

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

**YAPAY ZEKA YÖNTEMLERİYLE BİTKİ YAPRAK İMGELERİNDE**  
**PAS HASTALIKLARININ TESPİTİ**

**Emrullah ACAR**

**YÜKSEK LİSANS TEZİ**

**ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**

**DİYARBAKIR**

**Şubat-2012**

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

**YAPAY ZEKA YÖNTEMLERİYLE BİTKİ YAPRAK İMGELERİNDE**  
**PAS HASTALIKLARININ TESPİTİ**

**Emrullah ACAR**

**YÜKSEK LİSANS TEZİ**

**ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI**

**DİYARBAKIR**

**Şubat-2012**

T.C. DİCLE ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ MÜDÜRLÜĞÜ  
DİYARBAKIR

Emrullah ACAR tarafından yapılan “Yapay Zeka Yöntemleriyle Bitki Yaprak İmgelerinde Pas Hastalıklarının Tespiti” konulu bu çalışma, jürimiz tarafından Elektrik-Elektronik Mühendisliği Anabilim Dalında YÜKSEK LİSANS tezi olarak kabul edilmiştir.

Jüri Üyeleri

Başkan : Prof. Dr. Mehmet AKIN

Üye : Doç. Dr. Z. Fuat TOPRAK

Üye : Yrd. Doç. Dr. M. Sıraç ÖZERDEM

Tez Savunma Sınavı Tarihi: 01/02/ 2012

Yukarıdaki bilgilerin doğruluğunu onaylarım.

.../...../.....

Prof. Dr. Hamdi TEMEL

Enstitü Müdürü

## TEŐEKKÜR

Bu alıŐma, Dicle Üniversitesi DÜBAP:11:MF:24 numaralı proje kapsamında desteklenmiŐtir. Desteklerinden ötürü DÜ Bilimsel AraŐtırma Projeleri Koordinatörlüğüne teŐekkür ederim.

Bu tez alıŐması süresince her anlamda yardımını, bilimsel katkılarını ve tecrübelerini hiçbir zaman esirgemeyen deęerli hocam Yrd. Do. Dr. Mehmet Sira ÖZERDEM'e, alıŐmalarım boyunca her konuda destek veren arkadaşlarım ArŐ. Gör. Ramazan TEKİN'e ve ArŐ. Gör. Ömer Faruk Ertuęrul'a teŐekkür ederim.

Tezin hazırlanması sırasında her konuda fedakârlığını gösteren, maddi ve manevi yardımlarını esirgemeyen aileme sonsuz Őükranlarımı sunarım.

## İÇİNDEKİLER

	Sayfa
TEŞEKKÜR.....	I
İÇİNDEKİLER.....	II
ÖZET.....	IV
ABSTRACT.....	V
ÇİZELGE LİSTESİ.....	VI
ŞEKİL LİSTESİ.....	VII
KISALTMA VE SİMGELER.....	VIII
<b>1. GİRİŞ</b> .....	1
1.1. Bitki Yapraklarında Pas Hastalıkları .....	1
1.2. Tezin Hedefi.....	3
1.3. Tezin İçeriği.....	4
<b>2. KAYNAK ÖZETLERİ</b> .....	7
<b>3. MATERYAL ve METOT</b> .....	11
3.1. Öznitelik Çıkarımı.....	11
3.1.1. Gri Seviye Eş-Oluşum Matrisleri (GLCM).....	12
3.1.1.1. Haralick Doku Öznitelikleri.....	15
3.1.2. Gabor Dalgacık Dönüşümü.....	18
3.2. Sınıflandırma Yöntemleri.....	22
3.2. 1. Yapay Sinir Ağları.....	23
3.2.1.1. Yapay Sinir Ağlarının Genel Özellikleri.....	28
3.2.1.2. Yapay Sinir Ağlarının Avantaj ve Dezavantajları.....	29
3.2.1.3. Yapay Sinir Ağlarının Kullanım Alanları.....	31
3.2.1.4. Yapay Sinir Ağlarının Sınıflandırılması.....	31
- İleri Beslemeli Ağlar.....	31
- Geri Beslemeli Ağlar.....	32
3.2.1.5. Çok Katmanlı Algılayıcılar (MLP).....	33
3.2. 2. Destek Vektör Makineleri (DVM).....	35

	<b>Sayfa</b>
3.2.2.1. Doğrusal Destek Vektör Makinesi Sınıflandırıcıları.....	36
- Doğrusal Olarak Ayrılabilme Durumu.....	36
- Doğrusal Olarak Ayrılmama Durumu.....	40
3.2.2.2. Doğrusal Olmayan Destek Vektör Makinesi Sınıflandırıcıları.....	42
3.2.2.3. En Küçük Kareler Destek Vektör Makinesi (EKK-DVM) Sınıflandırıcıları.....	44
3.2.2.4. Bire-Karşı-Biri Metodu .....	45
3.2.2.5. Bire-Karşı-Diğerleri Metodu.....	46
3.2. 3. K En Yakın Komşu (k-NN) Algoritması.....	46
<b>4. BULGULAR VE TARTIŞMA.....</b>	<b>49</b>
4.1. Tespit Kapsamında Kullanılan İmgeler ve Önişlemler.....	49
4.2. Tespit Sisteminin Yapılandırılması.....	51
4.3. GWT Tabanlı Uygulama Sonuçları.....	52
4.4. GLCM Tabanlı Uygulama Sonuçları.....	54
<b>5. SONUÇ VE ÖNERİLER.....</b>	<b>59</b>
<b>6. KAYNAKLAR.....</b>	<b>61</b>
ÖZGEÇMİŞ.....	67

## ÖZET

### YAPAY ZEKA YÖNTEMLERİYLE BİTKİ YAPRAK İMGELERİNDE PAS HASTALIKLARININ TESPİTİ

YÜKSEK LİSANS TEZİ

Emrullah ACAR

DİCLE ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ  
ELEKTRİK-ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI

2012

Bitkilerdeki hastalıklar hasadı ve dolayısıyla verimi etkilemektedir. Hastalıkların önceden kestirilmesi, çiftçilerin alacağı önlemler ile verimi artıracaktır. Verimi etkileyen önemli hastalıkların başında pas hastalığı gelmektedir. Bu çalışmada bitki örneği olarak zambak çiçeği yaprak imgeleri, zirai uygulamalarla ilgili farklı zirai sitelerden bir uzman yardımıyla elde edilmiş, GLCM ve GWT tabanlı farklı sınıflandırıcı teknikleri kullanılarak yaprağında oluşan değişimin pas hastalığı olup olmadığını kestiren bir sistem tasarlanmıştır.

İlk aşamada, imgelere ilişkin gri seviyeli eş-oluşum matrisleri elde edilip matrislerin kontrast, korelasyon, enerji, homojenlik ve entropi değerleri hesaplanmıştır. İki boyutlu imgelere ilişkin matrislerden hesaplanan ortalama ve standart sapma değerleri öznitelik vektörüne eklenerek, her imge için toplamda 7 parametre içeren öznitelik vektörü oluşturulmuştur. İkinci aşamada, Gabor dalgacık dönüşümü kullanılarak her bir imgeden ayrı bir özellik matrisi elde edilip, matrislerin ortalama, standart sapma ve entropi değerleri hesaplanmıştır. Bu değerler ayrı bir öznitelik vektörüne eklenerek, her imge için 3 parametre içeren yeni bir öznitelik vektörü oluşturulmuştur.

Son aşamada, GLCM ve GWT tabanlı öznitelik vektörleri ayrı ayrı sınıflandırıcıların girişlerine verilerek performansı en iyi sistem belirlenmeye çalışılmıştır. Bu sistemlerde sınıflandırıcı olarak Çok Katmanlı Algılayıcı, k-En Yakın Komşu (k-NN) ve en küçük kareler Destek Vektör Makineleri (LS-SVM) yöntemleri kullanılmıştır. Zambak çiçeği yaprak imgeleri, 32 sağlıklı imge ve 21 hastalıklı imge olmak üzere toplam 53 imge verisinden oluşup, iki (1- sağlıklı, 2- hastalıklı) grupta sınıflandırılmıştır. Sınıflandırma çalışmaları sonucunda, en iyi performans %88,9 başarı ile GLCM tabanlı k-NN ve çok katmanlı yapay sinir ağının (7-5-1) ulaştığı gözlemlenmiştir. Bulduğumuz sonuçlar bu yöntemin pas hastalığını tespit etmede doğru ve etkili çalıştığını göstermektedir.

**Anahtar Kelimeler :** Zambak, Sınıflandırma, İmge İşleme, Gabor Dalgacık Dönüşümü, Gri Seviye Eş-Oluşum Matrisleri (GLCM), Yapay sinir ağları, En yakın k- komşu algoritması (k-NN), En küçük kareler destek vektör makineleri (LS-SVM)

## ABSTRACT

### THE ESTIMATION OF RUST DISEASES ON CROP LEAVE IMAGES BY USING ARTIFICIAL INTELLIGENCE METHODS

MSc THESIS

Emrullah ACAR

DEPARTMENT OF ELECTRICAL AND ELECTRONICS ENGINEERING  
INSTITUTE OF NATURAL AND APPLIED SCIENCES  
UNIVERSITY OF DICLE

2012

Crop diseases can affect yield and/or quality of the harvested commodity. This can influence profitability and raise the risks of farming. When the diseases are estimated early, the yield will increase by taking measures thanks to farmers. The rust disease is one of the most major crop diseases that affect crop yield. In this paper, daylily leaf images are used as crop sample and derived from different agricultural sites under expert control and daylily rust disease is estimated by using GLCM and GWT based different classifier techniques.

In the first stage, the features are extracted from images with using Gray Level Co-Occurrence Matrix (GLCM) method and 7 parameters are derived by this method for each digital camera image. These parameters are contrast, correlation, energy, homogeneity, entropy, standard deviation and mean for first texture feature vector. In the second stage, a new texture feature vector is based on Gabor Wavelet Transform (GWT) and 3 parameters are derived by this method for each digital image. These are mean, standart deviation and entropy.

In the final stage, GLCM and GWT based feature vectors are applied to different type of classifiers and these vectors are used as inputs in classification systems, respectively. The Multilayer Perceptron neural network (MLP),  $k$ -Nearest Neighbor ( $k$ -NN) and Least Squares Support Vector Machine (LS-SVM) classifiers have been chosen for learning and testing of 53 image data where 32 of them belongs to class I (normal), 21 of them belongs to class II (rust diseased). Different structures of networks are tested and the results are compared in terms of testing performance for each network model. The main purpose of this study is to estimate whether the crop is normal or rust diseased. Finally, the best performance was observed as 88.90 % in the  $k$ -NN and MLP network with 7-5-1 structure. Our results suggest this method is an accurate and efficient means of estimating daylily rust disease.

**Key Words:** Daylily, Classification, Image processing, Gabor Wavelet Transform, Multilayer Perceptron neural network (MLPNN),  $k$ -Nearest Neighbor ( $k$ -NN), Least Squares Support Vector Machine (LS-SVM)



## ÇİZELGE LİSTESİ

<u>Çizelge No</u>		<u>Sayfa</u>
<b>Çizelge 4.1.</b>	GWT tabanlı farklı MLP ağ sınıflandırıcısının performans değerleri	53
<b>Çizelge 4.2.</b>	GWT tabanlı farklı $k$ -NN sınıflandırıcısının performans değerleri	53
<b>Çizelge 4.3.</b>	GWT tabanlı farklı LS-SVM sınıflandırıcısının performans değerleri	54
<b>Çizelge 4.4.</b>	GLCM tabanlı farklı MLP ağ sınıflandırıcısının performans değerleri	55
<b>Çizelge 4.5.</b>	GLCM tabanlı farklı $k$ -NN sınıflandırıcısının performans değerleri	56
<b>Çizelge 4.6.</b>	GLCM tabanlı farklı LS-SVM sınıflandırıcısının performans değerleri	56

## ŞEKİL LİSTESİ

<u>Şekil No</u>		<u>Sayfa</u>
Şekil 3.1.	Tez kapsamında kullanılan örüntü tanıma sistemi	12
Şekil 3.2.	Eş-oluşum matrisinin elde edilmesi	14
Şekil 3.3.	Açı Diyagramı	15
Şekil 3.4.	Gabor dalgacık kümesi	20
Şekil 3.5.	Gabor dalgacık kümesinin uzaysal frekans düzlemindeki kapsamı	21
Şekil 3.6.	Miyelinli bir nöron yapısı	24
Şekil 3.7.	Biyolojik nöron şematik yapısı	25
Şekil 3.8.	Yapay nöron modeli	26
Şekil 3.9.	Hard-Limit Transfer Fonksiyonu	26
Şekil 3.10.	Lineer Transfer Fonksiyonu	27
Şekil 3.11.	Log-Sigmoid Transfer Fonksiyonu	27
Şekil 3.12.	Tan-Sigmoid Transfer Fonksiyonu	28
Şekil 3.13.	İleri beslemeli ve çok katmanlı YSA yapısı	32
Şekil 3.14.	Geri beslemeli yapay sinir ağı yapısı	33
Şekil 3.15.	Çok katmanlı algılayıcı için örnek ağ mimarisi	34
Şekil 3.16.	Doğrusal olarak ayrılabilen durum için optimal hiperdüzlem	37
Şekil 3.17.	Doğrusal ayrılabilme durumunda en uygun hiperdüzlem	38
Şekil 3.18.	Doğrusal ayrılmama durumunda en uygun hiperdüzlem	41
Şekil 3.19.	Giriş uzayındaki örneklerin özellik uzayına taşınması	42
Şekil 3.20.	Örnek veri kümesi ve sınıflandırma noktası	47
Şekil 4.1.	Zambak çiçeğine ait hastalıklı örnek yaprak imgeleri	50
Şekil 4.2.	Zambak çiçeğine ait sağlıklı örnek yaprak imgeleri	50
Şekil 4.3.	GLCM tabanlı sınıflandırma sisteminin yapısı	51
Şekil 4.4.	GWT tabanlı sınıflandırma sisteminin yapısı	51
Şekil 4.5.	GWT tabanlı LS-SVM sınıflandırıcısının ROC eğrisi	54
Şekil 4.6.	GLCM tabanlı LS-SVM sınıflandırıcısının ROC eğrisi	56

## KISALTMA VE SİMGELER

1-B	: 1 Boyutlu
2-B	: 2 Boyutlu
AİM	: Açısal İkinci Moment
$\alpha$	: Çıkıştaki Nöron Değeri
$b$	: Bias Değeri
B1	: Hiperdüzlem 1
B2	: Hiperdüzlem 2
BP	: Geriye Yayılma
CRF	: Koşullu Rastgele Alanlar
$d$	: MLP Ağında İstenilen Çıkış
$dv$	: Destek Vektörleri Sayısı
$\delta_i$	: Lagrange Çarpanları
$\delta_j$	: MLP Ağ Giriş, Çıkış ve Gizli Katmanlardaki Ağırlıkların Ayarlanması
DA	: Diskriminant Analizi
$e$	: MLP Ağ Çıkışındaki Hata İşareti
$\phi$	: Taşıma Fonksiyonu
$G(\vec{x})$	: İmge Yoğunluğu
$g(n)$	: Ağ Gradyeni
GKMTBS	: Gauss Karma Modeli Tabanlı Bayes Sınıflandırma
GLCM	: Gri Seviye Eş-Oluşum Matrisi
GWT	: Gabor Dalgacık Dönüşümü
$J_i(\vec{x})$	: Gabor Dalgacık Dönüşümü Matrisi
K	: SVM Çıkış Uzayı
KKT	: Karush-Kuhn-Tucker
k-NN	: K En Yakın Komşu Algoritması
$L$	: Lagrange Fonksiyonu
LR	: Lojistik Regresyon
LS-SVM	: En Küçük Kareler Destek Vektör Makinesi
$\mu$	: Ortalama
MLP	: Çok Katmanlı Algılayıcı
MSE	: Ortalama Karesel Hata

$n$	: MLP Ağı İterasyon Sayısı
$\eta$	: Öğrenme Oranı
$N_g$	: Gri Seviye Sayısı
P	: MLP Ağı Toplam Örüntü Sayısı
$p$	: Girişteki Nöron Değeri
$p(i, j)$	: Normalize Edilmiş Eş-Oluşum Matrisi
$P(x_i, x_j)$	: SVM Çekirdek Fonksiyonu
PCA	: Temel Bileşenler Analizi
$\psi$	: SVM Hata Terimi
$\psi_i(x)$	: Düzlem Dalgası
RBF	: Radyal Tabanlı Çekirdek Fonksiyonu
R	: Girişteki Nöron Sayısı
$\mathfrak{R}^d$	: SVM Giriş Uzayı
RSI	: Pas Şiddeti İndeksi
S	: SVM Üst Sınırı
$\sigma$	: Standart Sapma
S1	: Sınıf 1
S2	: Sınıf 2
SOM	: Kendinden Düzenlenen Haritalar
SOS	: Öz Örgütlemeli Sistemler
SVM	: Destek Vektör Makinesi
$\theta^o$	: Pikseller Arasındaki Açık
TFM	: Ters Fark Momenti
$w$	: Nöron Ağırlığı
$\ w\ $	: Norm Operatörü
$\Delta w_{ji}$	: MLP Ağı Sinaptik Ağırlıkların Ayarlanması
$x_i$	: MLP Ağındaki Nöron Girişi
$x_r$	: k-NN'de Sınıflandırılmak İstenen Örnekler
$x_s$	: k-NN Öğrenme Noktaları
$y_i$	: MLP Ağındaki Nöron Çıkışı
YSA	: Yapay Sinir Ağları

## 1. Giriş

Bu bölümde tezin amacı, hedefi ve kapsamı genel hatları ile anlatılmıştır. Tezin konusu bilgi teknolojilerini kullanarak bitki yapraklarında görülen pas hastalıklarının tespit edilmesidir. Pas hastalıkları doğal olarak Ziraat Mühendisliği kapsamında araştırılan bir konudur, ancak disiplinler arası çalışmalar ile konu farklı bir boyuttan değerlendirilebilmektedir.

Giriş bölümünde uzmanlık alanımıza girmeyen bitki yapraklarında pas hastalıkları konusu bizim gibi bilişim alanında çalışanların anlayabilmesi için kısaca tanıtılmış, tezin hedefi ve çalışmaların içeriğinden bahsedilmiştir.

### 1.1. Bitki Yapraklarında Pas Hastalıkları

“Pas hastalığı nedir?” sorusuna, hastalığın ne olduğu tanımlanarak cevap verilebilir. Pas hastalığı; bitkilerin köklerine ve yapraklarına gıda alımını durdurarak, bitkilerin zayıflamasına neden olan bir mantardır. Bu hastalığa “pas” hastalığı denmesinin sebebi, noktaların pas parçacıkları şeklinde görünerek sarı renkten parlak turuncuya, kırmızıya veya kahverengi rengine doğru değişen bir renk aralığında olmasıdır. Burada bazı lekeler düz olmalarına rağmen, bazıları da kabarık olabilirler. Bu hastalık, bitkilerde bulaşıcıyken çiçekler ve sebzeler arasında bulaşıcı değildirler. Pas lekesi, ilk başlarda hafif turuncu rengindedir. Ancak zaman ilerledikçe, koyu kahverengine dönüşür ve hastalığın bulaştığı yapraklar yere düşer. Hastalık yayıldıkça, ana gövde üzerinde pas lekeleri görülerek, sonunda bitki ölmeye başlar (Çoklu 2011, Dauber 2008).

Genel olarak ekim alanlarında 3 çeşit pas hastalığına rastlanır. Bunlar sırasıyla kahverengi pas, sarı pas ve siyah pastır. En sık görülen çeşidi, *yaprak veya kahverengi pas* olarak adlandırılır. Bu hastalık genellikle sulu tip uzun yapraklarda görülür. Etmenin yazlık ve kışlık sporları sarı pasla benzerlik gösterir. Hastalığa yol açan etmen; kışın ılıman geçtiği yerlerde, misel veya yazlık spor halinde güzlük mahsullerde bulunurken, yaz aylarında ise yabancı graminelerin taze yaprak veren kısımlarında bulunur. Yaşam alanlarında kışlık sporların bir etkisi olmayıp, ilkbahar aylarında yazlık sporlarını oluştururlar. Rüzgarla yayılmaya başlayan bu sporlar, uygun koşullarda (yüksek nem, 10<sup>0</sup>C - 18<sup>0</sup>C sıcaklık) başka hastalıklara yol açarlar. Bu hastalıkta, yazlık

## 1. GİRİŞ

---

sporların içinde bulunan püstüller noktacıklar şeklinde rastgele yaprak üzerinde dağılmıştır. Bu noktacıklar genellikle yanık kahverengi veya portakal sarısı rengindedir. Hastalığın ilerlemesine bağlı olarak püstüllerdeki epidermis parçalanır. Bazen ana püstül etrafında çevrelenmiş bir veya iki tane halka şeklinde küçük püstüller oluşabilir. Bu püstüller kahverengi pasın tanımlanması için önemli bir özellik olup, bu hastalık genellikle bitkilerde kara pastan önce sarı pastan sonra görülür (Çoklu 2011, Dauber 2008).

Bitkilerde görülen bir diğer pas hastalığı çeşidi, *şerit veya sarı pas* olarak adlandırılır. Bu hastalık tipinde etmenin yazlık sporları oval, çeperleri dikenli ve porları yüzeye dağılmış bir şekildedir. Kışlık sporlarında ise hücre çeperleri kalın iki hücreli ve düzdür. Bu hücreler arasında hafif boğum vardır. Hastalık ilkbahar aylarında hava sıcaklığının 10°C -15°C olduğu zamanda buğday tarlarında ortaya çıkmaya başlar. Yaprığın üst yüzeyinde sıracıklar halinde makine dikişine benzeyen püstüller meydana gelerek, bu püstüllerin içinde hastalık etmeninin yazlık sporları oluşur. Bu püstüller limon veya portakal renginde olup, ilkbahar aylarında bu püstüllerden oluşan yazlık sporlar rüzgarla dağılmaya başlarlar. Hastalığın oluşmasında sıcaklık ve nem önemlidir. Uygun şartlarda (%100 nem) ve yağışlarda yazlık sporlar çimlenerek konukçuya giriş yapar ve mevsim sonunda yazlık sporların yataklarında kışlık spor yatakları oluşmaya başlar. Hastalığa yol açan etmen; kışın ılıman geçtiği yerlerde günlük ekilen buğdaylarda misel veya yazlık spor halinde bulunurken yaz aylarında ise yüksek rakımlı bölgelerde canlı kalan yabancı buğdaygillerde yaşamını sürdürür. Bu hastalıkta, kışlık sporların yaşam çemberlerinde bir etkileri bulunmayıp, mantarın inkübasyon süresi genelde 12-14 gün arasındadır. Sarı pas hastalığı, bitkilerde en erken görülen pas çeşididir. En sık yapraklarda görülür fakat hastalığın şiddetli olduğu dönemlerde ürünlerin başak kılçıkları ve saplarında da görülebilir. Püstüller yaprak yüzeyinde sarı renkte ve bir şerit şeklinde dizilirler. Bu yüzden şerit pas olarak ta adlandırılırlar. Mevsim sonlarına doğru yazlık sporların yerini siyah renkteki kışlık sporlar alır (Çoklu 2011, Dauber 2008, McIntosh ve ark.1995).

Bitkilerde görülen diğer pas çeşidi gövde pası olup sezon sonunda veya elverişsiz koşullarda yazlık sporlar şeklini alan parlak siyah kışlık sporların aşırı üretiminden dolayı *siyah pas veya yaz pası* olarak adlandırılır. Siyah pas, pas hastalıklarının en yıkıcı çeşidi olup hastalığın gelişimi için uygun koşullar

sağlandığında bir ay içinde %50'ye varan kayıplara sebep olurlar. Bununla beraber yüksek nem oranına sahip ve 15<sup>0</sup>C -35<sup>0</sup> C sıcaklıktaki koşullarda başka enfeksiyonlara yol açarlar (Dauber 2008, McIntosh ve ark.1995).

## 1.2. Tezin Hedefi

Bu çalışmanın amacı, sahaya gitmeden; imge işleme tabanlı sınıflandırıcı modellerini kullanarak tarımsal ürünlerin pas hastalığına yakalanıp yakalanmadığının kontrol edilmesi ve eğer pas hastalığına yakalanmış ise tarımsal ilaçlama teknikleri kullanılarak erken tedavisinin sağlanıp ekonomik kayıpların önlenmesidir. Tezin kapsamında uygulama olarak, zambak çiçeğine ilişkin sayısal imgeler işlenip sınıflandırılarak pas hastalıklarının tespiti yapılmıştır.

Bu çalışmanın diğer çalışmalara göre farklılıkları şöyle özetlenebilir.

- Pas hastalıklarının tespitinde ilk kez zambak çiçeği yaprağı değerlendirilmiştir.
- Bir ziraatçının sahaya gitmeden imge işleme tabanlı sınıflandırıcı sistemini kullanarak, bu çiçeklerin pas hastası olup olmadığı hakkında karar vermesini sağlayan bir model tasarlanmıştır.

Sayısal imge işleme teknolojisi tıp, gıda mühendisliği, biyoloji, ziraat vb. birçok alanda geniş bir şekilde kullanılmaktadır. Sayısal imge işleme, imge işlemek için çok daha fazla karmaşık algoritmaların kullanımına imkan vermekte ve bu nedenle, hem basit işlerde daha iyi performans hem de analog araçlarla uygulanması imkansız olan yöntemlerin uygulanmasında kolaylık sağlayabilmektedir (Vajda 1994). Tarımsal alanlarda elde edilen sayısal kamera imgeleri, ürün hakkında gerekli bilgileri içermektedir. Bu yaklaşımla arazinin yüksek uzaysal çözünürlüklerde dış yüzeyindeki değişimlerini görüntülemek, hem daha esnek hem de düşük maliyetlidir. Ayrıca, bitki dinamiklerinin zaman üzerinden belirlenmesi için imgeler hızlı bir şekilde elde edilebilmekte ve kolay bir şekilde arşivlenebilmektedir.

Uygulamalarda, sayısal imgelerin önışlemeden geçirilerek öznitelik verilerinin elde edilmesi ve farklı sınıflandırma yöntemleri ile elde edilen sonuçların değerlendirilmesi, genel olarak izlenen bir yoldur. Bu tez çalışmasında, Gabor Dalgacık Dönüşümü (GWT) ve Gri Seviye Eş-Oluşum Matrisleri (GLCM) metotları öznitelik çıkarma metodu olarak kullanılmıştır. İmgeleri daha küçük bir boyutta ifade

## 1. GİRİŞ

---

edebilmek ve doku analizi çalışmalarında yüksek performans sergilemeleri nedeniyle, GLCM ve GWT yöntemleri tez kapsamında çalışılmıştır. Bu metotlarla ilgili detaylı bilgiler tez kapsamında anlatılmıştır.

Bu çalışmada kullanılan sınıflandırıcılar; Çok Katmanlı Algılayıcı (MLP), k-En Yakın Komşu (k-NN) ve en küçük kareler Destek Vektör Makinesi (LS-SVM) yöntemleridir. YSA'lar, ağırlıklandırılmış şekilde birbirlerine bağlanmış birçok işlem (nöron) biriminden oluşan matematiksel sistemlerdir. Aynı zamanda, giriş ve çıkışlar arasındaki karmaşık ilişkileri modellemek ve veri içindeki örüntüleri bulmak için kullanılabilir. Örüntü tanıma sistemlerinde k-NN yöntemi, sınıfları belli olan bir örnek kümesindeki gözlem değerlerinden yararlanarak, örneğe katılacak yeni gözlemin hangi sınıfa ait olduğunu belirlemek amacıyla kullanılır. LS-SVM ise, sınıflandırma ve regresyon analizi için kullanılan, verileri analiz etme ve örüntüleri tanımayla ilgili denetimli öğrenme yöntemlerini içeren yaklaşımdır. MLP ve LS-SVM, ikili sınıflandırma problemlerinde yüksek başarı performansına sahip olduğundan, k-NN algoritması ise mesafe ölçümü ve komşuluk sayısına bağlı olarak düzgün kümelenmiş verilerin ayrıştırılmasında yüksek başarı performansına sahip olmasından ötürü söz konusu yöntemler tez kapsamında kullanılmıştır.

Zambak çiçeğine ilişkin yaprak imgeleri, 32 sağlıklı imge ve 21 hastalıklı imge olmak üzere toplam 53 imgeden oluşup, iki (1-Sağlıklı veya normal, 2- Hastalıklı) grupta sınıflandırılmıştır. Bitki yaprağındaki pas hastalığının tespiti, imgelerin sınıflandırılmasıyla gerçekleştirilmiştir.

Özetle, sayısal imge işleme teknikleri ile birlikte, GWT ve GLCM özellik çıkartma metotları kullanılarak zambak çiçek imgelerinden elde edilen özellik vektörleri, sınıflandırıcıların giriş örüntüleri olarak kullanılmıştır. Farklı (MLP, k-NN, LS-SVM) sınıflandırıcıların girişlerine uygulanan giriş örüntüleri ile pas hastalığının tespitinde performansı en iyi sistemin belirlenmesine çalışılmıştır.

### 1.3. Tezin İçeriği

İmge işleme ve sınıflandırma teknikleri kullanılarak pas hastalığının tespiti amacıyla hazırlanan bu tez aşağıda anlatıldığı şekilde organize edilmiştir.

Bu tezin ikinci bölümünde, bitkilerde pas hastalığının tespiti ile ilgili geçmişten



günümüze yapılmış çalışmalar ve bununla ilgili literatür taramalarından bahsedilmiştir.

Bu tezin üçüncü bölümünde, öznelik çıkarma kapsamında GLCM, GWT metotları ve sınıflandırma işlemi kapsamında MLP, k-NN, LS-SVM tekniklerinden bahsedilmiştir.

Bu tezin dördüncü bölümünde, bitki yapraklarında görülen pas hastalıklarının tespiti için yapılan uygulamanın detayları anlatılmıştır. Uygulama olarak zambak çiçeğinin yaprağı temel alınmış ve iki aşamada konunun detayı açıklanmıştır. Birinci aşamada tespit sisteminin yapılandırılması, ikinci aşamada ise matlab uygulama sonuçları verilmiştir.

Bu tezin son bölümünde ise, elde edilen sonuçlar irdelenerek katkılar vurgulanmıştır. Ayrıca ileride yapılabilecek benzeri çalışmalar ve uygulama alanları için öneriler tartışılmıştır.



## 2. Kaynak Özetleri

Bu bölümde konuya ilişkin kaynaklar araştırılmış, genel olarak çalışmaların 3 grupta toplandığı saptanmıştır. Bu gruplar; fiziksel parametreler ve yapay zeka yöntemleriyle pas hastalığının tespiti (1.Grup), imge işleme ve yapay zeka yöntemleriyle pas hastalığının tespiti (2.Grup) ve istatistiksel yöntemlerle pas hastalığının tespitidir (3.Grup).

### 1.Grup çalışmalarına ilişkin örnek çalışmalar

Lei ve Shu-qin (2010), Çin'in Han Zhong bölgesinde sarı pas hastalığını tespit etmek için bir model önermişlerdir. Bu modelde, temel bileşenler analizi (PCA) yöntemi kullanılarak 1974 ve 1997 yılları arasındaki bölgedeki veriler analiz edilip sarı pas hastalığını etkileyen ana faktörler çıkarılmıştır. Daha sonra, PCA ile elde edilen ana parametreler yapay sinir ağının girişine verilerek Han Zhong bölgesinde 1994 ve 1997 yılları arasındaki sarı pas hastalığı başarılı bir şekilde tahmin edilmiştir.

Liu ve ark. (2010), çeltik panikülleri içinde meydana gelen farklı mantar enfeksiyonu seviyelerini ayırmak ve sınıflandırmak için temel bileşenler analizi (PCA) ve yapay sinir ağları tekniklerini kullanmışlardır.

Mo (2010), Çin'de buğday bitkilerinde sarı pas hastalığına sebep olan bakteri miktarını ve bunların ortaya çıkma aşamalarını yaz, kış ve ilkbahar aylarında tahmin etmek için üç ayrı geri yayımlı yapay sinir ağı modeli üzerinde çalışmıştır.

Rumpf ve ark. (2010), şeker pancarı yapraklarındaki yaprak benek patojeni *Cercospora beticola*, pas patojeni *Uromyces betae* veya tozlu küf patojeni *Erysiphe betae* hastalıklarının tespiti için Kernel radyal tabanlı destek vektör makineleri modelini ve spektral bitki indisleri (spectral vegetation indices) yöntemini kullanmışlardır. Fiziksel parametrelere bağlı oluşturulan 9 bitki indisi bir öznelik vektörü oluşturarak sınıflandırma için sistemin girişine verilmiştir. Sonuç olarak, şeker pancarı yapraklarındaki hastalıklı ve sağlıklı bitkilerin ayrıştırılmasında %97 performans sağlanmıştır. Yapraklardaki hastalıklara sebep veren üç hastalığın semptomlarının belirlenmesinde ise çoklu sınıflandırma yöntemi kullanılarak %86 doğruluk değeri hesaplanmıştır.

## 2. KAYNAK ÖZETLERİ

---

Bauer ve ark. (2011), şeker pancarı yapraklarındaki yaprak benek patojeni *Cercospora beticola* veya pas patojeni *Uromyces betae* adında farklı iki hastalığın tespiti için üç farklı sınıflandırma modeli kullanmışlardır. Bu modeller sırasıyla k en yakın komşu, Gauss karma modeli kullanarak Bayes sınıflandırma ve Koşullu Rastgele Alanlar (Conditional Random Field) sınıflandırıcılarından oluşmaktadır. Sonuç olarak sınıflandırma performansı *Cercospora beticola* hastalığı için %91 ve *Uromyces betae* hastalığı için %86 bulunmuştur.

Wang ve Ma (2011), Çin'de buğday bitkilerinde en çok görülen hastalık türü olan sarı pas hastalığının destek vektör makineleri ve regresyon analizi modelleriyle tahmini için bir çalışma yapmışlardır. Bu çalışmada, hastalığın epidemisini ve hastalığı etkileyen diğer faktörleri içeren toplam 3 veri kümesi için yapılan tahminler sonucunda, destek vektör makinelerinin regresyon analizi modeline göre daha yüksek tahmin ile çalıştığı ve her bir veri kümesi için %100 performansa ulaşmıştır.

Alves ve ark. (2011), yılında soyada pas monosiklik sürecinin şiddetini tarif etmek için yaprak nemi, sıcaklık ve mantar aşılama sonrası günlerden oluşan 3 adet giriş parametreleriyle, uyarlamalı sinirsel-bulanık bir sınıflandırıcı sistemi üzerinde çalışmışlardır.

Bu grup kapsamında değerlendirilebilecek benzer literatür çalışmaları Ping ve ark. (2000), Takanashi ve ark. (2007)'dir.

### **2.Grup çalışmalarına ilişkin örnek çalışmalar**

Moshou ve ark. (2004), buğday bitkilerinde sarı pas hastalığının otomatik uzaktan algılanması için sağlıklı ve hastalıklı buğday bitkileri arasındaki spektral yansıma farklarını dikkate alarak, bir algoritma önermişlerdir. Bu çalışmada temel alınan yöntemler, öz-örgütlenme sistemleri (Self-Organizing Systems) ve çok katmanlı yapay sinir ağları olup sınıflandırma performansı %99 bulunmuştur.

Sanyal ve ark. (2007), dünya genelinde pirinç üretimine ciddi anlamda zarar veren kahverengi benek ve Blast hastalıklarını tespit etmek için örüntü tanıma metodlarından biri olan çok katmanlı algılayıcı yapay sinir ağı modelini önermişlerdir. Bu çalışmada, 400 pirinç yaprağına ait imgeler işlenerek, kameralı telefonların içine

yerleştirilecek gömülü bir sistemle, çiftçilerin bitki pataloğuna ihtiyaç duymadan hastalığı erkenden tespit edebilmeleri amaçlanmıştır.

Li ve ark. (2010), Çin’de tozlu küf, sarı pas ve kuşgözü lekesi gibi üç farklı hastalığı otomatik tanımak için imgelere ait bazı morfolojik özellikler çıkarmışlardır. Daha sonra morfolojik veriler, temel bileşenler analizi (PCA) ve diskriminant analiz yöntemleriyle analiz edilip toplamda 14 adet morfolojik öznelik çıkarılmıştır. Bu parametrelerden 5 tanesi tanıma faktörü olarak seçilip, üç farklı hastalık örnekleri arasındaki tanınabilirlik oranı sırasıyla %96,7, %93,3 ve %86,7 olarak hesaplanmıştır.

Römer ve ark. (2011), buğday yapraklarında pas hastalığına sebep olan patojen enfeksiyonunu önceden tahmin edebilmek için imge işleme ve sınıflandırma yöntemlerini kullanmışlardır. Bu çalışmada imgelerden elde edilen öznelik vektörleri, destek vektör makineleri sınıflandırıcılarının girişinde kullanılarak %93 sınıflandırma performansı sağlanmıştır. Buğday bitkilerinde pas hastalığını tespit etmek için benzer çalışmalar (Wei 2007, Ping ve ark. 2000, Keqiang ve ark. 1995) tarafından yapılmıştır.

Bu grup kapsamında değerlendirilebilecek benzer literatür çalışmaları Youwen ve ark. (2004), Yanli ve ark. (2005) ‘tir.

### **3.Grup çalışmalarına ilişkin örnek çalışmalar**

Coakley ve ark. (1982), Amerika’da kış buğdayları üzerinde görülen sarı pas hastalığını tespit etmek için istatistiksel yöntemlerden faydalanmışlardır.

Eddy (2009), Amerika’da buğday bitkilerinde sarı pas hastalığını tahmin etmek için Lojistik Regresyon modelini kullanmışlardır.

Cui ve ark. (2009), soyada pas hastalığını tespit etmek için farklı pas şiddetindeki multispektral yaprak imgelerini kullanmışlardır. Daha sonra farklı formlardaki bitki indisleri kullanılarak geniş bant spektrumu içindeki yüzey yansımaları pas şiddetine bağlı olarak değişip, bitki fark indeksi (Difference Vegetation Index) yansıma farklarıyla pozitif bir ilişki göstermiştir. Ancak, bu indeks sadece pas tarafından kaynaklanan yansıma değişimlerini tanımlayamadığından alternatif olarak multispektral imgelerden, enfeksiyon kapmış bölge oranı (Ratio of Infected Area), doku bozulması renk indeksi (Lesion Color Index) ve pas şiddeti indeksi (Rust Severe Index)

## 2. KAYNAK ÖZETLERİ

---

gibi parametreler çıkartılıp, yaprak enfeksiyonu ve enfeksiyonun şiddeti belirlenmeye çalışılmıştır. Sonuçta, laboratuvar test sonuçları ile 3 parametrenin özellikle de pas şiddeti indeksi (RSI) parametresinin pas enfeksiyonuyla pozitif olarak bağlantılı olduğu kanıtlanmıştır.

Luaces ve ark. (2011), dünyadaki kahve bitkileri hastalığının başında gelen kahve pası hastalığına karşı alarm veren bir model önermişlerdir. Bu modelde, belirli bir eşik değeri aşıldığında sistemin alarm vererek çiftçileri uyarması amaçlanmıştır. Bunun için, bir regresyon ve iki deterministik olmayan sınıflandırıcı yöntemleri kullanılmıştır. Regresyon modeli olarak SVM regresyon, sınıflandırıcı olarakta SVM çoklu sınıf sınıflandırıcı ve Lojistik Regresyon sınıflandırıcıları kullanılmıştır. Sonuç olarak, regresyon değeri 0.95 bulunup, sınıflandırıcı performansları sırasıyla %90 ve %78 olarak hesaplanmıştır.

Bu grup kapsamında değerlendirilebilecek benzer literatür çalışmaları Keqiang ve ark. (1995), Wei (2007), Gang ve ark. (2006) 'dır.

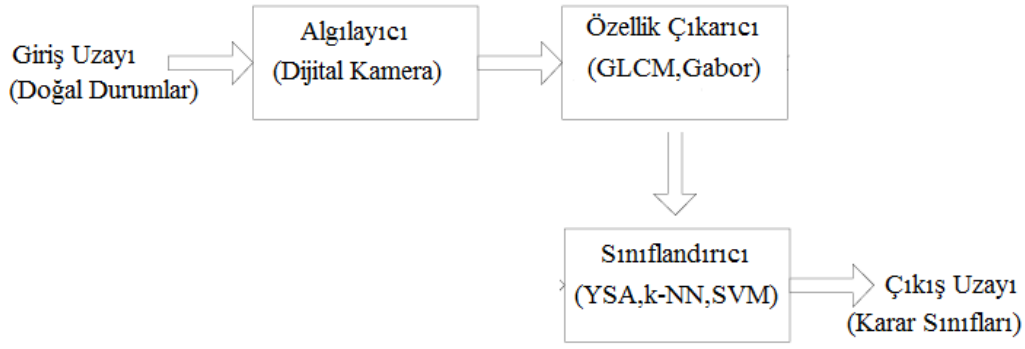
Bu tez çalışmasında, sayısal yaprak imgelerine ilişkin öznelik vektörleri yapay zeka yöntemleri ile sınıflandırılarak, bitki yapraklarında pas hastalıklarının tespiti yapılmıştır. Yapılan çalışma, kapsam olarak 2. grup çalışmaları arasında yer almaktadır.

### 3. Materyal ve Metot

Bitki yapraklarında görülen pas hastalıklarının tespitinde kullanılan yöntemler bu bölümde tanıtılmıştır. Tespit analizi iki aşamadan oluşmaktadır. Birinci aşamada bitki yaprak imgelerinden öznitelik vektörlerinin elde edilmesi, ikinci aşamada ise elde edilen öznitelik vektörleri ile sınıflandırılmasıdır.

#### 3.1. Öznitelik Çıkarımı

Örüntü tanıma uygulamalarının büyük bir çoğunluğunda, eğitim örnekleri çeşitli (boyut, gürültü, parametrelerin değişim aralığı, sistemin çalışma frekans aralığı, vb.) nedenlerden ötürü sınıflandırıcı girişine doğrudan uygulanmaz. Örneğin, nesne tanıma problemlerinde kameralar sayesinde elde edilen imgeler sayısal hale dönüştürülmelidir. Kamera vasıtasıyla gerçek dünyadan elde edilen bir imge bilgisayar ortamında sayısal bir imge olup, ham imge olarak adlandırılır. Bu imgelerin sayısal karşılığı çok boyutlu bir uzayda matris oluşturur. Elde edilen ham imgeler çoğu zaman çok boyutlu olduğundan, örneklerin bu halleriyle sınıflandırıcılara verilmesi sorun oluşturabilir. Bu sorunların başında, örnekte artan boyut sayısına paralel olarak, oluşturulacak eğitim setinin de üstel olarak artan sayıda örnekten oluşması gereğidir (curse of dimensionality). Diğer önemli problem ise çok boyutlu ham imgede bilgiyi işlemek için çok fazla zamana ihtiyaç duyulmasıdır. Bu nedenle imgeleri iyi temsil edecek öznitelikler çıkartılıp, daha küçük boyuttaki bir uzaya haritalanmaları sağlanır. Öznitelik çıkarımında farklı grupta bulunan örneklerin arasındaki korelasyonun mümkün olduğunca az olduğu, aynı grupta bulunan örneklerin arasındaki korelasyonun mümkün olduğunca yüksek olduğu bir yapı istenir. Bu sayede daha basit sınıflandırıcılar tasarlanabilir ve daha az işlem yükü oluşur. Şekil 3.1.'de, bu çalışmada kullanılan örüntü tanıma sisteminin blok diyagramı verilmiştir.



Şekil 3.1. Tez kapsamında kullanılan örüntü tanıma sistemi

Doğal imgelerin en önemli özelliklerinden birisi dokudur. Doku, bir desen veya desenlerin bir bölgede tekrarlanmasıdır. İmgeleri temsil eden öznitelikler, imgelerin kendilerinden doğrudan elde edilebildiği gibi değişik istatistiksel yöntemler veya boyut indirgeme yöntemleriyle de elde edilebilir. Bu çalışmada, imgeleri daha küçük bir boyutta ifade edebilmek ve imge dokularından öznitelik çıkarmak için, doku analizi çalışmalarında en başarılı iki yöntemden bahsedilmiştir. Bu yöntemler sırasıyla, GLCM ve GWT olmak üzere iki kısımda incelenecektir.

#### 3.1.1. Gri Seviye Eş-Oluşum Matrisleri (GLCM)

GLCM, M. Haralick tarafından ortaya atılmış bir öznitelik çıkarma yöntemi olup, gri tonlu bir imajın özniteliğini çıkarmaya yarar. GLCM, iki komşu piksel arasındaki ilişkiyi tanımlar. Bu piksellerden birincisi referans pikseli, ikincisi de komşu piksel olarak bilinir (Horng ve ark. 2003). GLCM, bir imgede art arda gelen iki piksel arasındaki frekans matrisini oluşturur. Matristeki dağılım, pikseller arasındaki mesafe ve açığa göre ayarlanır. GLCM,  $N_g$  boyutlu bir kare matris olup, matrisin her bir elemanı  $d$  mesafesindeki  $i$  ve  $j$  piksel değerlikli çiftin oluşum sayısını belirtir (Roumi 2009).

Öznitelik ölçümleri aşağıdaki adımları içerir:

- Gri seviye eş-oluşum matrisinin simetrik yapılması
- Her bir kombinasyonun olasılığının hesaplanması. Olasılık hesabı Eşitlik (3.1)'de gösterilmiştir.



$$\begin{aligned}
P(i, j, d, \varphi^o) = & \#\{(k, l), (m, n) \in D, \\
& ((k - m), (l - n)) \in \{-d, 0, d\}, I(k, l) = i, \\
& I(m, n) = j, \angle((k, l), (m, n)) = \varphi^o\}
\end{aligned} \tag{3.1}$$

Eşitlik (3.1), yöne bağlı olarak dört farklı matrisin elde edilmesine izin verir.  $P(0^0, d), P(45^0, d), P(90^0, d), P(135^0, d)$ .

Normalize edilmiş eş-oluşum matrisinin  $(i, j)$ . elemanı,  $p(i, j)$  fonksiyonu olarak Eşitlik (3.2)'de gösterilmiştir.

$$p_{d, \varphi^o}(i, j) = \frac{P(i, j, d, \varphi^o)}{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} P(i, j, d, \varphi^o)} \tag{3.2}$$

Eğer eş-oluşum matrisi simetrik ise,  $p(i, j) = (p(i, j) + p(i, j)^T) / 2$  olur. Eşitlikte  $T$  transpoz matrisini gösterir ve  $\vartheta = 0, 45, 90, 135$  olur (Bariamis ve ark. 2007). Eşitlik (3.1) ve (3.2)'de ifade söz edilen parametreler aşağıda tanımlanmıştır.

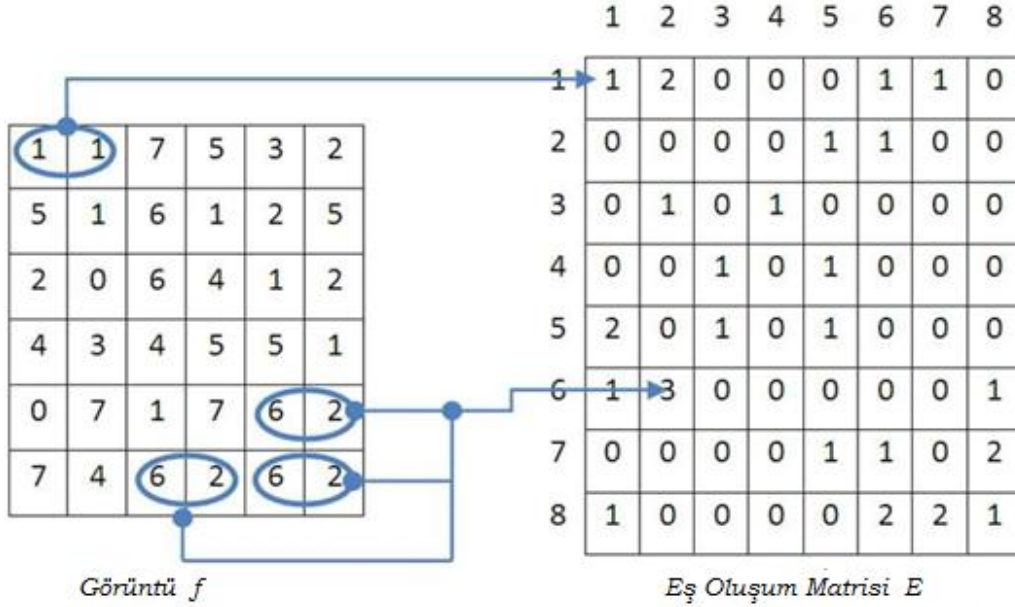
*Gri seviyelerin sayısı ( $N_g$ ):* Normalde bu parametre, 256 gri seviyeli bir gri ton imgesi için kullanılır. Ancak böyle bir imgede olası bütün piksel çiftleri dikkate alınmak zorunda olduğundan dolayı, yüksek hesaplama maliyetine ihtiyaç vardır. Bu durumun önlenmesi için gri seviyelerin sayısı azaltılmalıdır, böylece piksel kombinasyonları azalmış bir matris elde edilebilir. Eş oluşum matrisi daima bir kare matris olup boyutları gri seviye sayısına eşittir.

*Pikseller arası uzaklık ( $d$ ):* Eş-oluşum matrisi, bir imge içinde bulunan belli bir piksel çiftinin kaç defa tekrarlandığını kaydeder. Piksel çiftleri normalde komşudurlar ancak, eş oluşum matrisi ardışık olmayan pikseller arasındaki ilişkiyi analiz etmek için de hesaplanmalıdır. Bu nedenle, pikseller arasındaki uzaklığın daha önceden tanımlı olması gerekmektedir.

*Açı ( $\vartheta^o$ ):* Pikseller arası uzaklığın yanı sıra, piksel çiftlerinin yönlerinin bilinmesine de gerek vardır. En çok ortak bilinen yönler  $\vartheta = 0, 45, 90, 135$  ve bunların simetrik benzerleridir. Şekil 3.2.'de gri seviye sayısı 8, pikseller arası uzaklık  $d = 1$  ve yön açısı  $\vartheta^o = 0$  olarak hesaplanmış bir eş-oluşum matrisinin örneği verilmiştir.

### 3. MATERYAL ve METOT

Burada,  $f$  imge matrisi içindeki (1,1). ve (1,2) koordinatındaki (1,1)'lik piksel çifti bir defa tekrarlandığından dolayı, bu piksel çiftinin eş oluşum matrisi  $E$  deki (1,1). koordinatındaki elemanı, 1'e eşit olur. Benzer şekilde (6,2) piksel çifti,  $f$  imge matrisi içinde 3 defa tekrarlandığından dolayı,  $E$  eş-oluşum matrisindeki karşılığı 3'e eşit olur. Bu adımlar  $f$  imge matrisi içindeki diğer piksel çiftleri için de tekrarlanıp, imgeye ait  $E$  eş oluşum matrisi hesaplanmıştır.

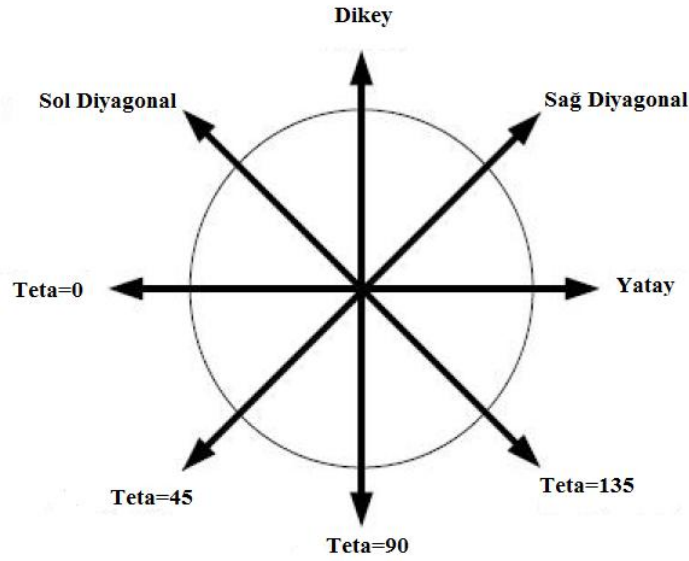


Şekil 3.2. Eş-oluşum matrisinin elde edilmesi

Eş-oluşum matrisi, doku imgesi içerisindeki gri seviyelerin uzaysal dağılımları hakkında bazı özniteliklere sahiptir (Miquel 2009). Dokuya ait bu öznitelikler  $N * N$  boyutlu bir eş-oluşum matrisinin tanımlanması için Haralick tarafından oluşturulmuş olup, Bölüm 3.1.1.1.'de ayrıntılı bir şekilde gösterilmiştir.

#### - İmgenin doku özniteliklerinin hesaplanması

M. Haralick tarafından imge hakkında bilgi içeren 14 tane doku özelliği (AİM, Kontrast, Korelasyon, Korelasyon toplamı: Varyans, TFM, Toplam Ortalama, Toplam Entropi, Toplam Varyans, Entropi, Varyans Farkı, Entropi Farkı, Korelasyon Bilgi Ölçüsü 1, Korelasyon Bilgi Ölçüsü 2 ve Azami Korelasyon Katsayısı) tanımlanmıştır (Roumi 2009). Bu doku öznitelikleri yatay, dikey, sağ diyagonal ve sol diyagonal yönlerindeki açılara göre oluşmaktadır. Bu yönler Şekil 3.3.'te gösterilmiştir.



Şekil 3.3. Doku özneliklerinin hesabı için kullanılan açılı diyagramı

Haralick doku özneliklerinin hesabı için kullanılan eşitlikler aşağıda belirtilmiştir.

$$p_x(i) = \sum_{j=1}^{N_g} p_{d,g^\circ}(i, j), \quad (3.3)$$

$$p_y(j) = \sum_{i=1}^{N_g} p_{d,g^\circ}(i, j), \quad (3.4)$$

$$p_{x+y}(k) = \sum_{i=1}^{N_g} \sum_{j=0}^{N_g} p_{d,g^\circ}(i, j), \quad k = \{2, 3, \dots, 2N_g\}, \quad k = i + j, \quad (3.5)$$

$$p_{x-y}(k) = \sum_{i=1}^{N_g} \sum_{j=0}^{N_g} p_{d,g^\circ}(i, j), \quad k = \{1, 2, \dots, N_g\}, \quad k = |i - j|, \quad (3.6)$$

### 3.1.1.1. Haralick Doku Öznelikleri

Bu yöntem ile her bir imgeye ait 14 adet öznelik çıkarılmış olup, detayları aşağıda belirtilmiştir (Roumi 2009).

#### 1. Açılmalı İkinci Moment (AİM)

Açılmalı ikinci moment, tekdüzellik veya enerji olarak da bilinir ve imgenin homojenliğini ölçer. AİM, pikseller çok benzer olduğu zaman yüksektir.

$$f_1 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_{d,g^o}(i, j)^2 \quad (3.7)$$

#### 2. Kontrast

Kontrast, referans piksel ve onun komşusu arasındaki, yoğunluk veya gri seviye varyasyonlarının ölçümüdür.

$$f_2 = \sum_{m=0}^{N_g-1} m^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_{d,g^o}(i, j) \right\}, |i-j|=m \quad (3.8)$$

#### 3. Korelasyon

Korelasyon, eş oluşum matrisi içindeki gri seviye değerlerinin lineer bağımlılığını hesaplar. Bunun yanında, referans pikselinin komşusuyla nasıl bağlantılı olduğunu gösterir.

$$f_3 = \frac{\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (ij) p_{d,g^o}(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (3.9)$$

Eşitlik (3.9)'da kullanılan  $\mu_x, \mu_y, \sigma_x, \sigma_y$  sırasıyla  $p_x$  ve  $p_y$  fonksiyonlarına ait ortalama ve standart sapma değerlerini gösterir.

#### 4. Kareler Toplamı: Varyans

Bu öznelik gri tonlu varyansın bir ölçüsüdür.

$$f_4 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_{d,g^o}(i, j) (i - \mu)^2 \quad (3.10)$$

#### 5. Ters Fark Momenti (TFM)

Ters fark momenti, bazen de homojenlik olarak adlandırılır ve sayısal bir imgenin yerel homojenliğini hesaplar. TFM, GLCM elemanlarının dağılımına ait yakınlık ölçümlerini diyagonal GLCM'ye dönüştürür.

$$f_5 = \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} \frac{1}{1 + (i-j)^2} p_{d,g^o}(i, j) \quad (3.11)$$

6. *Toplam Ortalama (Sum Average)*

$$f_6 = \sum_{i=2}^{2N_g} ip_{x+y}(i) \quad (3.12)$$

7. *Toplam Entropi (Sum Entropy)*

$$f_7 = -\sum_{i=2}^{2N_g} p_{x+y}(i) \log p_{x+y}(i) \quad (3.13)$$

Eşitlik (3.13)' te olasılık 0'a eşit ise,  $\log p_{x+y}(0)$  tanımlı değildir. Bunu önlemek için,  $\log(p)$  yerine  $\log(p + \varepsilon)$  kullanılmalıdır. Burada  $\varepsilon$  rastgele tanımlanmış küçük pozitif bir sabittir.

8. *Toplam Varyans (Sum Variance)*

$$f_8 = \sum_{i=2}^{2N_g} (i - f_7)^2 p_{x+y}(i) \quad (3.14)$$

9. *Entropi*

Entropi, imge sıkıştırma için gerekli olan imgenin bilgi miktarını gösterir.

$$f_9 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_{d,g^o}(i, j) \log p_{d,g^o}(i, j) \quad (3.15)$$

10. *Varyans Farkı (Difference Variance)*

$$f_{10} = \text{varyans}(p_{x-y}) \quad (3.16)$$

11. *Entropi Farkı (Difference Entropy)*

$$f_{11} = -\sum_{i=0}^{N_g-1} p_{x-y}(i) \log p_{x-y}(i) \quad (3.17)$$

12. *Korelasyon Bilgi Ölçüsü 1 (Information Measures of Correlation 1)*

$$f_{12} = \frac{GXY - GXY1}{\max(GX, GY)} \quad (3.18)$$

Eşitlik (3.18)'de bulunan parametreler aşağıda tanımlanmıştır.

$$GXY = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_{d,g^o}(i, j) \log(p_{d,g^o}(i, j)) \quad (3.19)$$

$GX$  ve  $GY$  sırasıyla  $p_x$  ve  $p_y$ 'nin entropi değerleridir.

$$GXY1 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_{d,g^o}(i, j) \log(p_x(i)p_y(j)) \quad (3.20)$$

$$GXY2 = -\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p_x(i)p_y(j) \log(p_x(i)p_y(j)) \quad (3.21)$$

#### 13. Korelasyon Bilgi Ölçüsü 2 (Information Measures of Correlation 2)

$$f_{13} = (1 - \exp[-2.0(GXY2 - GXY)])^{1/2} \quad (3.22)$$

#### 14. Azami Korelasyon Katsayısı (Maximal Correlation Coefficient)

$$f_{14} = (Q \text{ matrisinin ikinci en büyük özdeğeri})^{1/2} \quad (3.23)$$

$$Q(i, j) = \sum_m \frac{p_{d,g^o}(i, m)p_{d,g^o}(j, m)}{p_x(i)p_y(m)} \quad (3.24)$$

$N \times N$  boyutlu bir imge için Haralick karışıklığı,  $Q(N^2)$ 'ye eşittir.

#### 3.1.2. Gabor Dalgacık Dönüşümü (GWT)

Gabor fonksiyonları gürültü içindeki sinyallerin algılanması için ilk olarak Dennis Gabor tarafından önerilmiştir. Gabor'a göre, bilgi için bir "kuantum ilkesi" vardır (Gabor 1946). Buna göre, 1 boyutlu (1-B) sinyaller için birleşik zaman-frekans bölgesi sayısal olarak zorunlu bir şekilde belirtilmelidir, bu yüzden herhangi bir sinyal veya filtre bu bölge içindeki belirli en düşük alandan daha az yer tutamamaktadır. Zaman ve frekans çözünürlüğü arasında bir denge olup en iyi dengenin Gauss'un modüle edilmiş karmaşık üssel fonksiyonlar tarafından sağlandığı ortaya çıkarılmıştır. Böyle bir durumda, orijinal temel Gabor fonksiyonları, kipleyci dalganın frekansı değişirken sabitleştirilmiş Gauss fonksiyonları ile oluşturulmaktadır.

GWT, ana dalgacığı Gabor fonksiyonu olan dalgacık dönüşümü olarak ifade

edilebilir. Frekans ve yönelim karakteristiklerine göre insan görsel sistemiyle büyük benzerlik gösteren Gabor dalgacıkları bilgisayar vizyon uygulamalarında ve biyolojik vizyonların modellenmesinde, özellikle doku betimlemeleri, yüz tanıma, parmak izi tanıma ve ayırma algoritmalarında kullanılmaktadır (Buciu ve Gacsadi 2009). Gabor dalgacıkları gerek yönelim gerekse uzaysal lokalizasyon için iyi nitelikli bir filtre oluşturmaktadır (Ghosal ve ark. 2009). GWT yönteminin söz konusu özelliklerinden dolayı, tez kapsamında kullanılmıştır.

1-B sinyalleri frekans ve zaman düzlemlerinde en az belirsizlikle D.Gabor tarafından tanımlanmıştır ve orijinal Gabor filtre yapısı Daugman tarafından iki boyutlu (2-B) hale getirilmiştir. Daugman'ın geliştirmiş olduğu 2-B Gabor filtresi, 2-B Gauss zarfı ile modüle edilmiş, belirli bir frekans ve yönelimdeki karmaşık sinusoidal dalgadır. Daugman, yön seçici basit hücrelerin alıcı kısımlarını modellemek için Gabor fonksiyonunu 2-B formunda Eşitlik (3.25)'teki gibi genellemiştir (Daugman 1980; Kepenekçi 2001).

$$\psi_i(x) = \frac{\|\vec{k}_i\|}{\sigma^2} \exp\left(-\frac{\|\vec{k}_i\|^2 \|\vec{x}\|^2}{2\sigma^2}\right) \left[ \exp\left(j \vec{k}_i \cdot \vec{x}\right) - \exp\left(-\frac{\sigma^2}{2}\right) \right] \quad (3.25)$$

Eşitlik (3.25)'te tanımlı her bir  $\psi_i(x)$ ,  $k_i$  vektörü ile tanımlanmış ve bir Gauss fonksiyonu tarafından zarflandırılmış bir düzlem dalgasını ifade etmektedir. Burada  $\sigma$ , Gauss fonksiyonun standart sapması olup  $i$ 'nci filtrenin merkez frekansının ölçü ve yönelimi  $(\vec{k}_u, \mathcal{G}_v)$  belirten karakteristik dalga vektörü Eşitlik (3.26)'da ifade edilmiştir.

$$\vec{k}_i = \begin{pmatrix} \vec{k}_{ix} \\ \vec{k}_{iy} \end{pmatrix} = \begin{pmatrix} \vec{k}_u \cos \mathcal{G}_v \\ \vec{k}_u \sin \mathcal{G}_v \end{pmatrix} \quad (3.26)$$

Bunun yanında Eşitlik (3.25)'te köşeli parantez içindeki ilk terim, çekirdek (kernel) 'in salınım yapan kısmını ifade ederken ikinci terim de çekirdeğin DC değerini karşılamaktadır. Bu eşitlikte DC değeri çıkarıldığı zaman Gabor filtresi aydınlatmanın

### 3. MATERYAL ve METOT

tüm seviyelerine karşı duyarsız olmaktadır.

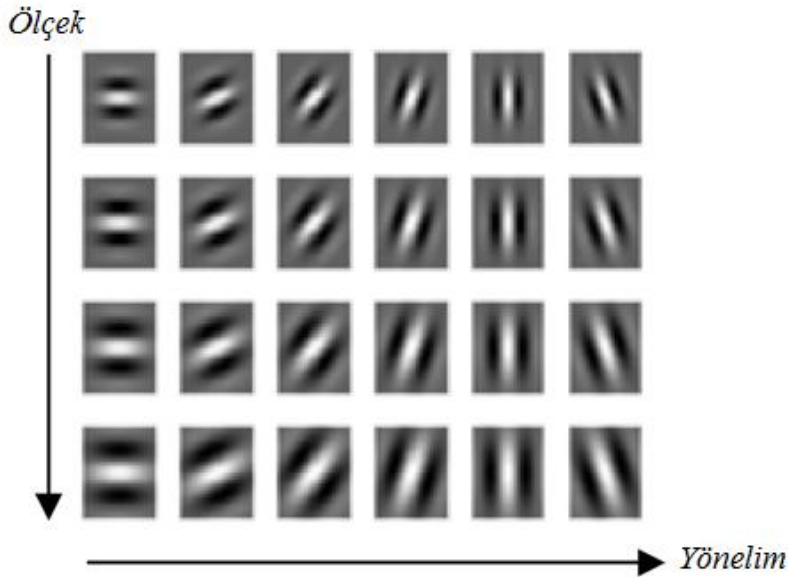
Son zamanlardaki nörofizyolojik kanıtlar, farklı boyutlardaki basit hücrelerin alıcı kısımlarına ait uzaysal yapısının gerçekte değişken olmadığını göstermiştir. Daugman (Lee 1996) ve diğerleri (Hubel ve Wisel 1978), (Marcelja 1980) bir grup basit hücrenin, 2-B Gabor dalgacıklarının bir ailesi olarak en iyi şekilde modellendiğini ileri sürmüştür. Bu sınıf, dönme ve genişleme ile üretilen birbirlerine çok yakın durumların bir ailesine eşdeğerdir.

Bir  $G$  imgesinin bu durumların içine doğru ayrışması, bu imgenin *dalgacık dönüşümü* olarak adlandırılır.

Kompleks bir Gabor filtresi, Gauss Kernel ile karmaşık bir sinüsoidin çarpımı şeklinde tanımlanır. 2 boyutlu bir GWT,  $G(\vec{x})$  imgesinin konvolüsyonuyla ifade edilir (Buciu ve Gacsadi 2009):

$$J_i(\vec{x}) = \int G(\vec{x}') \psi_i(\vec{x} - \vec{x}') d\vec{x}' \quad (3.27)$$

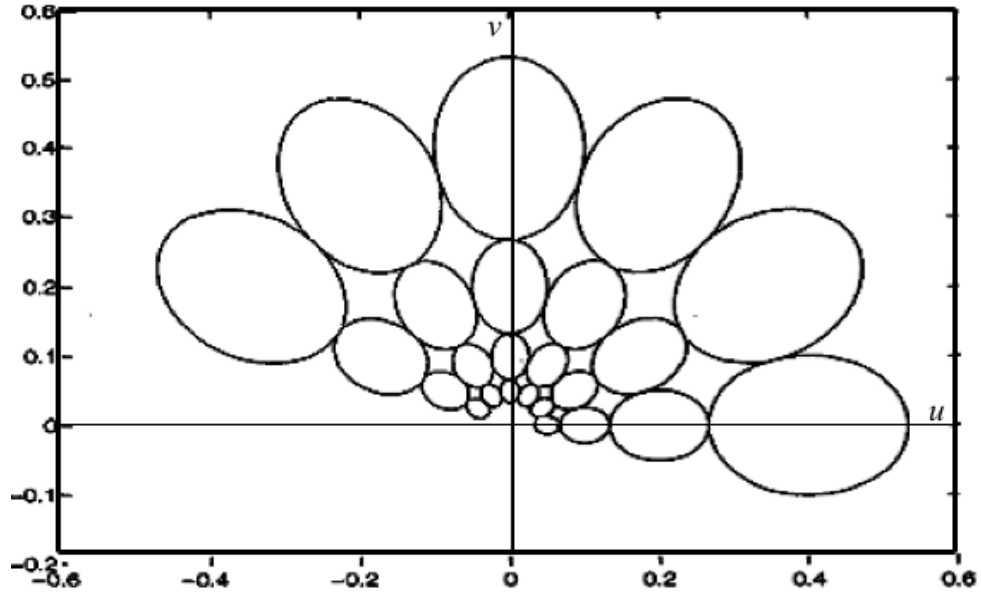
Eşitlikte  $G(\vec{x})$ ,  $\vec{x}$  değerindeki imge yoğunluğunu,  $\psi(x)$  ise Gabor filtrelerinin bir kümesini oluşturmaktadır. Şekil 3.4.'te gri kodlanmış farklı oryantasyonlara sahip Gabor dalgacık kümesinin bir örneği gösterilmiştir.



Şekil 3.4. Gabor dalgacık kümesi



Üst merkez frekansı  $u = 0.4$ , alt merkez frekansı  $u = 0.05$ , 6 yönlü ve 4 ölçekli Gabor filtre seti içindeki filtre cevaplarının yarım tepe noktası büyüklüğüne karşılık gelen çizgiler Şekil.3.5.'te gösterilmiştir.



Şekil 3.5. Gabor dalgacık kümesinin uzaysal frekans düzlemindeki kapsamı

Gabor dalgacık sayısı, her bir ölçek ve yönelim çiftine ait bir dalgacık olmak üzere hesaplanmaktadır. Bu çalışmada Gabor öznitelik vektörleri, imgenin 8 dalgacık (2 ölçek ve 4 yönelim) matrislerinin her birine ait istatistiksel değişkenlerinin bir vektör oluşturacak şekilde art arda bağlanmasıyla oluşturulmuştur. Her bir dalgacık dönüşümü matrisinden sırasıyla standart sapma, ortalama ve entropi değerleri hesaplanmıştır. Sonuç olarak her bir imgeye ait 8 dalgacık dönüşümü matrisinin istatistiksel değerleri art arda eklenerek toplamda  $8 \times 3 = 24$  uzunluğunda öznitelik vektörü sınıflandırıcının girişine verilmek üzere hesaplanmıştır.

### GWT tabanlı parametreler:

Gri tonlamalı imgelere ait her bir imgeden elde edilen matrix  $M \times N$  boyutlu  $J(i,j)$  matrisi olsun (Manian ve Vasquez 1997). Buna göre:

#### 1.Ortalama

$$m_{ij} = \frac{1}{MN} \sum_i^N \sum_j^M J(i, j) \quad (3.28)$$

#### 2. Standart Sapma

$$s_{ij} = \frac{1}{MN} \sum_i^M \sum_j^N (J_{ij} - m_{ij})^2 J_{ij} \quad (3.29)$$

#### 3. Entropi

$$ent_{ij} = - \sum_{i,j} J(i, j) \cdot \log(J(i, j)) \quad (3.30)$$

### 3.2. Sınıflandırma Yöntemleri

Sınıflandırma basamağında, örüntülerin özellik uzaylarına göre kendilerine en yakın sınıflara en az hata ile eşleştirilmesi hedeflenir. Sınıflandırıcının performansı iyi belirlenmiş özelliklere bağlıdır.

Sınıflandırıcılar, geleneksel ve akıllı olmak üzere iki grupta incelenebilir. Geleneksel sınıflandırıcılar, istatistiksel bir yöntem olan Bayes karar teorisi üzerine inşa edilmiştir. Geleneksel sınıflandırıcılara örnek olarak; en yakın komşu (k-NN), Fisher'in doğrusal sınıflandırıcıları, maksimum olabilirlik, ikili ağaç yöntemi ve çok değişkenli Gauss modelleri verilebilir. Bu sınıflandırıcılarda, özellik uzayı sınıflandırma uzayına dönüştürülürken bazı gürültülerin ortaya çıkması ve her bir sınıf için hata yönteminin belli olmaması önemli bir dezavantajdır. Akıllı sınıflandırıcılar ise YSA, SVM, genetik algoritma, bulanık mantık, gibi sınıflandırıcıları kapsar. Bunlar günümüzde başarılarını ispatlamış, sık kullanılan sınıflandırıcı türlerinden olup, özellikle genelleme yetenekleriyle örüntü tanıma problemlerinde çok büyük boyutlu verilerin sınıflandırılması için kolaylık sağlar (Kaymaz 2007).

Bu çalışmada akıllı sınıflandırıcı yöntemlerinden YSA ve SVM yöntemleri ile geleneksel sınıflandırıcılardan k-NN yöntemi kullanılmıştır. YSA ve SVM, ikili sınıflandırma problemlerinde genellikle yüksek başarılar sağladığından dolayı tercih edilmiş, k-NN algoritması ise mesafe ölçümü ve komşuluk sayısına bağlı olarak düzgün kümelenmiş verilerin ayrıştırılmasında yüksek başarı performansına sahip olmasından ötürü tez kapsamında kullanılmıştır.

### 3.2.1. Yapay Sinir Ağları (YSA)

YSA paralel dağılmış bir bilgi işleme sistemidir. “YSA, insan beyninin çalışma prensibini örnek alarak oluşturulmuş bir veri işleme yöntemidir. YSA’yı diğer yöntemlerden farklı kılan en önemli özelliklerden biri programlama yerine örneklerle öğrenme yöntemini izlemesidir. Bu özellik sayesinde diğer programların sahip olduğu çoğu olumsuzluk YSA da yoktur. YSA’nın çalışma prensibi ile insan beyninin çalışma prensibi arasında ortak yönler vardır (Öztemel 1996).

Bu sistem işaret kanalları (bağlantıları) ile birbirine bağlanan işlem elemanlarından oluşur. Çıkış işareti bir tane olup isteğe göre çoğaltılabilir. YSA yaklaşımının temel düşüncesiyle, insan beyninin fonksiyonları arasında benzerlik vardır. Bu yüzden YSA sistemine insan beyninin modeli denilebilir (Kohonen 1990). YSA, giriş veri setleri ve bunlara karşılık gelen hedef değerlerle ilgili örüntüleri hem tanıyabilir hem de öğrenebilir. Eğitimden sonra, YSA yeni bağımsız giriş verileri sonucunu tahmin etmek için kullanılabilir. YSA insan beyninin öğrenme sürecini taklit edebilir ve doğrusal olmayan karmaşık problemleri gürültülü ve özensiz olsa bile işleyebilir (Demirhan ve Güler 2010).

YSA çevre şartlarına göre davranışlarını şekillendirebilir. Girişler ve istenen çıkışların sisteme verilmesi ile kendisini farklı cevaplar verebilecek şekilde ayarlayabilir. Ancak son derece karmaşık bir içyapısı vardır. Onun için bugüne kadar gerçekleştirilen YSA; biyolojik fonksiyonların temel nöronlarını örnek alarak yerine getiren kompoze elemanlar olmuştur.

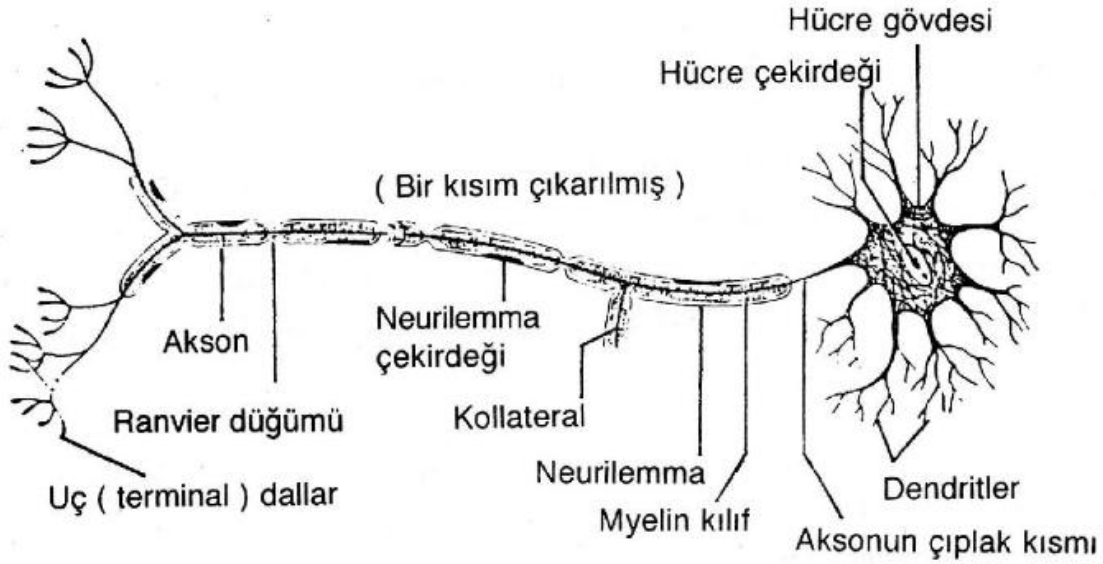
İnsan beyninin nasıl çalıştığı henüz mükemmel bir seviyede keşfedilmiş sayılmaz. Günümüzde bilgisayarlarla her ne kadar karmaşık matematiksel hesaplamaları hızlı ve doğru bir biçimde yapmak mümkün olsa bile, aynı bilgisayarlarla beynin birçok basit fonksiyonunu yerine getirmek kolay olmamaktadır. Benzer şekilde biyolojik beyin, eğitim ile öğrenme ve bilgiyi kendiliğinden yorumlama, hatta tamamlanmamış bilgilerden sonuçlar çıkartma yeteneğine sahiptir.

Bu, daha çok biyolojik sistemlerin, hücreler üzerinde dağıtılmış bilgiyi paralel olarak işleme özelliklerinden kaynaklanır. Hücreler birbirine bağlı ve paralel çalıştıklarından bazılarının işlevini yitirmesi halinde, diğerleri çalıştığı için sinir sistemi, fonksiyonunu tamamen yitirmez. YSA, bu özelliklere sahip olacak şekilde

### 3. MATERYAL ve METOT

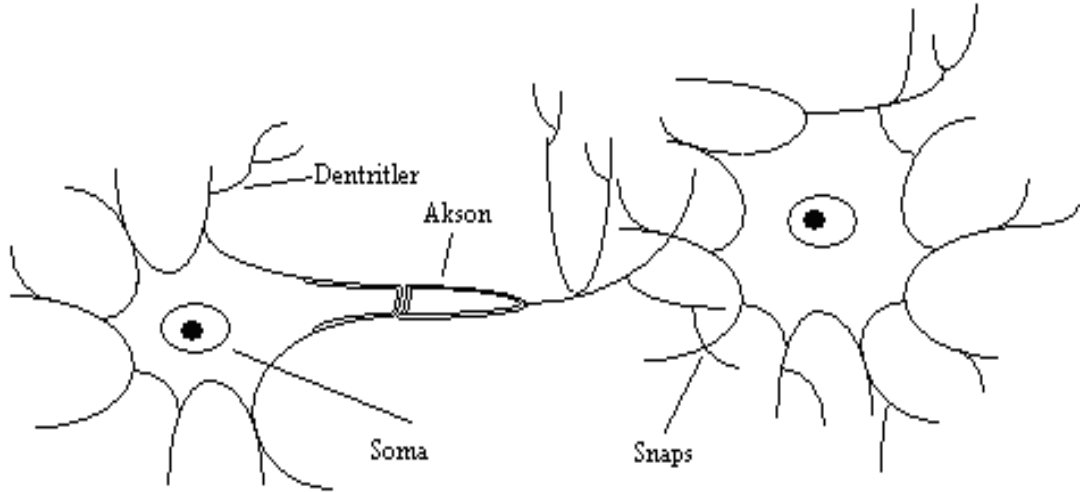
geliştirilmektedir. YSA'ları daha iyi anlamak için, biyolojik sinir hücrelerini gözlemlemek yararlı olacaktır (Taşkafa 2006; Acar 2010).

Biyolojik sinir ağlarında sinir sisteminin temel fonksiyonu nöron olarak adlandırılır. Nöronlar sinirsel uyarıları iletmeye yararlar. Bir nöron, gövde, gövdeye giren işaret alıcıları (dentrit) ve gövdeden çıkan işaret iletici (akson) olmak üzere üç kısımdan oluşur. Başka hücrelerden gelen uyarılar dentritlerin uçlarından alınır ve aksonların uçlarından diğer hücrelere iletilir. Aksonlar örtülerine göre miyelinli ya da miyelinsiz olarak sınıflandırılabilirler. Şekil 3.6.'da miyelinli bir nöronun yapısı gösterilmiştir.



Şekil 3.6. Miyelinli bir nöron yapısı

İki nöron arasındaki temas alanına sinaps adı verilir. Bu nöronlar arasında fiziksel bir temas yoktur. Elektrik sinyalleri kimyasal etkileşim sayesinde gönderilir. Şekil 3.7.'de bir biyolojik nöronun yapısı gösterilmiştir.

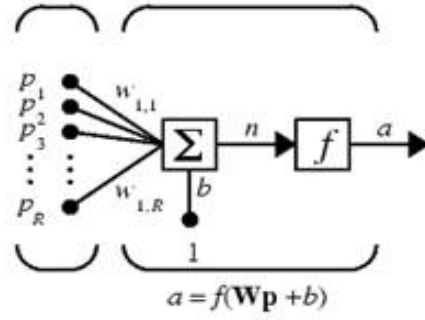


Şekil 3.7. Biyolojik nöron şematik yapısı

Nöronlar arasındaki sinaptik bağlantıların ayarlanması ile biyolojik sistemlerde öğrenme oluşturulur. İnsan yaşamı boyunca tecrübeler edindiğinden, bu tecrübelerin sinaptik bağlantılar üzerinde etkiye sahip olduğu ve öğrenmenin bu şekilde geliştiği düşünülmektedir.

Yapay sinir ağlarında bu işlemleri sağlamak için ağırlık fonksiyonları kullanılmaktadır, insanın deneme yanılma yoluyla öğrenmesi yapay sinir ağlarında yinelemeli eğitim sayesinde gerçekleştirilmektedir.

$p$  girişli ve  $\alpha$  çıkışlı bir yapay nöronun grafik gösterimi Şekil 3.8.'de verilmiştir. Bu gösterimlerde  $f$  aktivasyon fonksiyonunu,  $p$  nöron giriş sayısını,  $\alpha$  çıkışı,  $w$  ise ağırlıkları ifade etmektedir.  $p$  girişi önce  $w$  ağırlığı ile çarpılır. Ardından  $b$  bias değeri ile toplanır ve transfer fonksiyonundan geçirilerek  $\alpha$  çıkışı hesaplanır. Transfer fonksiyonu doğrusal ya da doğrusal olmayan türevlenebilir bir fonksiyondur. İşte bu basit nöron modelinde, giriş ile çıkış arasında basit bir matematiksel bağıntı kurulmuştur. Fakat nöronun arzu edilen çıkışı verebilmesi için  $w$  ve  $b$  ağırlık değerlerinin optimum olacak şekilde ayarlanması gerekmektedir. Birden fazla girişi olan nöron modeli aşağıda verilmiştir.



Şekil 3.8. Yapay nöron modeli

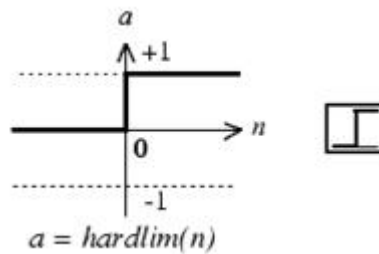
Burada,  $R$  nöronun giriş sayısıdır. Şekilde görüldüğü gibi giriş sayısı kadar da  $w$  ağırlık değeri vardır. Bu durumda yukarıdaki modelin matematiksel ifadesi aşağıdaki gibi olur:

$$\alpha = f\left(\sum_{i=1}^R w_i p_i + b\right) \quad (3.31)$$

Aktivasyon (eşik) fonksiyonları, nörona gelen net girdiyi işleyerek nöronun bu girdiye karşılık vereceği tepkiyi belirler. En çok kullanılan dört tane eşik (aktivasyon) fonksiyonu vardır. Bunlar sırasıyla hard-limit, lineer, log-sigmoid ve tan-sigmoid transfer fonksiyonları olup, aşağıda detayları anlatılmıştır (Acar 2010).

#### 1. Hard-limit transfer fonksiyonu

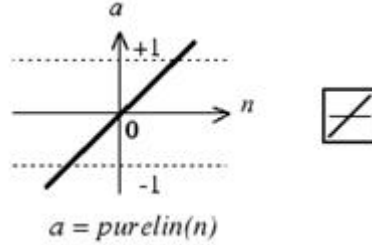
Keskin-limitli (hard-limit) transfer fonksiyonunun grafiği Şekil 3.9.' da verilmiştir. Burada, giriş fonksiyonu  $n$ , çıkış ise  $\alpha$ 'dır:  $\alpha = f(n)$ . Bu transfer fonksiyonunda, giriş değeri sıfırdan büyükse çıkış değeri bir, giriş değeri sıfırdan küçükse çıkış değeri sıfır olmaktadır. Sınıflandırma sistemlerinde bu fonksiyon türüne sık rastlanır.



Şekil 3.9. Hard-Limit Transfer Fonksiyonu

### 2. Lineer transfer fonksiyonu

Doğrusal (lineer) transfer fonksiyonunun grafiği Şekil 3.10.'da verilmiştir. Şekilde görüldüğü gibi giriş değeri hiçbir değişikliğe uğramadan aynen çıkmaktadır ( $a = n$ ). Doğrusal süzgeç problemlerinde bu fonksiyonun kullanımına sık rastlanır.

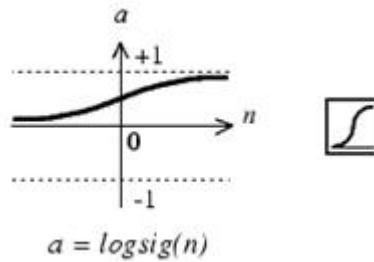


Şekil 3.10. Lineer Transfer Fonksiyonu

### 3. Log-sigmoid transfer fonksiyonu

Sigmoid fonksiyonunun grafiği Şekil 3.11.'de verilmiştir. Sigmoid, doğrusal olmayan logaritmik bir fonksiyondur. Çıkış değeri giriş değerinden bağımsız her zaman 0 ile 1 arasında olmaktadır. Geriye yayılım (Back propagation) algoritmalarında türevlenebilir olduğu için kullanılabilir. Doğrusal olmayan problemleri çözmek için kullanılır. Matematiksel ifadesi aşağıdaki gibidir:

$$a = \text{logsig}(n) = 1 / (1 + \exp(-n)) \quad (3.32)$$

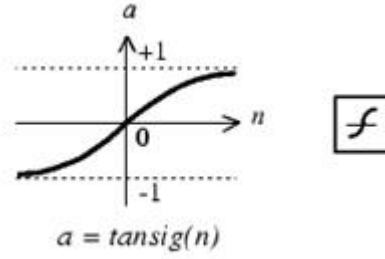


Şekil 3.11. Log-Sigmoid Transfer Fonksiyonu

### 4. Tan-sigmoid transfer fonksiyonu

Tanjant hiperbolik fonksiyonu, sigmoid fonksiyonuna benzer bir fonksiyondur. Sigmoid fonksiyonunda çıkışlar 0 ile 1 arasında değişirken hiperbolik tanjant fonksiyonunun çıkışları -1 ile 1 arasında değişmektedir. Hiperbolik tanjant sigmoid fonksiyonunun grafiği Şekil 3.12.'de verilmiştir.

$$a = \tan \text{sig}(n) = 2/(1 + \exp(-2 * n)) - 1 \quad (3.33)$$



Şekil 3.12. Tan-Sigmoid Transfer Fonksiyonu

(Toplama) fonksiyonu, işlem elemanına gelen net girişi hesaplayan bir fonksiyondur. Net giriş genellikle gelen bilgilerin ağırlıklarla çarpılıp toplanması ile belirlenir. Birleştirme (toplama) fonksiyonundan çıkan net toplam hücrenin çıktısını oluşturmak üzere aktivasyon fonksiyonuna iletilir. Aktivasyon fonksiyonu seçilirken fonksiyonun türevinin kolay hesaplanabilir olmasına dikkat edilmelidir. Çünkü geri beslemeli ağlarda aktivasyon fonksiyonunun türevi de kullanıldığından hesaplamaların hızlı gerçekleşmesi için türevi kolay alınan bir fonksiyon seçilmelidir (Kakıcı 2011).

Aktivasyon fonksiyonundan çıkan değer nöronun çıkışını oluşturur. Bu değer yapay sinir ağının bir çıktısı olarak hem dış dünyaya verilebilir hem de tekrardan ağın içinde kullanılabilir (Kakıcı 2011).

#### 3.2.1.1. Yapay Sinir Ağlarının Genel Özellikleri

Yukarıda genel olarak yapılan açıklamalardan YSA'nın karmaşık problemleri çözebilme yeteneğini ve bilgi işleme gücünü; öğrenebilme, yorumlama yeteneklerinin yanı sıra paralel dağılmış mimarisinden aldığı söylenebilir. Günümüzde birçok alanda YSA kendisini kanıtlamıştır ve aşağıdaki özellikleri sayesinde çoğu yerde uygulama alanı bulmuştur.

1. YSA doğrusal değildir. Bu özelliği bütün ağa yayılmış olduğundan doğrusal olmayan karmaşık problemlerin çözümünde önemli bir faktöre sahiptir.
2. YSA'nın istenilen davranışı gösterebilmesi için hedefe uygun olarak ayarlanması gerekir. Bunu yapmak için de hücrelerin uygun bağlantılara ve bağlantıların da uygun ağırlıklara sahip olması gerekir. Ancak YSA karmaşık bir yapıya sahip olduğundan ağırlıkları ve bağlantıları önceden ayarlamak güçtür.



Bu yüzden, YSA ilgilendiği problemle ilgili eğitim örneklerini alarak problemi öğrenmeye çalışır.

3. YSA problemle ilgili öğrenme işini gerçekleştirdikten sonra daha önce karşılaşmadığı test örnekleri için de genelleme yapabilir. Örneğin, YSA yüz tanıma problemlerinde eğitimle öğrenmeyi sağladıktan sonra daha önce sisteme verilmemiş farklı yüzleri tanımak için tahminde bulunabilir.
4. YSA ilgilendiği problemde meydana gelen değişikliğe göre ağırlıklarını ayarlar. Yani belirli bir problemi tanımak amacıyla eğitilmiş YSA, problemdeki değişikliklere göre tekrar eğitilebilir. Bu özelliği sayesinde sinyal işleme, örüntü tanıma ve sistem tanılama gibi konularda kullanımı etkindir.
5. YSA çok sayıda nöronun paralel bağlanmasından oluşan bir ağ yapısına sahip olduğundan, ağına sahip olduğu bilgi, ağın tüm bağlantıları üzerinde dağılmıştır. Ağ içindeki bazı bağlantıların ya da nöronların kaybolması, ağın doğru bilgi üretimini ciddi anlamda etkilemez. Bu yüzden, hatayı tolere etme yetenekleri önemli derecede yüksektir.
6. YSA'nın diğer bir özelliği donanım ve hızıdır. Paralel yapısı nedeniyle büyük ölçekli bütünleşmiş devre teknolojisiyle gerçekleştirilebilir. Bu özellik gerçek zamanlı uygulamalarda YSA'nın tercih edilmesini ve daha hızlı çalışmasını sağlar (Acar 2010).

### **3.2.1.2. Yapay Sinir Ağlarının Avantaj ve Dezavantajları**

Yapay sinir ağlarının sahip olduğu temel avantajların birkaçı aşağıda verilmiştir.

- YSA'nın temel işlevi bilgisayarın öğrenmesini sağlamaktır. Problemleri öğrenerek benzer problemlere mantıklı cevap verebilirler.
- Geleneksel programlara nazaran bilgi işleme yöntemleri daha farklıdır. Bu nedenle diğer programların sahip olduğu çoğu olumsuzluk YSA'da yoktur.
- Bilgiler ağına tamamına yayıldığından, nöronlardan bazılarının fonksiyonunu kaybetmesi hataya yol açmaz.

### 3. MATERYAL ve METOT

---

- YSA'nın öğrenebilmesi için örneklerin ağa gösterilerek istenilen hedeflere göre ağın eğitilmesi gerekmektedir. Ağın başarısı için problemin sisteme her yönlü tanıtılması gerekir. Ağın performansı seçilen örneklerle doğru orantılıdır.
- YSA'lar eğitildikten sonra daha önce hiç karşılaşmadıkları örnekler hakkında bilgi üretebilirler.
- Algılamaya yönelik uygulama alanları YSA'nın en başarılı olduğu kullanım alanlarıdır. Bu konuyla ilgili problemlerde başarıları kanıtlanmıştır.
- Örüntü tanıma problemleri ve sınıflandırmalarda kullanılabilirler. Kendisine verilen örneklerin kümelenmesi ve bir sonraki verinin hangi kümeye dahil olabileceği hakkında karar verebilirler.
- Ağa eksik verilen örüntüleri başarılı bir şekilde tamamlayabilirler.
- YSA'lar online olarak öğrenebilir ve kendi kendilerini eğitebilirler.
- Ağlar problemin ortaya çıkmasıyla derhal bozulmazlar, zaman içinde dereceli bozulma gösterirler.
- Bilgiler dağınık bir şekilde ağda yayılmıştır. Bu yüzden tek bir bağlantının kendi başına anlamı yoktur.

Yukarıda bahsedilen temel avantajların yanında YSA'nın daha çok avantajı vardır. Ancak bu avantajlarının yanında dezavantajları da vardır. Bazı temel dezavantajları aşağıdaki gibidir:

- Donanım bağımlı olduklarından paralel işlem yapabilme yetenekleri paralel çalışan işlemciye göre değişir.
- YSA'da problemin çözümüne yönelik uygun bir ağ yapısının oluşturulmasında belirli bir kural yoktur. Bu yapı deneme yanılma yöntemleriyle belirlenmektedir.
- Ağın eğitimini sağlayan momentum katsayısı, öğrenme katsayısı, ara katman nöron sayısı gibi parametre değerlerinin belirlenmesinde belirli bir kural yoktur.
- YSA'lar sayısal verilerle çalıştığından, problemler YSA'ya verilmeden önce mutlaka sayısal verilere dönüştürülmelidir. Bu mekanizmanın iyi bir şekilde

belirlenmesi kullanıcının yeteneğine bağlı olup, doğrudan ağın performansını etkiler.

- Ağın eğitiminin ne zaman bitirileceğine dair belirli bir metot yoktur. Belirli bir hata değeri referans alınarak eğitim tamamlanır (Öğücü 2006).

### **3.2.1.3. Yapay Sinir Ağlarının Kullanım Alanları**

Yapay sinir ağları sınıflandırma, tahmin ve modelleme amaçlı uygulamaların yanı sıra birçok alanda kullanılmaktadır. Metni konuşmaya çevirme, otomatik dil çevirme, makineler için sesli komutlar, sesle kilitlenmiş güvenlik sistemleri, sağır ve bedensel engelliler için yardımcı uygulamalar gibi lisan işleme alanlarında, metin tanıma, örüntü tanıma, karakter tanıma, sinyal işleme, mali alanlar ve güdümlerde kullanılmıştır.

Yapay sinir ağları farklı sektörlerde de geniş uygulama alanı bulmuştur. Bunlardan bazıları; otomatik pilot uygulamaları, uçuş simülasyonları uygulamaları gibi uzay alanında; yol durumuna göre sürüş analizi, otomatik yol izleme uygulamaları gibi otomotiv alanında; bütçe yatırım tahminleri, müşteri analizleri, kredi uygulamalarının geliştirilmesi uygulamaları gibi bankacılık alanında; sinyal işleme, imge işleme, radar, silah yönlendirme uygulamaları gibi savunma alanında, çip bozulma analizi, lineer olmayan modelleme uygulamaları gibi elektronik alanında; bunların yanında eğlence, üretim, sigortacılık, sağlık, robotik, telekomünikasyon ve güvenlik alanlarında kullanılmaktadır (Shahin ve ark. 2002).

### **3.2.1.4. Yapay Sinir Ağlarının Sınıflandırılması**

Yapay sinir ağları yapılarına ve öğrenme algoritmalarına göre sınıflandırılırlar. Nöronların bağlanma şekillerine göre yapay sinir ağları ileri beslemeli ve geri beslemeli ağlar olmak üzere ikiye ayrılır (Yaman ve Ballı 2009).

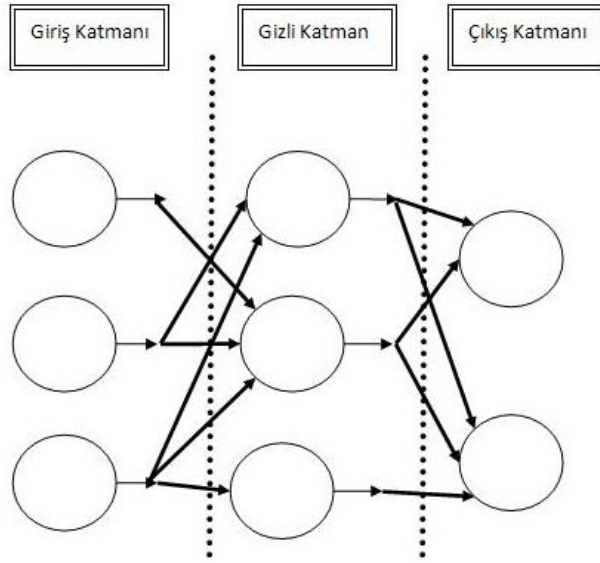
#### **- İleri Beslemeli Ağlar**

İleri beslemeli yapay sinir ağlarında katmanlardan birinin çıkışından alınan bilgi bir sonraki katmanın girişine ağırlıklar üzerinden verilir. Aynı katmandaki nöronlar arasında ya da bir önceki katman arasında bir bağlantı yoktur.

### 3. MATERYAL ve METOT

Dış dünyadan giriş yoluyla alınan bilgi hiçbir değişikliğe uğramadan gizli katmanlardaki nöronlara iletilir ve buradan alınan bilgi orta katmanlarda ve çıkış katmanında işlenerek ağ çıkışı belirlenir. En çok bilinen geriye yayılım öğrenme algoritması, bu tip YSA'ların öğrenmelerinde etkin bir şekilde kullanılır. Bu yüzden bu ağlara bazen geriye yayılım ağları da denir.

Şekil 3.13.'te ileri beslemeli çok katmanlı yapay sinir ağı yapısı gösterilmiştir.



Şekil 3.13. İleri Beslemeli ve Çok Katmanlı YSA Yapısı

#### - Geri Beslemeli Ağlar

Geri beslemeli yapay sinir ağlarında, çıkış ve ara katmanlardaki çıkışlardan ağ girişine ya da daha önceki ara katmanlara doğru geri bir besleme vardır. Geri besleme farklı katmanlardaki nöronlar arasında olduğu gibi, aynı katmandaki nöronlar arasında da olabilir. Bu tip yapay sinir ağları dinamik yapıda olduğundan, belli bir andaki çıkış hem o andaki çıkışı hem de daha önceki girişleri ifade eder. Bu yüzden, geri beslemeli ağlar önceden tahmin etme uygulamaları için kullanışlıdır. Elman, Hopfield, Jordan ve SOM ağları bu ağlara örnek verilebilir.

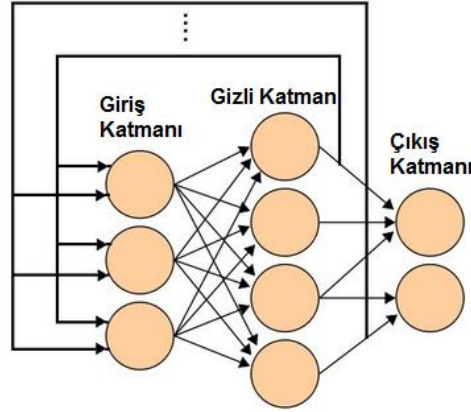
Geriye doğru hesaplamada, ağdan beklenen değer ile ağın çıkış değeri karşılaştırılır ve bu iki değer arasındaki fark alınıp hata olarak kabul edilir. Çıkış katmanında oluşan toplam hatayı hesaplamak için bütün hataların toplanması gerekir.

Hata değerleri negatif olabileceğinden, toplam hatanın sıfır olmasını engellemek için hataların kareleri ayrı ayrı alınarak toplama yapılır.

Sonraki aşamada, hesaplanan toplam hatanın karekökü alınarak çıkış katmanındaki hata miktarı belirlenir. Hatayı E ile gösterecek olursak toplam hata aşağıdaki eşitlikle ifade edilir.

$$\varepsilon = \frac{1}{2} \sum_m E_m^2 \quad (3.34)$$

Toplam hatayı minimum değere indirmek için hatanın kendisine neden olan işlem elemanlarının ağırlıklarının değiştirilmesi gerekir. Şekil 3.14.'te geri beslemeli bir ağ yapısı gösterilmiştir.



Şekil 3.14. Geri beslemeli yapay sinir ağı için örnek yapı

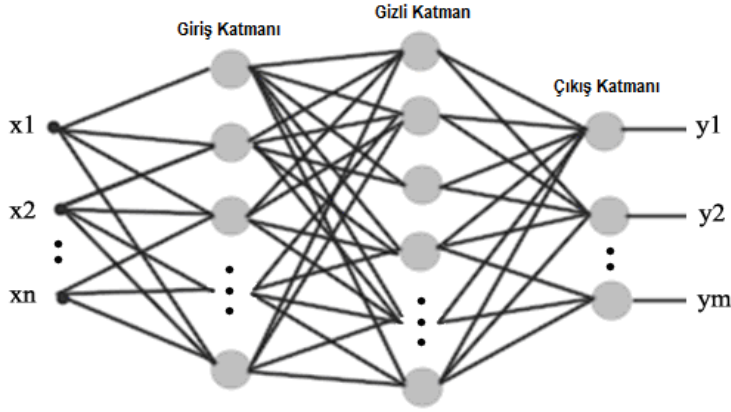
### 3.2.1.5. Çok Katmanlı Algılayıcılar (MLP)

Yapay sinir ağları, ağ mimarilerine göre tek katmanlı yapay sinir ağları ve çok katmanlı yapay sinir ağları olmak üzere ikiye ayrılır. Bu bölümde, yapay sinir ağının bir çeşidi olan çok katmanlı algılayıcılar (MLP) ve tez kapsamında kullanılan MLP'nin öğrenme algoritmasından bahsedilmiştir.

Yapılan çalışmalar sonucu çok katmanlı algılayıcı modeli Rumelhard ve arkadaşları tarafından geliştirilmiş ve XOR problemini çözmeyi başarmıştır (Anonim 2009). Temel amaç, ağın hedeflenen çıktısı ile üretilen çıktıları arasındaki hata farkını sıfıra indirmektir. Hatayı ağa yayarak gerçekleştirdiği için bu modele hata yayma modeli ya da geriye yayım modeli (backpropagation network) de denmektedir.

### 3. MATERYAL ve METOT

Yapay sinir ağlarında çok katmanlı algılayıcıya olan ilginin artmasından dolayı yeni bir dönem başlamıştır. Bu modelin yaygın kullanılmasındaki en önemli faktör, ağ eğitiminin birçok öğretim algoritması ile sağlanabilmesidir. Çok katmanlı algılayıcı ağ yapısında, nöronlar Şekil 3.15'te görüldüğü gibi farklı katmanlarda gruplanmıştır. Her bir katmanın çıkışı bağlı bulunduğu diğer katmanın giriş düğümlerine bağlıdır. Bu nedenle, giriş katmanının çıkışları aynı zamanda gizli katmanında girişlerini oluştururken, son katmanın çıkışları da ağın çıkışlarını oluşturur (Orhan ve ark.2010).



Şekil 3.15. Çok Katmanlı Algılayıcı için Örnek Ağ Mimarisi

MLP birçok tespit ve tahmin işlemlerini yerine getiren parametrik olmayan bir yapay sinir ağı tekniğidir. MLP' de gizli katmandaki her bir  $j$  nöronu,  $w_{ji}$  bağlantı ağırlığıyla giriş işaretlerinin çarpımlarının toplamını alır ve  $y_j$  çıkışını bu toplamın bir fonksiyonu olarak hesaplar (Acar ve ark.2011):

$$y_j = f\left(\sum w_{ji}x_i\right) \quad (3.35)$$

Burada  $f$  bir nörona etki eden işaretlerin ağırlıklı toplamını çıkış değerine dönüştüren bir aktivasyon fonksiyonudur. Aktivasyon fonksiyonu basit bir eşik fonksiyonu olup, sigmoidal veya hiperbolik tanjant fonksiyonu olabilir. Bu çalışmada, aktivasyon fonksiyonu olarak logsig fonksiyonu kullanılmıştır. YSA uygulamalarında en popüler öğrenme algoritması olan geriye yayılma (BP) öğrenme algoritması, MLP'nin eğitilmesi sürecinde kullanılmıştır. BP algoritması, ağ parametrelerinin ayarlanması ile kuadratik maliyet fonksiyonunun (cost function) mimimizasyonuna dayanmaktadır.

Ortalama karesel hata (MSE) ağın performansı olduğundan, eğitime işlemi ile maliyet fonksiyonunu minimize eden parametreler belirlenir. Ortalama karesel hata

(MSE) işlemleri, eşitlik (3.36) ve (3.37) 'de gösterilmiştir.

$$e_j(n) = d_j(n) - y_j(n) \quad (3.36)$$

$$\varepsilon(n) = \frac{1}{2P} \sum_{j \in C} \sum_{n=1}^P e_j^2(n) \quad (3.37)$$

Denklemlerde yer alan  $e$ ,  $n$ ,  $d$ ,  $y$ ,  $C$  ve  $P$  sırasıyla çıkıştaki hata işaretini, iterasyon sayısını, istenilen çıkışı, ağ tarafından üretilen çıkışı, çıkıştaki hata sinyalini ve eğitim setinde kullanılan toplam örüntü sayısını ifade etmektedir.

Sinaptik ağırlıkların güncellenmesinde (3.38)

$$\Delta w_{ji}(n) = \alpha \Delta w_{ji}(n-1) + \eta \delta_j(n) y_i(n) \quad (3.38)$$

Gizli ve çıkış katmanları arasındaki ağırlıklar için (3.39)

$$\delta_j(n) = e_j(n) \varphi' \left( \sum_{i=0}^m w_{ji}(n) y_i(n) \right) \quad (3.39)$$

Giriş ve gizli katmanlar arasındaki ağırlıklar için (3.40)

$$\delta_j(n) = \varphi' \left( \sum_{i=0}^m w_{ji}(n) y_i(n) \right) \sum_k \delta_k(n) w_{kj}(n) \quad (3.40)$$

bağıntıları kullanılır. Bağıntılarda yer alan  $\varphi$ ,  $j$ ,  $k$ ,  $\eta$  ve  $\alpha$  sırasıyla transfer fonksiyonunu, gizli katmandaki hücre sayısını, çıkış katmanındaki hücre sayısını, öğrenme oranını ve momentum katsayısını ifade etmektedir.

### 3.2.2. Destek Vektör Makineleri (SVM)

Destek Vektör Makineleri (SVM), istatistik ve bilgisayar bilimleri içinde bir dizi denetimli öğrenme yöntemleriyle bağlantılı bir kavram olup, örüntü tanıma, veri analizi, sınıflandırma ve regresyon analizi için kullanılmaktadır. SVM'nin sahip olduğu matematiksel modeller ilk olarak iki sınıflı doğrusal verilerin sınıflandırılması için ortaya çıkmış, zamanla doğrusal olmayan çok sınıflı verilerin sınıflandırılması için geliştirildi. SVM iki sınıfı birbirinden ayıran en uygun karar fonksiyonunu tahmin etmeye çalışır, başka bir ifadeyle iki sınıfı birbirinden en uygun şekilde ayıran hiperdüzlemi oluşturur (Huang ve ark.2002).

SVM sınıflandırıcılarında verilerin doğrusal yapıda olması ya da doğrusal olarak ayrılamayan bir yapıda olması şeklinde iki durumla karşılaşılabilir. Doğrusal olarak

ayrılabilen veriler sınıflandırılırken maksimum sınır (margin) kolay hesaplanabilir. Ancak günlük yaşantımızda karşılaştığımız çoğu problem doğrusal yapıda olmayabilir.

Bu nedenle doğrusal olmayan veriler üzerinden sınıflandırma yapılırken bu verilerin öncelikle doğrusal olarak ayrılacakları bir uzaya aktarılıp, bu uzay kümesinde sınıflandırmalarının yapılması gerekir (Eray 2008).

SVM sınıflandırıcıları, doğrusal SVM sınıflandırıcıları ve doğrusal olmayan SVM sınıflandırıcıları olmak üzere iki kısımda incelenecektir.

#### 3.2.2.1. Doğrusal Destek Vektör Makinesi Sınıflandırıcıları

Doğrusal SVM sınıflandırıcıları verilerin doğrusal olarak ayrılabilmesi ve ayrılamaması işlemini yapan SVM sınıflandırıcıları olarak iki grupta incelenecektir.

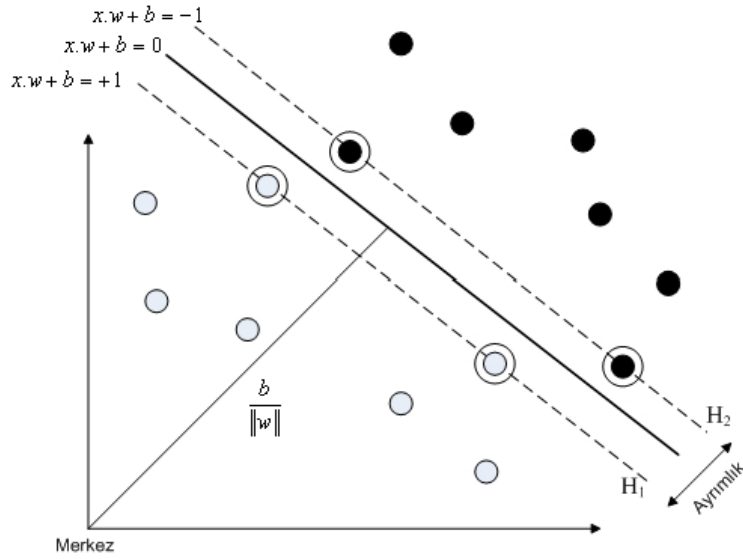
##### - Doğrusal Olarak Ayrılabilme Durumu

M elemandan oluşan eğitim veri kümesinin  $X = \{x_i, y_i\}, i = 1, 2, \dots, m$  olduğu kabul edilsin. Burada  $y_i \in \{-1, 1\}$  etiket değerlerini  $x_i \in \mathcal{R}^d$  öznitelik vektörünü gösterir. Bu iki değer, doğrusal olarak ayrılabilme durumunda hiperdüzlem ile ayrılacaktır. Bu hiperdüzlem aynı zamanda *Ayırıcı hiperdüzlem* olarak da adlandırılır. SVM'nin amacı bu hiperdüzlemlerle her iki örnek grubuna eşit mesafe sağlamaktır. Doğrusal olarak ayrılabilme durumu için ayırıcı hiperdüzlem şu şekilde tanımlanır.

$$f(x) = w^T x + b = \sum_{i=1}^m w_i x_i + b \quad (3.41)$$

Burada  $w$  ağırlık vektörünü,  $b$  değeri de bir sapma değerini ifade etmektedir ve hiperdüzlemin pozisyonu bu parametrelerin alacağı değerlerle belirlenir. Doğrusal ayrılabilme durumu için hiperdüzlem Şekil 3.16.'da gösterilmiştir.





Şekil 3.16. Doğrusal olarak ayrılabilen durum için optimal hiperdüzlem

Hiperdüzlem üzerindeki herhangi bir  $x$  noktası

$$wx + b = 0 \quad (3.42)$$

Eşitlik (3.42) koşulunu sağlar. SVM yönteminde örnek veriler doğrusal olarak ayrılabilirse  $y_i = +1$  ve  $y_i = -1$  şeklinde tanımlı örneklere en uygun ayırıcı hiperdüzlem bulunur. Eğitim veri kümelerinin aşağıda tanımlı denklemleri sağladığı kabul edilirse,

$$wx_i + b \geq +1, \quad y_i = +1 \text{ için}, \quad (3.43)$$

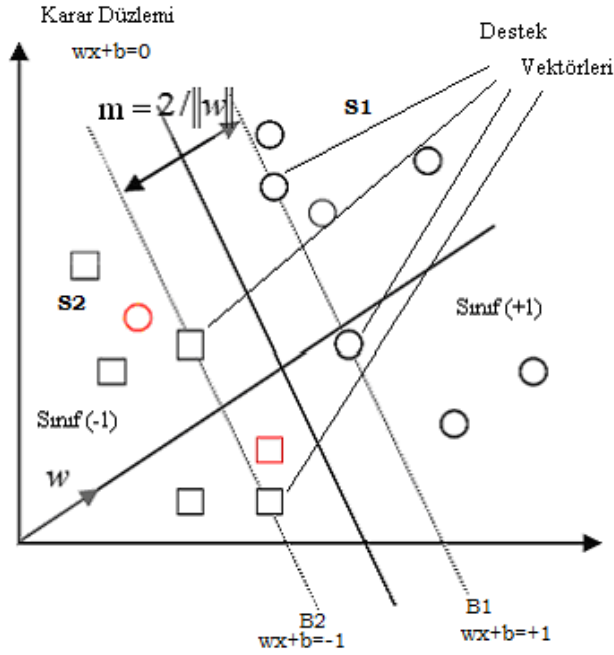
$$wx_i + b \leq -1, \quad y_i = -1 \text{ için}, \quad (3.44)$$

Bu durumda  $i = 1, 2, \dots, m$  için,

$$(w^T x_i - b)y_i \geq +1 \quad (3.45)$$

koşulu elde edilir. (3.45) eşitsizliğini sağlayan hiperdüzlemin her iki tarafında bulunan en yakın örneklerin hiperdüzleme olan dik uzaklıkları toplamı margin (aralık) olarak tanımlanır.

Şekil 3.17.'de görüldüğü gibi margin değerini en büyük yapan hiperdüzlem, en uygun ayırıcı hiperdüzlemdir.



Şekil 3.17. Doğrusal ayrılabilir durumda en uygun hiperdüzlem

Şekil 3.12. 'de B1 hiperdüzlemini oluşturan ve S1 sınıfını ayıran koşul, Eşitsizlik (3.43); B2 hiperdüzlemini oluşturan ve S2 sınıfını ayıran koşul, Eşitsizlik (3.44) kabul edilirse; B1 ve B2 hiperdüzlemlerinin orijine olan dik uzaklıkları sırasıyla  $|1-b|/\|w\|$  ve  $|-1-b|/\|w\|$  olur. Her iki hiperdüzlemin en uygun hiperdüzleme olan uzaklıkları da  $1/\|w\|$  olur. Bu durumda B1 ve B2 hiperdüzlemlerinde bulunan iki örnek kümesinin birbirlerine olan uzaklığı  $2/\|w\|$  olur. B1 ve B2 hiperdüzlemleri arasındaki maksimum uzaklık ağırlık değeri  $w$ 'un Euclid normunun yani ( $\|w\|$  değerinin) minimize edilmesiyle bulunur. Bu optimizasyon problemi, konveks karesel programlama problemi olduğundan Lagrange optimizasyon tekniği bu problemin çözümü için etkili bir yöntemdir. SVM'deki amaç bu iki hiperdüzlem arasındaki margin'in (uzaklığın) maksimum olmasını sağlamaktır. Bu nedenle  $\|w\|$  değeri minimum olmalıdır. Aralarındaki mesafe (margin) değeri en büyük olan iki hiperdüzlem üzerindeki noktalar *Destek Vektörleri* olarak adlandırılır.

En büyük mesafe (margin)'in bulunması için,  $\|w\|^2 = \sqrt{w^T w}$  ifadesi eşitlik (3.46)'da yerine konulursa;

$$f = \|w\|^2 / 2 = w^T w / 2 \quad (3.46)$$

$$g_i = -((w^T x_i + b)y_i) - 1 \leq 0 \quad i = 1, 2, \dots, m \quad (3.47)$$

Sırasıyla Minimizasyon (3.46) ve Kısıtlama (3.47) denklemleri elde edilir. Burada Eşitlik (3.46) çözülecek olan problemi, Eşitlik (3.47) ise bu problemin çözümünde kullanılacak olan koşulu belirtir. Bu eşitlikler ikinci dereceden bir optimizasyon problemi olduğundan problemin çözümü için standart karesel programlama yöntemlerine başvurulur. Doğrusal olarak ayrılmayan durumda problemin çözümünü kolaylaştırmak için Lagrange yöntemi kullanılarak, problemin Lagrange fonksiyonu bulunur. Bu fonksiyon,

$$L_{primal}(w, b, \delta) = \sum_{i=1}^m \delta_i g_i(x) + f(x) \quad (3.48)$$

$$L_{primal}(w, b, \delta) = -\sum_{i=1}^m \delta_i ((w^T x_i + b)y_i - 1) + \frac{w^T w}{2} \quad (3.49)$$

Bu ifadelerde kullanılan  $\delta = (\delta_1, \delta_2, \dots, \delta_m)$  çarpanların her biri eğitim kümesini oluşturur ve bu çarpanlara *Lagrange Çarpanları* denir.

Eşitlik (3.49)'da ifade edilen Lagrange fonksiyonunu çözmek için optimizasyon problemi, Karush-Kuhn-Tucker (KKT) koşulları kullanılarak dual hale dönüştürülür. Elde edilen dual problem denklemleri şöyledir:

$$L_{dual} = \sum_{i=1}^m \delta_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \delta_i \delta_j y_i y_j x_i^T x_j \quad (3.50)$$

$$\sum_{i=1}^m y_i \delta_i = 0, \quad \delta_i \geq 0, \quad i = 1, 2, \dots, m \quad (3.51)$$

Burada, Eşitlik (3.50) ve (3.51) problemin çözümü için sırasıyla maksimizasyon ve kısıtlama koşullarını belirtir.

Elde edilen dual problemin verilen koşullar altında çözümü ikinci dereceden optimizasyon problemi ile sağlanır. Eşitlik (3.50) karesel programlama ile çözümlenerek  $\delta$  değerleri bulunur, w KKT koşullarından Gradyen koşulu eşitliği (3.52) ve b ise KKT koşullarından Orthogonality koşulu Eşitlik (3.53) ile hesaplanır.

### 3. MATERYAL ve METOT

---

Bu koşullar sırasıyla:

$$\frac{\partial L_{primal}}{\partial w} = w^T - \sum_{i=0}^m \delta_i x_i y_i = 0, \quad \frac{\partial L_{primal}}{\partial w} = \left( \frac{\partial L_{primal}}{\partial w_1}, \frac{\partial L_{primal}}{\partial w_2}, \dots, \frac{\partial L_{primal}}{\partial w_m} \right) \quad (3.52)$$

$$\delta_i g_i = -((w^T x + b)y_i - 1)\delta_i = 0, \quad i = 1, 2, \dots, m \quad (3.53)$$

Problem çözümü sonucunda elde edilen Lagrange çarpanlarının büyük bir kısmının değeri sıfır olur. Burada eğitim örnekleri  $x_i$  'ye karşılık gelen Lagrange çarpanı  $\delta_i \geq 0$  olduğu zaman, bu örnekler *Destek Vektörleri* olarak adlandırılır ve bu noktalarda kısıtlamalar aktif haldedir. Destek vektörleri B1 ve B2 hiperdüzlemleri üzerinde bulunur.  $x_i$  eğitim örneklerine karşılık gelen Lagrange çarpanı  $\delta_i = 0$  olduğunda ise, bu örnekler destek vektörü değildir ve kısıtlamalar pasif haldedir. Bu örnekler B1 ve B2 hiperdüzlemleri üzerinde yer almayıp, bu düzlemlerin arka tarafında kalan örneklerden oluşur.

#### - Doğrusal Olarak Ayrılamama Durumu

Eğitim örneklerinin doğrusal olarak ayrılamadığı durumlarda doğrusal olmayan sınıflandırıcıların kullanılması, SVM'nin kötü bir genelleme yapmasına yol açar. Bunu engellemek için, Eşitlik (3.54) ve (3.55) biraz daha genişletilip belli bir hata payına izin verilerek düzenlenmelidir. Hata terimleri  $\psi$  ve  $i = 1, 2, \dots, m$  olarak gösterilirse;

$$w^T x_i + b \geq 1 - \psi_i, \quad y_i = +1, \quad \psi_i \geq 0 \quad (3.54)$$

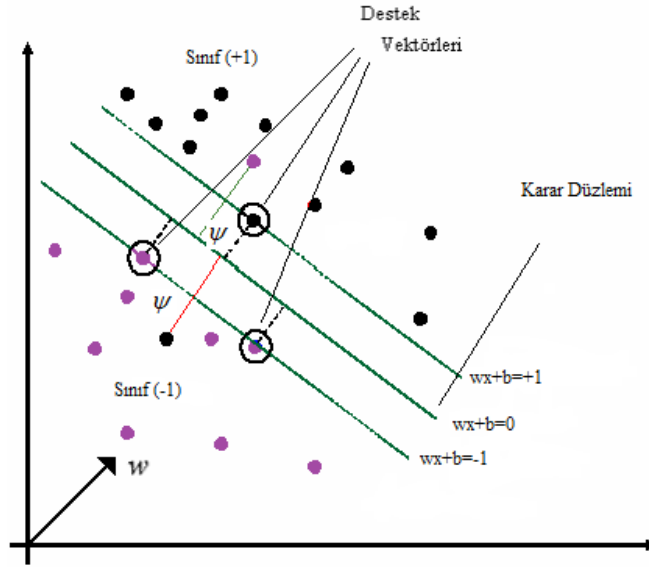
$$w^T x_i + b \leq -1 + \psi_i, \quad y_i = -1, \quad \psi_i \geq 0 \quad (3.55)$$

koşulları elde edilir. Burada  $\psi$  ve  $i = 1, 2, \dots, m$  pozitif zayıflık değişkenleri olarak adlandırılır. Bu denklemlere göre problem yeniden düzenlenecek olursa;

$$f(w, \psi) = S \sum_{i=1}^m \psi_i + \frac{1}{2} \|w\|^2, \quad i = 1, 2, \dots, m, \quad \psi_i \geq 0 \quad (3.56)$$

$$g_i = (w^T x_i + b)y_i \geq 1 - \psi_i, \quad i = 1, 2, \dots, m, \quad \psi_i \geq 0 \quad (3.57)$$

Minimizasyon (3.56) ve Kısıtlama (3.57) eşitlikleri tanımlanmış olur. Doğrusal ayrılmama durumunda en uygun ayırıcı hiperdüzlem Şekil 3.18.'de gösterilmiştir.



Şekil 3.18. Doğrusal ayrılmama durumunda en uygun hiperdüzlem

Doğrusal olarak verilerin ayrılamaması durumunda sistemin her koşulda bir çözüm üretmesini engellemek için, yani sistemin ezberlemesini önlemek için sistemde bir  $S$  üst sınırı tanımlanır. Bu parametre, mesafe (margin) için düzenleyici bir faktör ve hatalar için de önleyici bir faktör görevini alır. Belirlenen bu üst sınır Lagrange çarpanlarının alabileceği maksimum değeri ifade edip bu çarpanların  $0 \leq \delta_i \leq S$  değer aralığında kalmasına imkan verir. Problemin çözümü için Lagrange fonksiyonu yeniden düzenlenerek aşağıdaki gibi ifade edilebilir;

$$L_{primal} = S \sum_{i=1}^m \psi_i + \frac{1}{2} \|w\|^2 - \sum_{i=1}^m \delta_i ((x_i w + b) y_i - 1 + \psi_i) - \sum_{i=1}^m \delta_i \psi_i \quad (3.58)$$

Eşitlik (3.58)'de,  $\delta_i$  Lagrange parametrelerini ifade ederek,  $\psi_i$  hata teriminin pozitif olmasını sağlar. Bu eşitlikte ifade edilen Lagrange fonksiyonunu çözmek için optimizasyon problemi, Karush-Kuhn-Tucker (KKT) durumları kullanılarak dual hale dönüştürülür. Elde edilen dual problem denklemleri şöyledir:

$$L_{dual} = \sum_{i=1}^m \delta_i - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m \delta_i \delta_j y_i y_j x_i^T x_j \quad (3.59)$$

$$\sum_{i=1}^m y_i \delta_i = 0, \quad 0 \leq \delta_i \leq S, \quad i = 1, 2, \dots, m \quad (3.60)$$

### 3. MATERYAL ve METOT

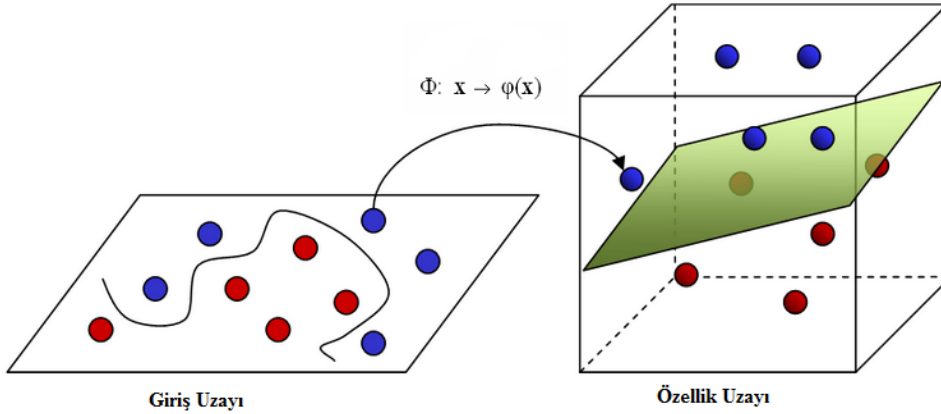
Burada, Eşitlik (3.59) ve (3.60) dual problemin çözümü için sırasıyla maksimizasyon ve kısıtlama koşullarını belirtir.  $0 \leq \delta_i \leq S$  aralığındaki Lagrange çarpanlarına karşılık gelen  $x_i$  değerleri *Destek Vektörleri* olarak adlandırılır.

#### 3.2.2.2. Doğrusal Olmayan Destek Vektör Makinesi Sınıflandırıcıları

Reel dünyada problemlerin birçoğu farklı bileşenlerden oluştuğundan, bu problemlerin çözümünde doğrusal olmayan karar destek makineleri kullanılabilir. Bu tür problemlerin çözümü için örneklerin taşıma fonksiyonları aracılığıyla, doğrusal olarak ayrılabilirleri daha büyük boyuttaki bir özellik uzayına taşınması gerekir. Başka bir deyişle;  $\phi$  taşıma fonksiyonu nu,  $\mathcal{R}^d$  giriş uzayında doğrusal olarak ayrılamayan örnekleri ve  $K$  çıkış uzayında doğrusal olarak ayrılması beklenen örnekleri temsil ederse;

$$\phi : \mathcal{R}^d \rightarrow K \quad (3.61)$$

olur. Şekil 3.19.'da giriş uzayındaki verilerin doğrusal olarak ayrılabilirliği daha yüksek boyutlu bir özellik uzayına taşınması görünmektedir.



Şekil 3.19. Giriş uzayındaki örneklerin özellik uzayına taşınması

Şekil 3.19.'da görüldüğü gibi, giriş uzayında doğrusal olarak ayrılamayan veriler çekirdek fonksiyonları kullanılarak bir üst boyuttaki özellik uzayına taşınır. Bu özellik uzayında, verileri doğrusal olarak ayrılabilir en uygun hiperdüzleme rastlamak mümkündür.

SVM'lerin eğitim aşaması, K özellik uzayındaki verilerin noktasal çarpımlarına (içsel çarpım) bağlı olup, bu nokta çarpımları Eşitlik (3.62)'de gösterilmiştir.

$$P(x_i, x_j) = \phi(x_i) \cdot \phi(x_j) \quad (3.62)$$

Bu eşitlikte, P çekirdek (kernel) fonksiyonunu belirtir. Çekirdek fonksiyonu olarak çeşitli fonksiyonlar kullanılabilir. Bu fonksiyonlardan bazıları şunlardır:

1. Sigmoid Çekirdek Fonksiyonu:

$$P(x_i, x_j) = \tanh(\kappa x_i^T x_j - \delta) \quad (3.63)$$

2. Polinom Çekirdek Fonksiyonu:

$$P(x_i, x_j) = (x_i^T x_j + 1)^p \quad (3.64)$$

3. Doğrusal Çekirdek Fonksiyonu:

$$P(x_i, x_j) = (x_i^T x_j) \quad (3.65)$$

4. Radyal Tabanlı Çekirdek Fonksiyonu (RBF):

$$P(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2\sigma^2}\right) \quad (3.66)$$

Çekirdek (kernel) fonksiyonlarına bağlı olarak özellik uzayında karar fonksiyonunun ifadesi, Eşitlik (3.67)'deki  $g(x)$  fonksiyonun işareti ile Eşitlik (3.68)'de gösterilmiştir.

$$g(x) = \sum_{i=1}^{dv} \delta_i y_i \phi(x_i) \phi(x_j) + b = \sum_{i=1}^{dv} \delta_i y_i P(x_i, x_j) + b \quad (3.67)$$

$$\text{sign}(g(x)) = \text{sign}\left(\sum_{i=1}^{dv} \delta_i y_i \phi(x_i) \phi(x_j) + b\right) = \text{sign}\left(\sum_{i=1}^{dv} \delta_i y_i P(x_i, x_j) + b\right) \quad (3.68)$$

Bu fonksiyonda  $dv$ , destek vektörlerinin sayısını,  $x_{ij}$  de destek vektörlerini ifade eder.

#### 3.2.2.3. En Küçük Kareler Destek Vektör Makinesi (LS-SVM) Sınıflandırıcıları

SVM, Suykens ve diğerleri (2002) tarafından önerilmiştir. Bu SVM çeşidi, standart SVM sınıflandırıcılarında olduğu gibi iki sınıf etiketli sınıflandırma işlemi düşünülerek oluşturulmuştur. SVM ikinci dereceden denklemlerin programlanması problemini ifade ederken, LS-SVM sınıflandırıcıları doğrusal bir takım denklemlerin çözülmesiyle elde edilir. Standart SVM sınıflandırıcılardaki optimizasyon işleminin çözümü amacıyla kullanılan ikinci dereceden programlama metotları bu teknikte kullanılmaz. Onun yerine doğrusal eşitlik kümesi (3.69) kullanılır. Bu eşitlikte;  $\|w\|$  Euclid formunu,  $n$  eğitim veri seti sayısını,  $\gamma$  regülarizasyon parametresini ve  $e_i$  hata miktarını ifade eder.

$$\frac{1}{2}\|w\|^2 + \frac{\gamma}{2} \sum_{i=1}^n e_i^2 \quad (3.69)$$

Bu eşitlik (3.74) koşuluna bağlı olarak minimize edilir. Eşitlik (3.70)'de yer alan  $y_i$  çıkış değerini,  $x_i$  giriş değerini,  $w$  ağırlık vektörünü,  $\varphi(\cdot)$  daha büyük boyuttaki özellik uzayına taşıma fonksiyonunu ve  $b$  bias değerini ifade eder.

$$y_i = w^T \varphi(x_i) + b + e_i, \quad i = 1, 2, \dots, n \quad (3.70)$$

Optimizasyon probleminin çözümünden sonra standart SVM sınıflandırıcılarda olduğu gibi optimizasyon eşitlikleri Eşitlik (3.71)'deki ikili problem yapısına dönüştürülür. Bu eşitlikte  $\alpha$  Lagrange çarpanlarını ifade eder.

$$L(w, b, e, \alpha) = \frac{1}{2}\|w\|^2 + \frac{\gamma}{2} \sum_{i=1}^n e_i^2 - \sum_{i=1}^n \alpha_i \{w^T \varphi(x_i) + b + e_i - y_i\} \quad (3.71)$$

Eşitlikte yer alan  $\alpha$  çarpanları standart SVM sınıflandırıcılarında pozitif olması gerekirken, LS-SVM sınıflandırıcılarında pozitif veya negatif olabilmektedir. LS-SVM'deki bu farklılık ve daha az eğitime zamanı gerektirmesi, LS-SVM'yi standart SVM sınıflandırıcılarından ayırır. Sınıflandırma performansı ise, bazı uygulamalarda standart SVM öğrenme metoduna yakın da olsa düşük çıkabilmektedir. Bunun sebebi de eğitime aşamasında daha kuvvetli ama zaman alan ikinci dereceden programlama yerine daha basit ama kısa sürede uygulanabilen doğrusal programlamayı kullanmasındandır (Li ve ark. 2011).



Eşitlik (3.71)'in optimizasyonundan sonra ve  $w, e$  parametrelerinin elenmesiyle; LS-SVM modeli için fonksiyon tahmini Eşitlik (3.72) ile ifade edilir.

$$y(x) = \sum_{i=1}^n \alpha_i K(x_i, x_j) + b \quad (3.72)$$

Burada  $K(\cdot)$  fonksiyonu kerneli ifade ederek,  $x$  giriş uzayının lineer olmayan bir şekilde çok boyutlu özellik uzayına haritalanmasını sağlar. LS-SVM; fonksiyonu Eşitlik (3.76)'daki ifadeyi kullanarak benzetmeye çalışır ve Radyal temel fonksiyonunu (RBF) kernel fonksiyonu olarak kullanır. RBF kernel, Eşitlik (3.73)'de gösterilmiştir.

$$K(x, x_i) = \exp(-|x - x_i|^2 / \sigma^2) \quad (3.73)$$

Burada  $\sigma$  pozitif reel bir sabittir. RBF kernel varlığında, Eşitlik (3.69)'daki  $\gamma$  parametresi ve Eşitlik (3.73)'deki  $\sigma$  parametresi olmak üzere iki tane ayarlamalı parametreye ihtiyaç vardır.

Bu çalışmada, LS-SVM probleminin eğitiminde iki tane bağımsız parametrenin olduğunu görebiliriz. Bu parametreler sırasıyla kernel genişlik parametresi sigma ( $\sigma$ ) ve regülarizasyon parametresi ( $\gamma$ ) olup LS-SVM sınıflandırıcısının genelleme performansını etkilemektedir. Bu nedenle, bu parametreler genelleme hatasını en aza indirecek şekilde ayarlanmalıdır (Sheng ve ark. 2011).

#### 3.2.2.4. Bire-Karşı-Biri Metodu

Bu metotta, çoklu sınıflandırma yapılırken  $r$  sınıf varsa  $r(r-1)/2$  tane sınıflandırıcı yapılandırılır ve aşağıdaki ikili sınıflandırma problemi çözülür (Pöyhönen ve ark.2003).

$$\min \frac{1}{2} (w^{ij})^T w^{ij} + S \sum_i^m \delta_m^{ij} (w^{ij})^T \quad (3.74)$$

$$(w^{ij})^T \phi(x_n) + b^{ij} \geq 1 - \psi_m^{ij} \quad \Leftarrow \quad y_m = i \quad (3.75)$$

$$(w^{ij})^T \phi(x_n) + b^{ij} \leq -1 - \psi_m^{ij} \quad \Leftarrow \quad y_m = j \quad (3.76)$$

#### 3.1.2.5. Bire-Karşı-Diğerleri Metodu

Bu yöntemde r sınıflı bir problem için r tane ikili sınıflandırıcı oluşturulur ve i.sınıftaki veri pozitif etiketli, geriye kalanlar ise negatif etiketli alınarak, i.SVM aşağıdaki denklemi çözer (Schwenker 2000; Yıldırım 2006).

$$\min \frac{1}{2} (w^i)^T w^i + S \sum_{j=1}^m \delta_j^i \quad (3.77)$$

$$(w^i)^T \phi(x_j) + b_i \geq 1 - \psi_j^i \quad \Leftarrow \quad y_j = i \quad (3.78)$$

$$(w^i)^T \phi(x_j) + b_i \leq -1 + \psi_m^i \quad \Leftarrow \quad y_j \neq i \quad (3.79)$$

#### 3.2.3. K En Yakın Komşu (k-NN) Algoritması

K en yakın komşu yöntemi öğrenmeye dayalı bir yöntem olmayıp, sınıflandırma problemlerini çözmeye yarayan bir modeldir. Bu yöntemde, sınıflandırma yapılacak verilerin örnek kümesindeki normal veri kümelerine yakınlıkları hesaplanarak; en yakın olduğu düşünülen n tane veride bulunan en fazla sınıfa göre sınıflandırma yapılır. Burada esas olan sınıflandırma yapılmadan önce, her bir sınıfın özelliklerinin önceden net bir şekilde belirtilmiş olmasıdır (Kaymaz 2007).

Sınıflandırma alanında iyi bilinen k-NN yaklaşımının birçok uygulamada başarılı olduğu bilinmektedir. İlk olarak, bir test seti elemanından her bir örnek seti elemanına olan mesafe ölçülür. Önceden belirtilmeyen bir öznitelik vektörü kendisine en yakın k komşularının çoğunluğunun sahip olduğu etiket ile etiketlenir. Bu yöntemde öznitelik vektörü istatistiksel dağılımdan bağımsızdır.

Bunun yanı sıra, sistemin performansını benzerlik ölçümü, örnek kümesindeki davranışların yeterli sayıda olması ve eşik değeri gibi faktörler etkilemektedir. Ancak en önemli kontrol parametresi komşu sayısı k'dır. Verilen örüntü en yakın k komşusuna olan uzaklıklara bakılarak sınıflandırılır (Buttrey ve Karo 2001).

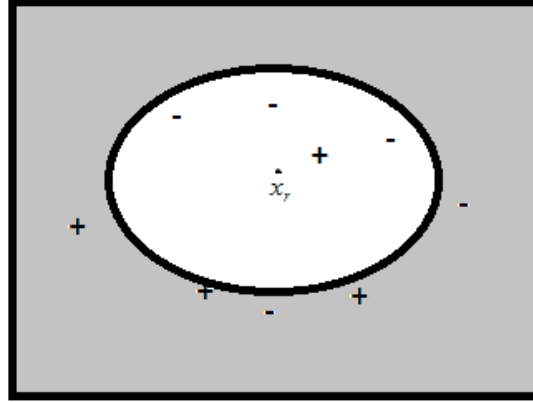
Bu algoritmada kullanılan farklı parametrelerden biri de uzaklık ölçme yöntemidir. Yapılan çalışmada, verilen noktaya göre en yakın komşular Öklid (Euclidean) uzaklığı kullanılarak hesaplanmıştır (Hilavin ve ark.2011).

İki örnek nokta arasındaki öklid mesafesi Eşitlik (3.80)'de gösterilmiştir.

$$d(x_r, x_s) = \left[ \sum_{i=1}^p c_i (x_{ri} - x_{si})^2 \right]^{1/2} \quad (3.80)$$

Sıradan bir öklid mesafesinde ağırlıklar 1 eşittir ( $c_i = 1, i = 1, 2, \dots, p$ ).

En yakın komşu algoritmasında öğrenilen fonksiyon ayrık ve reel değerli olabilir. Ayrık değerli fonksiyonlarda  $x_r$  sınıflandırılmak istenen örnek ve  $x_s$  değeri öğrenme noktaları olarak adlandırılınsın. Bu durumda eğer  $k=1$  seçilirse “1-NN” (1 En yakın komşu algoritması), örnek  $x_r$  için  $f(x_s)$  hesaplanır ve  $f(x_r)$ 'ye değer ataması yapılır. Şekil 3.20.'de örnek veriler boolean değerlerle gösterilmiştir ve eğitim örnekleri “+” ve “-“ şeklinde ifade edilmiştir. Burada  $x_r$  örneği 1-NN'e göre pozitif, 5-NN'e göre negatif olarak sınıflandırılır.



Şekil.3.20. Örnek veri kümesi ve sınıflandırma noktası



#### 4. Bulgular ve Tartışma

Bu bölümde bitki yapraklarında görülen pas hastalıklarının tespiti için yapılan uygulamanın detayları anlatılmıştır. Uygulama olarak zambak çiçeğinin yaprağı temel alınmış ve üç aşamada konunun detayı açıklanmıştır. Birinci aşamada tezde kullanılan imgeler ve önışlemler, ikinci aşamada tespit sisteminin yapılandırılması ve üçüncü aşamada ise uygulama sonuçları verilmiştir.

##### 4.1. Tez Kapsamında Kullanılan İmgeler ve Önışlemler

Yapılan çalışmada zambak çiçeğine ait sağlıklı ve hastalıklı imgelerden oluşan iki ayrı veri kümesi sayısal olarak işlenip GWT ve GLCM metotları ile öznitelik vektörleri oluşturulduktan sonra farklı (k-NN, MLP, LS-SVM) sınıflandırma teknikleri kullanılarak pas hastalığının tespiti yapılmıştır. Kayda alınan toplam 53 imge farklı zirai sitelerinin veri tabanlarından elde edilmiş (Anonim 2008; Bergeron 2009), DÜ Ziraat Fakültesinden çeşitli öğretim üyelerinden destek alınarak imgelerde yer alan ürünlerin hastalık ve sağlık durumu (1-Sağlıklı veya normal, 2- Hastalıklı) değerlendirilmiştir. Bu imgelerin 32 tanesi sağlıklı, 21 tanesi hastalıklı imgelerden oluşmaktadır. İki gruba ilişkin hastalıklı örnek imgeler Şekil 4.1.'de, sağlıklı örnek imgeler ise Şekil 4.2.'de gösterilmiştir.

Bu çalışmada kullanılan imgeler 3 boyutlu JPEG formatında olduğundan, imgeler MATLAB programı yardımıyla 2 boyutlu gri tonlu imge formatına dönüştürülmüştür. Uygulamada kullanılan imgelerin ölçekleri 320x480 pikseldir. İmgeler 'bicubic' interpolasyon yöntemi ile ölçeklendirilerek 16x16 piksellik imgelere dönüştürülmüş ve tüm imgelerde histogram eşitlemesi yapılmıştır.

#### 4. BULGULAR VE TARTIŞMA

---



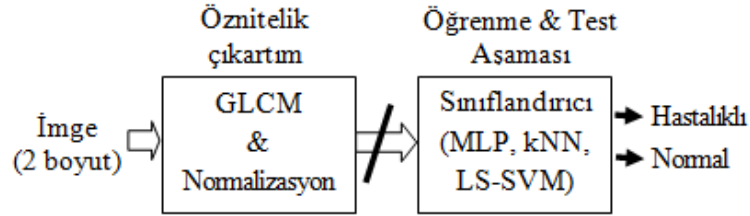
Şekil 4.1. Zambak çiçeğine ait hastalıklı örnek yaprak imgeleri



Şekil 4.2. Zambak çiçeğine ait sağlıklı örnek yaprak imgeleri

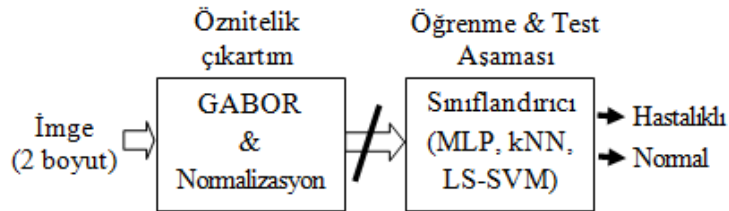
## 4.2. Tespit Sisteminin Yapılandırılması

İmgelerden elde edilen GLCM ve GWT matrislerinden ayrı ayrı istatistiksel değerler hesaplanarak her bir öznelik çıkarma yönteminden farklı vektörler elde edilmiştir. GLCM matrisinden elde edilen istatistiksel parametreler sınıflandırıcı girişine uygulanmış olup, oluşturulan sistem model yapısı Şekil 4.3.'te gösterilmiştir. Bu parametreler sırasıyla kontrast, korelasyon, enerji, homojenlik, entropi, standart sapma ve ortalama değerleridir. Ürün imgelerini sınıflandırmak için oluşturulan sistemin giriş parametreleri imgelere ilişkin öznelik vektörü olup sistemin girişine uygulanan parametreler, sınıflandırıcı türüne göre normalizasyona tabi tutularak, [0,1] değer aralığına dönüştürülmüştür. Sonraki aşamada MLP, k-NN ve LS-SVM sınıflandırıcıları eğitilip test edilerek pas hastalığının tespiti yapılmıştır.



Şekil 4.3. GLCM tabanlı sınıflandırma sisteminin yapısı

GLCM tabanlı sınıflandırma sistemine benzer bir yapı GWT tabanlı sınıflandırma sistemi oluşturulmuştur (Şekil 4.4). Gabor öznelik vektörleri, imgenin 8 dalgacık (2 ölçek ve 4 yönelim) matrislerinin her birine ait istatistiksel değişkenlerinin bir vektör oluşturacak şekilde art arda eklenmesiyle oluşturulmuştur. Her bir dalgacık dönüşümü matrisinden sırasıyla standart sapma, ortalama ve entropi değerleri hesaplanmıştır. Sonuç olarak her bir imgeye ait 8 dalgacık dönüşümü matrisinin istatistiksel değerleri art arda eklenerek toplamda  $8 \times 3 = 24$  uzunluğunda öznelik vektörü oluşturulmuştur. Elde edilen öznelik vektörleri sınıflandırıcı için giriş örüntüleri olarak kullanılmıştır.



Şekil 4.4. GWT tabanlı sınıflandırma sisteminin yapısı

## 4. BULGULAR VE TARTIŞMA

Her iki sınıfa ilişkin örnek örüntüler Şekil 4.1 ve Şekil 4.2’de gösterilmiştir. Toplam 53 örüntüden, 35 adet örüntü (21 adet normal, 14 adet hastalıklı) eğitim seti için kullanılmıştır. Kalan 18 adet örüntü ise (11 adet normal, 7 adet hastalıklı) test seti olarak kullanılmıştır. Eğitim ve test setlerinin örüntüleri rastgele seçilerek belirlenmiştir. Ürün imgelerinin sınıflandırılması için; MLP yapay sinir ağı,  $k$ -NN ve LS-SVM yöntemleri sırasıyla kullanılmıştır. Her bir yapının en iyi performansa erişimi için gerekli eğitim işlemleri yapılmıştır.

MLP,  $k$ -NN ve LS-SVM modellerinin eğitiminde, normalize edilmiş veriler sistemin girişine uygulanmıştır. Sınıflandırıcı modellerinin performansları;

$$Performans = \frac{\text{Doğru tahmin edilen örüntü sayısı}}{\text{Toplam test örüntü sayısı}} \cdot 100 \quad (4.1)$$

bağıntısı ile hesaplanmıştır. Öğrenme ve test aşamalarında ortaya çıkan hatalar, Ortalama Karesel Hata Karekökü (RMS) ;

$$RMS = \left( \frac{1}{p} \sum_j |t_j - y_j|^2 \right)^{1/2} \quad (4.2)$$

bağıntısı ile hesaplanmıştır. Bu bağıntıda yer alan  $p$ ,  $t$  ve  $y$  sırasıyla toplam örüntü sayısını, hedef çıkışı ve model tarafından üretilen çıkışı ifade etmektedir. Her modelin performans ve RMS değerleri hesaplanarak, modellerin performansları karşılaştırılmıştır.

### 4.3. GWT Tabanlı Uygulama Sonuçları

GWT tabanlı MLP sınıflandırıcısının farklı topolojileri için performans bilgileri Çizelge 4.1.’de verilmiştir. Kolmogorov’un 1957 yılında yaptığı çalışmada, tek gizli katmanlı bir ağın her türlü problemi sonlu adımda çözebileceğini matematiksel olarak kanıtlamış, ancak gizli katmanda bulunması gereken nöron sayısı belirsizliğini korumuştur. Bundan dolayı, MLP uygulamaları tek gizli katman üzerinden yapılmıştır. MLP ağ yapısı  $24-n-1$  olup, en iyi performansı sağlayacak yapının bulunabilmesi için  $n$  değeri değişken olarak ve [5-50] değer aralığında alınmıştır.

MLP performans değerlerine bakıldığında, Çizelge 4.1.’de görüldüğü gibi performans başarısı [61.1-83.3] aralığında gözlemlenmiş ve en yüksek performans değerinin 24-40-1 ağ topolojisi tarafından sağlandığı saptanmıştır. Ağırlıklar (0,1)



aralığında rastgele belirlenmiş, öğrenme faktörü 0.8, momentum katsayısı 0.1 ve tüm katmanlarda aktivasyon fonksiyonu olarak logsig kullanılmıştır.

**Çizelge 4.1.** GWT tabanlı farklı MLP ağ sınıflandırıcısının performans değerleri

MLP Yapısı	Performans (%)	RMS
24-5-1	66.7	0.5774
24-10-1	61.1	0.6236
24-15-1	72.2	0.5270
24-20-1	61.1	0.6236
24-25-1	77.8	0.4714
24-30-1	61.1	0.6236
24-35-1	66.7	0.5734
<b>24-40-1</b>	<b>83.3</b>	<b>0.4082</b>
24-45-1	77.8	0.4714
24-50-1	72.2	0.5270

GWT tabanlı  $k$ -NN sınıflandırıcısının farklı komşu ( $k$ ) sayıları için performans bilgileri Çizelge 4.2.'te verilmiştir. Bu sınıflandırıcıda, en yakın komşu sayısı ( $k$ ), [1-20] değer aralığında alınarak, her bir  $k$  sayısı için performans değeri hesaplanmıştır.  $k$ -NN yapısına bakıldığında Çizelge 4.2.'te görüldüğü gibi performans başarısı [55.6-72.2] aralığında gözlemlenmiştir.

**Çizelge 4.2.** GWT tabanlı  $k$ -NN sınıflandırıcısının performans değerleri.

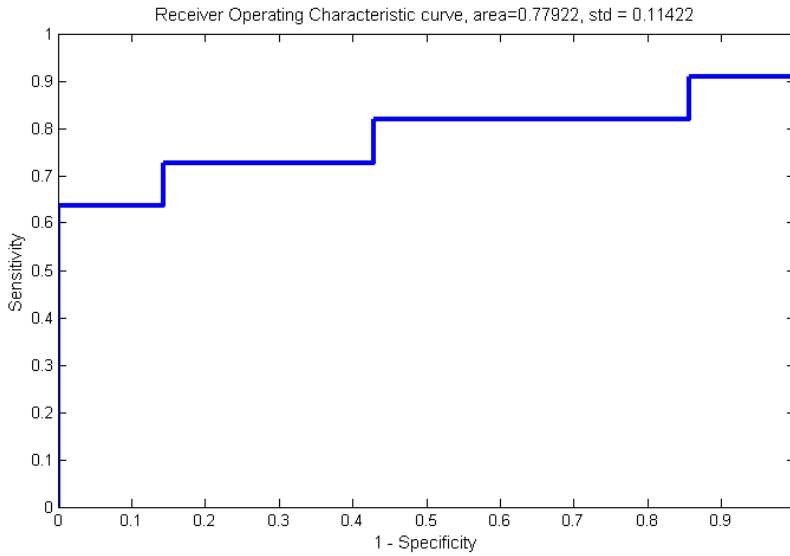
Komşu sayısı ( $k$ )	Eşlenmeyen örüntü oranı	Performans (%)	RMS
1	0.333	66.7	0.5774
2	0.333	66.7	0.5774
3	0.333	66.7	0.5774
4	0.333	66.7	0.5774
5	0.333	66.7	0.5774
<b>6</b>	<b>0.278</b>	<b>72.2</b>	<b>0.5270</b>
7	0.399	61.1	0.6236
8	0.399	61.1	0.6236
9	0.454	55.6	0.6667
10	0.333	66.7	0.5774
11	0.333	66.7	0.5774
12	0.333	66.7	0.5774
13	0.333	66.7	0.5774
14	0.333	66.7	0.5774
15	0.333	66.7	0.5774
16	0.333	66.7	0.5774
17	0.399	61.1	0.6236
18	0.399	61.1	0.6236
19	0.399	61.1	0.6236
20	0.399	61.1	0.6236

#### 4. BULGULAR VE TARTIŞMA

GWT tabanlı LS-SVM sınıflandırıcısının performansı Çizelge 4.3.'te verilmiştir. LS-SVM probleminin eğitiminde kernel genişlik parametresi sigma ( $\sigma=260$ ) ve regülarizasyon parametresi ( $\gamma=2246$ ) olarak seçilmiştir. Çizelge 4.3.'te görüldüğü gibi sistemin performansı %77.8 olarak gözlemlenmiş ve ROC eğrisi Şekil 4.5.'te gösterilmiştir.

**Çizelge 4.3.** GWT tabanlı LS-SVM sınıflandırıcısının performans değerleri.

Kernel Tipi	Eşlenmeyen örüntü oranı	Performans (%)	RMS
<b>RBF</b>	<b>0.222</b>	<b>77.8</b>	<b>0.4714</b>
Linear	0.333	66.7	0.5774



**Şekil 4.5.** GWT tabanlı LS-SVM sınıflandırıcısının ROC eğrisi

Bu aşamada, k-NN yönteminin LS-SVM yöntemine göre daha başarılı olduğu gözlemlenmiştir. Bunun sebebi olarak iki sınıfa ait verilerin düzlemde doğrusal olmayan bir şekilde dağılmış olmasından kaynaklı olabileceği düşünülmektedir.

#### 4.4. GLCM Tabanlı Uygulama Sonuçları

Bu kısımda, GWT tabanlı öznitelik vektörü yerine GLCM tabanlı öznitelik vektörü kullanılarak performansı en iyi öznitelik çıkarma metodu belirlenmeye çalışılmıştır. GLCM tabanlı MLP sınıflandırıcısının farklı topolojileri için performans bilgileri Çizelge 4.4.'te verilmiştir. MLP ağ yapısı 7-n-1 olup, en iyi performansı sağlayacak yapının bulunabilmesi için n değeri değişken olarak ve [5-50] değer

aralığında alınmıştır. MLP performans değerlerine bakıldığında, Çizelge 4.4.'te görüldüğü gibi performans başarısı [72.2-88.9] aralığında gözlemlenmiş ve en yüksek performans değerinin 7-5-1 ağ topolojisi tarafından sağlandığı saptanmıştır. Ağırlıklar (0,1) aralığında rastgele belirlenmiş, öğrenme faktörü 0.8, momentum katsayısı 0.1 ve tüm katmanlarda aktivasyon fonksiyonu olarak logsig kullanılmıştır.

**Çizelge 4.4.** GLCM tabanlı farklı MLP ağ sınıflandırıcısının performans değerleri.

MLP Yapısı	Performans (%)	RMS
<b>7-5-1</b>	<b>88.9</b>	<b>0.3333</b>
7-10-1	83.3	0.4082
7-15-1	72.2	0.5270
7-20-1	83.3	0.4082
7-25-1	77.8	0.4714
7-30-1	77.8	0.4714
7-35-1	77.8	0.4714
7-40-1	72.2	0.5270
7-45-1	83.3	0.4082
7-50-1	78.8	0.4714

GLCM tabanlı  $k$ -NN sınıflandırıcısının farklı komşu ( $k$ ) sayıları için performans bilgileri Çizelge 4.5.'te verilmiştir. Bu sınıflandırıcıda, en yakın komşu sayısı ( $k$ ), [1-20] değer aralığında alınarak, her bir  $k$  sayısı için performans değeri hesaplanmıştır.  $k$ -NN yapısına bakıldığında Çizelge 4.5.'te görüldüğü gibi performans başarısı [72.2-88.9] aralığında gözlemlenmiştir.

GLCM tabanlı en küçük kareler Destek Vektör Makineleri (LS-SVM) sınıflandırıcısının performansı Çizelge 4.6.'da verilmiştir. LS-SVM probleminin eğitiminde kernel genişlik parametresi sigma ( $\sigma=71$ ) ve regülerizasyon parametresi ( $\gamma=19.8$ ) olarak seçilmiştir. Çizelge 4.6.'da görüldüğü gibi sistemin performansı % 83.3 olarak gözlemlenmiş ve ROC eğrisi Şekil 4.6.'da gösterilmiştir.

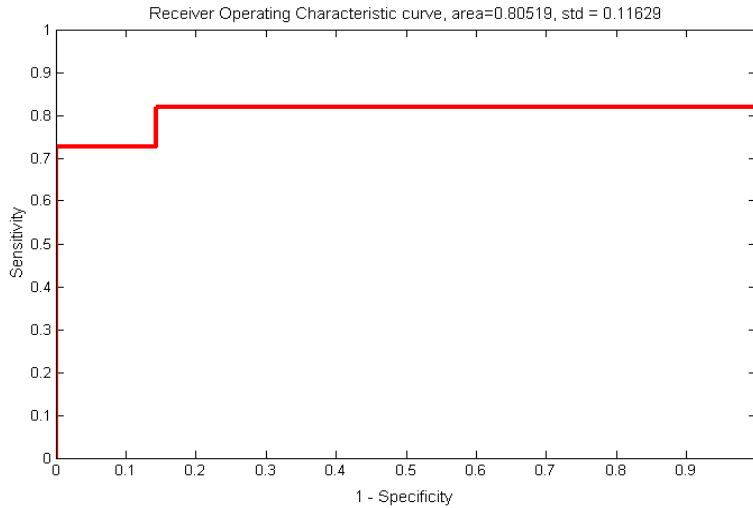
## 4. BULGULAR VE TARTIŞMA

Çizelge 4.5. GLCM tabanlı  $k$ -NN sınıflandırıcısının performans değerleri.

Komşu sayısı ( $k$ )	Eşlenmeyen örüntü oranı	Performans (%)	RMS
1	0.222	77.8	0.4714
2	0.222	77.8	0.4714
3	0.278	72.2	0.5270
4	0.222	77.8	0.4714
5	0.278	72.2	0.5270
6	0.222	77.8	0.4714
7	0.222	77.8	0.4714
8	0.222	77.8	0.4714
9	0.222	77.8	0.4714
10	0.278	72.2	0.5270
11	0.167	83.3	0.4082
12	0.167	83.3	0.4082
13	0.167	83.3	0.4082
<b>14</b>	<b>0.101</b>	<b>88.9</b>	<b>0.3333</b>
15	0.167	83.3	0.4082
<b>16</b>	<b>0.101</b>	<b>88.9</b>	<b>0.3333</b>
17	0.167	83.3	0.4082
18	0.222	77.8	0.4714
19	0.222	77.8	0.4714
20	0.222	77.8	0.4714

Çizelge 4.6. GLCM tabanlı LS-SVM sınıflandırıcısının performans değerleri.

Kernel Tipi	Eşlenmeyen örüntü oranı	Performans (%)	RMS
<b>RBF</b>	<b>0.167</b>	<b>83.3</b>	<b>0.4082</b>
Lineer	0.222	77.8	0.4714



Şekil 4.6. GLCM tabanlı LS-SVM sınıflandırıcısının ROC eğrisi

Sonuç olarak, elde edilen **%88.9** performans başarısı ile modeller içerisinde en yüksek performans GLCM tabanlı  $k$ -NN ve MLP sınıflandırıcılarında gözlemlenmiştir.

Farklı bitkilerin yapraklarına bakılarak hastalıkların saptanması niteliğinde literatürde birçok çalışma yapılmış olup, başarı oranı %86-%100 arasında olduğu gözlenmektedir. Bu çalışmalardan bazıları aşağıda özetlenmiştir.

Rumpf ve ark. (2010) yaptıkları çalışmada; kernel tabanlı SVM modelini kullanmış olup şeker pancarı yapraklarındaki hastalıklı ve sağlıklı bitkilerin ayrıştırılmasında % 97 performans sağlanmıştır (1. Grup).

Bauer ve ark. (2011) yaptıkları çalışmada; k-NN, GKMTBS ve CRF yöntemleri kullanılmış olup şeker pancarı yapraklarındaki hastalıklı bitkilerin ayrıştırılmasında sırasıyla % 86 ve % 91 performans sağlanmıştır (1. Grup).

Li ve ark. (2010) yaptıkları çalışmada; PCA ve DA yöntemleri kullanılmış olup üç farklı yaprak hastalığının ayrıştırılmasında sırasıyla %96.7, %93.3 ve %86.7 performans sağlanmıştır (2. Grup).

Luaces ve ark. (2011) yaptıkları çalışmada; SVM ve LR yöntemleri kullanılmış olup kahve bitkilerindeki pas hastalığının tespitinde sırasıyla %90 ve %78 performans sağlanmıştır (3. Grup).

Römer ve ark. (2011) yaptıkları çalışmada; SVM modeli kullanılmış olup buğday yapraklarındaki pas hastalığının tahmininde %93 performans sağlanmıştır (2. Grup).

Wang ve Ma. (2011) yaptıkları çalışmada; SVM ve RA yöntemleri kullanılmış olup buğday yapraklarındaki sarı pas hastalığının tespitinde % 100 performans sağlanmıştır (1. Grup).

Moshou ve ark. (2004) yaptıkları çalışmada; SOS ve MLP yöntemleri kullanılmış olup buğday bitkilerindeki sarı pas hastalığının tahmininde % 99 başarı sağlanmıştı (2. Grup).

Coakley ve ark. (1982) (3. Grup), Keqiang ve ark.(1995) (3. Grup), Wei(2007) (3. Grup), Eddy (2009) (3. Grup), Mo (2010) (1. Grup), Ping ve ark.(2000) (1. Grup), yaptıkları çalışmalarda YSA ve LR kullanılmış olup yüksek performanslar sağlanmıştır.

Yapılan bu tez çalışmasında ise zambak çiçeği yaprağındaki pas hastalıklarının tespiti için bir sistem tasarlanmıştır. Sistemin sağladığı en yüksek performans (%88.9) daha önce yapılan çalışmaların başarıları (%86-%100) ile kıyaslandığında başarı

#### **4. BULGULAR VE TARTIŞMA**

---

oranının kabul edilebilir aralıkta olduđu düşünölmektedir. Doğrudan zambak çiçeđi üzerine yapılmıř bir tespit sistemi bulunmadıđından ve ayrıca deđerlendirilecek ortak bir veri tabanı olmadıđından sistemin başarı performansı kıyaslanamamaktadır.

## 5. Sonuç ve Öneriler

Bitkilerdeki hastalıklar, hasadı ve dolayısıyla verimi etkilemektedir. Hastalıkların önceden tespiti, çiftçilerin alacağı önlemler ile verimi artıracaktır. Verimi etkileyen önemli hastalıkların başında pas hastalığı gelmektedir. Bu çalışmada bitki örneği olarak zambak çiçeğine ilişkin yaprak imgeleri kullanılarak, bitkide pas hastalığının tespiti amaçlanmıştır.

Bu çalışmada, zambak çiçeğine ilişkin sağlıklı ve hastalıklı sayısal ürün imgelerinin, farklı öznitelik çıkarma (GWT, GLCM) yöntemleri ve sınıflandırma (MLP, k-NN, LS-SVM) teknikleri kullanılarak, pas hastalığının tespiti için 2 farklı (1.Sağlıklı, 2.Hastalıklı) grupta sınıflandırılmıştır.

Analizin ilk aşamasında, öncelikle GWT öznitelik parametreleri kullanılarak sistem eğitilmiş ve sınıflandırıcıların performansları değerlendirilmiştir. İkinci aşamasında, GLCM parametreleri kullanılarak aynı adımlar tekrarlanıp performansı en iyi sistem modeli belirlenmeye çalışılmıştır. Sonuç olarak, en yüksek performansın %88.9 başarı yüzdesi ile GLCM tabanlı k-NN (k=14 ve 16) ve MLP yapay sinir ağının 7-5-1 ağ topolojisinde gözlemlendiği saptanmıştır.

Uygulamada, GLCM öznitelik vektörlerinin GWT öznitelik vektörlerine göre daha iyi performans gösterdiği gözlemlenmiştir. Çalışmada kullanılan imgelerin içerdiği pas hastalığının piksel piksel yayılması ve iki komşu piksel arasındaki bağlantıların ayarlanmasıyla oluşturulması, GLCM yönteminin daha başarılı olmasını sağlamıştır.

Sınıflandırıcılar arasında k-NN ve MLP yöntemlerinin LS-SVM yöntemine göre daha başarılı olduğu gözlemlenmiştir. Bunun sebebi olarak, çalışmada kullanılan verilerin doğrusala yakın ancak doğrusal olmayan bir şekilde dağılımının olduğu düşünülmektedir. k-NN yönteminde mesafe ölçümü ve komşuluk sayısı, sistemin çıkışı doğrudan etkilediğinden dolayı, bu yöntemde başarının yüksek olması muhtemeldir. YSA'da topolojinin, transfer fonksiyonlarının ve öğrenme algoritmasının doğru belirlenmiş olması, ağın nöron bağlantılarının doğru katsayılara ulaştırılmasına sebep olmuştur. Bu da test verilerinin sınıflandırılmasında performansı yükseltmiştir. LS-SVM'de ise problemlerin çözümünde doğrusal denklemler kullanıldığından dolayı çalışmada kullanılan veriler ayırıcı hiperdüzlemle tam olarak ayrıştırılmadığı

## 5.SONUÇ VE ÖNERİLER

---

düşünülmektedir. Bu durum, LS-SVM yönteminin başarı performansını negatif yönde etkilemektedir.

Yapılan çalışmada elde edilen sınıflandırma sonuçlarıyla, daha önce yapılan çalışmalar kıyaslandığında başarı oranının kabul edilebilir aralıkta olduğu düşünülmektedir. Ayrıca önerilen modelin test sonucunda elde edilen toplam sınıflama doğruluğu, modelin pas hastalığının tespit edilmesinde kullanılabileceğini göstermiştir. Pas hastalığının tespiti için oluşturulan sistemin geliştirilmesi ve başarı oranının artırılmasıyla çiftçilerin bitki pataloğuna ihtiyaç duymadan, hastalığın ilk aşamasından itibaren tespit edebileceği düşünülmektedir. Bu durum, bitki hastalıklarıyla mücadelede çiftçilere destek olma noktasında büyük bir aşama niteliğindedir.

Bu çalışmanın devamı olarak İTÜ-UHUZAM ve TAGEM eş yürütücülüğünde ve DPT desteği ile gerçekleştirilen 2000A020010 nolu “*Tarımsal Rekolte Tahmin ve Kuraklık İzleme*” projesi kapsamında elde edilecek sayısal uydu veya kamara imgelerinin sınıflandırılmasıyla belirli bir tarımsal alandaki ürün hastalığının otomatik tespit edilmesi düşünülmektedir.



## 6. Kaynaklar

Acar, H.2010. Uyanıklık Seviyesinin Kestiriminin DSP Tabanlı olarak Gerçekleştirilmesi. Yüksek Lisans Tezi, Dicle Üniversitesi Fen Bilimleri Enstitüsü, Diyarbakır.23.

Acar, E., Özerdem,M.S. ve Akpolat,V.2011. Forecasting Diabetes Mellitus with Biometric Measurements, *International Archives of Medical Research*, 1(1): 28-42.

Alves, M., Pozza, E., Costa,J, Carvalho, L. ve Alves,L. Adaptive neuro-fuzzy inference systems for epidemiological analysis of soybean rust. *Environmental Modelling & Software*, 26 (2011): 1089-1096.

Anonim, (2008). *Daylily Rust*. Erişim: [[http:// massnrc.org/ pests/ pestFAQsheets/ daylilyrust. html](http://massnrc.org/pests/pestFAQsheets/daylilyrust.html)]. *Erişim Tarihi: 21.12.2011*.

Anonim.2009. Çok Katmanlı Algılayıcı. Erişim: [[http:// www.zinderud.com/gelisenler/ cok-katmanli-algilayici-multi-layer-perceptron.html](http://www.zinderud.com/gelisenler/cok-katmanli-algilayici-multi-layer-perceptron.html)]. Erişim Tarihi: 22.11.2011.

Bariamis, D. G., Iakovidis, D. K. ve Maroulis, D. E. 2007. FPGA Architecture for Fast Parallel Computation of Co-occurrence Matrices. *Microprocessors and Microsystems*, 31 (2007): 160–165.

Bauer,S., Korc,F. ve Förstner W.2011. The potential of automatic methods of classification to identify leaf diseases from multispectral images. *Precision Agriculture*,12(3): 361-377.

Bergeron, S., (2009). *Daylily Rust Information Page*. Erişim: [[http:// web.ncf.ca /ah748/ rust. html](http://web.ncf.ca/ah748/rust.html)]. *Erişim Tarihi: 21.12.2011*.

Bilgin, S.2008,Kalp hızı değişkenliğinin dalgacık dönüşümü ve yapay sinir ağları kullanılarak analizi. Doktora tezi, Sakarya Üniversitesi Fen Bilimleri Enstitüsü, Sakarya.138.

Buciu,I. ve Gacsadi, A. 2009. Gabor Wavelet Based Features for Medical Image Analysis and Classification. Applied Sciences in Biomedical and Communication Technologies. ISABEL 2009. 2nd International Symposium, 24-27 Nov. 2009, Bratislava. S, 1-4.

Buttrely, S. and Karo, C.2001. Using k-nearest-neighbor classification in the leaves of a tree. *Computational Statistics & Data Analysis*, 40 (2002) 27-37.

Coakley, S. M., Boyd, W. S. ve Line, R. F. 1982. Statistical Models for Prediction of Stripe Rust on Winter Wheat in the Pacific Northwest. *Phytopathology*, 72: 539-1542.

## 6. KAYNAKLAR

---

Cui,D.,Zhan,Q.,Minzan,L.,Zhao,Y. ve Hartman,G.2009. Detection of soybean rust using a multispectral image sensor. *Sens. & Instrumen. Food Qual.*, 3(2009): 49-56.

Çoklu, O. 2011. Buğday Pas Hastalıkları. Erişim: [[http:// www.bahcebitkileri.org/bugday-pas-hastaliklari.html](http://www.bahcebitkileri.org/bugday-pas-hastaliklari.html)]. Erişim Tarihi: 21.12.2011.

Dauber, R. 2008. Rust Disease on Tomato Plants. Erişim: [[http://www.ehow.com/about\\_6514767\\_rust-disease-tomato-plants.html](http://www.ehow.com/about_6514767_rust-disease-tomato-plants.html) ]. Erişim Tarihi: 21.12.2011.

Daugman, J. G. Two dimensional spectral analysis of cortical receptive field profile. *Vision Research*, 20(1980): 847-856.

Demirhan, A. ve Güler, İ.2010. Özörgütlemeli Harita Ağları ve Gri Düzey Eş Oluşum Matrisleri ile Görüntü Bölütleme. *Gazi Üniv. Müh. Mim. Fak. Der.* 25(2): 285-291.

Eddy, R.2009. Logistic Regression Models to Predict Stripe Rust Infections on Wheat and Yield Response to Foliar Fungicide Application on Wheat in Kansas. Master thesis, Kansas State University, Department of Plant Pathology College of Agriculture, Kansas.53-67.

Eray, O.2008. Destek Vektör Makineleri İle Ses Tanıma Uygulaması. Yüksek Lisans tezi, Pamukkale Üniversitesi,Fen Bilimleri Enstitüsü,Denizli.44.

Gabor, D.1946. Theory of communication. *J. IEE*, 93(1946): 429-459.

Gang, C., Haiguang, W. ve Zhanhong, M. Forecasting Wheat Stripe Rust by Discrimination Analysis. Erişim: [[http:// en.cnki.com.cn/ Article\\_en/ CJFDTOTAL-ZWBH200604005.htm](http://en.cnki.com.cn/Article_en/CJFDTOTAL-ZWBH200604005.htm)]. Erişim Tarihi: 21.12.2011.

Ghosal,V., Tikmani,P. ve Gupta, P.2009. Face Classification Using Gabor Wavelets and Random Forest. Canadian Conference on Computer and Robot Vision, 25-27 May. 2009,Kelowna,BC. S, 68-73.

Hilavin, İ., Kuntalp, M. ve Kuntalp, D.2011. Aritmilerin Spektral Özellikler Kullanılarak En Yakın K Komşu Metoduyla Sınıflandırılması. 2011 IEEE 19th Signal Processing and Communications Applications Conference (SIU), 20-22 Nisan 2011,Antalya.S,411-414.

Hornig, M. H., Huang, X. J. ve Zhuang, J. H. 2003. Texture Feature Coding Method for Texture Analysis and It's Application. *Journal of Optical Engineering*, 42 (1): 228-238.

Huang, C., Davis, L.S., Townshed, J.R.G. 2002.An Assessment of Support Vector Machines for Land Cover Classification. *International Journal of Remote Sensing* 23(4): 725–749.

Hubel, D. H. and Wiesel, T. N. 1978. Functional architecture of macaque monkey visual cortex. *Proc. Royal Soc. B (London)*, 198(1978): 1-59.

Kakıcı, A. 2011. Yapay Sinir Ağlarına Giriş. Erişim: [[http:// www.ahmetkakici.com/yapay-sinir-aglari/ yapay-sinir-aglarinin-mimarisi-ve-yapi elemanlari/](http://www.ahmetkakici.com/yapay-sinir-aglari/yapay-sinir-aglarinin-mimarisi-ve-yapi-elemanlari/)]. Erişim Tarihi: 22.11.2011.

Kaymaz, E.2007. Yapay bağışıklık sistemi tabanlı k-NN sınıflandırma algoritması ile protein örüntülerinin hücredeki yerleşim yerlerinin belirlenmesi. Yüksek Lisans tezi, Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Elazığ.45.

Keqiang, C., Zhiyu, Z. ve Shuanma, M. 1995. The Establishment of a Mid-term Forecast Model for Leaf Rust of Wheat. Erişim: [[http://en.cnki.com.cn/ Article\\_en/CJFDTOTAL-ZWBF199501010.htm](http://en.cnki.com.cn/Article_en/CJFDTOTAL-ZWBF199501010.htm)]. Erişim Tarihi:20.12.2011.

Kepekci, B. 2001, Face Recognition Using Gabor Wavelet Transform. Yüksek Lisans tezi, ODTÜ Fen Bilimleri Enstitüsü, Ankara.65.

Kohonen, T. 1990. The Self-Organizing Map. *Proceedings of Institute of Electrical and Electronics Engineers*, 78(1990): 1464-1480.

Lee, T. S. 1996. Image representation using 2-d Gabor wavelets. *IEEE Trans.On Pattern Analysis Pattern Analysis and Machine Intelligence*, 18(10):

Lei, Y.ve Shuqin, L. 2009. Prediction of wheat stripe rust by wavelet neural network. *Micro-computer Information*, 25(12): 42-43.

Lei, Y.ve Shuqin, L. 2010. Application of principal component analysis of wheat stripe rust. Erişim: [[http://en.cnki.com.cn/Article\\_en/CJFDTOTAL-SJSJ201002061.htm](http://en.cnki.com.cn/Article_en/CJFDTOTAL-SJSJ201002061.htm)]. Erişim Tarihi: 21.12.2011.

Li, J., Gao, L. ve Shen, Z.2010. Extraction and analysis of digital images feature of three kinds of wheat diseases. 2010 3rd International Congress on Image and Signal Processing (CISP2010), 16-18 Oct. 2010, Yantai. 2543–2548.

Li, X., Nie, P., Jun,Z. and He, Y. 2011.Using wavelet transform and multi-class least square support vector machine in multi-spectral imaging classification of Chinese famous tea. *Expert Systems with Applications* 38(9):11149-11159.

Liu,Z., Wu, H. ve Huang J.2010.Application of neural networks to discriminate fungal infection levels in rice panicles using hyperspectral reflectance and principal components analysis. *Computers and Electronics in Agriculture*, 72 (2010): 99-106.

## 6. KAYNAKLAR

---

Lourakis, M.I.A.2005. A Brief Description of the Levenberg-Marquardt Algorithm Implemented by levmar.Eriřim:[ <http://www.ics.forth.gr/~lourakis/levmar>]. Erisim Tarihi :22.12.1996.

Luaces, O., Rodrigues, L., Meira,C. ve Bahomende, A.2011. Using nondeterministic learners to alert on coffee rust disease. *Expert Systems with Applications*, 38 (2011): 14276-14283.

Manian, V. ve Vasquez, R. A.1997. Framework for SAR Image Classification: Comparison of CO-occurrence and a Gabor based method. *Geoscience and Remote Sensing,1997. IGARS '97. Remote Sensing-A scientific Vision for Sustainable Development. 1997 IEEE International*, 1(1997): 335-337.

McIntosh, RA., Wellings, CR. Ve Park, RF.1995. Wheat Rusts. CSIRO Publications, Yayın No:45,Sayfa: 1-87. Sydney.

Marcelja, S. 1980.Mathematical description of the responses of simple corticalCells. *J. Optical Soc. Am.*, 70(1980): 1297-1300.

Miquel, C.M.2009. Texture Descriptors applied to Digital Mammography. Master thesis, University of Girona, Department of Computer Architecture and Technology, Girona.14-18.

Mo, L.2010. Prediction of Wheat Stripe Rust using Neural Network. *Intelligent Computing and Intelligent Systems (ICIS), 2010 IEEE International Conference*, 29-31 Oct. 2010, Xiamen.S,475-479.

Moshou, D., Bravo, C., West, J., Wahlen, S., McCartney, A. ve Ramon, H.2004. Automatic detection of 'yellow rust' in wheat using reflectance measurements and neural Networks. *Computers and Electronics in Agriculture*, 44(3): 173-188.

Orhan, U., Hekim, M.ve Özer, M.2010. EEG İşaretlerinin Çok-katmanlı Algılayıcı Yapay Sinir Ağı Modeli ile Sınıflandırılmasında Ayrıklaştırma Yaklaşımı. 15. Biyomedikal Mühendisliği Ulusal Sempozyumu - BİYOMUT 2010, 21-24 Nisan 2010, Antalya. S, 1-3.

Öğücü, M.O, 2006, Yapay Sinir Ağları ile Sistem Tanıma. Yüksek Lisans tezi, İTÜ. Fen Bilimleri Enstitüsü, İstanbul.10-11.

Öztemel, E. 1996. Bilgisayarda Öğrenme ve Yapay Nöral Ağları. Otomasyon, Yayın No:45, Sayfa: 134-140.İstanbul.

Ping, H., Wei, Y., Qi, L., Yong, D. ve Hua, K..2000. Prediction of Wheat Stripe Rust in Hanzhong Area by BP Neural Network. Eriřim:[ [http:// en.cnki.com.cn/ Article\\_en/CJFDTOTAL-XBNX200003007.htm](http://en.cnki.com.cn/Article_en/CJFDTOTAL-XBNX200003007.htm)]. Eriřim Tarihi:20.12.2011.

Pöyhönen, S., Negrea, M., Arkkio, A., Hyötyniemi, H. ve Koivo, H.2002. Fault Diagnostics of an Electrical Machine with Multiple Support Vector Classifiers. *Proc. of the 17th IEEE Int. Symp. on Intelligent Control, ISIC'02*, 1(2002): 373-378.

Roumi, M.2009. Implementing Texture Feature Extraction Algorithms on FPGA. Master thesis, Delft University of Technology, Faculty of Electrical Engineering, Mathematics and Computer Science,Delft,Netherlands.15.

Römer, C., Bürling, K., Hunsche,M., Rumpf, T., Noga,G. ve Plümer, L. Robust fitting of fluorescence spectra for pre-symptomatic wheat leaf rust detection with Support Vector Machines. *Computers and Electronics in Agriculture* 79 (2011): 180–188.

Rumpf, T., Mahlein, A.-K., Steiner, U., Oerke, E.-C., Dehne,H.-W. ve Plümer,L.2010. Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral. *Computers and Electronics in Agriculture*, 74(1): 91-99.

Sanyal, P., Bhattacharya, U., Parui, S. K., Bandyopadhyay, S. K. Ve Patel, S. Color Texture Analysis of Rice Leaves Diagnosing Deficiency in the Balance of Mineral Levels towards Improvement of Crop Productivity. 10th International Conference on Information Technology (ICIT 2007), 17-20 Dec, Orissa.85-90.

Schwenker, F.2000. Hierarchical Support Vector Machines for Multi-Class Pattern Recognition. *Proc. of 4th Int. Conference on Knowledge-Based Intelligent Engineering Systems and Allied Technologies*, 2(2000): 561-565.

Shahin, M. A., Tollner, E. W. and Gitaitis, R. D. Classification of sweet onions based on internal defects using image processing and neural network techniques. *Transactions of the ASAE*, 2002, 45(5): 1613-1618.

Sheng, H., Zhang,L., Tang,J.,Sun,Y. Ve Zhao,L. 2011. *Quality, Reliability, Risk, Maintenance, and Safety Engineering (ICQR2MSE)*, 17-19 Haziran 2011,Xi-an.673-677.

Suyskens, J., Gestel,T., Brabender, J., Moor, B. ve Vandewalle J.2002. Least Squares Support Vector Machines. © World Scientific Publishing Co. Pte. Ltd, Yayın No: (ISBN 981-238-151-1), Sayfa: 71. Leuven.

Taşkafa, Ş.2006. Asenkron Motorların Yapay Sinir Ağları ile Vektör Kontrolü. Yüksek Lisans tezi, Fırat Üniversitesi Fen Bilimleri Enstitüsü, Elazığ.25.

## 6. KAYNAKLAR

---

Vajda, F.1994. Techniques and trends in digital image processing and computer vision. Mathematical Modelling and Simulation of Industrial and Economic Processes, IEE Colloquium. 06 Ağustos 2002, Budapest.1.

Wang, H. ve Ma, Z. 2011. Prediction of Wheat Stripe Rust Based on Support Vector Machine. 2011 Seventh International Conference on Natural Computation, 26-28 July 2011, Shanghai.378-382.

Wei, Y.2007. Study on the Prediction Model of Wheat Rust in Shangqiu.Erişim: [[http://en.cnki.com.cn/Article\\_en/CJFDTOTAL-AHNY200736080.htm](http://en.cnki.com.cn/Article_en/CJFDTOTAL-AHNY200736080.htm)].Erişim Tarihi:20.12.2011.

Yaman, B. ve Ballı, E.2009. Yapay Sinir Ağlarının Sınıflandırılması. Erişim: [[http://www.controlworld.tk/web\\_project\\_files/002cbe462b70b9a6ebe7f314a616c067/project\\_y\\_sa\\_report.html](http://www.controlworld.tk/web_project_files/002cbe462b70b9a6ebe7f314a616c067/project_y_sa_report.html)]. Erişim Tarihi: 22.11.2011.

Yanli, C., Pengfei, C., Xiaozhi, D., Zhihua, L. ve Shuangxi, W. 2005. Image processing and extracting color features of greenhouse diseased leaf. Erişim: [[http://en.cnki.com.cn/Article\\_en/CJFDTOTAL-NYGU2005S2008.htm](http://en.cnki.com.cn/Article_en/CJFDTOTAL-NYGU2005S2008.htm)]. Erişim Tarihi:20.12.2011.

Yıldırım, S. 2006. Arıza Teşhisinde Destek Vektör Makinelerinin Kullanımı. Yüksek Lisans tezi, Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Elazığ.38.

Youwen, T. ve Chenghua, L. 2004. Study on Plant Disease Recognition Using Support Vector Machine and Chromaticity Moments. Erişim: [[http://en.cnki.com.cn/Article\\_en/CJFDTOTAL-NYJX200403025.htm](http://en.cnki.com.cn/Article_en/CJFDTOTAL-NYJX200403025.htm)]. Erişim Tarihi:20.12.2011.

Takanashi, H., Furuya, H. ve Chonan, S. Prediction of Disease Infection of Welsh Onions by Rust Fungus Based on Temperature and Wetness Duration. Control Applications, 2007. CCA 2007, 1-3 Oct, Akita. 325-330.

## ÖZGEÇMİŞ

Adı ve Soyadı: Emrullah ACAR

Doğum Yeri: Midyat

Doğum Tarihi: 01.02.1987

Medeni Hali: Bekar

Yabancı Dili: İngilizce, İsveççe

Eğitim Durumu (Kurum ve Yıl)

Lise: Midyat Aziz Önen Lisesi / Midyat, 2004

Lisans: Çukurova Üniversitesi, Elektrik Elektronik Mühendisliği (İng.) /ADANA,2009

Linköping University, Electrical Engineering / SWEDEN, 2008 (Değişim Öğrencisi)

Çalıştığı Kurum/Kurumlar ve Yıl:

FDL Enerji

Ar-Ge Mühendisi (2009-2010), Adana

Batman Üniversitesi

Araştırma Görevlisi (2010-Halen), Batman

Yayımları (SCI ve diğer makaleler)

1- Acar,E., Özerdem, M.S. and Akpolat, V. 2011. Diabetes Mellitus Forecast Using Various Types of Artificial Neural Networks, International Advanced Technologies Symposium 2011 (IATS 2011), 3(2011), 16-18 May 2011, Elazığ. S,196.

2- Acar, E., Özerdem, M.S. and Akpolat, V. 2011. Forecasting Diabetes Mellitus with Biometric Measurements, *International Archives of Medical Research*, 1(1). 28-42.