalışma ile karşı karşıya kalırlar. Gözlem sayısındaki artış ve bundan daha etkili olarak her gözlemle ilişkili değişken sayısındaki artış sebebiyle birçok geleneksel veri analiz yönteminin başarı oranı düşmüştür. Genellikle veriler, büyük bir kısmı anlama katkısı olmayan ve kayda değer bilgi kaybı yaşanmadan, efektif bir şekilde, daha küçük sayıda değişken ile temsil edilebilecek yapıdadır. Bu azalmayı mümkün kılan matematiksel işlemlere boyut değişim teknikleri denir. İstatistik ve yapay öğrenme vb. alanlar tarafından geliştirilen bu teknikler güncel araştırma konularındandır. Boyut değişimi, kullanılan algoritmanın çalışma performansını arttırmanın yanı sıra sınıflandırma, kümeleme doğruluğunu arttırmak içinde kullanılan bir tekniktir. Gürültülü ya da hatalı giriş verileri çoğu kez algoritmaların beklenen performansının düşmesine sebep olur. Bilgi taşımayan veri sütunlarının silinmesi algoritmaların daha genel sınıflandırma alanları, kuralları bulmasını ve yeni veri ile genel olarak daha iyi performansa ulaşmasını sağlayabilir. İdeal olarak azaltılmış gösterim, verinin anlamlı gerçek boyutuna karşılık gelecek kadar boyuta sahip olmalıdır. Verinin gerçek boyutu, veride gözlemlenen özellikleri açıklamak için ihtiyaç duyulan minimum sayıdaki parametrelerdir. Boyut değişimi, çok boyutlu uzayın istenmeyen özellikleri ve boyutluluk lanetini azaltacağı içinde önemlidir. Boyut değişimi, öznitelik dönüşümü ve öznitelik seçimi olarak iki gruba ayrılabilir (Imola K. Fodor, 2002, L.J.P. Van Der Maaten ve diğ., 2008, C.O.S. Sorzano ve diğ., 2014, Rosaria Silipo ve diğ., 2014).

3.4.1 Öznitelik dönüşümü

Öznitelik dönüşümü, yeni öznitelik vektörünün oluşturulduğu bir süreçtir. Bu süreçte çok boyutlu uzayda yer alan vektör daha az ya da daha çok boyutlu uzayda temsil edilecek şekilde dönüştürülür. Öznitelik dönüşümü lineer olabileceği gibi lineer olmayan tekniklerle de yapılabilir. Boyut değişimine göre öznitelik azaltımı ve öznitelik çoğaltımı biçimleri bulunmaktadır. Dönüştürülmemiş öznitelik vektörünün $A_1, A_2, A_3, \dots, A_n$ değişkenlerini içerdiğini kabul edersek öznitelik çoğaltımı, değişkenler arasındaki ilişkilerle alakalı eksik bilgilerin keşfedilmesi ve ilave değişkenler çıkarımı veya oluşturulmasıyla öznitelik vektör uzayının büyütülmesidir. Öznitelik çoğaltımından sonra vektörde m adet değişken ilave olabilir $A_1, A_2, A_3, \dots, A_n, A_{n+1}, A_{n+2}, \dots, A_{n+m}$. Örneğin A_k değişkeni ($n < k \leq n+m$) orijinal vektördeki A_i ve A_j ($i \leq n, j \leq n$) değişkenlerinin mantıksal bir işleme tabi tutulması ile oluşturulmuş olabilir, A_1 değişkeninin genişliği ve A_2 değişkeninin uzunluğu temsil ettiği bir vektörde $A_1 * A_2$ mantıksal işlemi ile oluşturulan A_3 değişkeni alanı temsil edebilir (3.6). Öznitelik azaltımı, dönüştürülmemiş öznitelik vektöründen bazı fonksiyonel haritalamalar ile daha küçük boyutlu öznitelik vektörünün oluşturulması işlemidir. Öznitelik değişken azaltımından sonra $m < n$ olmak üzere, B_1, B_2, \dots, B_m değişkenlerini içeren yeni bir öznitelik vektörü oluşturulur. Bu dönüşümde A_1, \dots, A_n F_i dönüşüm fonksiyonunun girdilerini, B_i de çıktısını oluşturmaktadır (3.7) (Huan Liu ve diğ., 1998, Chris Ding ve diğ., 2002, Hanan Samet, 2006).

$$A_i = F_i(A_1, A_2, \dots, A_n) \quad (i > n) \quad (3.6)$$

$$B_i = F_i(A_1, A_2, \dots, A_n) \quad (i < n) \quad (3.7)$$

3.4.1.1 Temel bileşenler analizi

TBA (Temel Bileşenler Analizi, Principal Components Analysis, PCA) verideki örüntünün tanımlanması ve verinin bezerlik ve farklılıklarını vurgulayacak şekilde ifade edilmesinin bir yoludur. Verideki örüntü bulunduğunda veri boyutunun sayısı azaltılarak, içerdiği bilgiden fazla kayıp yaşanmadan verinin sıkıştırılması (boyutunun azaltılması) mümkündür. TBA aşağıdaki adımlardan oluşmaktadır;

1. Verinin merkezlenmesi.
2. Verinin normalize edilmesi.

3. Eigen ayrışımını hesaplanması.
4. Eigen vektör projeksiyonu ile yeni veri setinin oluşturulması.

Boyut azaltımı, dördüncü adımda tüm eigen vektörlerinin kullanılmaması ve seçilen en büyük eigen vektörlere göre yeni veri setinin oluşturulması ile gerçekleştirilir (Lindsay I Smith, 2002, Url6, 2016, Url7, 2016).

3.4.1.2 Doğrusal ayrıştırma analizi

DAA (Doğrusal Ayrıştırma Analizi, Linear Discriminant Analysis, LDA), aynı sınıfa ait örüntüler arasındaki dağılımı olabildiğince en küçük, sınıflar arasındaki dağılımı olabildiğince en büyük yapmayı hedefler. Böylece verilerin dağılımı daha iyi anlaşılabilir. DAA ve TBA'nın temel farkı, DAA'nın verileri sınıflandırırken, TBA'nın öznitelik sınıflandırması yapmasıdır. DAA'da orijinal örüntülerin konumu değiştirilmeksizin örüntüler arasında ayırım yapılır. Fakat TBA'da örüntüler öznitelik uzayına indirildiğinde hem şekilleri hem de konumları değişir. TBA sınıf etiketlerini kullanmadan veri seti içerisindeki varyansı maksimum yapacak temel birleşenleri bulmayı hedeflediği için gözetimsiz algoritma olarak tanımlanabilir. TBA'nın aksine, DAA sınıf etiketlerini kullanarak sınıflar arasındaki mesafeyi maksimum yapacak eksenleri temsil edecek doğrusal ayırıcıları bulmayı hedeflediği için gözetimli bir algoritmadır. DAA aşağıdaki adımlardan oluşmaktadır;

1. D boyutlu ortalama vektörlerin hesaplanması.
2. Dağılım matrislerinin hesaplanması.
3. Dağılım matrisleri için genelleştirilmiş özdeğer (eigenvalue) probleminin çözülmesi.
4. Yeni öznitelik alt uzay için doğrusal ayırıcıların seçilmesi.
5. Verinin yeni altuzay üzerindeki projeksiyonunun oluşturulması ile yeni veri setinin oluşturulması.

Boyut azaltımı, dördüncü ve beşinci adımda tüm doğrusal ayırıcıların kullanılmaması ve seçilen doğrusal ayırıcılara göre yeni veri setinin oluşturulması ile gerçekleştirilir (Kardi Teknomo, 2015, Sebastian Raschka, 2014, Murat Gök, 2011).

3.4.2 Öznitelik seçimi

İlgisiz ve gereksiz verilerin mümkün olduğunca belirlenmesi ve silinmesi işlemidir. Öznitelik seçimi yeni öznitelik üretilmemesi ile öznitelik dönüşümünden farklıdır. Orijinal öznitelik vektörünün bir alt kümesi seçilerek öznitelik uzayı küçültülür, öznitelik dönüşümünde ise öznitelik uzayı büyüyebilir. Öznitelik seçimi verinin boyutunu azaltarak öğrenme algoritmasının daha hızlı ve daha etkin çalışmasını sağlayabilir. Bu etki bazı durumlarda öznitelik sınıflandırıcısının hassasiyet artışı olarak bazı durumlarda ise hedef temsilin kolayca yorumlanabileceği modeli olarak gözlenebilir. İdealde öznitelik seçim yöntemleri, özniteliklerin altkümeleri içerisindeki 2^m rakip arasından, bazı değerlendirme fonksiyonlarına göre en iyi olanı bulmak için arama yaparlar. Ancak en iyi olanın bulunmaya çalışıldığı bu işlem orta ölçekli boyuta sahip öznitelik vektörleri için bile zahmetli, masraflı ve pratikte uygulanamaz olabilir. Sezgisel veya rastgele arama yöntemlerine dayalı diğer yöntemler, seçim başarı performansından ödün vererek hesaplama karmaşıklığını azaltmaya çalışırlar. Bu tür yöntemlerin tüm alt kümelerin araştırılmasının önüne geçilecek şekilde arama yapabilmeleri için başlangıç noktası, arama organizasyonu, değerlendirme stratejisi, durdurma kriteri olarak adlandırılan dört temel unsurun tanımlanması gerekmektedir. Başlangıç noktası, arama uzayında hangi noktadan başlanacağını ifade eder, boş küme, evrensel küme gibi başlangıç kümeleri seçilebilir. Arama organizasyonu, tüm alt kümeler içerisinde bulunan noktadan bir sonraki adaya geçiş için kullanılacak stratejileri belirlemektedir. Değerlendirme stratejisi, seçilen altkümelere ait değerlerin hesaplanması ve kıyaslanması için gerekli olan yöntemlerdir. Durdurma kriteri, öznitelik altkümeleri uzayında yapılan arama işleminin ne zaman duracağına karar verilmesini sağlayan unsurdur. Tüm altkümelerin araştırılmadığı bu yöntemlerde en iyi olanın bulunması garanti edilmemekle birlikte, temel unsurların probleme uygun seçilmesi ile en iyiye yakın sonuçların elde edilebilmesi mümkündür (Huan Liu ve diğ., 1998, Mark a. Hall, 1999, Jiliang Tang ve diğ., 2014).

3.4.2.1 Ardışık öznelik seçimi

SİS (Sıralı İleri Seçim, SFS, Sequential Forward Selection) en basit greedy arama algoritmasıdır. Aday çözüm kümesi boş olarak işleme başlanır ve nitelikler kümeye sırasıyla ilave edilir. İşleme, yeni eklenen nitelik sonucunda oluşan kümenin değerlendirme değeri, kümenin bir önceki durumundaki değerlendirme değerine göre arttığı sürece devam edilir. Ölçüm sonucu aynı ya da daha küçük olur ise ekleme yapılmaz bir sonraki niteliğe geçilir. Değer yükselmiyor ve denenecek nitelik kalmadı ise işlem sonlandırılır. Optimum çözüm altkümesi küçük olduğunda SİS tercih edilir, eklenen bir nitelik daha sonra çıkarılamaz.

SGS (Sıralı Geriye Seçim, SBS, Sequential Backward Selection) SİS'e benzer şekilde çalışır. Aday çözüm kümesi tüm nitelikleri içerecek şekilde işleme başlanır ve nitelikler kümeden sırasıyla çıkarılır. İşleme, çıkarılan nitelik sonucunda oluşan kümenin değerlendirme değeri, kümenin bir önceki durumundaki değerlendirme değerine göre arttığı veya aynı kaldığı sürece devam edilir. Ölçüm sonucu daha küçük olur ise çıkarma yapılmaz bir sonraki niteliğe geçilir. Değer azalıyor ve denenecek nitelik kalmadı ise işlem sonlandırılır. Optimum çözüm altkümesi büyük olduğunda SGS tercih edilir, çıkarılan bir nitelik daha sonra eklenemez (Ricardo Gutierrez Osuna, 2013).

3.4.2.2 ReliefF algoritması

ReliefF nitelik önem belirleme algoritmasında anahtar fikir birbirine yakın örnekler arasındaki ayrımın, değerlerine göre ne kadar iyi yapabildiği temel alınarak niteliklerin kalitesini tahmin etmektir. Bu amaçla, ReliefF algoritmasında rastgele seçilen R_i örneği için biri aynı sınıftan olan ve en yakın eşleşme H olarak adlandırılan, diğeri farklı sınıftan ve en yakın eşleşmeme olarak adlandırılan M olmak üzere en yakın iki komşusu aranır. Tüm A öznelikleri için R_i , M ve H değerlerine göre $W[A]$ kalite tahmini güncelleştirilir. Eğer R_i ve H örneklerine ait A nitelik değeri farklı ise A aynı etiketlere sahip iki örneği ayırdığı için kalite tahmini $W[A]$ değeri azaltılır. Diğer yandan, eğer R_i ve M örneklerine ait A nitelik değeri farklı ise A aynı farklı etiketlere sahip iki örneği ayırdığı için kalite tahmini $W[A]$ değeri arttırılır. Tüm süreç kullanıcı tarafından tanımlanan m değeri kadar tekrarlanır. Oluşan dizide en yüksek değere sahip olan nitelik en önemli olarak değerlendirilir (Marko Robnik Sikonja ve Igor Kononenko, 2003).

3.5 Sınıflandırma ve Değerlendirme

Bir tahmin problemi genel olarak model seçimi ve model değerlendirmesi kısımlarından oluşur. Model seçiminde en iyi modelin seçilmesi amacı ile rekabet ettirdiğimiz çeşitli modellerin performansları ölçülür. Verilen model kümesi $M=M_1, M_2, \dots, M_n$ olmak üzere, veri seti üzerinde en iyi performansı elde edecek modelin seçilmesi için, aynı öğrenme modelinin farklı karmaşıklık parametreleri (K-EYK için farklı K komşuluk değerleri, K-Ortalama için farklı K küme sayısı, vb.) kullanılarak ya da farklı öğrenme modelleri (Bayes, K-EYK, YSA, vb.) ile eğitim ve ölçüm süreçleri gerçekleştirilir. Yapılan ölçüm ve değerlendirmeler sonucunda, veri seti üzerinde en iyi performans değerlerine ulaşan modelin, yeni veri setleri üzerinde de en iyi performans değerlerine ulaşacağı kabulü ile uygun model belirlenir. Ölçüm işlemi için doğruluk, hassasiyet, vb. metrikler, değerlendirme içinde çapraz-doğrulama gibi doğrulama yöntemleri kullanılmaktadır (Amit C. Patel ve diğ., 2005, Hector Corrada Bravo ve diğ., 2010, Piyush Rai, 2011).

3.5.1 Naive Bayes sınıflandırıcısı

Adını Matematikçi Thomas Bayes'den alan, Bayes karar teorisine dayanan, ortalama riskin en aza indirildiği istatistiksel bir sınıflandırıcıdır. Bayes Teoremi, olasılıkları ve gözleme dayanan olasılık dağılımlarındaki parametreleri saptama yöntemidir. Naive Bayes sınıflandırmasında sisteme hangi sınıflara ait oldukları bilinen belirli bir oranda öğretilmiş veri sunulur. Öğretilmiş veriler üzerinde yapılan olasılık hesapları ile sisteme sunulan yeni test verileri, daha önce elde edilmiş olasılık değerlerine göre işletilir ve verilen test verisinin hangi sınıfa ne kadar olasılıkla ait olduğu tespit edilmeye çalışılır. Öğrenme veri setimiz s , bilinmeyen öznitelik vektörümüz x ve probleme ait m adet sınıf etiketlerimiz w_1, w_2, \dots, w_m verildiğinde, x 'in hangi sınıf etiketine ait olduğuna dair bilgi içeren koşullu olasılıklar denklem (3.8) ile hesaplanabilir. Öğrenme veri setine ait gözlem sayısı n ve i 'inci sınıf etiketine ait gözlem sayısı n_i olmak üzere $P(w_i)$ denklem (3.9) ve $P(x)$ denklem (3.10) ile hesaplanabilmektedir.

$$P(w_i | x) = \frac{P(x | w_i) \cdot P(w_i)}{P(x)}, \quad i=1, 2, \dots, m \quad (3.8)$$

$$P(w_i) = \frac{n_i}{n}, i=1, 2, \dots, m \quad (3.9)$$

$$P(x) = \sum_{i=1}^m P(x \setminus w_i).P(w_i) \quad (3.10)$$

Bulunan koşullu olasılıklardan denklem (3.11) yardımı ile bilinmeyen öznitelik vektörü için hedef sınıf etiketi tahmini yapılır.

$$\text{Maksimum}(P(w_i \setminus x)), i=1, \dots, m \{ x, i \text{ sınıfına aittir} \quad (3.11)$$

Bayes sınıflandırıcılarından en basiti sayılan Basit (Naive) Bayes Sınıflandırıcısında, eğitim öznitelik vektöründe yer alan tüm niteliklerin verilen sınıf bağlamında birbirinden bağımsız olduğu ve normal dağılıma sahip olduğu varsayılmaktadır. Bu kabul ile $P(x \setminus w_j)$ denklem (3.12) ile ve olasılık yoğunluk fonksiyonu denklem (3.13) ve (3.14) ile hesaplanmaktadır, k adet nitelik içeren öznitelik vektöründe x_i i numaralı niteliği temsil etmektedir (Rahman Usta, 2014, John M. McNamara ve diğ., 2006, Sergios Theodoridis ve diğ., 2006, Richard O. Duda ve diğ., 1997).

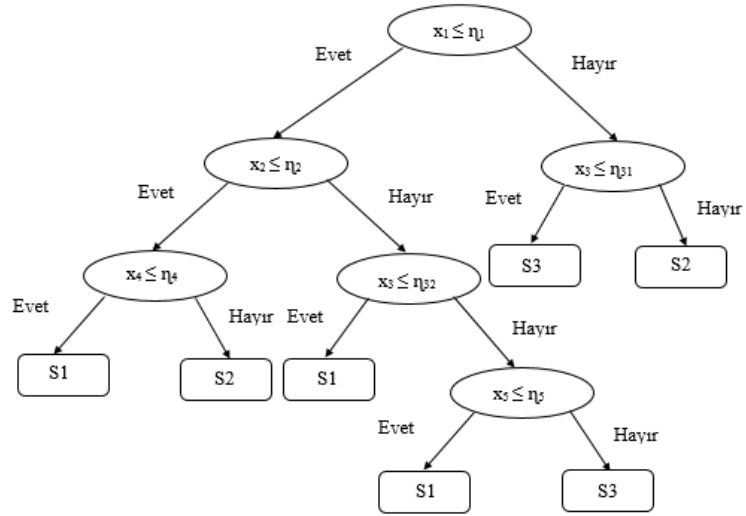
$$P(x \setminus w_j) = \prod_{i=1}^k P(x_i \setminus w_j), j=1, \dots, m \quad (3.12)$$

$$P(x_i \setminus y_i) = g(x_i; \mu_{c_j}, \sigma_{c_j}) \quad (3.13)$$

$$g(x_i; \mu_{c_j}, \sigma_{c_j}) = \frac{1}{\sigma_{c_j} \sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x_i - \mu_{c_j}}{\sigma_{c_j}} \right)^2} \quad (3.14)$$

3.5.2 C 4.5 sınıflandırıcısı

Karar ağacı sınıflandırıcıları sinyal sınıflandırma, karakter sınıflandırma, uzaktan algılama, tıbbi teşhis, uzman sistemler ve konuşma sınıflandırma gibi birçok geniş alanda başarılı bir şekilde kullanılmaktadır. Karar ağacı sınıflandırıcılarının en önemli özelliği, karışık karar verme süreçlerini daha basit kararlar topluluğuna dönüştürmektir ve böylece bu yolla alınan sonuç çözümünün istenilen çözüme benzemesi istenir. Örnekler Şekil 3.2 de gösterildiği gibi ağacın kökünden başlayıp yaprağa doğru gezinerek sınıflandırılır. Burada yol ise her bir düğümdeki bölünme şartının sonucuna göre belirlenir. Ağacın kökünden başlanır ve köke karşılık gelen öznitelik ele alınır. Ardından bu öznitelikğin gözlenen değerinin hangi çıkış dalına karşılık geldiği bulunur. Yol üzerindeki bir sonraki düğüm seçilmiş olan dalın sonunda olan düğümdür. Aynı işlemler bu düğüm içinde tekrarlanır ve yaprağa ulaşıncaya kadar ağaçta gezinilir. Kategorik veriler ile çalışan ID3 algoritmasından farklı olarak C4.5 algoritması sayısal veriler ile çalışma olanağı sağlar (S. Rasoul Safavian ve diğ., 1991, Lior Rokach, 2010).



Şekil 3.2 : Örnek bir karar ağacı.

3.5.3 K-EYK sınıflandırıcısı

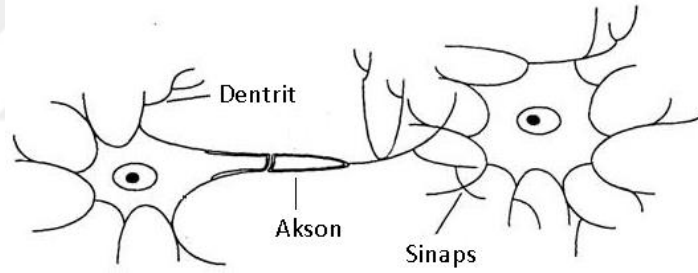
K-EYK danışmanlı, parametrik olmayan tembel bir öğrenme algoritmasıdır. Parametrik olmayan öğrenme algoritmalarında veri dağılımı üzerinde herhangi bir varsayım yapılmaz. Bu yüzden veri dağılımı hakkında az veya hiç ön bilgi olmadığı durumlarda tercih edilir. Gerçek dünyadaki birçok uygulamaya yönelik veri setinin, teorik varsayımlara uymayı bu tür yöntemlerin önemini arttırmaktadır. K-EYK algoritmasında eğitim veri setinden bir genelleme, çıkarım yapılmadığı için eğitim safhası bulunmamaktadır, tembel olarak adlandırılma sebebi de bu özelliğidir. Eğitim safhasının çok hızlı olmasına karşın test safhasında tüm veri setine ihtiyaç duyulmaktadır ve tahmin, veri setinin tamamı kullanılarak yapılabilmektedir. Buda test safhasında tüm veri setini tutabilecek kadar hafıza, işleyecek kadar zaman ve işlem gücü maliyeti anlamına gelmektedir. K-EYK algoritmasında veri setinde bulunan noktaların metrik uzayda bulunduğu varsayılır ve çoğunlukla Öklid olmak üzere noktalar arasındaki mesafeyi tanımlayan Manhattan, Hamming, vb. gibi uzaklık veya benzerlik ölçüleri ile kullanılır. Algoritmanın adına da yerleşmiş olan K değeri tahmin işlemine etki edecek komşu sayısını ifade eder. Farklı karar verme yöntemlerine göre K değerinin değişebilmesine karşın yaygın kabul K'nın tek sayı olarak seçilmesidir. K'nın 1 olarak seçildiği özel durumunda algoritma "En Yakın Komşu Algoritması" olarak adlandırılmaktadır. Sınıflandırma için kullanılan yaklaşım; sınıflandırılmak istenen yeni örneğin daha önceden hangi sınıfa ait olduğu bilinen örneklerden K tanesine yakınlığına bakılır. K sayısının değeri sınıflandırma açısından önemlidir. K değeri küçük olursa model gürültüden etkilenir, büyük olursa da değerlendirilemeyen örnekler oluşur, sınıflandırma performansına doğrudan etki etmektedir. Algoritma aşağıdaki şekilde özetlenebilir;

- K değerini belirle, yeni öznitelik vektörü için tüm vektörler ile uzaklık değerlerini hesapla.
- En yakın K tanesini belirle ve belirlenen K tanesinin sınıf etiketleri arasından oylama ile seçim yap.
- Etiket seçim için, en çok tekrar eden sınıf, en yakın vektöre ait sınıf, ağırlıklı oylama ile en yüksek puanı alan sınıf, vb. yöntemler kullanılabilir.

(Richard O. Duda ve diğ., 1997, Ola Söder, 2008, Saravanan Thirumuruganathan, 2010, Oliver Sutton, 2012).

3.5.4 YSA sınıflandırıcısı

Genellikle sinir ağırları olarak adlandırılan yapay sinir ağırları üzerine çalışmalar beyinin geleneksel dijital bilgisayarlardan tamamen farklı bir şekilde işlem yaptığının keşfedilmesi ile başlamıştır. Beyinin organik elemanı olan nöronlar silikon mantık kapılarına kıyasla daha yavaştır, silikon yonga içerisinde olaylar 10^{-9} saniyede gerçekleşirken nöron içerisinde 10^{-3} saniyede meydana gelmektedir. Buna rağmen şaşırtıcı sayıdaki nöron ve nöronlar arası bağlantılar sayesinde, kompleks, doğrusal olmayan ve paralel bilgi işleme yeteneğine sahip olan beyin, örüntü tanıma gibi belirli hesaplamaları gerçekleştirmek üzere nöronları organize ederek en hızlı dijital bilgisayardan bile birçok kat daha hızlı işlem yapabilmektedir. Yapay nöron, doğal nörondan esinlenilerek oluşturulmuş bir işlem birimidir. Doğal nöronlar (Şekil 3.3) sinyalleri zarlarına veya dentritlerine yerleşmiş sinapslar yoluyla alırlar. Alınan sinyaller belirli bir eşiği aşacak kadar güçlü olduğunda nöron aktive olur ve sinyali aksonu boyunca yayar. Oluşturulan bu sinyal diğer sinapslara gönderilecek başka nöronları da aktive edebilir.

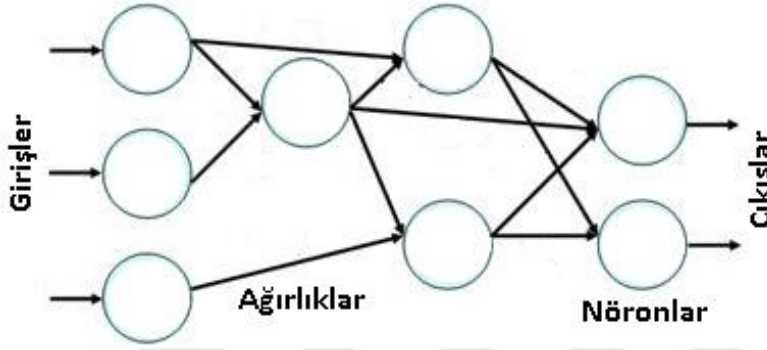


Şekil 3.3 : Doğal nöronun yapısı.

Yapay nöron modellenirken doğal nöronların karmaşıklığı oldukça azaltılmıştır. Yapay nöron (Şekil 3.4) temelde, sinapsları temsil eden girişleri, sinyal gücünü temsil eden giriş ağırlıklarını, nöron eşik seviyesini belirleyen aktivasyon fonksiyonunu ve nöronun aksona gönderdiği çıkış sinyalinin gücünü belirleyen çıkış fonksiyonunu içermektedir. YSA lar bilgiyi işlemek amacı ile yapay nöronları birleştirirler. Yapay sinir ağırları (Şekil 3.5), basit işlem ünitesi olan nöronlardan ve nöronlar arasındaki yönlü ve ağırlıklı bağlantılardan oluşur. Bağlantının ağırlığı ne kadar büyük olursa iletilen sinyalde o kadar güçlü olur. Ağırlıklar negatife olabilir, bu durumda sinyal ağırlık tarafından zayıflatılır. Ağırlık ve eşik değerlerine bağlı olarak nöronlarda yapılan hesaplamalar farklılık gösterecektir, bu özellik kullanılarak belirli giriş değerleri için istenen çıkış değerini elde edebilmek için ağırlıklar veya eşik değerleri ayarlanabilir.



Şekil 3.4 : Yapay nöronun yapısı.



Şekil 3.5 : Yapay sinir ağının yapısı.

Verilen eğitim veri setine karşılık istenen çıkış değerlerinin elde edilmesi için ağda bulunan tüm ağırlıkları veya eşik değerlerini değiştirecek olan algoritmanın bulunması ve kullanılmasına ağın eğitimi denir. Ağ giriş ve çıkışlarına göre ağırlıklara veya eşik değerlerine tekrarlı uygulanan ayarlamalar ile ağ bulunduğu ortamı öğrenir ve tahmin doğruluğunu artırır. Öğrenme süreci aşağıdaki şekilde özetlenebilir;

- YSA bir ortam tarafından uyarılır.
- Yapılan uyarım sonucunda YSA değişime uğrar.
- Değişim sonucunda YSA ortama farklı bir şekilde yanıt verir.

YSA'ların eğitimi için kullanılacak çeşitli algoritmalar bulunmaktadır. Hata değerinin hesaplanması, ağırlık ve eşik değerlerinin değiştirilmesi işlemlerinde birbirinden farklılık gösteren algoritmaların farklı avantaj ve dezavantajları bulunabilmektedir. Deneysel bilginin depolanması ve kullanılması için birçok işlem ünitesinin kullanılması, bilginin tüm ağa dağıtılarak saklanması yönleri ile YSA'ların çalışma şekli beyinin çalışma şekline benzerlik göstermektedir (Simon Haykin, 1999, Carlos Gershenson, 2003, David Kriesel, 2005, M. Hajek, 2005).

3.5.5 Performans değerlendirme

Seçilen modelin eğitimi işlemi ardından, sahadaki gerçek performansının tahmin edilebilmesi için daha önce kullanılmamış veri setleri ile modelin genelleştirme performansının ölçülmesi gerekmektedir. Gerçek performansın tahmin edilmesinde oldukça iyi sonuçlar verdiği doğrulanmış ve kullanımı yaygın metotlardan biri çapraz doğrulamadır. Eğitim kümesi için verilmiş m adet örnek için k kat çapraz doğrulama için adımlar;

1. Eğitim kümesi örneklerini rastgele sırala.
2. Kümeyi eşdeğer k alt kümeye böl (her birine yaklaşık m/k eleman düşmeli).
3. $i = 1, \dots, k$
 - i alt kümesi haricindeki tüm örnekler ile sınıflandırıcıyı eğit.
 - i alt kümesindeki örnekler ile sınıflandırıcıyı test et.
 - Yanlış sınıflandırılan eleman sayısı n_i yi hesapla.
4. Sınıflandırıcının genelleştirilmiş tahmini hata oranı olarak **(3.15)** hesapla.

$$E = \frac{\sum_{i=1}^k n_i}{m} \quad (3.15)$$

Daha doğru bir tahmin elde etmek için k -kat çapraz doğrulama işlemi rastgele düzenleme ile birkaç kez (t) tekrarlanabilir ve ortalama hata (e), standart sapma (σ) **(3.16)**, **(3.17)** formülleri ile hesaplanarak kullanılabilir.

$$e = \frac{\sum_{j=1}^t E_j}{t} \quad (3.16)$$

$$V = \frac{\sum_{j=1}^t (E_j - e)^2}{t-1}, \quad \sigma = \sqrt{V} \quad (3.17)$$

Öznitelik vektörleri x_1, \dots, x_n için c_1, \dots, c_m sınıf etiketi eşleştirmesi yapan bir sınıflandırma işlemi aşağıdaki farklı kategorilere ayrılabilir. İkili sınıflandırma: giriş, birbiri ile örtüşmeyen c_1 ve c_2 sınıflarından sadece birine atanır, ikili sınıflandırma en popüler sınıflandırma problemlerindedir. Çok sınıflı sınıflandırma: giriş, birbiri ile örtüşmeyen n sınıftan sadece birine atanır. İkili sınıflandırmadan farklı olarak sınıf sayısı n ikiden büyüktür. Çok etiketli sınıflandırma: giriş birbiri ile örtüşmeyen n sınıftan birkaç tanesine atanabilir. Hiyerarşik sınıflandırma: giriş, alt sınıflara veya guruplara bölünerek hiyerarşik bir yapı oluşturmuş n sınıftan sadece birine atanır. Oluşturulan hiyerarşi sınıflandırma esnasında değiştirilemez. Çok sınıflı bir sınıflandırmada kullanılacak ölçümler ve hesaplama yöntemleri çizelge 3.1'de gösterilmektedir.

Çizelge 3.1 : Çok sınıflı bir sınıflandırıcı için performans ölçümleri.

Ölçüm	Formül
Ortalama Doğruluk	$\frac{\sum_{i=1}^l \frac{da_i + de_i}{da_i + de_i + ya_i + ye_i}}{l}$
Hata Oranı	$\frac{\sum_{i=1}^l \frac{ya_i + ye_i}{da_i + de_i + ya_i + ye_i}}{l}$
Kesinlik $_{\mu}$	$\frac{\sum_{i=1}^l da_i}{\sum_{i=1}^l (da_i + ya_i)}$
Duyarlık $_{\mu}$	$\frac{\sum_{i=1}^l da_i}{\sum_{i=1}^l (da_i + ye_i)}$
Kesinlik $_M$	$\frac{\sum_{i=1}^l \frac{da_i}{(da_i + ya_i)}}{l}$
Duyarlık $_M$	$\frac{\sum_{i=1}^l \frac{da_i}{(da_i + ye_i)}}{l}$

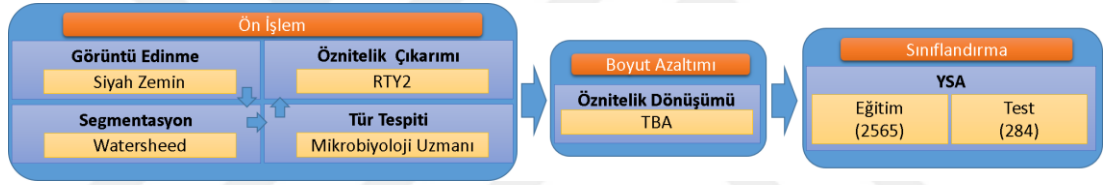
Belirlenen c_i sınıfı için doğru artı (da), doğru eksi(de), yanlış artı (ya) ve yanlış eksi (ye) değerleri sınıflandırmaya ait karmaşıklık matrisi oluşturularak hesaplanır. Kesinlik ve duyarlılık ölçümlerinin hesaplanmasında mikro ve makro olmak üzere iki farklı yaklaşım bulunmaktadır. Makro ölçümlerde (formüllerde M harfi ile gösterilmiştir) her sınıf için ilgili ölçüm hesaplanarak ortalaması alınır, mikro ölçümde ise (formüllerde μ harfi ile gösterilmiştir) ilgili ölçüme ait değerler tüm sınıfların değerlerinin kümülatif toplamı alınarak hesaplama yapılır. Makro hesaplama yaklaşımına ait sonuçlarda tüm sınıflar eşit temsil edilirken mikro hesaplama yaklaşımında büyük sınıflar daha fazla temsil oranına sahiptir (Marina Sokolova ve diğ., 2009, Vincent Labatut ve diğ., 2011).





4. GELİŞTİRİLEN SİSTEM

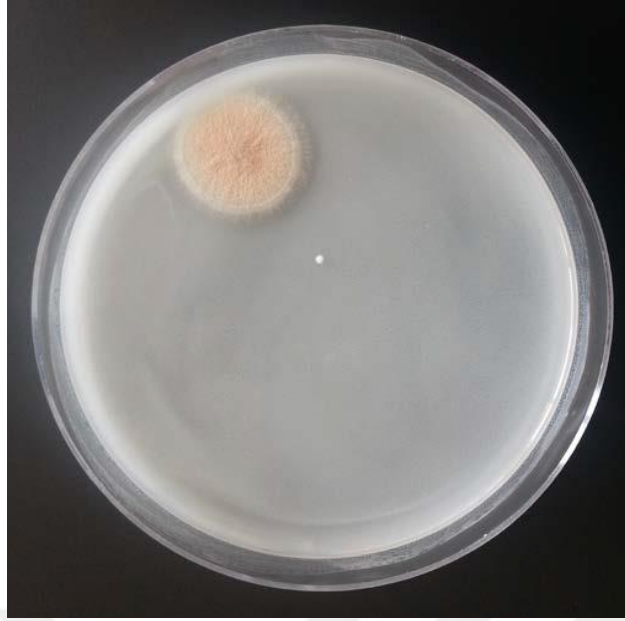
Yapay öğrenme modeli temel olarak veri toplama, ön işleme, öznitelik çıkarımı ve sınıflandırma ana başlıklarından oluşmaktadır. Koloni görüntülerinden morfolojik özellikler kullanılarak makine öğrenmesi teknikleri ile tür tahminlerinin yapılabilmesi için geliştirilen, siyah zeminde görüntü edinme, watershed ile segmentasyon, RTY2 ile öznitelik çıkarımı, TBA öznitelik dönüşümü ve YSA ile sınıflandırma aşamalarından oluşan sisteme ait mimari görüntüsü Şekil 4.1 de gösterilmektedir.



Şekil 4.1 : Sistem mimarisi.

4.1 Görüntülerin Elde Edilmesi

Uygulamada kullanılacak paylaşılan bir veri tabanı bulunmadığından veri tabanının oluşturulması için Marmara bölgesinde yer alan çeşitli süt işleme fabrikalarında petri görüntüsü toplama çalışması yapılmıştır. Örnek numuneler uygun şartlarda, ISO 6611/ IDF 94:2004 ve IDF 73B:1998 standartları kullanılarak oluşturulmuştur. Fotoğraflar 8mp kamera ile gün ışığı ve 3 noktadan açılmalık led kaynak ile aydınlatılmış ortamda, siyah zemin kullanılarak çekilmiştir. Farklı aydınlatma (üstten, alttan, led ışık kaynak, floresan ışık kaynağı, gün ışığı) ve zemin (siyah, beyaz, beyaz ışıklı) alternatifleri ile denenerek, Şekil 4.2 de görüldüğü gibi en az gürültülü fotoğrafların gün ışığında 3 noktadan aydınlatma ve siyah zemin ile elde edildiği tespit edilmiştir. NikonD80 kamera ve 850 nmdalga boyu (yakın kızıl ötesi) aydınlatma ile çekimler yapılmıştır fakat bu tür aydınlatmanın küf, maya, bakteri görüntülerinde farklılık oluşturmadığı tespit edilmiş ve başarılı sonuç alınamamıştır.

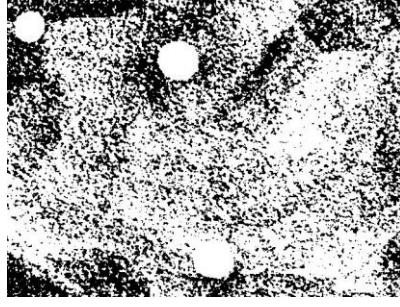


Şekil 4.2 : Fotoğraf örneği.

4.2 Nesnelerin Ayrıştırılması

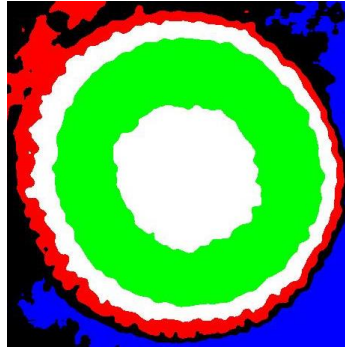
Nesne segmentasyon algoritmalarının kodlanması ve uygulanması için MATLAB ortamı kullanılmıştır. Algoritmaların kıyaslanmasında insan müdahalesi olmadan, görüntüde bulunan nesnelere en çoğunu ve hedef nesnenin görüntüsünü en iyi kapsayacak olması kriterlerini sağlayarak olmaları değerlendirilmiştir. Çeşitli segmentasyon yöntemleri küf, maya ve bakteri kolonileri içeren görüntüler üzerinde, algoritma için gerekli parametrelerin farklı değerleri ve otomatik parametre seçimi ile deneyerek, bazı segmentasyon yöntemlerinin diğerlerine kıyasla daha yüksek başarı oranlarına sahip olduğu tespit edilmiştir.

Maya ve bakteri kolonileri besiyeri ile benzer görüntüler oluşturabilmektedir. Bu benzerlik eşikleme tabanlı segmentasyonun performansını düşürmesine sebep olmaktadır. Otsu algoritması ile yaptığımız çalışmalar sonucunda besiyeri görüntüsüne benzeyen nesnelerin olduğu imgelerde eşik seviyesinin küçük olmasının gürültü seviyesini arttırdığı, Şekil 4.3 de görüldüğü gibi besiyeri ile nesne görüntülerinde birleşmelerinin meydana gelerek nesne görüntülerini bozduğu tespit edilmiştir. Eşik seviyesinin büyük olması da koloni görüntülerinin tespit edilememesine sebep olmaktadır. Yapılan çalışmalar farklı renk ve tonların bulunduğu imgelerde insan müdahalesi olmadan eşik seviyesi tespiti ve imge segmentasyonu başarısının düşük olduğunu göstermektedir.



Şekil 4.3 : Otsu algoritması maya-zemin eşikleme problemi görüntüsü.

Gabor filtreleri ile doku segmentasyonu tekniğini kullanarak yaptığımız çalışmalar, tekniğin iyi sonuçlar verdiğini göstermesinin yanında, farklı dokulara sahip imgelerde, insan müdahalesi ile küme sayısının belirlenmesi gerektiği, imgede bulunan küme sayısından daha az küme sayısı ile kümeleme algoritması çalıştırıldığında küf, maya ve bakteri görüntülerinde birleşme, daha fazla küme sayısı ile işlem yapıldığında ise nesne görüntüsünün parçalanarak bozulduğunun tespit edilmesini sağlamıştır. Ayrıca, farklı doku bölgeleri oluşturabilen küf görüntülerinin Şekil 4.4 de görüldüğü gibi parçalara ayrılarak birden fazla küf görüntüsü olarak algılandığı görülmüştür. Nesne görüntülerinin insan müdahalesi olmadan ve herhangi bir bozulmaya uğramaksızın besiyeri görüntüsünden ayrılması, yapılacak olan tür tahmini ve elde edilecek tür sayısını direk olarak etkilediği için farklı dokulu nesne ve çoklu tür içeren imgelerde Gabor filtreleri ile imge segmentasyonu sistem performansını düşürücü etkiye sahiptir.



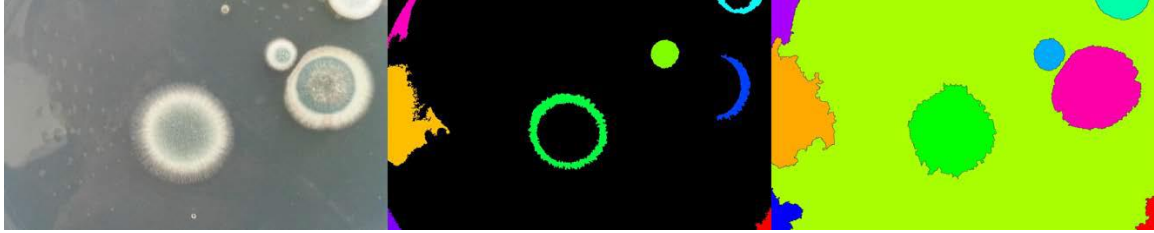
Şekil 4.4 : Çok dokulu küf nesnesi Gabor segmentasyon görüntüsü.

K-ortalamlar ve renk tabanlı algoritma ile yaptığımız çalışmalar ile farklı renk bölgeleri içerebilen küf kolonilerinin, renk tabanlı segmentasyon sonucunda Şekil 4.4 de görüldüğü gibi görüntülerinin bozulduğu tespit edilmiştir. Kolonilerin içerebileceği renk sayısı farklılık gösterdiği için küme sayısı tahmin edilememekte ve insan müdahalesi gerekmektedir.



Şekil 4.5 : Çoklu dokulu küf nesnesi renk tabanlı segmentasyon görüntüsü.

Objelerin ayrıştırılması için yapılan çalışmalarda, en başarılı sonuçlar Şekil 4.7 de görüldüğü gibi "Watershed Segmentasyon" yöntemi ile elde edilmiştir. Yapılan denemelerde algoritmanın Şekil 4.6 da görüldüğü gibi küf, maya, bakteri nesneleri haricinde sayım dışı olarak değerlendirilecek çok sayıda alan tespit ettiği görülmüştür. Oluşturulan bu nesnelere küf, maya, bakteri nesnelere tespitini engellemekle birlikte işleme tabi tutulması gereken gereksiz nesnelere işlem yükünü arttırmakta ve sayım dışı nesnelere değerlendirme doğruluğunun önemini arttırmaktadır.



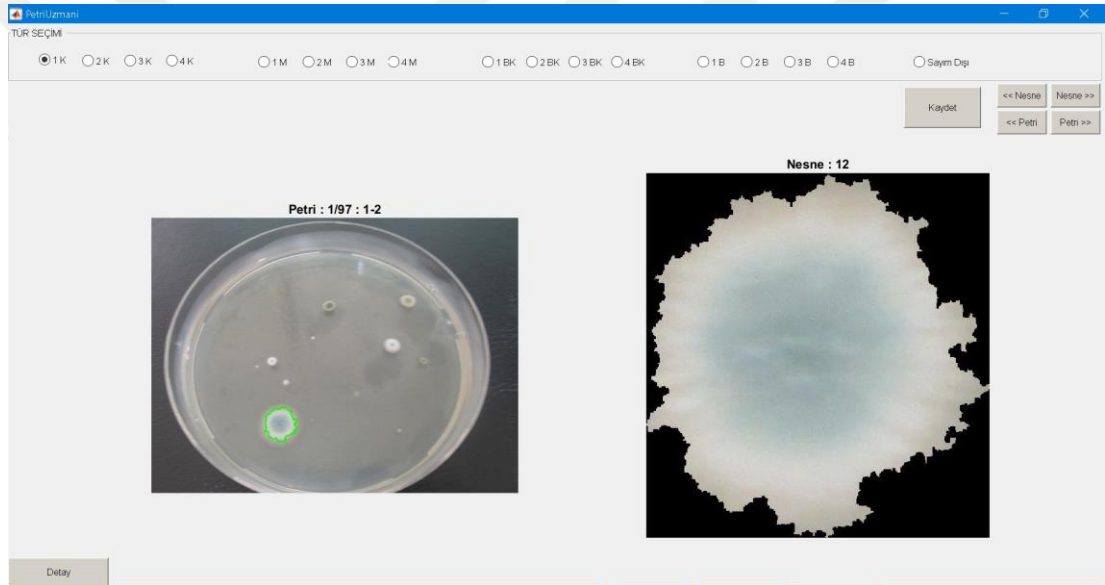
Şekil 4.6 : Watershed segmentasyon.



Şekil 4.7 : Ayrıştırılmış küf kolonileri.

4.3 Nesne Özniteliklerinin Çıkarılması

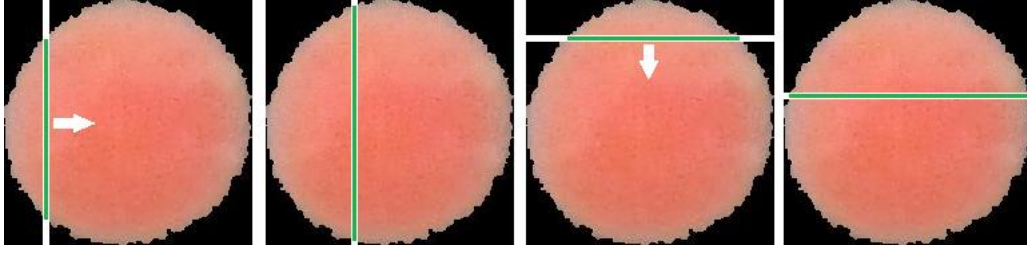
Nesne özniteliklerinin çıkarılması için algoritmaların kodlanması ve uygulanması MATLAB ortamında yapılmıştır. Küf maya ve bakteri imgelerinden özniteliklerin oluşturulması için tüm imgelere Watersheed segmentasyonu uygulanarak oluşturulan 2849 adet nesne görüntüsü kullanılmıştır. Nesne görüntülerine ait sınıf etiketlerinin mikrobiyoloji uzmanları tarafından belirlenerek veri tabanının oluşturulması için MATLAB ortamında “Petri Uzmanı” olarak adlandırılan program geliştirilmiştir. Öznitelik çıkarım tekniği olarak fraktal geometri, HOG gibi mevcut yöntemler ve yeni geliştirilen yöntemler kullanılmıştır.



Şekil 4.8 : Petri Uzmanı ekran görüntüsü.

4.3.1 Şekil tabanlı yöntem 1

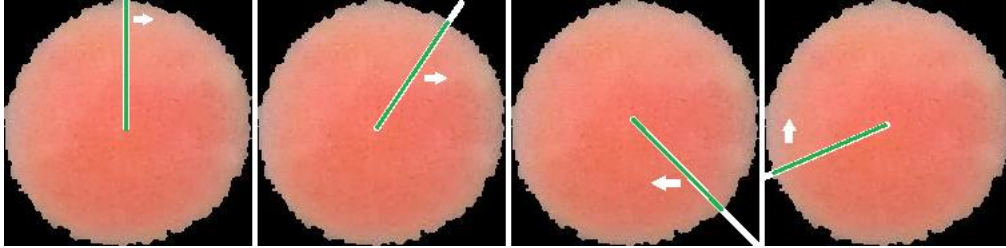
Küf, maya ve bakterilerin sınıflandırılmasında şekil bilgisi önemli bir sınıflandırma aracıdır. İnsan gözü değerlendirme yaparken, koloni görüntüsündeki farklı geometrik şekiller ve kenar şekli değişimleri, verilecek kararı etkilemektedir. Şekil bilgisinden yararlanarak öznitelik geliştirimi için koloni görüntüsü üzerinde, belirlenen sütun için tüm satırları kapsayan bir doğru belirledik (Şekil 4.9), doğruyu görüntünün ilk sütunundan son sütununa kadar soldan sağa doğru hareket ettirerek içerdiği siyah harici piksellerin sayısını ardışık bir dizi haline getirdik. Aynı işlemi ilk satırdan son satıra kadar yukarıdan aşağıya doğru tekrarladık ve oluşan iki diziyi birleştirdik. Elde ettiğimiz sinyal ile ilgili görüntüye ait öznitelik vektörümüzü oluşturduk.



Şekil 4.9 : Şekil tabanlı yöntem 1 ile öznelik oluşturma.

4.3.2 Şekil tabanlı yöntem 2

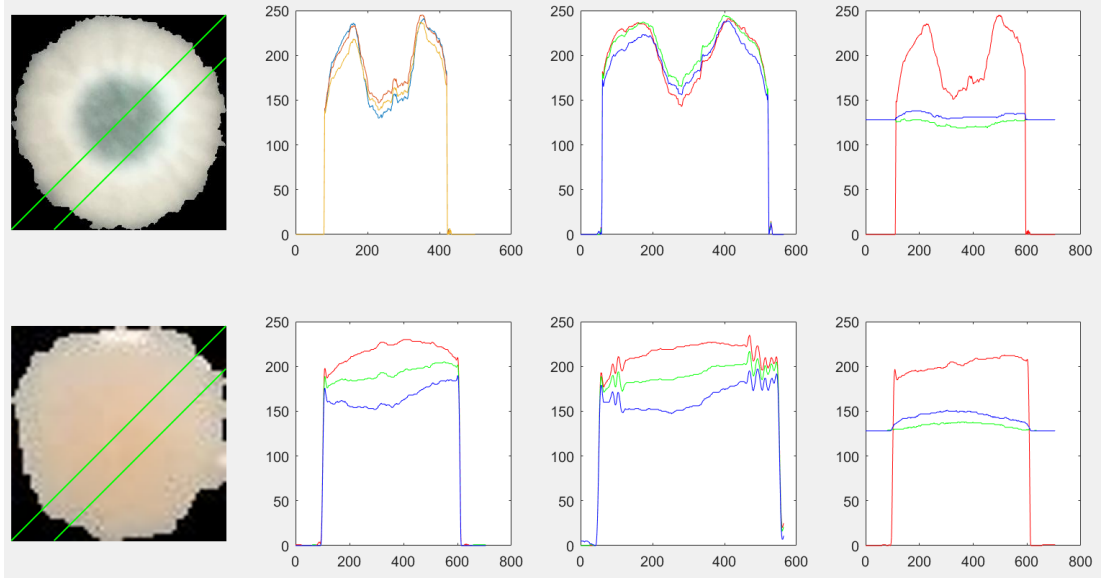
Küf, maya ve bakterilerin sınıflandırılmasında şekil bilgisi önemli bir sınıflandırma aracıdır. İnsan gözü değerlendirme yaparken, koloni görüntüsündeki farklı geometrik şekiller ve kenar şekli değişimleri verilecek kararı etkilemektedir. Şekil bilgisinden yararlanarak öznelik geliştirimi için koloni görüntüsü üzerinde, görüntü merkezinden kenar noktasına uzanan bir doğru belirledik (Şekil 4.10). Doğruyu 1'er derecelik açılar ile 360 derece saat yönünde döndürerek içerdiği siyah harici piksellerin sayısını ardışık bir dizi haline getirdik. Elde ettiğimiz sinyal ile ilgili görüntüye ait öznelik vektörümüzü oluşturduk.



Şekil 4.10 : Şekil tabanlı yöntem 2 ile öznelik oluşturma.

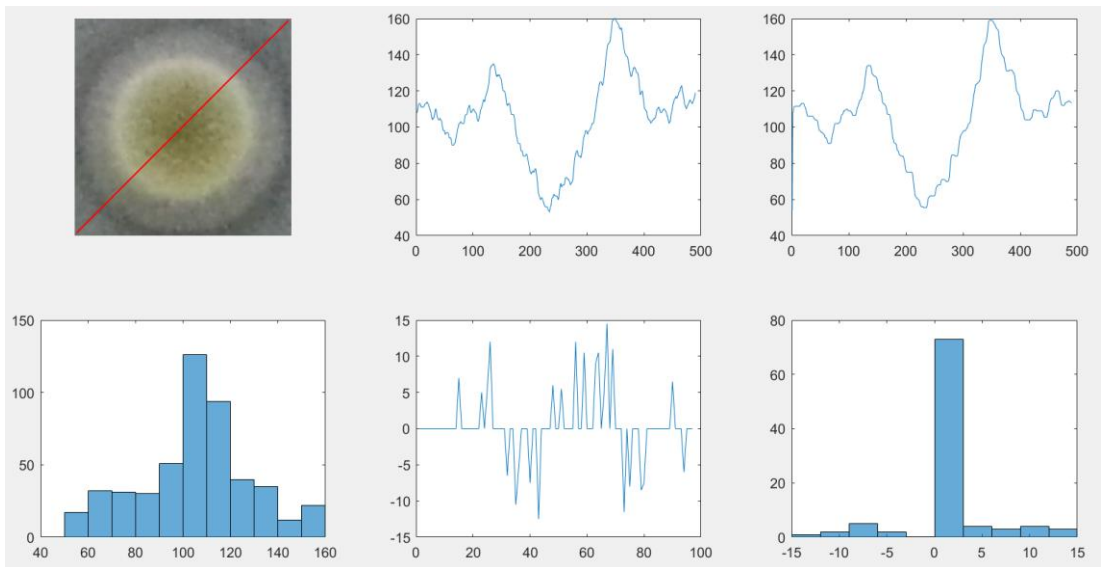
4.3.3 Renk tabanlı yöntem 1

Küf, maya ve bakterilerin sınıflandırılmasında renk bilgisi önemli bir sınıflandırma aracıdır. İnsan gözü değerlendirme yaparken, koloni görüntüsündeki farklı renk sayısı, renk değişim ve geçişleri verilecek kararı etkilemektedir. Küf ve maya koloni görüntüsünde belirlenen iki farklı doğru boyunca RGB ve LAB renk uzaylarındaki değişim grafikleri Şekil 4.11'de gösterilmektedir.



Şekil 4.11 : Farklı konum ve renk uzaylarındaki renk değişim eğrileri.

Renk bilgisinden yararlanarak öznelik geliştirimi için koloni görüntüsü üzerinde bir doğru belirledik, belirlediğimiz doğru boyunca var olan ve R, G, B olarak adlandırılan üç kanal renk bilgisi ayrıştırdık. Ayrıştırılan her kanalın renk sinyalindeki aşırı değişim gösteren gürültülü veriden kurtulmak için onuncu dereceden orta değer filtresi uyguladık. Elde ettiğimiz sinyalin türevini alarak eşik seviyesi uyguladık. Eşik seviyesinin altında kalan değerleri sıfıra, üstünde kalan değerleri ise kendi değerlerine eşitledik ve sinyalin histogram eğrisini oluşturduk. Bir renk kanalına ait histogram bilgisinin oluşturulması için işlem basamaklarına ait görüntü Şekil 4.12’de yer almaktadır.

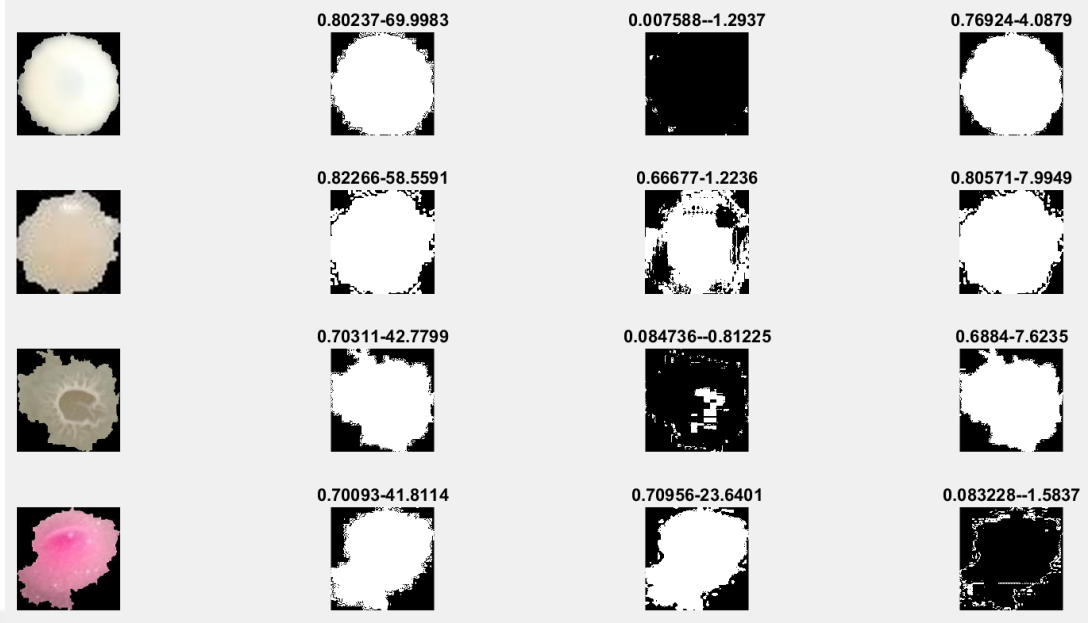


Şekil 4.12 : Renk tabanlı yöntem 1 ile öznelik oluşturma.

Her renk kanalına ait histogram bilgilerini birleştirerek belirlediğimiz doğru boyunca var olan renk değişimlerini kullanarak özneliği oluşturduk. Her koloni görüntüsü için belirlenen farklı doğruların özneliklerinin birleştirilmesi ile görüntüye ait öznelik vektörü oluşturma işlemi tamamlanmıştır. Geliştirdiğimiz renk tabanlı yöntem 1 e ait iki farklı öznelik vektörü oluşturulmuştur. İlkinde köşegen doğruları (RTY1.1), ikincisinde köşegen doğruları ile birlikte kenar orta nokta birleşim doğruları birlikte (RTY1.2) kullanılmıştır.

4.3.4 Renk tabanlı yöntem 2

Küf, maya ve bakterilerin sınıflandırılmasında renk bilgisi önemli bir sınıflandırma aracıdır. İnsan gözü değerlendirme yaparken, koloni görüntüsündeki farklı renk sayısı, renk değişim ve geçişleri verilecek kararı etkilemektedir. İnsan gözünde algılamaya etkisi olan renk, parlaklık, doygunluk, vb. özelliklerin küf, maya, bakteri, bakteri koliform koloni görüntülerinde etkisinin araştırılması için imgeler farklı renk uzaylarına dönüştürülerek oluşan sinyaller incelenmiştir. Yapılan çalışma sonucunda LAB, HSV, NTSC renk uzaylarındaki farklı bileşenlere ait sinyallerin farklı mikroorganizma koloni görüntüleri için ayırt edici değişimlere sahip olduğu tespit edilmiştir. Renk uzaylarından elde edilen bilgiler ile öznelik vektörünün oluşturulması için her renk uzayının her bileşeninden iki farklı değişken türetilmiştir. Bunlardan ilki, imgeye ait sinyal (parlaklık, renk, ...) gücünü temsil eden piksel ortalamasıdır. İlgili renk uzayının her kanalı için ortalama değer alınarak bir dizide birleştirilmiştir. İkinci hesaplanan değer, eşik seviyesini aşan piksellerin alanını temsil eden, %50 eşik seviyesini aşan piksel sayısının tüm piksel sayısına oranıdır. İlgili renk uzayının her kanalı için alan oranı hesaplanarak bir dizide birleştirilmiştir. Sayıların hesaplanması için eşik seviyesinin altında kalan pikseller sıfır ve eşik seviyesi üstünde kalan pikseller bir olarak kabul edilerek işlem yapılmıştır. Küf, maya, bakteri ve bakteri koliform görüntüleri için LAB renk uzayında yapılan işlemin görselleştirilmiş hali Şekil 4.13'de gösterilmektedir. Oluşturulan iki dizi birleştirilerek renk uzayına ait öznelik ve üç renk uzayına ait öznelikler birleştirilerek imgeye ait öznelik vektörü oluşturulmuştur.



Şekil 4.13 : Renk tabanlı yöntem 2 ile öznitelik oluşturma.

4.3.5 Öznitelik seçim ve öznitelik dönüşüm yöntemleri

Elde edilen öznitelik vektörlerine uygulanacak öznitelik seçim ve öznitelik dönüşüm algoritmalarının sınıflandırma performanslarına etkilerinin bulunarak daha yüksek sınıflandırma başarımı elde edilmesi amacı ile üç adet öznitelik seçim ve iki adet öznitelik dönüşüm tekniği uygulanmıştır. Her üç öznitelik seçim tekniğinde de öznitelik vektöründen değişken çıkartmadan ölçüm ve kıyaslamalara başlanarak, sırası ile değişken sayısı azaltılarak kıyaslamalara devam edilmiş ve devam koşulu olarak sınıflandırma başarımlarının azalmaması şartı kullanılmıştır. Kullanılan ilk yöntemde performans ölçümü için diskriminant analiz fonksiyonu ile sınıflandırma ve sınıflandırma hata değerlerinin 10 kat çapraz doğrulama sonuçları dikkate alınmıştır. Kullanılan ikinci yöntemde performans ölçüm ve kıyaslaması ilk yöntem ile aynı olmakla birlikte işlem yapılacak öznitelik vektöründe mutlaka bulunması ve bulunmaması gereken değişkenlerin tespit edilmesi için önem dereceleri hesaplanarak minimum ve maksimum olmak üzere iki eşik değeri belirlenmiştir. Değerlendirme aşamasında önem derecesi minimum eşik değerinin altında olan değişkenler etkileri hesaplanmadan öznitelik vektöründen çıkartılmış, maksimum eşik seviyesinin üzerinde olan değişkenlerde kalıcı duruma getirilerek çıkarılmaları durumunda oluşacak başarımların değişimleri ölçülmemiştir. Kullanılan son yöntemde performans ölçümü için YSA ile sınıflandırma ve sınıflandırma hata değerlerinin 10 kat çapraz doğrulama sonuçları dikkate alınmıştır. Öznitelik vektörlerine öznitelik dönüşüm tekniği olarak TBA ve DAA yöntemi kullanılmıştır.

4.4 Performans Değerlendirme

Oluşturulan özniteliklerin kullanılarak yapılacak olan sınıflandırma ve performans değerlendirme işlemi algoritmalarının kodlanması ve uygulanması MATLAB ortamında yapılmıştır. Küf, maya ve bakterilerin sınıflandırılması için Naive Bayes, k-EYK, C4.5 Karar Ağacı ve YSA sınıflandırıcıları kullanılmıştır. Naive Bayes, k-EYK ve C4.5 sınıflandırma algoritma başarımları birbirine yakın (± 3) olmakla beraber YSA sınıflandırma algoritmasının tüm öznitelikler için diğer algoritmalara kıyasla yüksek başarımlara sahip olduğu tespit edilmiştir. Sınıflandırma performanslarının kıyaslanması için 10 kat çapraz doğrulama test tekniği ve ortalama sınıf doğruluğu ölçümü kullanılmıştır. 10 kat çapraz doğrulama için veri seti 10 eşit kümeye bölünür, 9 küme eğitim için, 1 küme test için kullanılır ve her küme test kümesi olarak kullanılacak şekilde işlem tekrarlanır. Öznitelik çıkarım teknikleri için TBA ve DAA öznitelik dönüşüm tekniklerinin uygulanmasının bazı öznitelikler için YSA sınıflandırıcısı altında başarımlarının arttığını göstermektedir. Öznitelik seçiminde diskriminant analiz fonksiyonu kullanılan yöntemlerde, etkisinin az olduğu ölçülerek çıkartılan değişkenlerin, YSA sınıflandırmasında etkisinin olduğu ve sınıflandırma başarımlarını çok düşürdüğü tespit edilmiştir. Oluşan bu durum sebebi ile diskriminant analizine göre başarımların artırılmasına rağmen YSA sınıflandırma başarımları düştüğü için kullanılan ilk iki öznitelik seçim tekniği ile elde edilen öznitelik vektörleri elenerek değerlendirme dışı bırakılmıştır. Çizelge 4.1’de yer alan beş öznitelik çıkarım tekniği için YSA etki ölçümü ile öznitelik seçim yöntemi uygulanabilmiştir. Maksimum değişken sayısı azaltımı iki değişken ile GLCM öznitelik vektöründe gerçekleşmiştir. “1 Küf”, “2 Küf”, “3 Küf”, “4 Küf”, “1 Maya”, “2 Maya”, “1 Bakteri”, “2 Bakteri”, “3 Bakteri”, “4 Bakteri”, “1 Bakteri Koliform”, “2 Bakteri Koliform” ve “Sayım Dışı” olmak üzere 13 sınıf etiketinden oluşan ve YSA sınıflandırıcısı kullanılarak yapılan tahminin performans değerleri Çizelge 4.2’de yer almaktadır.

Çizelge 4.1 : Öznitelik seçimi sonrası değişken sayıları.

Öznitelik	Seçim Öncesi	Seçim Sonrası
RTY2	18	17
GLCM	16	14
HU	7	6
FRAC36	36	35

Çizelge 4.2 : Tür ve adet tahmini için başarımlar sonuçları.

Öznitelik	Doğruluk (%)	Doğruluk (%) (TBA)	Doğruluk (%) (DAA)	Doğruluk (%) (SGS)
RTY2	84.52	86.77	69.92	85.05
FRAC	83.22	85.75	73.36	82.94
RTY1.2	74.41	-	-	-
HOG	68.66	45.45	73.71	-
GLCM	64.29	69.46	55.04	58.23
RTY1.1	64.09	-	-	-
HIST	62.51	59.85	49.46	-
SURF	45.42	29.13	77.11	-
ŞTY1	42.80	27.73	72.62	-
HARRIS	42.51	27.76	76.48	-
MINEIGEN	41.70	26.11	74.55	-
FAST	39.84	25.66	76.62	-
ŞTY2	36.75	23.69	42.23	-
HU	32.68	23.59	26.47	32.01
BRISK	30.92	22.15	71.46	-
SHAPE	25.83	26.04	-	-

Deneysel sonuçlar yeni geliştirilen yöntemlerden RTY2 öznitelik çıkarım tekniğinin YSA sınıflandırıcısı altında %84 ortalama sınıf doğruluğu ile en iyi başarımları elde ettiğini, RTY1.2'nin %74 ile üçüncü, RTY1.1'in %64 ile altıncı, ŞTY1'in %42 ile dokuzuncu ve ŞTY2'nin %36 ile onüçüncü sırada yer aldığını göstermektedir.

Öznitelik dönüşüm uygulanarak üretilen öznitelik vektörleri ile sınıflandırmada TBA uygulanmış RTY2 tekniği YSA sınıflandırıcısı altında %86 ortalama sınıf doğruluğu ile en iyi başarımları elde ettiği görülmektedir. Ölçülen bu değer TBA kullanımının başarımları arttırdığını ve TBA kullanılmadan elde edilen %84 lük başarımlar oranına kıyasla % 2 lük bir artış elde edildiğini göstermektedir. Geliştirdiğimiz RTY1.1, RTY1.2 yöntemlerine TBA uygulanamamış olup, ŞTY1 ve ŞTY2 yöntemlerinin başarımları TBA kullanımı ile düşüş göstermiştir. RTY1.1, RTY1.2 yöntemlerine DAA uygulanamamış olup, ŞTY1, ŞTY2 ve RTY2 yöntemlerinin başarımları DAA kullanımı ile artış göstermiştir. Alınan sonuçlar TBA öznitelik dönüşüm tekniğinin, problemimizin çözümünde kullanılabileceği, DAA öznitelik dönüşüm tekniğinin ise uygun olmadığını göstermektedir.

Deneyisel sonuçlar bazı öznelik çıkarım teknikleri için öznelik seçim tekniği ile elde edilen öznelik vektörünün YSA sınıflandırıcısı altında başarı performansının arttığını göstermektedir. SGS öznelik seçim tekniği ile üretilen öznelik vektörleri ile sınıflandırmada yeni geliştirdiğimiz RTY2 tekniğinin YSA sınıflandırıcısı altında %85 ortalama sınıf doğruluğu ile en iyi başarıyı elde ettiği görülmektedir. Ölçülen bu değer öznelik seçimi kullanımının başarıyı arttırdığını ve öznelik seçimi kullanılmadan elde edilen %84'lik başarı oranına kıyasla %1'lik bir artış elde edildiğini göstermektedir. RTY1.1, RTY1.2, ŞTY1 ve ŞTY2 öznelik vektörlerinin büyüklüğü sebebi ile öznelik seçim işlemi uygulanamamıştır. Diskriminant analiz yöntemi ile etki ölçümü yapılan SGS yöntemlerinde başarı dramatik bir şekilde düşmüştür. Alınan sonuçlar YSA etki ölçümü ile öznelik dönüşüm tekniğinin problemimizin çözümünde kullanılabilceği, diskriminant analiz etki ölçümü ile öznelik dönüşüm tekniğinin ise uygun olmadığını göstermektedir.

Yüksek başarı oranı ile RTY2 öznelik çıkarım tekniğinin, TBA öznelik dönüşüm tekniği ile birlikte YSA sınıflandırıcısı altında kullanımı, tür ve adet tahmini probleminin çözümü için en iyi performans değerlerine sahiptir.

Yapılan ölçümlere ait karmaşıklık matrisleri incelendiğinde, bazı tekniklerin çoğunlukla mikroorganizmanın türünü yanlış sınıflandırmasına rağmen adet tahminin daha iyi olduğu, bazı tekniklerin ise bunun tam tersine, tür tahmininde daha az hata oranına sahip olmasının yanında adet tahmininde ise daha fazla hata oranına sahip olduğu tespit edilmiştir. Tespit edilen bu değerler, kullanılan öznelik çıkarım yöntemlerinin, koloni görüntülerinden çıkarım yapmaya çalıştığımız tür ve adet tahminleri için farklı başarı oranlarına sahip olduğunu göstermektedir. Özneliklerin tür ve adet tahmini üzerindeki başarılarının ve içerdikleri bilgilerin kıyaslanması için tahmin problemi tür ve sayı tahmini olarak iki farklı alt kategoriye bölünerek sınıflandırma ve performans değerlendirmeleri gerçekleştirilmiştir. “Küf”, “Maya”, “Bakteri” , “Bakteri Koliform” ve “Sayım Dışı” olmak üzere 5 sınıf etiketinden oluşan ve YSA sınıflandırıcısı kullanılarak yapılan tür tahmininin performans değerleri Çizelge 4.3’de yer almaktadır.

Çizelge 4.3 : Tür tahmini için başarıml sonuçları.

Öznitelik	Doğruluk (%)	Doğruluk (%) (TBA)	Doğruluk (%) (DAA)	Doğruluk (%) (SGS)
RTY2	92.77	93.96	82.80	91.75
FRAC	89.47	90.28	80.63	89.72
RTY1.2	87.47	-	-	-
RTY1.1	86.00	-	-	-
HOG	80.86	65.78	88.45	-
GLCM	75.50	82.59	71.08	78.13
HIST	73.53	71.50	61.57	-
HARRIS	67.18	48.12	84.87	-
MINEIGEN	64.90	49.77	85.57	-
SURF	64.79	43.45	83.61	-
FAST	63.22	46.37	82.87	-
HU	53.00	37.45	46.93	52.61
BRISK	52.54	37.42	68.48	-
ŞTY1	51.63	45.91	43.86	-
ŞTY2	45.84	36.29	54.05	-
SHAPE	27.48	27.13	-	-

Deneysel sonuçlar yeni geliştirdiğimiz RTY2 öznitelik çıkarım tekniğinin tür tahmini için YSA sınıflandırıcısı altında %92 ortalama sınıf doğruluğu ile en iyi başarıml elde ettiğini, RTY1.2'nin %87 ile üçüncü, RTY1.1'in %86 ile dördüncü, ŞTY1'in %51 ile ondördüncü ve ŞTY2'nin %45 ile onbeşinci sırada yer aldığını göstermektedir. Genel tahmin sonuçları ile kıyaslandığında sıralamada değişiklikler olmasına karşın en iyi başarımla sahip teknikte değişiklik görülmemiştir.

Boyut azaltımı kullanılarak üretilen öznitelik vektörleri ile sınıflandırmada TBA boyut dönüşümü uygulanan RTY2 tekniği, YSA sınıflandırıcısı altında %93 sınıf doğruluğu ile en iyi başarıml elde ettiği görülmektedir.

Yüksek başarıml oranı ile RTY2 öznitelik çıkarım tekniğinin, TBA öznitelik dönüşüm tekniği ile birlikte YSA sınıflandırıcısı altında kullanımı, tür tahmini probleminin çözümü için en iyi performans değerlerine sahiptir.

Türü bilinen koloni görüntülerinden içerdiği koloni sayısının tahmini için "1 Adet", "2 Adet", "3 Adet", ve "4 Adet" olmak üzere 4 sınıf etiketinden oluşan ve YSA sınıflandırıcısı kullanılarak yapılan adet tahmininin performans değerleri Çizelge 4.4'de yer almaktadır.

Çizelge 4.4 : Adet tahmini için başarımları sonuçları.

Öznitelik	Doğruluk (%)	Doğruluk (%) (TBA)	Doğruluk (%) (DAA)	Doğruluk (%) (SGS)
FRAC	92.77	93.40	85.96	92.56
RTY2	90.49	93.34	85.22	90.77
HOG	84.84	70.09	90.49	-
RTY1.2	83.63	-	-	-
HIST	81.57	81.05	76.97	-
RTY1.1	77.17	-	-	-
ŞTY2	74.20	57.60	68.62	-
ŞTY1	73.18	54.02	78.73	-
GLCM	69.88	79.61	76.20	71.29
SURF	63.74	56.05	84.20	-
HU	61.88	52.47	57.60	61.71
SHAPE	61.18	61.46	-	-
FAST	59.74	51.11	84.31	-
HARRIS	58.34	52.76	83.08	-
MINEIGEN	57.60	53.56	81.68	-
BRISK	54.05	53.25	78.87	-

Deneysel sonuçlar FRAC öznitelik çıkarım tekniğinin adet tahmini için YSA sınıflandırıcısı altında %92 sınıf doğruluğu ile en iyi başarımları elde ettiğini, RTY2'nin %90 ile ikinci, RTY1.2'nin %83 ile dördüncü, RTY1.1'in %77 ile altıncı, ŞTY2'nin %74 ile yedinci ve ŞTY1'in %73 ile sekizinci sırada yer aldığını göstermektedir. Genel tahmin sonuçları ile kıyaslandığında en iyi başarımları sahip teknik dâhil olmak üzere sıralamada değişiklikler olduğu görülmektedir.

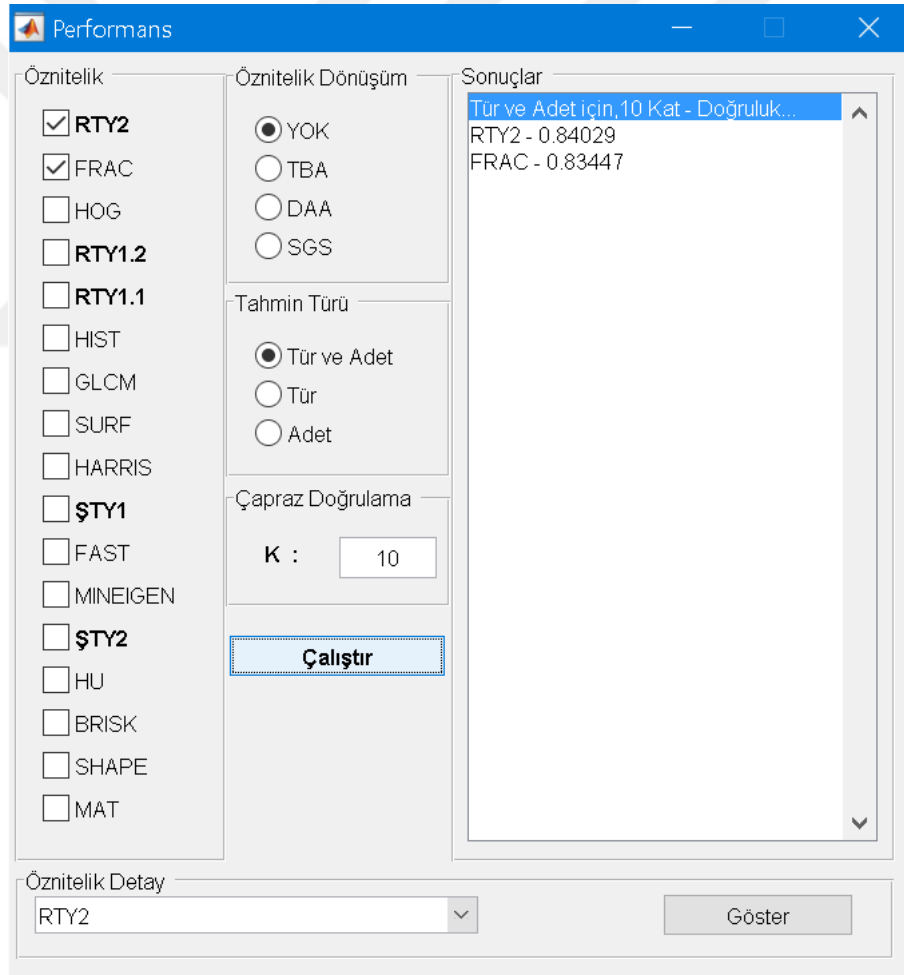
Boyut azaltımı kullanılarak üretilen öznitelik vektörleri ile sınıflandırmada TBA öznitelik dönüşümü uygulanan FRAC tekniği, YSA sınıflandırıcısı altında %93 ortalama sınıf doğruluğu ile en iyi başarımları elde ettiği görülmektedir.

Yüksek başarımları oranı ile FRAC öznitelik çıkarım tekniğinin, TBA öznitelik dönüşüm tekniği ile birlikte YSA sınıflandırıcısı altında kullanımı, tür tahmini probleminin çözümü için en iyi performans değerlerine sahiptir.

Yapılan çalışmalar sonucunda öznitelik seçim ve öznitelik dönüşüm yöntemlerinin sınıflandırma ortalama doğruluğu ve çalışma zamanlarına genel olarak olumlu etkiler sağladığı, öznitelik seçim işleminin doğası gereği öznitelik dönüşüm işleminden daha fazla zaman aldığı ve bazı durumlarda uygulanmasının mümkün olmadığı gözlenmiştir.

Genel, tür ve adet sınıflandırması için yapılan deneysel sonuçlar YSA sınıflandırıcısı altında genel ve tür tahmini için RTY2, adet tahmini için ise FRAC öznitelik çıkarım tekniklerinin TBA öznitelik dönüşüm tekniği ile birlikte kullanımının başarılı performans değerlerine ulaşıldığını göstermektedir. Geliştirilen yöntemlerden RTY1 ve RTY2'nin kolonilerin türleri hakkında, ŞTY1 ve ŞTY2'nin de kolonilerin adetleri hakkında daha fazla bilgi içerdiği ve RTY2 öznitelik çıkarım tekniği, TBA öznitelik dönüşümü ve YSA sınıflandırıcısının birlikte kullanıldığı modelin genel (tür ve adet) tahmin probleminin çözümü için en uygun sistem olduğu tespit edilmiştir.

Farklı özelliklere sahip sistem seçim ve performans kıyaslaması, öznitelik vektör detay görüntüleme işlemleri için MATLAB ortamında "Performans" olarak adlandırılan arabirim (Şekil 4.14) geliştirilmiştir.



Şekil 4.14 : Performans arabirimi ekran görüntüsü.



5. SONUÇLAR VE DEĞERLENDİRME

5.1 Sonuç

Endüstride birçok uygulamada ihtiyaç duyulan koloni sayımı, büyük sayıdaki kolonilerin elle sayılmasında sıkıcı, zaman alıcı ve insan hatası meyilli bir işlemdir. Uzun ve zahmetli bir işlem olmasının yanında kolonilerinin tespiti ve sayılması laborantın yorgunluğu, eğitimi ve tecrübesi ile ilişkili olarak değişkenlik gösterebilen karar ve yeteneklerine bağlıdır. Koloni sayımının yapılabilmesi için uzman personelin koloni morfolojileri üzerine eğitim alması ve tecrübe etmesi gerekmektedir. Alınan bu eğitimler ve tecrübeler insandan insana aktarılamamakta ve işlemi yapacak her personel için eğitim ve tecrübe safhalarının tekrarlanması gerekmektedir. Koloni tanımlama ve sayımının gerektirdiği bu durumlar, işlemin bir sistem tarafından otomatik olarak yapılma ihtiyacını ortaya koymaktadır. Yapılan bu çalışma ile koloni görüntülerinden yapay öğrenme sistemleri ile otomatik olarak tür tahmininin yapılabileceği gösterilmiştir ve devamında yapılacak çalışmalara katkı sağlanması umulmaktadır.

5.2 Tavsiyeler ve Gelecek Çalışmalar

Geliştirilen sistemde kullanılan veriler uygun ortam oluşturularak manuel olarak çekilmiştir. Veri tabanının büyütülmesi ve sistemin endüstriyel şartlarda kullanılması için otomatik fotoğraf çekim ve depolama sistemi tasarlanabilir. Tasarlanan sistem farklı lokasyonlara yerleştirilerek genel veri tabanı oluşturulabilir ve paylaşımına açılabilir. Sistemde mevcut segmentasyon yöntemlerinden Watersheed segmentasyon tekniği kullanılmıştır, daha hızlı ve daha iyi segmentasyon işlemi gerçekleştirecek yöntemler geliştirilebilir. Sınıflandırma performansının artırılması amacı ile farklı boyut değişim yöntemleri ve hibrit sınıflandırma algoritmaları geliştirilebilir. Geliştirilen sistem ile koloni görüntülerinin yüksek doğruluk oranı ile otomatik tanınması sağlanmıştır, sistem endüstride kullanılacak portatif ve kolay kullanılabilir bir cihaza dönüştürülerek işlem insan müdahalesinden çıkarılarak otomatikleştirilebilir.

Sonuç olarak, koloni morfolojisi ile tür tahmini üzerinde çalışılabilecek birçok konu için ön çalışma niteliğinde bir çalışma gerçekleştirilmiş olup, bu çalışmanın devamı niteliğindeki yeni araştırma konularına yol göstermesi bakımından gerçekleştirilen bu tez çalışmasının önü açıktır.



KAYNAKLAR

- Albayrak, N.B., Oktay, A.B., Akgül, Y.S.**, 2012: Bel Omurları Arası Disklerin Yerlerinin Belirlenmesi. *Signal Processing and Communications Applications Conference*.
- Albregtsen, F.**, 2010: INF 4300 – Digital Image Analysis. *University of Oslo*.
- Alt, F.L.**, 1968: Digital pattern recognition by moments. *Journal of the ACM, Volume 9 Issue 2, pages 240-258*.
- Altuntaş, S., Meral, H.**, 2015: Süt Endüstrisinde Moleküler Metotlar ile Mikrobiyel Tanı. *9.Gıda Mühendisliği Kongresi*.
- Altuntaş, V., Altuntaş, S., Gök, M.**, 2016: Agar Surface Colony Morphology Classification with Pattern Recognition. *International Multidisciplinary Congress of Eurasia*.
- Altuntaş, V., Altuntaş, S., Gök, M.**, 2016: Machine Learning Techniques for Colony Classification. *International Conference on Education in Mathematics, Science & Technology*.
- Balaban, M.E.**, 2015: Veri Madenciliği ve Makine Öğrenmesi Temel Algoritmaları ve R Dili ile Uygulamaları. *Çağlayan Kitabevi*, ISBN: 9789754360899.
- Bay, H., Ess, A., Tuytelaars, T., Gool, L.V.**, 2006: SURF : Speeded-Up Robust Features. *Computer Vision – ECCV, Volume 3951 of the series Lecture Notes in Computer Science*.
- Bovik, A.**, 2009: The Essential Guide to Image Processing. *Academic Press*, ISBN: 978-0-12-374457-9.
- Bravo, H.C. and Irizarry, R.A.**, 2010: Lecture 5: Model selection and assessment. *University of Maryland, Center for Bioinformatics and Computational Biology*.
- Buruş, M.E., Kuzu, R.S., Beyazıt, S., Varol, G.**, 2014: Market Raflarında Dönüşüm ve Ölçeklendirmeye Dayanıklı Nesne Tanıma. *Akademik Bilişim*.
- Chitade, A.Z., Katıyar, S.K.**, 2010: Color Based Image Segmentation using K-Means Clustering.
- Costa, A.F., Humpire-Mamani, G., Traina, A.J.M.**, 2012: An Efficient Algorithm for Fractal Analysis of Textures. *Graphics, Patterns and Images 25th SIBGRAPI Conference*, doi: 10.1109/SIBGRAPI.2012.15.
- Ding, C., He, X., Zha, H., Simon, H.D.**, 2002: Adaptive Dimension Reduction for Clustering High Dimensional Data. *Proceedings of International Conference on Data Mining*.

- Duda, R.O., Hart, P.E. and Stork, D.G.**, 1997: Pattern Classification (2nd ed.). *Wiley-Interscience*, ISBN: 978-0-471-05669-0.
- Erhan, C., Tazehkandi, A.A., Yeşilyurt, H.Y., Bayram, İ.**, 2013: Öznitelik Betimleyicileri Füzyonu ile Trafik İşaretlerinin Tespit Edilmesi ve Tanınması. *Signal Processing and Communications Applications Conference*.
- Genel Mikrobiyoloji**, 2006. *Gıda Teknolojisi Genel Mikrobiyoloji Milli Eğitim Bakanlığı Ankara*.
- Gershenson, C.**, 2003: Artificial neural networks for beginners. *arXiv*, cs/0308031 [cs.NE].
- Gök, M.**, 2011: HIV-1 Proteaz Enziminin Kesme Konumlarının Tespitinde Yeni Öznitelik Vektörleri. *Sakarya Üniversitesi Fen Bilimleri Enstitüsü Doktora Tezi*.
- Halkman, A. K.(Anonim)**, 2005: Merck Gıda Mikrobiyolojisi Uygulamaları. *Başak Matbaacılık ve Tanıtım Hizmetleri Ltd. Şti*.
- Hajek, M.**, 2005: Neural Networks. *University of KwaZulu-Natal*.
- Haykin, S.S.**, 1999: Neural Networks and Learning Machines Third Edition. *Prentice Hall*, ISBN-13: 978-0131471399.
- Hu, M.K.**, 1962: Visual Pattern Recognition by Moment Invariants. *IRE Transactions on Information Theory*, volume 8, issue 2, doi: 10.1109/TIT.1962.1057692.
- Jain, A.K., Farrokhnia, F.**, 1990: Unsupervised texture segmentation using Gabor filters. *Systems, Man and Cybernetics Conference Proceedings, IEEE International Conference*, doi: 10.1109/ICSMC.1990.142050.
- Juan, L., Gwun, O.**, 2009: A coparison of SIFT, PCA-SIFT and SURF. *Internetal Journal of Image Processing*.
- Jumb, V., Sohani, M., Shrivasa, A.**, 2014: Color Image Segmentation Using K-Means Clustering and Otsu's Adaptive Thresholding. *International Journal of Innovative Technology and Exploring Engineering*, ISSN: 2278-3075
- Karakaya, F., Altun, H., ve Çavuşlu, M.A.**, 2009: Gerçek Zamanlı Nesne Tanıma Uygulamaları için HOG Algoritmasının FPGA Tabanlı Gömülü Sistem Uyarlaması. *Signal Processing and Communications Applications Conference*.
- Kocadere, G.**, 2010: Video Görüntülerinin Sabitlemesi. *Ankara Üniversitesi Fen Bilimleri Enstitüsü Yüksek Lisans Tezi*.
- Kriesel, D.**, 2005: A Brief Introduction to Neural Networks David Kriesel. *adres: http://www.dkriesel.com*.
- Labatut, V., Cherifi, H.**, 2011: Accuracy Measures for the Comparison of Classifiers. *The 5th International Conference on Information Technology*, arXiv:1207.3790 [cs.LG].
- Liao, P.S., Chen, T.S, Chung, P.C.**, 2001: A Fast Algorithm for Multilevel Thresholding. *Inf. Sci. Eng.* 17 (5): 713–727.

- Liu, H. And Motoda, H.**, 1998: Feature Transformation and Subset Selection. *IEEE Intelligent Systems 13*, pp. 26-28 doi:10.1109/MIS.1998.671088.
- Mandelbrot, B.B.**, 1983: Fractal Geometry of Nature. *New York, ABD: W. H. Freeman and Company*, ISBN 978-0-7167-1186-5.
- McNamara, J.M., Green, R.F. and Olsson, O.**, 2006: Bayes' theorem and its applications in animal behaviour. *Wiley on behalf of Nordic Society Oikos*, doi:10.1111/j.0030-1299.2006.14228.x.
- Meyer F. ve Beucher S.**, 1990: Morphological Segmentation. *Journal of Visual Communication and Representation*, 1, s.21-46.
- Mohanaiah, P., Sathyanarayana, P., GuruKumar, L.**, 2013: Image Texture Feature Extraction Using GLCM Approach. *International Journal of Scientific and Research Publications*.
- Morphology and Classification of Bacteria.** (t.y.). Alındığı tarih: 07.01.2016, adres: <http://www.nos.org/media/documents/dmlt/Microbiology/Lesson-01.pdf>
- Morphology and General Properties of Fung.** (t.y.). Alındığı tarih: 07.01.2016, adres: <http://www.nos.org/media/documents/dmlt/Microbiology/Lesson-51.pdf>.
- Morse, B.S.**, 2000: Lecture 18: Segmentation (Region Based). *Brigham Young University*.
- Nixon, M.S., Aguado, A.S.**, 2002: Feature Extraction and Image Processing Second Edition. *Academic Press*, ISBN: 978-0-12372-538-7.
- Osuna, R.G.**, 2013: CSCE 666: Pattern Analysis. *Texas A&M University Department of Computer Science and Engineering*.
- Otsu, N.**, 1979: A threshold selection method from gray-level histograms. *IEEE Trans. Sys., Man., Cyber*, doi:10.1109/TSMC.1979.4310076.
- Patel, A.C., Markey, M.K.**, 2005: Comparison of three-class classification performance metrics: a case study in breast cancer CAD. *Proc. SPIE 5749, Medical Imaging: Image Perception, Observer Performance, and Technology Assessment*, doi: 10.1117/12.595763.
- Polikar, P.**, 2006: Pattern Recognition. *Wiley Encyclopedia of Biomedical Engineering*.
- Rai, P.**, 2011: CS5350/6350: Machine Learning. *The University of Utah, School of Computing*.
- Raschka, S.**, 2014: Linear Discriminant Analysis. http://sebastianraschka.com/Articles/2014_python_lda.html.
- Rokach, R.**, 2010: Pattern Classification Using Ensemble Methods. *ol. 75, World Scientific Publishing Co. Pte. Ltd., Singapore, ISBN: 978-981-4271-06-6*.
- Safavian, S.R.**, 1991: A Survey of Decision Tree Classifier Methodology. *IEEE Transaction on Systems, Man and Cybernetics*, 21, 3, 660-674.

- Samet, H.**, 2006: Foundations of Multidimensional and Metric Data Structures. *Morgan Kaufmann*, ISBN-13: 978-0123694461.
- Seo, N.**, 2006: Texture Segmentation using Gabor Filters. *The University of Maryland, ENEE731 Project*.
- Shah, M.**, 1997: Fundamentals Of Computer Vision. *Department of Computer Science, University of Central Florida*.
- Shanmugavadivu, P., Sivakumar, V.**, 2012: Fractal Dimension Based Texture Analysis of Digital Images. *Procedia Engineering Volume 38*, doi:10.1016/j.proeng.2012.06.348.
- Smith, L.I.**, 2002: A tutorial on Principal Components Analysis.
- Sikonja, M.R., Kononenko, I.**, 2003: Theoretical and Empirical Analysis of ReliefF and RReliefF. *Machine Learning Journal vol 53, issue 1-2, pp 23-69*, doi:10.1023/A:1025667309714.
- Sivalingamaiah, M., Reddy, B.D.V.**, 2012: Texture Segmentation Using Multichannel Gabor Filtering. *IOSR Journal of Electronics and Communication Engineering*, issn: 2278-2834.
- Soille P. ve Vincent, L.**, 1990: Determining Watersheds in Digital Pictures via Flooding Simulations. *Visual Communications and Image Processing*, vol. 1360, 1990.
- Sokolova, M., Lapalme, G.**, 2009: A systematic analysis of performance measures for classification tasks. *Information Processing and Management: an International Journal*, 45, 427-437, doi: 10.1016/j.ipm.2009.03.002.
- Söder, O.**, 2008: kNN classifiers 1. What is a kNN classifier?. *Phonetic Sciences, Faculty of Humanities, University of Amsterdam*, adres: http://www.fon.hum.uva.nl/praat/manual/kNN_classifiers_1__What_is_a_kNN_classifier_.html.
- Sutton, O.**, 2012: Introduction to k Nearest Neighbour Classification and Condensed Nearest Neighbour Data Reduction. *The Pennsylvania State University CiteSeerX*, oai:CiteSeerX psu:10.1.1.296.87652013-07-17, doi: 10.1.1.296.8765.
- Tang, J., Alelyani, S. and Liu, H.**, 2014: Feature Selection for Classification: A Review. in *Data Classification: Algorithms and Applications*, Chapman and Hall/CRC, pp. 37-64, ISBN: 978-1-4665-8674-1.
- Teknomo, K.**, 2015: Discriminant Analysis Tutorial. <http://people.revoledu.com/kardi/tutorial/LDA/index.html>.
- Theodoridis, S. and Koutroumbas, K.**, 2006: Pattern Recognition (Third Edition). *Academic Press*, ISBN: 978-0-12-369531-4.
- Thirumuruganathan, S.**, 2010: A Detailed Introduction to K-Nearest Neighbor (KNN) Algorithm. *Wordpress*, adres: <https://saravananthirumuruganathan.wordpress.com/2010/05/17/a-detailed-introduction-to-k-nearest-neighbor-knn-algorithm>.
- Topaloğlu, M., Gangal, A.**, 2006: Watershed Dönüşümü Kullanılarak Corpus Callosumun Bölütlenmesi. *URSI – Türkiye 2006 3. Bilimsel Kongresi*.

- Url-1** <<https://en.wikipedia.org/wiki/Microorganism>>, alındığı tarih: 29.06.2015.
- Url-2** <<https://www.le.ac.uk/se/centres/sci/selfstudy/eco7.htm>>, alındığı tarih: 15.07.2015.
- Url-3** <<http://www.mikrobiyoloji.org/TR/Genel/BelgeGoster.aspx?F6E10F8892433CFFAAF6AA849816B2EF34B8FF67F6C5B31A>>, alındığı tarih: 29.06.2015.
- Url-4** <<http://kuf.nedir.com/>>, alındığı tarih: 19.01.2016.
- Url-5** <<http://www.biyodoc.com/Canlilarin-siniflandirilmesi-mantarlar-alemi.html>>, alındığı tarih: 29.06.2015.
- Url-6** <<http://www.visiondummy.com/2014/05/feature-extraction-using-pca/>>, alındığı tarih: 21.03.2016.
- Url-7** <https://en.wikipedia.org/wiki/Principal_component_analysis>, alındığı tarih: 21.03.2016.
- Usta, R.**, 2014: Naive Bayes Sınıflandırma Algoritması. Alındığı tarih: 04.12.2015
adres: <http://kodcu.com/2014/05/naive-bayes-siniflandirma-algoritmasi>.
- Wu, M.N., Lin, C.C., Chang, C.C.**, 2007: Brain Tumor Detection Using Color-Based K-Means Clustering Segmentation. *Intelligent Information Hiding and Multimedia Signal Processing*, doi: 10.1109/IIHMSP.2007.4457697.
- Yoldaş, M., Şakar, M.O., Dirlikli, M., Kılınç, O., Ergin, S.**, 2013: Mamografi İmgelerinden HOG Öznelikleri Çıkartılarak Hastaların Kanseri Seviyelerinin Belirlenmesi. *Tmmob Emo Ankara Şubesi Haber Bülteni, İlk Bildiriler Konferansı*.
- Youse, A.E., Carlstrom, C.** 2003: Food Microbiology: A Laboratory Manual. *Wiley-Interscience*.

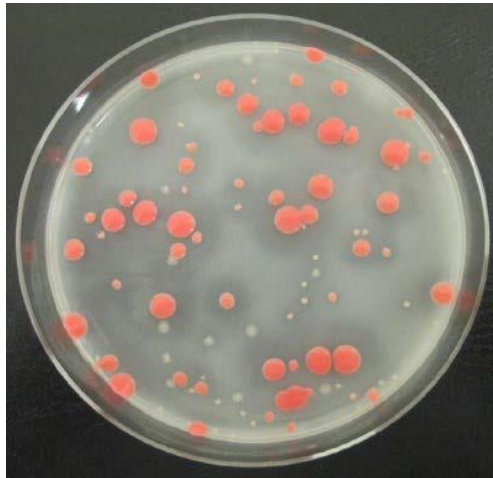
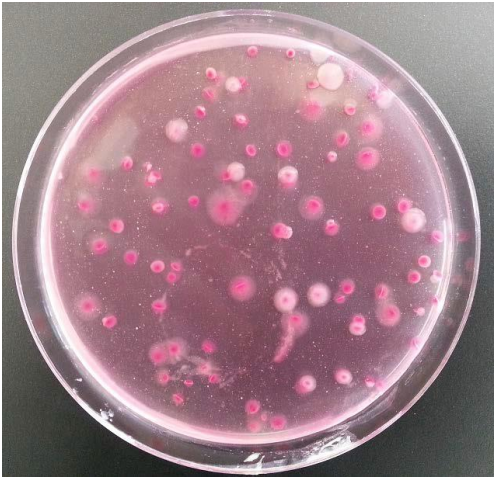
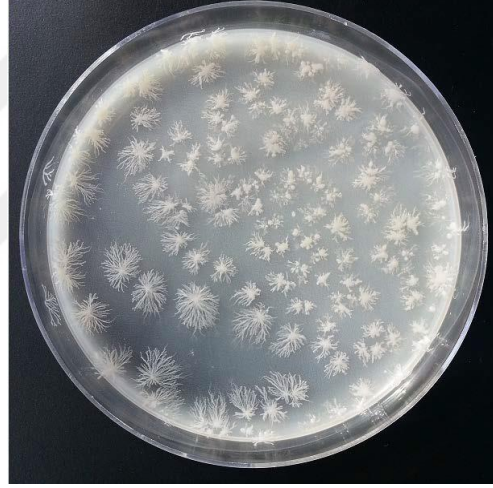


EKLER

EK A.1 : Petri Görüntüleri

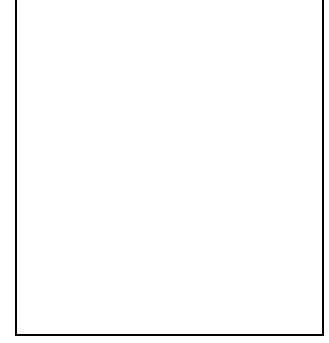


EK A.1



Şekil A.1 : Petri görüntüleri.

ÖZGEÇMİŞ



Ad Soyad: Volkan ALTUNTAŞ

Doğum Yeri ve Tarihi: Zonguldak - 1982

Adres: Bursa, Türkiye

E-Posta: valtuntas at gmail.com

Lisans: Gebze Yüksek Teknoloji Enstitüsü - Bilgisayar
Mühendisliği

Mesleki Deneyim ve Ödüller:

- **2015 - Halen:** Uzman, Bursa Teknik Üniversitesi
- **2011 - 2015:** Arge Proje Yöneticisi, Süttaş
- **2007 - 2011:** Proje Yöneticisi, Stok
- **2005 - 2007:** Arge Proje Yöneticisi, Enmos
- **Bölüm Birincisi,** Gebze Yüksek Teknoloji Enstitüsü
- **Üstün Başarı Ödülü,** Türk Eğitim Vakfı

TEZDEN TÜRETİLEN YAYINLAR/SUNUMLAR

- **Altuntas V.,** Altuntaş S., Gök M., 2016: Machine Learning Techniques for Colony Classification. *International Conference on Education in Mathematics, Science & Technology (ICRES 2016)*, May 19-22, 2016 Bodrum, Turkey.
- **Altuntas V.,** Gök M., Aktaş T., 2016: Agar Surface Colony Morphology Classification with Pattern Recognition. *International Multidisciplinary Congress of Eurasia (IMCOFE 2016)*, July 11-13, 2016 Odessa, Ukraine.
- **Altuntas V.,** Altuntaş S., Gök M., 2016: Automatic Colony Segmentation on Agar Surface by Image Processing. *Journal of Scientific and Industrial Research*.