

**ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

YÜKSEK LİSANS TEZİ

Hatice KILINÇ

**DERİN ÖĞRENMELİ KONVOLÜSYONEL SİNİR AĞLARI
(DEEP LEARNING CONVOLUTIONAL NEURAL NETWORK)
KULLANARAK FOTOĞRAFTAN TRİZOMİ 21(DOWN
SENDROMU) TESPİTİ**

BİYOMEDİKAL MÜHENDİSLİĞİ ANABİLİM DALI

ADANA-2018

**ÇUKUROVA ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**DERİN ÖĞRENME Lİ KONVOLÜSYONEL SİNİR AĞLARI (DEEP
LEARNING CONVOLUTIONAL NEURAL NETWORK) KULLANILARAK
FOTOĞRAFTAN TRİZOMİ 21 (DOWN SENDROMU) TESPİTİ**

Hatice KILINÇ

YÜKSEK LİSANS TEZİ

BİYOMEDİKAL MÜHENDİSLİĞİ ANABİLİMDALI

Bu tez 19.9/2018 Tarihinde Aşağıdaki Jüri Üyeleri Tarafından
Oybirliği/Oyçokluğu İle Kabul Edilmiştir.

Dr. Öğr. Üyesi C. Veysel BAYSAL
DANIŞMAN

Doç. Dr. Mutlu AVCI
JÜRİ

Dr. Öğr. Üyesi Onur ÜLGİN
JÜRİ

Bu tez Enstitümüz Biyomedikal Mühendisliği Anabilim Dalında hazırlanmıştır.
Kod No :

Prof. Dr. Mustafa GÖK
Enstitü Müdürü

Not: Bu tezde kullanılan özgün ve başka kaynaktan yapılan bildirişlerin, çizelge, şekil ve fotoğrafların kaynak gösterilmeden kullanımı 5846 sayılı Fikir ve Sanat Eserleri Kanunundaki hükümlere tabidir.

ÖZ

YÜKSEK LİSANS TEZİ

DERİN ÖĞRENMELİ KONVOLÜSYONEL SİNİR AĞLARI (DEEP LEARNING CONVOLUTIONAL NEURAL NETWORK) KULLANARAK FOTOĞRAFTAN TRİZOMİ 21(DOWN SENDROMU) TESPİTİ

Hatice KILINÇ

ÇUKUROVA ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ BİYOMEDİKAL MÜHENDİSLİĞİ ANABİLİM DALI

Danışman : Dr. Öğr. Üyesi C. Veysel BAYSAL
Yıl: 2018, Sayfa: 81
Jüri : Dr. Öğr. Üyesi C. Veysel BAYSAL
: Doç. Dr. Mutlu AVCI
: Dr. Öğr. Üyesi. Onur ÜLGİN

Down sendromu kromozomal bir bozukluktur, kişinin fiziksel görünümü etkilemektedir. Down Sendromunun teşhisi, kesin sonuç veren pahalı genetik testlerin uygulanması ile gerçekleştirilmektedir. Bu nedenle uzmanlar, daha basit ve ekonomik olan fenotipteki fiziksel görünüm değişimlerini kullanarak teşhis koymaktadırlar. Bu çalışmanın amacı, fenotip değişimlerini kullanarak, fotoğraftan down sendromlu kişileri tespit edecek, yapay zekaya dayalı bir sınıflandırma ve karar destek sistemi oluşturmaktır. Çalışmada downlu kişilerle normal kişilerin fotoğraflarının sınıflandırılması, Matlab-Alexnet yazılımı ile derin öğrenme yöntemi (CNN) kullanılarak yapılmıştır. Fotoğraf verisetleri internetten açık kaynak lisanslama ile elde edilmiştir. Resimlere gürültü eklenerek de performans testleri gerçekleştirilmiştir. Sonuçlar incelendiğinde, hem eğitim, hem de performans analizi süreçlerinde, tasarlanan sistemin büyük oranda(%85-90 oranında) başarılı olduğu görülmüştür.

Anahtar Kelimeler: Yapay Zeka, Makine öğrenmesi, Derin öğrenme, Down sendromu, Down sendromlu kişilerin yüzlerini tespit etme.

ABSTRACT

MSc THESIS

DETECTION OF TRISOMY 21 (DOWN SYNDROME) FROM IMAGE USING DEEP LEARNING CONVOLUTIONAL NEURAL NETWORK METHOD

Hatice KILINÇ

**CUKUROVA UNIVERSITY
INSTITUTE OF NATURAL AND APPLIED SCIENCES
DEPARTMENT OF BIOMEDICAL ENGINEERING**

Supervisor : Asst. Prof. Dr. C. Veysel BAYSAL
Year: 2018, Pages: 81
Jury : Asst. Prof. Dr. C. Veysel BAYSAL
: Assoc Prof. Dr. Mutlu AVCI
: Asst. Prof. Dr. Onur ÜLGEN

Down syndrome is a chromosomal disorder which affects physical appearance of the person. Diagnosis of Down Syndrome is performed using expensive but accurate genetic methods. For this reason, experts diagnose the syndrome using physical appearance change of the phenotype, which is simpler and cheaper. The objective of this study is to design and implement an artificial intelligence based classification and decision support system, in order to identify people with down syndrome from the photographs of phenotypes. In the study, the classification of photos with the downies and normal persons was done by using the Matlab-Alexnet software and the Deep Learning Method (CNN). Photo dataset was obtained from the internet by open source licensing. Performance tests for the system is also performed by adding noisy photos. According to the results obtained, it was concluded that the classifier system executed tasks with a success rate of 85-90% ,in the training phase and also in performance analysis processes.

Key Words: Deep learning, Machine learning, Down syndrome, Face Recognition of Downies

GENİŞLETİLMİŞ ÖZET

Günümüzde dismorfik hastalıkların tanısı yüze bakılarak koyulabilmektedir. Bunlar arasında en bilinenlerinden biri de Down Sendromu (Trizomi 21)'dur. Hastalığın genetik bir hastalık ve kromozomal kaynaklı olduğu 1959 yılında Lejeune ve arkadaşları tarafından rapor edilmiştir[1]. Down Sendromunun teşhisi laboratuvar ortamında kromozomların incelenmesi ve genetik testlerin uygulanması ile mümkün olmaktadır. Genetik testler kesin sonuç verir fakat pahalıdır. (M.E. Sipahi, 2008). Bu nedenle uzmanlar ilk aşamada, genetik testleri genetik testleri gerçekleştirmeden önce, fenotipteki değişimleri göz önünde bulundurarak teşhis koymaktadırlar. Down Sendromunun fenotiplerinde belirgin özellikler; göz şekilleri, anormal dil büyüklükleri, kafa şekilleri ve ense kalınlıklarıdır ve normal bir insana göre farklıdır. Bu sebeple fotoğraftan tanı koymak da mümkündür.

Son zamanlarda oldukça popüler hale gelen deep learning (derin öğrenme) hayatımızın birçok alanında kullanılmaktadır. E-mail filtreleme, arama motorlarında aranan eşleştirme, akıllı telefonlarda, sosyal medyada, e-ticaret gibi alanlar bunlara yazılabilir. Akademik çalışmalar bu alanlarda kullanımı için öncü olmuştur. Deep learning günlük yaşam dışında akademik çalışmalarda yüz tanıma, nesne tanıma, nesne detekte etme, text ayırt etme, ses tanıma amaçlı kullanılır. Derin öğrenme bir yapay sinir ağı çeşididir ve çok katmanlı yapıdadır. Ne kadar katman artırılırsa derinlik o kadar artar ve dolayısıyla doğru öğrenme oranı artar. Derin konvolüsyonel ağlar görüntü, video, konuşma ve ses işlemede başarılı olarak kullanılıyorken, tekrarlayan sinir ağları text ve konuşma gibi ardışık verilerde kullanılır.

Yapılan çalışmada down sendromlu kişilerin fotoğrafları kullanılarak sınıflandırma yapma amaçlanmıştır. Toplamda 460 adet resim kullanılmıştır. Bunun içerisinde 20 adet normal 20 adet down sendromlu kişilerin resimleri seçilerek 230 adet resim LFW data setinden kullanılmıştır. 230 adet down

sendromlu kişilerin resmi birkaç internet(açık kaynak lisanslı) sitesinden indirilmiştir. 230 adet resim matlabda görüntü işlemeyle motion ve salt-pepper filtreleri kullanılarak bozulmuş ve eğitilmiş, sonrasında performansı test edilmiştir. Derin öğrenmede nitelik çıkarma mühendisliği yapılmamaktadır. Nitelik çıkarma işlemi konvolüsyon katmanlarında gerçekleşmektedir. Yani derin öğrenme ağlarında nitelik çıkarma ve sonrasında sınıflandırma yaparak kendi kendine öğrenme sağlanır. Deep learning yapısının avantajı bu olmasına rağmen fazla sayıda resim gerektirmesi bir dezavantajdır. Çünkü fotoğraf bulup veri seti oluşturmak ciddi ve zor bir iştir. Bazı çalışmalar yapabilmek için veri tabanı oluşturma çalışmaları yürütülmektedir. Örneğin ImageNet Büyük Ölçekli Görsel Tanıma Yarışması (ImageNet Large Scale Visual Recognition Challenge – ILSVRC) yüzlerce obje kategorisi ve milyonlarca resimden obje sınıflandırmaya yönelik yapılan bir faaliyettir. Bu faaliyet ile birlikte büyük veri setleri ve bazı derin öğrenme modelleri oluşturulmuştur. Bu modellerden biri de Alexnet olarak adlandırılan bir derin öğrenme ağ modelidir. Alexnet Alex Krizhevsky ve arkadaşları ile birlikte oluşturduğu bir Imagenet modelidir. Alexnet bir konvolüsyonel sinir ağıdır. 5 adet konvolüsyon katmanı 3 adet de tamamen bağlı katman bulunmakta ve son katmanda hangi sınıfa ne olasılıkla ait olduğunu hesaplayan softmax katmanı bulunmaktadır.

Derin öğrenmede en büyük sorun veri seti büyüklüğü ve ezberleme sorunudur. Bu sorunlar için data artırma ve yeni bir yöntem olan “dropout” yöntemleri ileri besleme döngüsünde kullanılır. Yapay sinir ağlarında öğrenme sağlanması için her döngüde ne kadar öğrendiğini test etmek için doğrulama yapılır. Hatayı hesaplamak için tahmin edilen ile doğrusu arasındaki farkların karesi alınarak hesaplama yapılır. Bu da parabolik bir fonksiyon oluşur. Bu fonksiyon maliyet, kayıp ya da hata fonksiyonu olarak adlandırılır ve öğrenme sağlanması için bu fonksiyonun minimuma indirilmesi gerekmektedir. Bu fonksiyon parabolik bir fonksiyon olduğundan yerel minimuma ulaşıldığında hata da minimuma inmiş olur ve bu hatayı azaltma işlemi kısmi türevi alınarak yapılır

ve zincir kuralı kullanılarak hata tüm katmanlara yayılır. Fakat bu işlem küçük adımlarla yapılmalıdır. Bu da küçük öğrenme oranı belirlenmesiyle olur. Küçük adımlarla minimuma yavaş yavaş ulaşmaya çalışır. Öğrenme oranı büyük olduğunda minimuma ulaşması zorlaşır ve sistemin kafası karıştırarak hatanın büyümesine sebep olur. Alexnet dropout oranı 0.5 öğrenme oranı çalışma esnasında belirlenir. 0.9 momentum değeri 0.0005 ağırlık sönümleme değeridir. Mini batch size değeri de 128'dir.

Önerilen çalışmada alexnet modeli kullanılarak sınıflandırma yapılmıştır. Matlab paket programı kullanılmıştır. Bu çalışmayla down sendromlu ve normal kişilerin yüzlerini iki sınıfa ayırma amaçlanmıştır. Çalışmada sınıflandırma aşamasında eklenen gürültülerin öğrenme performansına etkisi incelenmiştir. Sonuçlarda meydana gelen değişiklikler tablolar ve grafikler halinde sonuçlar kısmında verilmiştir.



TEŞEKKÜR

Çalışmam süresince bana yol gösteren ve her türlü desteğini esirgemeyen danışmanım Dr. Öğretim Üyesi Cabbar Veysel Baysal' a teşekkürü bir borç bilirim.

Manevi olarak ve iş paylaşımı olarak anlayışlarından ötürü bölümdeki araştırma görevlisi arkadaşlarıma ve bölüm öğretim üyesi hocalarıma çok teşekkür ediyorum.

Hayatım boyunca benden desteklerini esirgememiş olan en büyük destekçim, aileme teşekkürü borç bilirim.

İÇİNDEKİLER

SAYFA

ÖZ.....	I
ABSTRACT	II
GENİŞLETİLMİŞ ÖZET	III
TEŞEKKÜR	VII
İÇİNDEKİLER.....	VIII
ÇİZELGELER DİZİNİ.....	X
ŞEKİLLER DİZİNİ	XII
SEMBOLLER VE KISALTMALAR	XVI
1. GİRİŞ.....	1
2. LİTERATÜR ARAŞTIRMASI.....	3
2.1. Down Sendromu ve Teşhis Yöntemleri.....	3
2.2. Yapay Zeka Yöntemleri ile Yüz Tanıma Sistemleri.....	4
2.3. Down Sendromlu Yüz Tanıma için Yapay Zeka Uygulamaları.....	7
3. MATERYAL VE METOT	11
3.1. Giriş	11
3.2. YSA(Yapay Sinir Ağları)	12
3.3. Derin Öğrenme Yöntemleri	18
3.3.1. Konvolüsyonel Sinir Ağları.....	20
3.3.1. Konvolüsyon.....	21
3.3.2. Normalizasyon.....	25
3.3.3. Pooling.....	26
3.3.4. ReLU	27
3.3.5. Dropout.....	28
3.3.6. Softmax.....	29
3.4. Öğrenme Aktarma (Transfer Learning).....	35
3.5. Alexnet	38
3.6. Fotoğraf Veri Setinin Oluşturulması	46

3.7. Özniteliklerin Çıkarılması	47
3.8. Öğrenme Verileri ve Ön eğitilmiş Ağ kullanımı.....	48
3.8.1. Normal veri seti ile eğitim.....	48
3.8.2. Bozuk veri seti ile eğitim.....	51
3.8.3. Az Sayıda Resimle Yapılan Eğitim	57
3.8.3.1. 50 Adet Resimle Yapılan Eğitim.....	57
3.8.3.2. 100 Adet Resimle Yapılan Eğitim.....	59
4. SONUÇLAR	63
4.1. Normal Eğitim Setiyle Yapılan Eğitim sonuçları.....	64
5. TARTIŞMA.....	69
KAYNAKLAR.....	71
ÖZGEÇMİŞ.....	81

ÇİZELGELER DİZİNİ

SAYFA

Çizelge 3.1. Optimum eğitimde kullanılan değiştirilebilir hiperparametreler.	50
Çizelge 4.1. Data set boyutunun eğitim performansına etkisi.....	66
Çizelge 4.2. Kalitesi bozulmuş ve karıştırılmış resimlerle yapılan eğitimin performansı.	68
Çizelge 4.3. Makine öğrenmesi yöntemleri kullanılarak yapılan down sendromlu resmi sınıflandırma ve mevcut çalışma.....	68





ŞEKİLLER DİZİNİ

SAYFA

Şekil.2.1.	Trizomi 21 kromozom yapısı ¹	4
Şekil.2.2.	Aktif Yüz Tanıma blok diyagramı	6
Şekil.2.3.	VGG-Face ağ yapısı.....	7
Şekil.3.1.	Yapay sinir ağı topolojisi	14
Şekil.3.2.	Sigmoid ve ReLU fonksiyonları	15
Şekil.3.3.	Derin İnanç Ağları Ağ Yapısı.....	20
Şekil.3.4.	Örnek Bir CNN Yapısı ²	20
Şekil.3.5.	Konvolüsyon matris işlem gösterimi	23
Şekil.3.6.	Düzenli 3 katmanlı bir Sinir Ağı.....	25
Şekil.3.7.	Bir ConvNet, nöronlarını katmanlardan birinde görselleştirildiği gibi üç boyutta (genişlik, yükseklik, derinlik) düzenler.....	25
Şekil.3.8.	max-pooling ve mean-pooling	26
Şekil.3.9.	Transfer Learning(Aktarmalı Öğrenme) Aşamaları.....	37
Şekil.3.10.	Mnist data setinin tabi tutulduğu eğitim modeli Resim	39
Şekil.3.11.	Yerel Tepki Normalizasyon şeması	42
Şekil.3.12.	AlexNet Şeması	44
Şekil.3.13.	Dropouttan önce.....	45
Şekil.3.14.	Dropouttan sonra.....	45
Şekil.3.15.	Dropout Açık Şekil	45
Şekil.3.16.	Gürültü eklenmiş resimler.....	47
Şekil.3.17.	Bulanıklaştırılmış resimler.....	47
Şekil.3.18.	Eğitim sürecinde elde edilen performans başarı grafiği.....	48
Şekil.3.19.	Optimum değerlerde eğitimin sınıflandırması %97.62.....	49
Şekil.3.20.	Eğitim setinden seçilen rastgele 16 görüntü	49
Şekil.3.21.	Validasyon setinden seçilen rastgele 4 görüntü	50
Şekil.3.22.	Eğitim süresince elde edilen grafik(bulanık resimlerle)	51

Şekil.3.23. Bulanık resimlerle yapılan eğitim sonucu %86.96doğruluk oranı.	52
Şekil.3.24. Eğitim setinden rastgele seçilen bulanık resimler.....	52
Şekil.3.25. Validasyon setinden seçilen rastgele 4 görüntü	53
Şekil.3.26. Eğitim süresince elde edilen grafik(gürültülü resimlerle).....	53
Şekil.3.27. Gürültülü resimle eğitim sonucu sınıflandırma doğruluğu %91.30	54
Şekil.3.28. Gürültülü resimlerde eğitimde kullanılan resimlerden rastgele seçilmiş resimler	54
Şekil.3.29. Sınıflandırma sonucu validasyon resimlerinde rastgele seçilmiş 4 resim.....	55
Şekil.3.30. Eğitim süresince elde edilen grafik(karışık resimlerle)	55
Şekil.3.31. Gürültülü resimle eğitim sonucu sınıflandırma doğruluğu %91.30	56
Şekil.3.32. Karışık resimlerde eğitimde kullanılan resimlerden rastgele seçilmiş resimler	56
Şekil.3.33. Validasyon setinden seçilen rastgele 4 görüntü	57
Şekil.3.34. Eğitim süresince elde edilen grafik(50 adet resimle yapılan eğitim).....	57
Şekil.3.35. Aynı resimlerden 50 adet seçilip aynı hiperparametrelerle eğitim sonucu	58
Şekil.3.36. 50 adet resim bulunan eğitim seti için eğitimde kullanılan resimlerden rastgele seçilmiş resimler	58
Şekil.3.37. Sınıflandırma sonucu Validasyon resimlerinden rastgele seçilen 4 adet resim.....	59
Şekil.3.38. Eğitim süresince elde edilen grafik(100 adet resimle yapılan eğitim).....	59
Şekil.3.39. Aynı resimlerden 100 adet seçilip aynı hiper parametrelerle eğitim sonucu	60

Şekil.3.40.	100 adet eğitim veri setinden rastgele seçilmiş resimler.....	60
Şekil.3.41.	Sınıflandırma sonucu validasyon veri setinden rastgele seçilmiş 4 resim.....	61
Şekil 4.1.	Karışık resimlerle yapılan test sonucu %85 başarı elde edilmiştir.	64
Şekil 4.2.	Down Sendromlu resimlerle yapılan test sonucu %100 başarı elde edilmiştir.....	65
Şekil 4.3.	Normal resimlerle yapılan test sonucu %75 başarı elde edilmiştir.	66



SEMBOLLER VE KISALTMALAR

YSA	: Yapay Sinir Ağları
MLP	: MultiLayerPerceptron, çok katmanlı perseptron
K-NN	: K- Nearest Neighbour, K En Yakın Komşu
SVM	: Support Vector Machine, Destekçi Vektör Makinesi
IQ	: IntelligenceQuotient – Zeka derecesi
LFW	: Labeled Face in Wild,
NCC	: En Yakın Merkez Sınıflandırma
GWT	: Gabor Dalgacık Dönüşümü
PCA	: Temel Bileşen Analizi
ILSVRC	: ImageNet Large Scale Visual Recognition Challenge -ImageNet Büyük Ölçekli Görsel Tanıma Yarışması
X	: Giriş
W	: Ağırlık
Σ	: Transfer Fonksiyon
Φ	: Aktivasyon Fonksiyon
ReLU	: Rectified Linear Unit
DBN	: Derin İnanç Ağları
CNN	: Konvolüsyonel Sinir Ağları
t	: zaman
X(a)	: a ya bağlı giriş fonksiyon
W(t-a)	: Kernel fonksiyon
*(yıldız)	: Konvolüsyon işlemi
I	: Image-Fotoğraf
K	: kernel
(i,j)	: Fotoğrafın pixel sıralı ikilisi
(m,n)	: Kernelin pixel sıralı ikilisi
S	: Konvolüsyon çıkışı

$H \times W \times D$	: Yükseklik x Genişlik x Derinlik
P	: Padding (Doldurma)
S	: Stride (Adım)
$3D$	: 3 Boyutlu
μ	: ortalama
σ	: ortalamadan standart sapma
Z	: Normalizasyonun sonucu
T	: pooling boyutu
f_j	: yapay sinir ağı aktivasyon çıkışı
L_i	: Loss Function-Kayıp Fonksiyonu
ΔW	: W 'nin kısmi türevi
$a_{x,y}^i$	: i . kernelin (x,y) pozisyonu için kaynak çıkışı
N	: kernel sayısı
n	: normalizasyon komşuluk boyutu
α, β, k	: hiperparametre
LRN	: Yerel Tepki Normalizasyonu
GPU	: Grafik İşleme Birimi
P_i ve λ_i	: i ninci özvektörler
i	: İterasyon indisi
V	: Momentum değişkeni
e	: Öğrenme oranı

1. GİRİŞ

Günümüzde dismorfik hastalıkların tanısı yüze bakılarak koyulabilmektedir. Bunlar arasında en bilinenlerinden biri de Down Sendromu(Trizomi 21)'dur. Hastalığın genetik bir hastalık ve kromozomal kaynaklı olduğu 1959 yılında Lejeune ve arkadaşları tarafından rapor edilmiştir(Nikawa ve ark.,1977). Down Sendromunun teşhisi laboratuvar ortamında kromozomların incelenmesi ve genetik testlerin uygulanması ile mümkün olmaktadır. Genetik testler kesin sonuç verir fakat pahalıdır.(M.E. Sipahi, 2008). Bu nedenle uzmanlar ilk aşamada, genetik testleri genetik testleri gerçekleştirmeden önce, fenotipteki değişimleri göz önünde bulundurarak teşhis koymaktadırlar. Down Sendromunun fenotiplerinde belirgin özellikler; göz şekilleri, anormal dil büyüklükleri, kafa şekilleri ve ense kalınlıklarıdır ve normal bir insana göre farklıdır. Bu sebeple fotoğraftan tanı koymak da mümkündür.

Literatürde fotoğraf üzerinden bilgisayarlı yüz tanıma sistemleri mevcuttur.(Kurban ve ark.,2017),(Nakada ve ark.2017). Bu tip sistemler teşhis destek amaçlı hasta kayıt sistemlerinden yararlanabilir ve temel faydaları şunlardır:

1. Down sendromu hastalarının uzmanları bir bakışta çocuğun hasta olup olmadığını anlayabilir fakat her zaman, her yerde bir uzman tanı koymak için bulunamayabilir. Uzman yoksa böyle bir sistemle tanı koymak hem daha hızlı hem de daha basit olacaktır.
2. Ortak kullanım alanlarında böyle kişilerin yüzlerinden rahatsızlığı teşhis edilip onların gereksinim ve ihtiyaçlarını karşılamak kolay olacaktır.
3. Teşhis koyma, İnsan Bilgisayar Etkileşimi alanında uygulanabilir ve böylece robotların zihinsel ve gelişimsel bozuklukları olan insanlara karşı daha duyarlı bir tutuma sahip olmalarına yardımcı olur.

Bu çalışmanın amacı, fenotip değişimlerini kullanarak, fotoğraftandown sendromlu kişileri tespit edecek yapay zekaya dayalı bir sınıflandırma ve karar destek yapısı oluşturmaktır. Bu konudaki benzer bir çalışma YSA(Yapay Sinir Ağları)nın MLP(Multi Layer Perceptron, çok katmanlı perseptron) Geri Yayılım Algoritması kullanılarak yapılmıştır.(O.Eroğul ve ark.,2009). Başka bir çalışmada Gabor Wavelet Transform(Gabor Dalgacık Dönüşümü) algoritması ile nitelik çıkarma işlemi yapıp K-NN(K- Nearest Neighbour, K En Yakın Komşu)ve SVM(Support Vector Machine, Destekçi Vektör Makinesi) istatistiksel sınıflandırıcılar kullanılarak Down Sendromunun yüz niteliklerinin geleneksel YSA ile tanınması sağlanmıştır.(Saraydemir ve ark.,2014).

Bu tez çalışmasında, önceki çalışmalardan farklı olarak, gürültü eklenmiş resimler kullanılarak downlu kişilerle normal kişilerin fotoğraflarının sınıflandırılması, Matlab-Alexnet yazılımı ile derin öğrenme yöntemi (Convolutional Neural Network) kullanılarak gerçekleştirilmiştir. Fotoğraf verisetleri internetten açık kaynak lisanslama ile elde edilmiştir. Derin öğrenmenin, yapay sinir ağının gürültülü ve gürültüsüz resimlere karşılık başarı performansı verilmiştir.

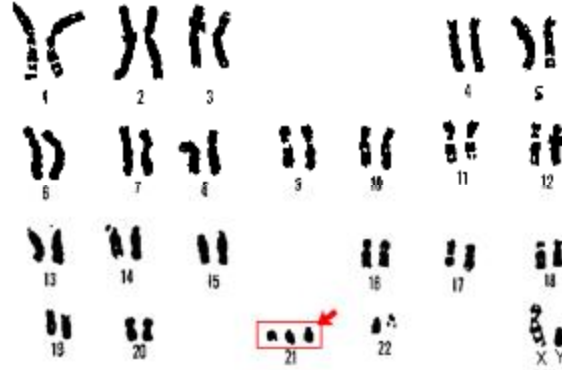
2. LİTERATÜR ARAŞTIRMASI

2.1. Down Sendromu ve Teşhis Yöntemleri

Down sendromu dismorfik bir hastalıktır. Dismorfik hastalıklar gen dizilim bozukluğu ile meydana gelmektedir. Gen dizilimi bozukluğuna ise çeşitli dış etkenler sebep olmaktadır. Örneğin, annenin yaşının büyük olması, akraba evliliği, hamilelik esnasında annenin radyasyon maruz kalması neticesinde genlerin mutasyona uğraması vb. sebepler etkili olmaktadır. Gen dizilimi bozukluğu ise gebelikte bebeğin gelişimi esnasında görünüm bozukluklarına ve mental retardasyona sebep olabilmektedir.

Literatürde Dismorfik bozukluk olarak kabul edilen bir çok sendrom bulunmaktadır. Bunların başlıcaları; Trizomi 18 sendromu (Edwards Sendromu) (Laos ve ark.,2003), Trizomi 13 sendromu (Patau Sendromu) (Lynch ve ark., 1989), Trizomi 8 sendromu (C sendromu) (Rygl ve ark.,2004), Trizomi 22 sendromu (Virtaneva ve ark.,2001), Trizomi 21 Down sendromudur (Mongolizm) (Nikawa ve ark.,1977). Bu çalışma ise trizomi 21, down sendromu veya mongolizm üzerinde durmaktadır.

Down sendromu, trizomi 21 veya mongolizm 21. kromozomun üç tane bulunmasından kaynaklı bir dismorfik bir hastalıktır. Normal insanlarda her kromozom 2 adet bulunur ve 46 adettir. Fakat bu kişilerde 21. kromozomun 3 tane olması kromozom sayılarının 47 tane olmasına neden olmakta ve onların dış görünümünde ciddi farklılıklara neden olmaktadır. Bu dismorfik bozukluk 1866 yılında John Langdon Down tarafından klinik olarak isimlendirilmiştir (Nikawa ve ark.,1977).Yüzlerine ait nitelikler ise normal insanlardan farklı olarak, daha düz burun, daha küçük ağız, dışarıda dil, daha küçük kulaklar, ve çekik gözlere sahiptirler. Boyunları kalındır, elleri küçük, parmakları kısa kısadır. Down sendromlu bebeklerin sıklıkla kas tonusları doğum esnasında azalır. Ortalama bir yetişkinin gelişim seviyesine ve kilosuna ulaşmazlar.

Şekil.2.1. Trizomi 21 kromozom yapısı¹

Doğum öncesi veya doğum sonrası teşhis konabilmektedir. Doğum öncesi kan testleri, ense kalınlığı; doğum sonrası ise yüz şeklinden ellerinden boyun kalınlığından anlaşılır. Down sendromluların elleri küçüktür, parmakları kısadır, ya bir elinde ya da her iki elinde avuç içinde tek çizgi “Simian çizgisi” bulunur. Yaş itibariyle geçen yüzyılda 20’li yaşlara ancak yaşayabilmekte iken günümüz şartlarında bakım standartları ve özeni yüksek olduğundan yaş daha yüksek seviyeye çıkmıştır. Downsendromlu olan kişilerde aynı zamanda “zeka geriliği” de görülmektedir. Down sendromunun getirilerinden biri de orta düzeyli 50–70 IQ (Intelligence Quotient – Zeka derecesi) ile 35–50 IQ arasında değişebilen zeka geriliği olarak görülmektedir(Sipahi, 2008).

Daha önce de belirtildiği gibi sadece yüz özelliklerine bakarak fotoğraftan teşhis etmektir amaç. Kan testleri pahalı, doktorun da bizzati bakarak incelemesi gerekmektedir. Haliyle bu iş uzmanlık gerektirir. Hem hekimi zorluktan kurtarmak, hem de hızlı tanı koymak amacıyla bilgisayardan ve yapay zekadan yararlanılabilir.

2.2. Yapay Zeka Yöntemleri ile Yüz Tanıma Sistemleri

Yapay zeka bir çok öğrenme işlemini gerçekleştirmeye başladığında yüz tanıma için de kullanılabilir mi? sorusuna cevap aranmaya başladı.

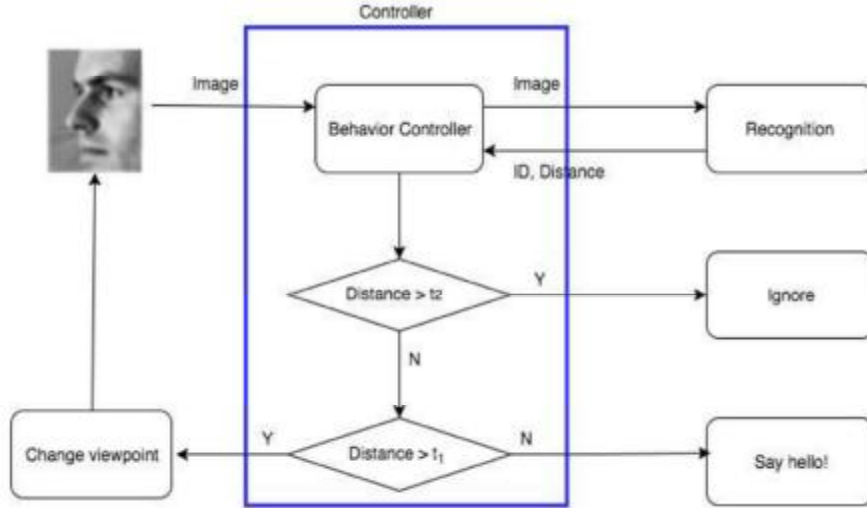
Her insanın yüzünde burun, göz, çene bulunmaktadır. İnsanların yüzlerini tanımda belirleyici unsurlardır. Sadece genişlikleri uzunlukları farklıdır.

Çocuklukla başlayan öğrenme sürecinde insanlar bunu öğrenerek büyümektedir. Yüze ait farklılıklar nitelik çıkartılıp sınıflandırılmak suretiyle yüz tanıma sistemleri oluşturulur. Yatayda ve dikeyde belli sayıda piksellerin toplanması suretiyle nitelik çıkarma işlemine dayanan bu yöntemde bir sonraki aşama sınıflandırmadır. Belirlenen nitelikler sonrasında Bayes sınıflandırıcı ile sınıflandırılmıştır(Brunelli ve ark.,1993).

Yüz tanıma, yüz tanımayı çalışmak için üretilmiş olan birçok veritabanından büyük ölçüde yararlanmıştır. Bu veri tabanlarının çoğu, yüz tanıma probleminde belirli parametrelerin çalışmasını kolaylaştırmak için kontrollü koşullar altında oluşturulmuştur. Bu parametreler, pozisyon, poz, aydınlatma, ifade, arka plan, kamera kalitesi, oklüzyon, yaş ve cinsiyet gibi değişkenleri içerir. Bu çalışmanın yapılabilmesi için insanların doğal pozlarından resimlerinden faydalanılmıştır. Bu çalışmayı yapabilmek için oluşturulan veri tabanı LFW web sayfasında bulunmaktadır(Huang ve ark.,2008). Daha önce kullanılmamış bir eşleşme yöntemi kullanılmıştır. Eğitim setinde 1100 eşleşmiş çift ve 1100 eşleşmemiş çift test setinde ise 500 eşleşmiş 500 eşleşmemiş çift kullanılmıştır(Huang ve ark.,2008).

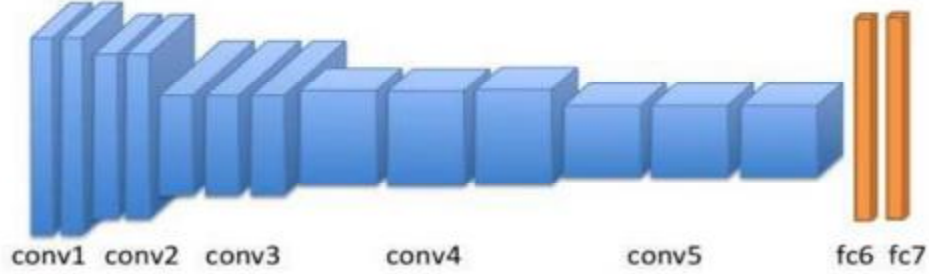
Gerçek dünya, karmaşık görüntülerde yüz tanıma, poz tahmini ve yer işareti tahmini için birleşik bir model sunuluyor. Model, ortak bir parça havuzla sahip ağaçların karışımlarına dayanmaktadır. Her yüz simgesi bir parça olarak modelleniyor ve bakış açısına göre topolojik değişimleri yakalamak için küresel karışımlar kullanılıyor. Ağaç yapılı modellerin küresel elastik deformasyonu yakalamada şaşırtıcı derecede etkili olduğunu ve yoğun grafik yapılarının aksine optimize etmenin kolay olduğu gösterilmektedir. Standart yüz kriterleri ve yeni bir “normal ortamlarında” açıklamalı veri kümesinde kapsamlı sonuçlar sunuyoruz, bu da sistemin, bazen de dikkate değer şekilde, en iyi uygulamalardan biri olan durumun ilerlemesini öneriyor. Model yüzlerce yüzle ustaca eğitilmiş olsa da, milyarlarca örnekle (Google Picasa ve face.com gibi) zorlanan ticari sisteme göre karşılaştırılabilir(Zhu ve ark.,2012).

Yüz tanıma problemini çözmek için ikili ağaç tanıma stratejisine sahip SVM'ler kullanılmaktadır. İfade, poz ve yüzey detaylarında oldukça yüksek değişkenlik derecesine sahip 40 kişiden oluşan 400 görüntüden oluşan Cambridge ORLface veri tabanındaki SVM'lerin potansiyelini gösterilir. Ayrıca tanıma deneyini, 137 bireyden oluşan 1079 görüntüden oluşan daha geniş bir yüz veritabanında sunulmaktadır. SVM temelli tanıma, En Yakın Merkez Sınıflandırma (NCC) kriterlerini kullanarak standart özyüz yaklaşımı ile karşılaştırır. SVM'nin minimum hata oranı %8.79, ki hata oranı %15.14 olan NCC'den çok daha iyidir(Gua ve ark.,2000).



Şekil.2.2. Aktif Yüz Tanıma blok diyagramı(Nakada ve ark.2017)

Nitelik çıkarma işlemi VGG-Face(Parkhi ve ark.,2015) kullanılarak yapılan çalışmanın diyagramı yukarıda anlatılmıştır.



Şekil.2.3. VGG-Face ağ yapısı(Nakada ve ark.,2017)

Nitelik çıkarma işlemi Caffé(Jia ve ark.,2014) Zoo Kütüphanesi kullanılarak yapılmıştır. Resimler PIE(Taigman ve ark.,2014) datasetinden kullanılmıştır. Bu data seti oluşturulurken -90 ve +90 dereceler arasında 13 farklı açıdan kameralarla çekilerek oluşturulmuştur. İnsanlar sosyal hayatlarının içerisinde dışarıda bir tanıdığını gördüğü zamana selam verirler. Tanımyorlarsa tepkisiz olarak yanından geçip giderler. Bu çalışmada amaçlanan işlem budur. Sınıflandırıcı olarak KNN kullanılmış ve %90 başarı elde edilmiştir(Nakada ve ark.,2017).

2.3. Down Sendromlu Yüz Tanıma için Yapay Zeka Uygulamaları

Down sendromlu kişilerin yüzleri diğer insanların yüzlerinden bazı farklılıkları var. 2.1. bölümde anlatılan farklılıklardan yararlanarak nitelik çıkarma işlemi ve sonrasında sınıflandırma yöntemleri kullanılarak down sendromluların yüz tanıma işlemleri sağlanır.

Dismorfik yüz özelliklerinin ayırt edici özelliklerle ayırt edilmesi için çeşitli eğitim veri setleri kullanılarak bir değerlendirme ortaya konmuştur. Gabor Dalgacık Dönüşümü'nü (GWT) özellik çıkarıcı, K-En Yakın Komşu (K-NN) ve Destek Vektör Makineleri (SVM) olarak istatistiksel sınıflandırıcı olarak kullanılmaktadır. Eğitim veri setinin artan boyutuna göre, SVM için çekirdek fonksiyonunu ve K-NN için mesafe ölçümünü seçerek sınıflandırma doğruluğu analiz edildi. Genel sınıflandırma görevinin sonunda, Radial Basis Function (RBF)

kernel tipli GWT-SVM yaklaşımı, eğitim verileri setinde 400 resim ile en iyi sınıflandırma doğruluğu oranını% 97,5 olarak elde etmiştir(Saraydemir ve ark.,2015).

Bir destek vektör makinesi (SVM) ve çok sınıflı SVM tabanlı görüntü sınıflandırıcılarının çekirdek ve parametrelerini optimize etmek için yapısal risk minimizasyonu ve çapraz validasyon uygulanmıştır, böylece genetik anormalliklerin teşhisini mümkün kılar. Sınıfları ayıran hipotezden kalıpların mesafesini eşleştirerek, beklenen riski azaltan yanlış sınıflandırılmış kalıpların yüzdesini reddedilmektedir. SVM'nin diğer son teknoloji ürünü sınıflandırıcılara göre doğru performansı SVM-bazlı genetik sendrom tanısının yararını göstermektedir(David&Lerner,2004).

Kriyofasiyal özellikler, genetik hastalıkları teşhis ederken klinik genetikçiler için son derece bilgilendiricidir. Ultra-nadir gelişimsel hastalıkların yüksek verimli tanılarına yönelik ilk adım olarak, bilgisayar görüşündeki son gelişmeleri uygulayan otomatik bir yaklaşım geliştirildi. Bu algoritma, normal olmayan klinik fotoğraflardan elde edilen fenotipik bilgileri çıkarır ve makine öğrenimini kullanarak, çok boyutlu "Klinik Yüz Fenotip Uzayında" insan yüz dismorfizmelerini modeller. Bu alan hastaları bilinen sendromlar bağlamında bulur ve böylece tanısal hipotezlerin oluşumunu kolaylaştırır. Sonuç olarak, yaklaşım, gelişimsel bozukluklardan şüphelenilen hastalar için olası tanıların arama alanını büyük ölçüde (27,6 kat) daraltarak klinisyenlere yardımcı olacaktır. Dahası, bu Klinik Yüz Fenotip Alanı, bilinen bir sendrom tanısı olmasa bile, hastalığın tanımlanmasına yardımcı olarak, hastaların fenotip ile kümelenmesini sağlar. Bu yaklaşımın, fonksiyonel genetik yol karşılaştırmaları yoluyla klinik sıralama verilerinden nedensel genetik varyantların çıkarılması için yeni bir yöntem sunulduğu görülmektedir(Fery ve ark.,2014).

Down sendromlu hastaların yüz fotoğraflarına görüntü işleme yöntemleri uygulanarak bir ön tanı aracı geliştirilmesi hedeflenmiştir. Çalışmada değerlendirilen veriler; 5-6 yaş grubundaki çocuklardan alınmıştır. Hasta ve normal

denek gruplarının her birinde 18 çocuğa ait fotoğraf kullanılmıştır. Fotoğrafların tamamında elastik yüz demet grafik yöntem kullanılarak yüzdeki kritik noktalar tespit edilmiştir. Daha sonra elde edilen kritik noktalardan klinik ön tanı için 10'ar adet öz nitelik vektörü sağlanmıştır. Öz nitelik vektörleri, yapay sinir ağlarının girişine eğitim ve sınıflandırma için uygulanmıştır. Sonuç olarak, kullanılan yapay sinir ağının başarımı ortalama %68,7 olup bu yöntemle Down sendromlu hastanın klinik ön tanısı konulabilmektedir(Eroğul ve ark.,2009).

Dismorfik hastalıkların en büyük özelliği yüzden anlaşılabilir olmasıdır. Down Sendromu, serebral palsi, otizm spektrum bozukluk, fetal alkol sendromu, zihinsel gerilik, erken yaşlanma gibi hastalıklar yüzden anlaşılabilir. Bu hastalığı olan çocukların hastalıklarının teşhisinin sadece uzmanlar tarafından yapılabilir olması bir sorundur. Bunun için bilgisayar ve yapay zeka uygulamasından yararlanılarak bu hastalıklara sahip çocukların fotoğraflarından yararlanılarak bir sınıflandırma yapılmıştır. Çalışmada hazır eğitilmiş olan Alexnet (Krizhevsky ve ark.,2012) üzerinde başlangıç öğrenme oranı ve epok sayısı değiştirilerek nitelik çıkarma işlemi yapılmıştır. Sınıflandırıcı olarak destek vektör makineleri(SVM) kullanılmıştır. %98.80 oranında başarılı bir sonuç elde edilmiştir(Shukla ve ark.,2017).

Bu tez çalışmasında kullanılan ağ, bir derin öğrenme ağıdır ve önceden kullanılmış YSA modellerinden çok daha hızlı, çok daha fazla veri setiyle çalışabilen, daha yüksek bir doğruluk oranıyla sonuç veren bir yapay sinir ağıdır ve önerilen çalışmayla ilgili çalışmalar mevcuttur(Shukla ve ark.,2017). Önceki çalışmalardan farklı olarak gürültü eklenmiş görüntüler hem eğitimde hem de test süreçlerinde kullanılmıştır.



3. MATERYAL VE METOT

3.1. Giriş

Yapılan çalışmada down sendromlu kişilerin fotoğrafları ile yapay zeka yöntemlerini kullanarak, özellikle Derin Öğrenme Yöntemi (Deep Learning) ile sınıflandırma yapma amaçlanmıştır. Derin öğrenme için, Öğrenme Aktarma (Transfer Learning) yöntemi uygulanmıştır. Veri seti oluşturmak için, toplamda 460 adet resim kullanılmıştır. Bunun içerisinde 20 adet normal 20 adet down sendromlu kişilerin resimleri seçilerek performans testleri için kullanılmıştır. 230 adet resim LFW data setinden kullanılmıştır. 230 adet down sendromlu kişilerin resimleri birkaç internet sitesinden açık kaynak lisansı ile indirilmiştir (pixabay.com, pinterest.com). 230 adet resim matlabda görüntü işlemeyle motion ve salt-pepper filtreleri kullanılarak bozulmuş ve eğitilmiş, sonrasında performansı test edilmiştir. Derin öğrenmede nitelik çıkarma mühendisliği yapılmamaktadır. Nitelik çıkarma işlemi konvolüsyon katmanlarında gerçekleşmektedir. Yani derin öğrenme ağlarında nitelik çıkarma ve sonrasında sınıflandırma yaparak kendi kendine öğrenme sağlanır. Deep learning yapısının avantajı bu olmasına rağmen fazla sayıda resim gerektirmesi bir dezavantajdır. Bazı çalışmaları yapabilmek için veri tabanı oluşturma çalışmaları yürütülmektedir. Örneğin ImageNet Büyük Ölçekli Görsel Tanıma Yarışması (ImageNet Large Scale Visual Recognition Challenge – ILSVRC) yüzlerce obje kategorisi ve milyonlarca resimden obje sınıflandırmaya yönelik yapılan bir faaliyettir. Bu faaliyet ile birlikte büyük veri setleri ve bazı derin öğrenme modelleri oluşturulmuştur. Bu modellerden biri de Alexnet olarak adlandırılan bir derin öğrenme ağ modelidir. Alexnet Alex Krizhevsky ve arkadaşları (2012) ile birlikte oluşturduğu bir Imagenet modelidir. Alexnet bir konvolüsyonel sinir ağıdır. 5 adet konvolüsyon katmanı 3 adet de tamamen bağlı katman bulunmakta ve son katmanda hangi sınıfa ne olasılıkla ait olduğunu hesaplayan softmax katmanı bulunmaktadır.

Derin öğrenmede en büyük sorun veri seti büyüklüğü ve ezberleme sorunudur. Bu sorunlar için data artırma ve yeni bir yöntem olan “dropout” yöntemleri ileri besleme döngüsünde kullanılır. Yapay sinir ağlarında öğrenme sağlanması için her döngüde ne kadar öğrendiğini test etmek için doğrulama yapılır. Hatayı hesaplamak için tahmin edilen ile doğrusu arasındaki farkların karesi alınarak hesaplama yapılır. Bu sebeple de parabolik bir fonksiyon oluşur. Bu fonksiyon maliyet, kayıp ya da hata fonksiyonu olarak adlandırılır ve öğrenme sağlanması için bu fonksiyonun minimuma indirilmesi gerekmektedir. Bu fonksiyon parabolik bir fonksiyon olduğundan yerel minimuma ulaşıldığında hata da minimuma inmiş olur ve bu hatayı azaltma işlemi kısmi türevi alınarak yapılır ve zincir kuralı kullanılarak hata tüm katmanlara yayılır. Fakat bu işlem küçük adımlarla yapılmalıdır. Bu da küçük öğrenme oranı belirlenmesiyle olur. Küçük adımlarla minimuma yavaş yavaş ulaşmaya çalışılır. Öğrenme oranı büyük olduğunda minimuma ulaşması zorlaşır ve sistemin kafasını karıştırarak hatanın büyümesine sebep olur. Veri setinin büyük olması geri yayılım algoritmasını maliyetli hale getirdiğinden deep learning ağlarında olasılıksal kademeli azalma kullanılır ki bu da mini-batch kavramını doğurmuştur(LeCun ve ark.,2015). Alexnet dropout oranı 0.5 öğrenme oranı çalışma esnasında belirlenir. 0.9 momentum değeri 0.0005 ağırlık sönümlleme değeridir. Mini batch size değeri de 128’dir. Bu çalışmada Alexnet modeli kullanılarak sınıflandırma yapılmıştır. Matlab paket programı kullanılmıştır. Bu çalışmayla down sendromlu ve normal kişilerin yüzlerini içeren fotoğrafları iki sınıfa ayırma amaçlanmıştır. Çalışmada sınıflandırmanın başarı oranının yüksek olması hedeflenmektedir.

Derin öğrenme bir yapay sinir ağı çeşididir ve bu şekilde öğrenme sağlanacaktır.

3.2. YSA(Yapay Sinir Ağları)

Yapay Sinir Ağları (YSA) sayısal veriler ile hesaplama yapma, bilgileri saklama, sisteme sunulan örnekleri kullanarak problemi öğrenebilme ve bu sayede daha önce karşısına çıkmayan durumlara çözüm üretebilme özelliklerine sahiptir.

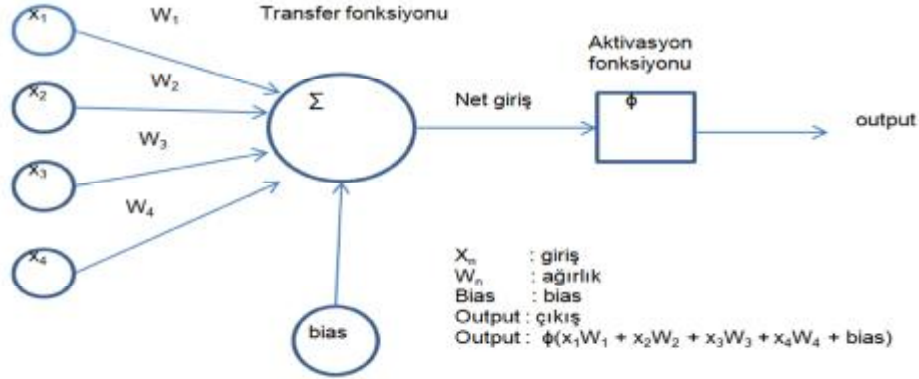
YSA'lar, ağırlıklandırılmış şekilde birbirlerine bağlanmış birçok işlem biriminden (nöron) oluşan matematiksel sistemlerdir. Bir işlem birimi, sık sık transfer fonksiyonu olarak anılan bir denklemdir. Bu işlem birimi, diğer nöronlardan gelen sinyalleri alır; bunları bir araya getirir, dönüştürür ve dijital bir sonuç ortaya çıkarır. Genelde işlem kısımları kabaca gerçek nöronlara(biyolojik) karşılık gelirler ve bir ağ içinde birbirlerine bağlanırlar; bu yapı da sinir ağını oluşturmaktadır. YSA, mühendislik, fen bilimleri, satış vb. alanlarda kullanılmakta; tahmin, sınıflandırma, veri ilişkilendirme, veri yorumlama ve veri filtreleme problemlerinde uygulanmaktadır.

3 tür öğrenme mevcuttur: denetimli öğrenme, denetimsiz öğrenme ve takviyeli öğrenme.

Denetimli öğrenmede işlemle etiketli verilerle yapılır yani doğruluğu kontrol ederek öğrenme vardır. Denetimsiz öğrenmede veriler etiketlenmemiş olarak verilir. Makine kendi kendine öğrenir. Takviyeli öğrenmede ise yine bir öğretici bulunmaktadır.

Yapay sinir ağları biyolojik sinir sisteminden etkilenecek geliştirilmiştir. Biyolojik sinir hücreleri birbirleri ile sinapslar vasıtası ile iletişim kurarlar. Bir sinir hücresi işlediği bilgileri akson'ları yolu ile diğer hücrelere gönderirler. Benzer şekilde yapay sinir hücreleri dışarıdan gelen bilgileri bir toplama fonksiyonu ile toplar ve aktivasyon fonksiyonundan geçirerek çıktıyı üretip ağırlık bağlantılarının üzerinden diğer hücrelere (proses elemanlarına) gönderir. Değişik toplama ve aktivasyon fonksiyonları vardır. Yapay sinir ağlarını birbirlerine bağlayan bağlantıların değerlerine ağırlık değerleri denmektedir. Proses elemanları birbirlerine paralel olarak 3 katman halinde bir araya gelerek bir ağ oluştururlar. Bunlar;

- Girdi katmanı
- Ara katmanlar
- Çıktı katmanı



Şekil.3.1. Yapay sinir ağı topolojisi

Toplama Fonksiyonu: Yapay sinir ağlarında bu işlem her bir girdinin ağırlıklarla çarpılıp toplanmasıyla bir sonraki katmanın girişini belirleyen fonksiyondur.

Aktivasyon Fonksiyonu: Toplama fonksiyonu tarafından hücreye gönderilen net bilgiyi işleyerek hücrenin bu girdiye karşılık üreteceği çıktıyı belirleyen fonksiyondur. Toplama fonksiyonunun sonucu aktivasyon fonksiyonunda işletilip çıkışa iletilir. Aktivasyon fonksiyonunun kullanılma nedeni zamanla ilgili durumlarda toplama fonksiyonunun çıkışının değişimine olanak sağlamaktır. Doğrusal olmayan problemler için de sonuç üretebilmesi için genellikle aktivasyon fonksiyonu doğrusal olmayan bir fonksiyon seçilir. Doğrusal olmayan problemlerin çözümü YSA'nın güçlü özelliklerinden biridir ve bu durum aktivasyon fonksiyonları sayesinde sağlanmaktadır. Doğrusal olmayan problemlerin çözülmesi için doğru aktivasyon fonksiyonu seçilmelidir. Burada dikkat edilmesi gereken kısıt, fonksiyonun türevinin alınabilir olması gerekliliğidir. Geri beslemeli ağlar üzerinde geriye yönelik hesaplama esnasında aktivasyon fonksiyonunun türevi alındığı için hesaplama maliyetinin artmaması adına fonksiyonun türevi kolay hesaplanan aktivasyon fonksiyonu tercih edilmelidir. Çok katmanlı ağ modelinde genellikle aktivasyon fonksiyonu olarak sigmoid fonksiyonu kullanılmaktadır, fakat türevinin daha kolay alınabilmesinden ötürü

derin öğrenme literatüründe ReLU fonksiyonu tercih edilmektedir. Çünkü veri seti normalize edilse de edilmese de daha geç doyuma ulaşması nedeniyle tercih edilir.

Sigmoid:

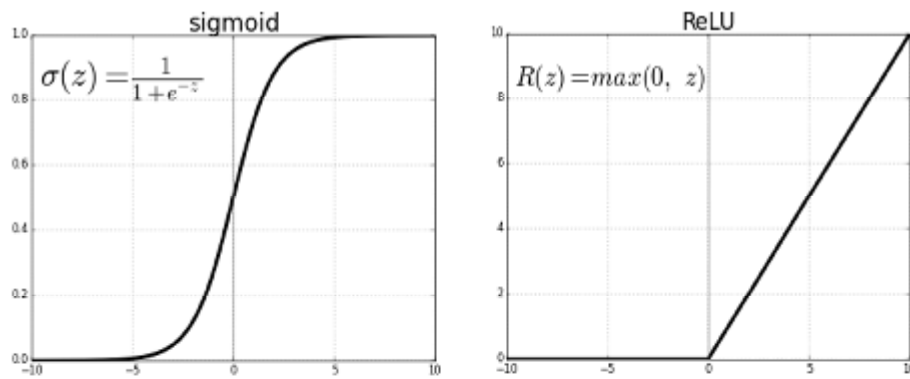
$$F(NEt) = \frac{1}{1 + e^{-NEt}}$$

Bu fonksiyon sürekli ve türevi alınabilir bir fonksiyondur. Doğrusal olmayışı dolayısıyla yapay yapay sinir ağı uygulamalarında en sık kullanılan fonksiyondur. Bu fonksiyon girdi değerlerinin her biri için 0-1 arasında bir değer üretir.

ReLU: Gelen bilgiler sıfırdan küçük ise sıfır, büyük ise fonksiyonun çıktısını çıkış olarak verir.

$$F(NEt) = \begin{cases} NEt, & NEt > 0 \\ 0, & NEt \leq 0 \end{cases}$$

Bu iki fonksiyonun farkı ve ReLU'nun avantajı 3.3.4. bölümde açıklanmıştır.



Şekil.3.2. Sigmoid ve ReLU fonksiyonları

Bilgiler ağa girdi katmanından iletilir. Ara katmanlarda işlenerek oradan çıktı katmanına gönderilirler. Bilgi işlemeden kasıt ağa gelen bilgilerin ağırlık değerleri kullanılarak çıktıya dönüştürülmesidir. Ağırlık girdiler için doğru çıktıları üretebilmesi için ağırlıkların doğru değerlerinin olması gerekmektedir. Doğru ağırlıkların bulunması işlemine ağırlık eğitilmesi denmektedir. Bu değerler başlangıçta rastgele atanırlar. Daha sonra eğitim sırasında her örnek ağa gösterildiğinde ağırlık öğrenme kuralına göre ağırlıklar değiştirilir. Daha sonra başka bir örnek ağa sunularak ağırlıklar yine değiştirilir ve en doğru değerleri bulunmaya çalışılır. Bu işlemler ağ eğitim setindeki örneklerin tamamı için doğru çıktılar üretinceye kadar tekrarlanır. Bu sağlandıktan sonra test setindeki örnekler ağa gösterilir. Eğer ağ test setindeki örneklere doğru cevaplar verirse ağ eğitilmiş kabul edilmektedir. Ağırlıklar belirlendikten sonra her bir ağırlığın ne anlama geldiği bilinmemektedir. O nedenle yapay sinir ağlarına “kara kutu” yakıştırması yapılmaktadır. Ağırlıkların tek tek ne anlama geldikleri bilinmemekle birlikte ağırlık girdiler hakkındaki kararını bu ağırlıkları kullanarak vermesi, ağırlık zekasının bu ağırlıklarda saklandığı söylenebilir. Ağırlık bir olayı öğrenmesi o olay için en doğru yapay sinir ağı modelini seçmekle mümkündür. Şu ana kadar bir çok yapay sinir ağı modeli geliştirilmiştir. Bir yapay sinir ağının modelini şu bilgiler karakterize etmektedir:

- Ağırlık topolojisi
- Kullanılan toplama fonksiyonu
- Kullanılan aktivasyon fonksiyonu
- Öğrenme stratejisi
- Öğrenme kuralı

Yapay sinir ağlarında Öğrenme:

Yapay sinir ağlarında öğrenme işleminde giriş verileri dış dünyadan sisteme verilir. Ardından veriler işlenir ve işlenen veriler aktivasyon fonksiyonel

ulaştıktan sonra sistemin çıkış değeri üretilir. Sistemin çıkışı eğitim setinden girişe karşılık gelmesi gereken beklenen değer ile karşılaştırılarak hata bulunur. Farklı öğrenme algoritmalarıyla hata değeri azaltılarak beklenen çıkışa yaklaşılmaya çalışılır. Bu süreçte güncellenen ağırlıkların ağırlıklarıdır. Ağırlıklar her iterasyonda güncellenerek hedef olan beklenen çıkışa yakınlaşmaya yani hatayı en aza indirme ölçütündeki en az miktarının ne olacağı ise dışarıdan verilen bir değerdir. Eğer yapay sinir ağı verilen giriş-çıkış ile hedefe ulaşmış ise ağırlık değerleri güncellenmez. Bu ağırlıklar tutulur. Ağırlıkların güncellenerek istenilen duruma gelmesine kadar geçen zamana “öğrenme” adı verilir.

Yapay sinir ağı öğrendikten sonra ağı daha önce verilmeyen girişler sunularak sinir ağı test edilir. Ağ daha önce karşılaşmadığı bu örneklerle de iyi sonuç verebiliyorsa ağ problemi öğrenmiş anlamını taşımaktadır. Bu durumu sınamak için probleme ait veriler eğitim seti ve test seti olarak iki gruba ayrılmalıdır. Eğitim seti öğrenmesinin sağlanması için giriş ve çıkışları sisteme sunulan verilerdir. Test seti eğitim setinden farklı olarak sisteme daha önce sunulmamış, çıktısı bilinmesine rağmen sisteme sadece girdilerin sunulduğu verilerdir. Probleme ait verilerin ne kadarının eğitim seti ne kadarının test seti olması gerektiği konusunda belirli bir kural bulunmamaktadır.

Ağın öğrenmesi istenilen probleme ait veriler, problem uzayını kapsayacak şekilde bulunmalıdır. Eğitim setinde bulunan veriler sisteme sırayla gösterilerek ağın problemi öğrenmesi sağlanır. Ağ problemi öğrendikten sonra test setindeki veriler ağı sunularak ağın performansı ölçülür. Hiç karşılaşmadığı örnekler karşısındaki davranışı ağın problemi ne kadar öğrendiğini ortaya koyar. Öğrenilmesi istenen problem için modellenilecek olan ağın topolojik yapısı belirlenir. Ağ için toplama ve aktivasyon fonksiyonları, öğrenme katsayısı, momentum katsayısı gibi parametreler belirlenir. Nöronları birbirine bağlayan ağırlık değerlerinin ataması yapılır. Sistemin problemi öğrenmesi esnasında kullanacak olan ilk değerler genellikle rastgele seçilir ve öğrenme sırasında hesaplanan hata değerine göre bu değerler güncellenir. Ağın öğrenmeye başlaması

ve öğrenme kuralına uygun olarak ağırlıkların güncellenmesi için ağa eğitim setindeki veriler sunulur. Eğitim setinden sunulan girdiler için ileri yönde hesap ile ağın çıktısı hesaplanır. Ağın çıktısı beklenen çıktı ile karşılaştırılarak hata değerleri hesaplanır. Geriye doğru hesaplama ile hesaplanan hatanın azalması için ağırlıkların güncellenmesi gerçekleştirilir. Ağ öğrenmeyi tamamlayıncaya kadar işlemler devam eder.

1950’li yıllarda başlayan yapay zeka çalışmaları gelişerek devam etmektedir. Teknolojinin gelişmesine paralel olarak gelişim gösteren yapay zeka için günümüzün en popüler yapay sinir ağı çeşidi derin öğrenmeli sinir ağları çeşididir. Geleneksel yapay sinir ağlarında ve makine öğrenmesinde görüntü tanıma nitelik çıkarımından sonra sınıflandırmaya dayalı olarak yapılmaktadır. Derin öğrenmede öğrenme ham data ile gerçekleşmektedir. Aşağıdaki şekilde yapay zekanın zamana göre değişimi görülmektedir.

3.3. Derin Öğrenme Yöntemleri

2006 yılında Geoffrey Hinton, derin sinir ağlarının açgözlü katmanlı ön eğitim yöntemi ile etkili bir şekilde eğitilebileceğini göstermiştir(Hinton, Osindero ve ark.,2006). Diğer araştırma grupları, aynı stratejiyi birçok başka derin ağları eğitmek için kullanmıştır(Bengio, Lamblin ve ark. 2007, Ranzato, Poultney ve ark. 2007). Daha iyi performans sergileyen sinir ağlarını tasarlayanın yolu daha derin ağların kurulması gerektiği ve derinliklerin teorik önemine dikkat çekilmesi için “Derin Öğrenme” teriminin kullanılması yaygınlaştırılmıştır(Bengio and LeCun 2007, Delalleau and Bengio 2011, Montufar 2014, Pascanu, Gülçehre ve ark. 2014).

Son zamanlarda oldukça popüler hale gelen deep learning(derin öğrenme) hayatımızın birçok alanında kullanılmaktadır. E-mail filtreleme, arama motorlarında aranan eşleştirme, akıllı telefonlarda, sosyal medyada, e-ticaret gibi alanlar bunlara yazılabilir(LeCun ve ark.,2015). Akademik çalışmalar bu alanlarda kullanımı için öncü olmuştur. Deep learning yüz tanıma, nesne tanıma, nesne

detekte etme, text ayırt etme, ses tanıma amaçlı da kullanılır. Derin öğrenme bir yapay sinir ağı çeşididir ve çok katmanlı yapıdadır. Ne kadar katman artırılırsa derinlik o kadar artar ve dolayısıyla doğru öğrenme oranı artar. Derin konvolüsyonel ağlar görüntü, video, konuşma ve ses işlemede başarılı olarak kullanılıyorken, tekrarlayan sinir ağları text ve konuşma gibi ardışık verilerde kullanılır.

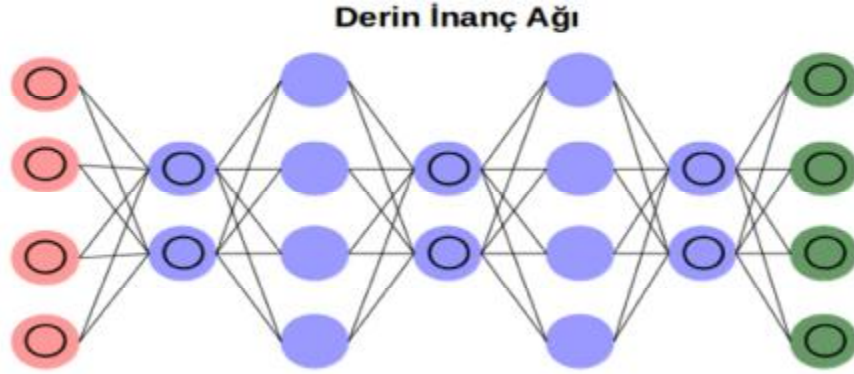
Derin öğrenme, 2010'lu yıllarda kullanılmaya başlanmış, büyük veri denizi ile tek bir katmanda değil, birçok katmanda makine öğreniminde kullanılan hesapları tek bir seferde yapan, makine öğreniminde tanımlanması gereken parametreleri bile kendisi keşfeden, belki de daha iyi parametreler ile değerlendirmelerde bulunabilen bir sistemdir. Derin öğrenme yapay sinir ağları beynin yapısına ve şekline benzer fonksiyonlarından yararlanarak oluşturulmuş algoritmalarıdır. Bu algoritmaları şöyle sıralayabiliriz:

- Derin İnanç Ağları (DBN)
- Konvolüsyonel Sinir Ağları(CNN)

Makine öğrenmesinde, Derin İnanç Ağı (Deep Belief Networks), üretken grafiksel bir modeldir ya da alternatif olarak, gizli düğümlerin birden fazla katmanından oluşan, katmanlar arasında bağlantı olan fakat düğümler arasında olmayan bir derin sinir ağı sınıfıdır.

Denetimsiz bir dizi örnek üzerinde eğitim gördüğünde, DBN, girişlerini olasılıksal olarak yeniden yapılandırmayı öğrenebilir. Katmanlar daha sonra özellik dedektörleri gibi davranırlar. Bu öğrenme aşamasından sonra, sınıflandırma yapmak için bir DBN daha fazla denetimle eğitilebilir.

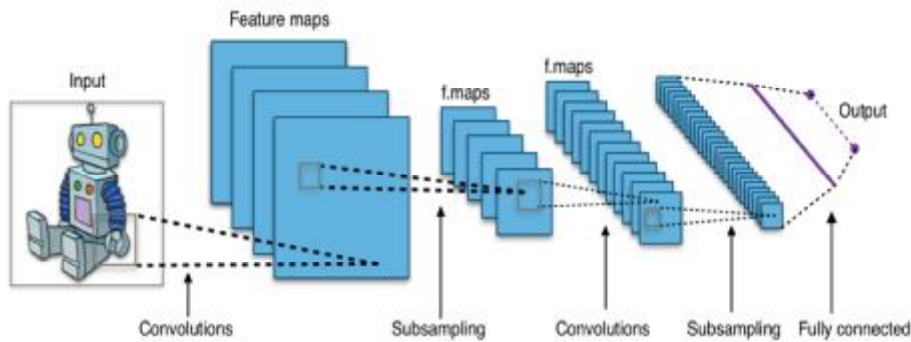
DBN'ler, her bir alt ağı gizli katmanı, bir sonraki katmanın görünür katmanı olarak işlev gören, kısıtlı Boltzmann makineleri (RBM) veya otomatik kodlayıcılar gibi basit, denetimsiz ağların bir bileşimi olarak görülebilir.



Şekil.3.3. Derin İnanç Ağları Ağ Yapısı

3.3.1. Konvolüsyonel Sinir Ağları

İsminden de anlaşılacağı üzere matematiksel işlem olarak konvolüsyon kullanılır. Konvolüsyon lineer işlemlerin özel bir türüdür. Konvolüsyonel sinir ağları ise en az bir katmanında konvolüsyon işlemi yapılan bir sinir ağı türüdür. Fakat buradaki konvolüsyon işlemi normal ya da mühendislik matematiğindeki konvolüsyon işleminden farklıdır. Konvolüsyonel sinir ağları, konvolüsyon katmanı, ReLU katmanı, Pooling(havuzlama) katmanı, normalizasyon katmanı, tamamen bağlı katmanlar ve softmax katmanı bulunmaktadır. Resimlerden nitelik çıkarma işlemi konvolüsyon katmanlarında, sınıflandırma işlemi tamamen bağlı katmanlarda, ne kadar olasılıkla eşleştiği ise softmax katmanında gerçekleşir.

Şekil.3.4. Örnek Bir CNN Yapısı²

3.3.1. Konvolüsyon

En genel haliyle, konvolüsyon, gerçek değerli iki fonksiyon üzerinde gerçekleşen bir işlemdir. Konvolüsyonun tanımını anlatmak için, kullanabileceğimiz iki fonksiyonun örnekleri ile başlarız.

Örneğin lazerle bir uzay mekiğinin yeri izleniyor. Lazer sensörü uzay mekiğinin t zamanında yeri olan basit $X(t)$ çıkışını üretir. Burada X ve t gerçek değerlerdir, örneğin, herhangi t anlık bir zamanda alınan farklı bir değerdir.

Ayrıca bu sensör biraz gürültülü. Daha az gürültülü tahmin yürütmek için birkaç ölçümü birlikte alarak ortalamasını almak isteniyor. Doğal olarak son ölçümler daha yakındır, böylelikle son ölçümlere daha fazla ağırlık veren ağırlıklandırılmış ortalama olmasını istenmektedir. Bunu ağırlıklandırma fonksiyonu olan $w(a)$ ile yapılabilir ki a bir ölçüm dönemidir. Her an bu şekilde ağırlıklandırılmış ortalama işlem uygulanırsa uzay mekiğinin pozisyonunu daha düzgün tahmin etmeyi sağlayan yeni bir fonksiyon elde edilir:

$$S(t) = \int X(a)W(t - a)da$$

Yukarıdaki işlem konvolüsyondur ve yıldızla da şöyle gösterilir:

$$S(t) = (X * W)(t)$$

Konvolüsyonel ağ terminolojisinde ilk argüman (bu örnekte X fonksiyonu) konvolüsyona giriş olarak isimlendirilir ve ikinci argüman (bu örnekte W fonksiyonu) **kernel** olarak adlandırılır. Çıkış ise **nitelik haritası** olarak adlandırılır. Yukarıdaki örnekte ölçüm aralıksız olarak yapılır fakat bu gerçekçi değildir. Bilgisayar üzerinde çalışırken zaman ayrıştırılmıştır. Gerçekçi ölçüm üzerine düşünersek saniye başına bir ölçüm alınarak işlem yapılır. Burada t zaman indisidir ve tam sayıdır dolayısıyla X ve W de tamsayı ifade eder.

$$S(t) = (X*W)(t) = \sum_{a=-\infty}^{\infty} X(a)W(t-a)$$

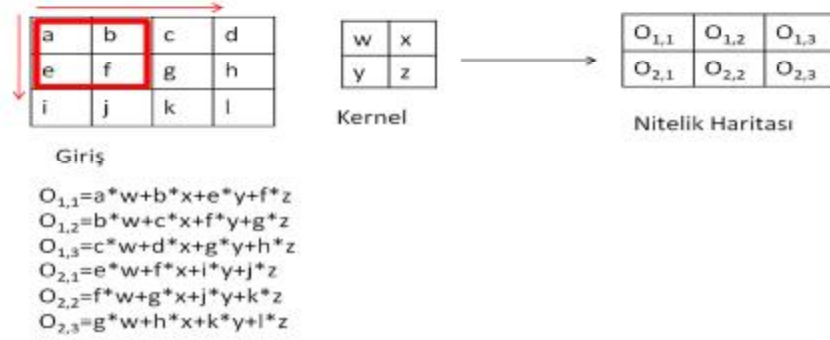
Makine öğrenmesi uygulamalarında giriş fonksiyonu çok boyutlu bir dizi seti ve kernel fonksiyonu da bir sürü parametreden oluşan çok boyutlu bir diziden oluşur. Sıklıkla bir seferde birden fazla eksende konvolüsyon işlemi yapılır. Yani giriş olarak iki boyutlu bir resimse kernel de iki boyutlu bir matris olur.

$$S(i,j) = (I*K)(i,j) = \sum_m \sum_n I(i-m, j-n)K(m,n)$$

Yukarıdaki denklem kerneli girişe göre kaydırmak demektir. Bu da konvolüsyonun değişmez özelliğini artırır. Fakat bu özellik makine öğrenmesi kütüphaneleri için çok önemli değildir. Bunun yerine bir çok makine öğrenmesi kütüphanesi kerneli ters çevirmeden işlem yapar ki bunun adı konvolüsyonla ilişkili olan “çapraz korelasyon” dur. Fakat konvolüsyona benzediği için konvolüsyonel sinir ağı olarak anılmaktadır:

$$S(i,j) = (I*K)(i,j) = \sum_m \sum_n I(i+m, j+n)K(m,n)$$

Ayrık konvolüsyon bir matris çarpımı olarak görülür. Tipik konvolüsyonel sinir ağlar büyük girişlerle etkili bir şekilde başa çıkmak için daha ileri uzmanlıklardan yararlanır. Aşağıdaki şekilde konvolüsyonel sinir ağlarında işlemin nasıl gerçekleştiğinin şekli bulunmaktadır:



Şekil.3.5. Konvolüsyon matris işlem gösterimi

Konvolüsyon, bir makine öğrenim sisteminin geliştirilmesine yardımcı olabilecek üç önemli fikirden yararlanır: seyrek etkileşimler, parametre paylaşımı ve eşdeğişkenli gösterimler. Ayrıca, konvolüsyon, değişken büyüklükteki girdilerle çalışmak için bir araç sağlar. Geleneksel sinir ağı katmanları, her bir giriş birimi ve her bir çıkış birimi arasındaki etkileşimi açıklayan ayrı bir parametreye sahip bir matris parametresi ile matris çarpmasını kullanır. Bu, her çıkış biriminin her giriş birimi ile etkileştiği anlamına gelir. Fakat konvolüsyonel ağlar, tipik olarak seyrek etkileşimlere sahiptirler (ayrıca seyrek bağlantı veya seyrek ağırlıklar olarak da adlandırılır). Bu, kerneli girişten daha küçük yaparak gerçekleştirilir. Her bir konvolüsyon işleminden sonra çıkan sonuçların piksel sayısı azaldığı için eğer kenarlarda gözden kaçırılmaması gereken bir nitelik varsa satırların ve sütunların sonuna sıfır eklenerek kenar ve köşe nitelikleri de korunmuş olur. Bu işleme padding(doldurma) işlemi denmektedir.

Örneğin, bir görüntüyü işlerken, giriş görüntüsünde binlerce veya milyonlarca piksel olabilir, ancak yalnızca on veya yüzlerce pikselden oluşan kernelli kenarlar gibi küçük ve anlamlı özellikleri tespit edilebilir. Bu, hem modelin bellek gereksinimlerini azaltan hem de istatistiksel verimliliğini artıran daha az parametre kaydetmemiz gerektiği anlamına gelir. Ayrıca çıktığı

hesaplamanın daha az işlem gerektirdiği anlamına gelir. Verimlilikteki bu gelişmeler genellikle oldukça büyüktür.

Parametre paylaşımı, bir modelde birden fazla işlev için aynı parametrenin kullanılmasını ifade eder. Geleneksel bir sinir ağında, ağırlık matrisinin her bir elemanı, bir katmanın çıktısını hesaplarken tam olarak bir kez kullanılır. Bu, girdinin bir ögesi ile çarpılır ve daha sonra yeniden gözden geçirilmez. Parametre paylaşımı için eşanlamlı olarak, bir ağırlık ağırlıkları bağladığı söylenebilir, çünkü bir girdiye uygulanan ağırlığın değeri başka bir yerde uygulanan ağırlığın değerine bağlıdır. Bir konvolüsyonel sinir ağında, çekirdeğin her bir üyesi girdinin her pozisyonunda kullanılır (sınır ile ilgili tasarım kararlarına bağlı olarak, sınır piksellerinin bir kısmı hariç).

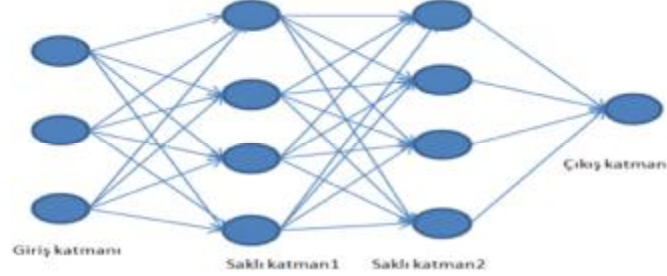
Konvolüsyon işlemi tarafından kullanılan parametre paylaşımı, her konum için ayrı bir parametre kümesini öğrenmek yerine, yalnızca bir set öğrenileceği anlamına gelir.

Bir ConvNet'in her katmanı, farklılaşabilir bir fonksiyon vasıtasıyla bir hacim aktivasyonunu diğerine dönüştürür. ConvNet mimarilerini oluşturmak için üç ana katman türü kullanılır. Konvolüsyonel Katman, Havuz Katmanı ve Tam Bağlı Katman (tam olarak Normal Sinir Ağlarında görüldüğü gibi). Bu katmanları tam bir ConvNet mimarisi oluşturmak için toplanır. Resimlerin üç boyutlu olduğu düşünüldüğünde $H \times W \times D$ boyutunda şeklinde adlandıırırsak $K \times K$ ise kernel boyutu denilirse konvolüsyon çıktısının kaç piksel olacağı şu şekilde hesaplanır:

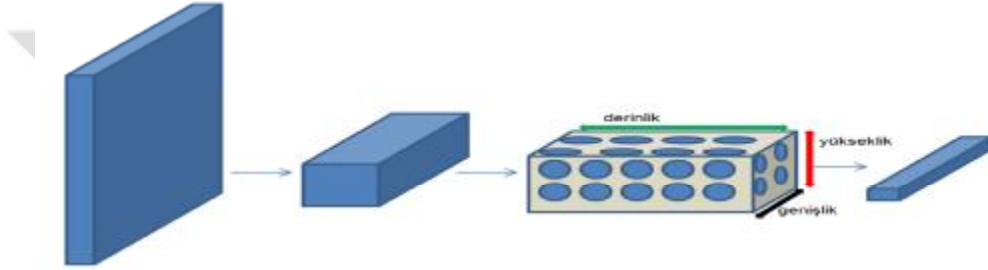
$$\frac{H + 2P - K}{S} + 1$$

P : Padding (satırların ve sütunların sonuna sıfır ekleme)

S : Stride(adım belirtir, kernel kayarken kaç adımla kayacağı belirlenir).



Şekil.3.6. Düzenli 3 katmanlı bir Sinir Ağı.



Şekil.3.7. Bir ConvNet, nöronlarını katmanlardan birinde görselleştirildiği gibi üç boyutta (genişlik, yükseklik, derinlik) düzenler. Bir ConvNet'in her katmanı, 3D giriş hacmini 3 boyutlu çıkış hacmindeki nöron aktivasyonlarına dönüştürür. Bu örnekte, siyah giriş katmanı görüntüyü tutar, böylece genişliği ve yüksekliği görüntünün boyutları olur ve derinlik 3 olur (Kırmızı, Yeşil, Siyah kanallar).

3.3.2. Normalizasyon

Kabaca normalleştirme demektir. Yapay sinir ağlarında verilerin büyüklüğü önemlidir. Veriler büyüdükçe kapladıkları hafıza artar ve bu da yapay sinir ağının hem verimini düşürür hem de çalışma hızını azaltır. Tüm veri seti değerini 0-1 arasına sıkıştırarak işlemler kolay hale getirilir. Bu işlemi tüm veri setinin ortalamasını alarak her bir veriden çıkarır ve böylelikle veriler 0-1 aralığına alınmış olur. Standardizasyonun (veya Z Puanı Normalizasyonunun) sonucu, özelliklerin standart bir normal dağılım özelliklerine sahip olacak şekilde yeniden ölçeklendirilmesidir.

$$\mu = 0, \sigma = 1$$

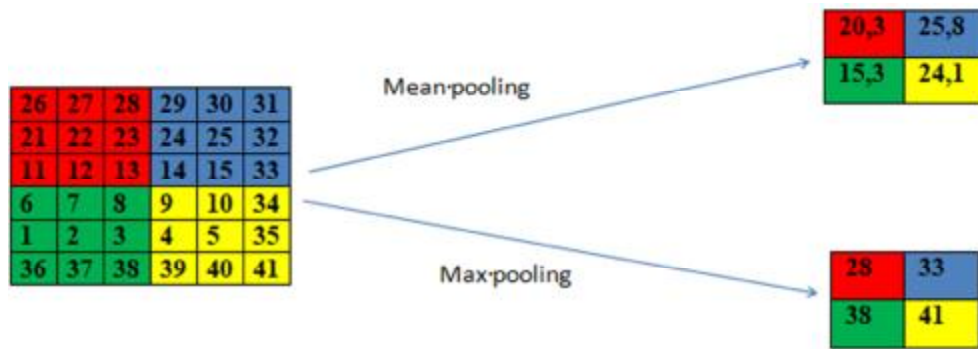
Burada μ ortalama (ortalama) ve σ ortalamadan standart sapma; Örneklerin standart puanları (z puanları) şu şekilde hesaplanır:

$$z = \frac{X - \mu}{\sigma}$$

Özelliklerin standart sapması 1 ile 0 civarında ortalananak şekilde standartlaştırılması, farklı birimleri olan ölçümleri karşılaştırdığımızda değil, aynı zamanda birçok makine öğrenme algoritması için genel bir gereklilik olarak da önemlidir.

3.3.3. Pooling

Bir havuzlama fonksiyonu, ağın çıkışı, yakındaki çıkışların bir özet istatistiğine sahip belirli bir konumda değiştirir. Örneğin, maksimum havuzlama(max-pooling) işlemi dörtgen mahaldeki en büyüğünü çıktı olarak(Zhou ve ark.,1988) verir. Diğer popüler havuzlama fonksiyonları; ortalama(mean-pooling) ve minimum havuzlama fonksiyonlarıdır.



Şekil.3.8. max-pooling ve mean-pooling

Yukarıdaki şekilde 3*3 maxpooling ve mean pooling işlemleri gösterilmektedir. Havuzlama işlemi belirli bölgelerden alınan çıktılarından ibarettir. Örneğin, bir yüz resmi seçildiğinde. Yüzde belirgin olan gözler, burun ve dudaklardır. Burada en belirgin olan ise gözlerdir. Buradaki öznelilikler yüzün diğer kısımlarının özneliliklerine göre farklıdır. Buradaki nitelikleri seçmek için pooling(havuzlama) işlemi kullanılır. Bu işlem sırasında matris boyutları küçülür.

Bir sonraki katmandaki parametrelerin sayısı giriş boyutunun bir fonksiyonu olduğunda (bir sonraki katmanın tam olarak bağlandığı ve matris çarpımına dayalı olduğu gibi), giriş büyüklüğündeki bu azalma, aynı zamanda istatistiksel verimliliğin artmasını ve parametreleri saklamak için bellek gereksinimlerinin azalmasını sağlar. Pooling çıktısının piksel sayısı şu şekilde hesaplanır:

$$\frac{H + 2P - T}{S} + 1$$

T : pooling boyutu

3.3.4. ReLU

Rectified Linear Unit(Düzleştirilmiş Doğrusal Birim) bir aktivasyon fonksiyonu çeşididir. Rectified Linear Unit, son zamanlarda popüler hale geldi. $F(x) = \max(0, x)$ işlevini hesaplar. Diğer bir deyişle, aktivasyon sıfıra eşit olarak eşiklenir. ReLU kullanımının çeşitleri artıları ve eksileri vardır.

Artıları:

1. Sigmoid/tanh fonksiyonlarına kıyasla stokastik gradyan inişinin yakınsamasını büyük ölçüde hızlandırdığı bulunmuştur(Krizhevsky ve ark.,2009). Bunun doğrusal, doyurucu olmayan formundan kaynaklandığı iddia edilmektedir.

2. Maliyetli işlemler içeren nöronlar tanh/sigmoid ile karşılaştırıldığında, relu basitçe bir aktivasyon matrisini sıfıra eşikleyerek uygulanabilir.

Eksi yönleri: Ne yazık ki, ReLU birimleri eğitim sırasında kırılabilir ve “ölebilir”. Örneğin, bir ReLU nöronu boyunca akan büyük bir gradyan, ağırlıkların, nöronun herhangi bir veri noktasında tekrar aktif olmayacağı şekilde güncellenmesine neden olabilir. Bu olursa, birim boyunca akan gradyan, o noktadan sonsuza kadar sıfır olacaktır. Yani, ReLU birimleri eğitim sırasında geri dönülemez bir şekilde öldürebilirler çünkü veri çoğaltma devreden çıkarılabilir. Örneğin, öğrenme oranı çok yüksekse, ağırlıkların % 40'ının “ölü” olabileceği (yani tüm eğitim veri setinde asla aktif olmayan nöronlar) görülebilir. Öğrenme oranının uygun bir şekilde ayarlanmasıyla, bu daha az sıklıkla görülen bir durumdur.

3.3.5. Dropout

Tam bağlı (fully connected) katmanlarda belli eşik değerin altındaki düğümlerin seyreltilmesinin başarıyı arttırdığı gözlenmiştir. Yani zayıf bilgilerin unutulması öğrenimi arttırdığı gözlenmiştir.

Seyreltme (Dropout) değerinin bazı özellikleri aşağıdaki gibidir;

- Seyreltme (Dropout) değeri olarak genelde 0,5 kullanılmaktadır. Farklı kullanıldığı durumlar da yaygındır. Probleme ve veri setine göre değişiklik göstermektedir.
- Seyreltme (Dropout) için rastgele eleme yöntemi de kullanılabilir.
- Seyreltme (Dropout) değeri eşik değeri olarak kullanıldığında [0,1] aralığında bir değer olarak tanımlanmaktadır.
- Tüm katmanlarda aynı dropout değeri kullanılması zorunlu değildir; farklı seyreltme (Dropout) değerleri de kullanılabilir.

3.3.6. Softmax

Softmax fonksiyonu bir çeşit sınıflandırıcıdır. Lojistik regresyon isimli sınıflandırıcı ikili sınıflandırma yaparken softmax fonksiyonu lojistik regresyonun çok sınıflıdır.

$\frac{1}{\sum_j e^{f_j}}$ terimi dağılımı normalleştirir. Yani değerlerin toplamlarını 1'e eşitler. Dolayısıyla bu işlem sonucunda hangi sınıfa ne kadar olasılıkla ait olduğunu hesaplamaktadır.

Bir test girişi x verildiğinde, aktivasyon fonksiyonu sonucu $j = 1, \dots, k$ her bir değeri için $p(y=j \mid x)$ olasılığını tahmin etmesini ister. Örneğin, sınıf etiketinin, k farklı olası değerlerin her birine sahip olma ihtimalini tahmin etmek isteniyor. Böylece, aktivasyon fonksiyonu sonucu bize k tahmini olasılıklarımızı veren bir k boyutlu vektörü (elemanların toplamı 1) üretir.

Öğrenmenin gerçekleşmesi için hata değerinin hesaplanması gerekir ve softmax fonksiyonu için hata değeri softmaxloss fonksiyonu ile hesaplanır.

Softmax sınıflandırıcısında, $f(x_i; W) = Wx_i$ fonksiyon eşlemesi değişmeden kalır, fakat şimdi bu skorları her bir sınıf için normalize edilmemiş log olasılıkları olarak yorumluyoruz ve aşağıdaki form olan çapraz entropi kaybı kullanılır(github.com):

$$L_i = -\log\left(\frac{e^{f_{yi}}}{\sum_j e^{f_j}}\right) \quad \text{ya da} \quad L_i = -f_{yi} + \log\left(\sum_j e^{f_j}\right)$$

Yapılan çalışmada ise eğitim esnasında kayıp fonksiyonu olarak L2 norm dediğimiz farkların karesi kullanılmıştır. Aşağıdaki örnek geri yayılım çözümlemesinde kullanılan yöntem de bu şekildedir.

Konvolüsyonel sinir ağlarında da öğrenme işlemi geri yayılım algoritması ile gerçekleşir. Örnek bir konvolüsyonel sinir ağında geri yayılım algoritması uygulaması yapılırsa işlem adımları ve ağırlık güncellemeleri şu şekilde gerçekleşir:

Başlangıç parametreleri:

- C1 katmanı, $k_{1,p}^1$ (size 5×5) ve b_1^1 (boyut 1×1), $p = 1, 2, \dots, 6$
- C2 katmanı, $k_{p,q}^2$ (size 5×5) ve b_q^2 (boyut 1×1), $q = 1, 2, \dots, 12$
- FC katmanı, W (size 10×192) ve b (boyut 10×1)

Konvolüsyon Katmanı, aktivasyon fonksiyonu olarak sigmoid fonksiyon kullanıldığını kabul edersek;

$$C_p^1 = \sigma(I * k_{1,p}^1 + b_p^1), \sigma = \frac{1}{1+e^{-x}}$$

$$C_p^1(i,j) = \left(\sum_{u=-2}^2 \sum_{v=-2}^2 I(i-u, j-v) \cdot k_{1,p}^1(u,v) + b_p^1 \right)$$

Burada $p = 1, 2, 6$ olduğu için, C1 katmanında 6 özellik haritası vardır, “*” konvolüsyon anlamındadır ve i, j, özellik haritasının satır ve sütun indisleridir. Sadece o kısımlarını tutmak sıfır doldurulmuş kenarlar olmadan hesaplanan konvolüsyon, C_p^1 boyutu 24×24 , oldukça 28×28 gibi

Pooling Katmanı S1:

$$S_p^1(i,j) = \frac{1}{4} \left(\sum_{u=0}^1 \sum_{v=0}^1 C_p^1(2i-u, 2j-v) \right) i, j = 1, 2, \dots, 12$$

Konvolüsyon Katmanı C2:

$$C_q^2 = \sigma(I * S_{1,p}^1 * k_{p,q}^2 + b_q^2)$$

$$C_q^2(i,j) = \left(\sum_{p=1}^6 \sum_{u=-2}^2 \sum_{v=-2}^2 S_p^1(i-u, j-v) \cdot k_{p,q}^2(u,v) + b_q^2 \right), i, j = 1, 2, \dots, 12$$

Burada $q = 1, 2, \dots, 12$ çünkü C2 katmanında 12 özellik haritası vardır. Sadece sıfır-yastıklı kenarlar olmadan hesaplanan konvolüsyonun parçalarını koruyarak, S'nin boyutu 12×12 yerine S_p^1 'nin boyutu 8×8 'dir.

Pooling Katmanı S2:

$$S_q^2(i, j) = \frac{1}{4} \left(\sum_{u=0}^1 \sum_{v=0}^1 C_q^2(2i - u, 2j - v) \right) \quad i, j = 1, 2, \dots, 4$$

Vektörizasyon ve Sıralama:

Her S_q^2 bir 4×4 matristir ve S2 katmanında 12 adet matris mevcuttur. İlk olarak, her S_q^2 sütun taraması ile vektör haline getirilir, daha sonra $4 \times 4 \times 12 = 192$ uzunluğunda uzun bir vektör oluşturmak için bir araya getirilir. Bunun ifadesi ise:

$$f = F(\{S_q^2\}_{q=1, 2, \dots, 12})$$

ve tersinin süreci,

$$\{S_q^2\}_{q=1, 2, \dots, 12} = F^{-1}(f)$$

Tamamen Bağlı Katman(Fully Connected Layer) FC:

$$y' = \sigma(W * f + b)$$

Kayıp Fonksiyonu(Loss Function) :

Doğru etiketi y kabul edelim, loss fonksiyonu

$$L = \frac{1}{2} \sum_{i=1}^{10} (y'(i) - y(i))^2$$

Geri Yayılım: Geri yayılımda parametreler sondan başa doğru güncellenir. Yani $W, b, k_{p,q}^2, b_q^2, k_{1,p}^1, b_p^1$.

ΔW (size 10×192) :

$$\begin{aligned}\Delta W(i, j) &= \frac{\partial L}{\partial W(i, j)} \\ &= \frac{\partial L}{\partial y'(i)} \cdot \frac{\partial y'(i)}{\partial W(i, j)} \\ &= (y'(i) - y(i)) \cdot \frac{\partial}{\partial W(i, j)} \sigma(\sum_{j=1}^{192} W(i, j) * f(j) + b(i)) \\ &= (y'(i) - y(i)) \cdot y'(i) \cdot (1 - y'(i)) \cdot f(j) \\ \Delta y'(i) &= (y'(i) - y(i)) \cdot y'(i) \cdot (1 - y'(i)) \text{ boyutları } 10 \times 1 \text{ olan} \\ \Delta W(i, j) &= \Delta y'(i) \cdot f(j) \\ \Delta W &= \Delta y' \times f^T\end{aligned}$$

Δb (size 10×1) :

$$\begin{aligned}\Delta W(i, j) &= \frac{\partial L}{\partial b(i, j)} \\ &= \frac{\partial L}{\partial y'(i)} \cdot \frac{\partial y'(i)}{\partial b(i)} \\ &= (y'(i) - y(i)) \cdot \frac{\partial}{\partial b(i)} \sigma(\sum_{j=1}^{192} W(i, j) * f(j) + b(i)) \\ &= (\hat{y}(i) - y(i)) \cdot y'(i)(1 - y'(i)) \\ \Delta b &= \Delta y'\end{aligned}$$

$\Delta k_{p,q}^2$ (size 5×5) :

$$\begin{aligned}\Delta f(j) &= \frac{\partial L}{\partial f} \\ &= \sum_{i=1}^{10} \frac{\partial L}{\partial y'(i)} \cdot \frac{\partial y'(i)}{\partial f(j)} \\ &= (y'(i) - y(i)) \cdot \frac{\partial}{\partial f(j)} \sigma(\sum_{j=1}^{192} W(i, j) * f(j) + b(i)) \\ &= \sum_{i=1}^{10} (y'(i) - y(i)) \cdot y'(i)(1 - y'(i)) \cdot W(i, j) \\ &= \sum_{i=1}^{10} \Delta y'(i) \cdot W(i, j) \\ \Delta f &= W^T \times \Delta y'\end{aligned}$$

Uzun hata vektörünü Δf (boyut 192×1) ile yeniden şekillendiriyoruz:

$$\{\Delta S_q^2\}_{q=1,2,\dots,12} = F^{-1}(\Delta f)$$

S2 katmanı üzerinde hata alır ($12 \times 4 \times 4$ hata haritaları). S2 katmanında herhangi bir parametre olmadığından, herhangi bir türev işlemi yapmamız gerekmemektedir. Daha sonra, C2 katmanındaki hatayı elde etmek için örnekleme yapılır.

$$\begin{aligned} \Delta C_q^2(i, j) &= \frac{1}{4} \Delta S_q^2([i/2], [j/2]), i, j = 1, 2, \dots, 8 \\ \Delta k_{p,q}^2(u, v) &= \frac{\partial L}{\partial k_{p,q}^2(u, v)} \\ &= \left(\sum_{i=1}^8 \sum_{j=1}^8 \frac{\partial L}{\partial C_{p,q}^2(i, j)} \frac{\partial C_{p,q}^2(i, j)}{\partial k_{p,q}^2(u, v)} \right) \\ &= \sum_{i=1}^8 \sum_{j=1}^8 \Delta C_{p,q}^2(i, j) \cdot \frac{\partial}{\partial k_{p,q}^2(u, v)} \sigma \left(\sum_{p=1}^6 \sum_{u=-2}^2 \sum_{v=-2}^2 S_p^1(i - u, j - v) \cdot k_{p,q}^2(u, v) + b_q^2 \right) \\ &= \sum_{i=1}^8 \sum_{j=1}^8 \Delta C_{p,q}^2(i, j) \cdot C_{p,q}^2(i, j) \cdot \left(1 - C_{p,q}^2(i, j) \right) \cdot S_p^1(i - u, j - v) \\ \Delta C_{q,\sigma}^2(i, j) &= \Delta C_{p,q}^2(i, j) \cdot C_{p,q}^2(i, j) \cdot \left(1 - C_{p,q}^2(i, j) \right) \\ \text{aslında C2 katmanında sigmoid fonksiyonundan önceki hatadır. Bu nedenle,} \\ \Delta C_{q,\sigma}^2(i, j) &= \sum_{p=1}^6 \sum_{u=-2}^2 \sum_{v=-2}^2 S_p^1(i - u, j - v) \cdot k_{p,q}^2(u, v) + b_q^2 \\ \Delta k_{p,q}^2(u, v) &= \sum_{i=1}^8 \sum_{j=1}^8 S_{p,rot180}^1(u - i, v - j) \cdot \Delta C_{q,\sigma}^2(i, j) \\ \Delta k_{p,q}^2(u, v) &= S_{p,rot180}^1 * \Delta C_{q,\sigma}^2, \\ \Delta b_q^2 &(\text{boyut } 1 \times 1) : \\ \Delta b_q^2(u, v) &= \frac{\partial L}{\partial b_q^2} \\ &= \left(\sum_{i=1}^8 \sum_{j=1}^8 \frac{\partial L}{\partial C_{p,q}^2(i, j)} \frac{\partial C_{p,q}^2(i, j)}{\partial b_q^2(u, v)} \right) \end{aligned}$$

$$= \sum_{i=1}^8 \sum_{j=1}^8 \Delta C_{p,q}^2(i,j) \cdot \frac{\partial}{\partial b_q^2(u,v)} \sigma(\sum_{p=1}^6 \sum_{u=-2}^2 \sum_{v=-2}^2 S_p^1(i - u, j - v) \cdot k_{p,q}^2(u,v) + b_q^2)$$

$$= \sum_{i=1}^8 \sum_{j=1}^8 \Delta C_{p,q}^2(i,j) \cdot C_{p,q}^2(i,j) \cdot (1 - C_{p,q}^2(i,j))$$

$$\mathbf{\Delta} \sum_{i=1}^8 \sum_{j=1}^8 \Delta C_{p,q}^2(i,j)$$

$$\Delta k_{1,p}^1 \text{ (size } 5 \times 5 \text{):}$$

$$\Delta S_p^1(i,j) = \frac{\partial L}{\partial S_p^1(i,j)}$$

$$\begin{aligned} \Delta S_p^1(i,j) &= \sum_{q=1}^{12} \sum_{u=-2}^2 \sum_{v=-2}^2 \frac{\partial L}{\partial C_{p,\sigma}^2(i+u,j+v)} \frac{\partial C_{p,q}^2(i+u,j+v)}{\partial S_p^1(i,j)} \\ &= \sum_{q=1}^{12} \sum_{u=-2}^2 \sum_{v=-2}^2 \Delta C_{p,\sigma}^2(i+u,j+v) \\ &\quad v) \frac{\partial}{\partial S_p^1(i,j)} \sigma(\sum_{p=1}^6 \sum_{u=-2}^2 \sum_{v=-2}^2 S_p^1(i,j) \cdot k_{p,q}^2(u,v) + b_q^2) \\ \Delta S_p^1(i,j) &= \sum_{q=1}^{12} \sum_{u=-2}^2 \sum_{v=-2}^2 \Delta C_{p,\sigma}^2(i+u,j+v) \cdot k_{p,q}^2(u,v) \\ &= \sum_{q=1}^{12} \sum_{u=-2}^2 \sum_{v=-2}^2 \Delta C_{p,\sigma}^2(i - (-u), j - (-v)) \cdot k_{p,q,rot180}^2(-u, -v) \\ &= \Delta S_p^1 \sum_{q=1}^{12} \Delta C_{p,\sigma}^2 * k_{p,q,rot180}^2 \end{aligned}$$

Örnekleme ile C1 katmanındaki hata alınır.

$$\Delta C_p^1(i,j) = \frac{1}{4} \Delta S_p^1([i/2], [j/2]), i, j = 1, 2, \dots, 24$$

$$\Delta k_{1,p}^1(u,v) = \frac{\partial L}{\partial k_{1,p}^1(u,v)}$$

$$= \sum_{i=1}^{24} \sum_{j=1}^{24} \frac{\partial L}{\partial C_p^1(i,j)} \frac{\partial C_p^1(i,j)}{\partial k_{1,p}^1(u,v)}$$

$$= \sum_{i=1}^{24} \sum_{j=1}^{24} \frac{\partial L}{\partial C_p^1(i,j)} \frac{\partial}{\partial k_{1,p}^1(u,v)} \sigma(\sum_{i=-2}^2 \sum_{j=-2}^2 I(i - u, j - v) \cdot k_{1,p}^1(u,v) + b_p^1)$$

$$= \sum_{i=1}^{24} \sum_{j=1}^{24} \Delta C_p^1(i,j) \cdot C_p^1(i,j) (1 - C_p^1(i,j)) \cdot I(i - u, j - v)$$

$$= \Delta C_{p,\sigma}^1(i,j) = \Delta C_p^1(i,j) \cdot C_p^1(i,j) (1 - C_p^1(i,j))$$

$$\Delta k_{1,p}^1(u,v) = \sum_{i=1}^{24} \sum_{j=1}^{24} I_{rot180^0}(u - i, v - j) \cdot \Delta C_{p,\sigma}^1(i,j)$$

$$\Delta k_{1,p}^1 = I_{rot180^0} * \Delta C_{p,\sigma}^1$$

Δb_p^1 (size 1×1) :

$$\begin{aligned}
 \Delta b_p^1 &= \frac{\partial L}{\partial b_p^1(u,v)} \\
 &= \sum_{i=1}^{24} \sum_{j=1}^{24} \frac{\partial L}{\partial C_p^1(i,j)} \frac{\partial C_p^1(i,j)}{\partial b_p^1(u,v)} \\
 &= \sum_{i=1}^{24} \sum_{j=1}^{24} \Delta C_p^1(i,j) \frac{\partial}{\partial b_p^1(u,v)} \sigma(\sum_{u=-2}^2 \sum_{v=-2}^2 I(i - u, j - v) \cdot k_{1,p}^1(u,v) + b_p^1) \\
 &= \sum_{i=1}^{24} \sum_{j=1}^{24} \Delta C_p^1(i,j) C_p^1(i,j) (1 - C_p^1(i,j)) \\
 &= \sum_{i=1}^{24} \sum_{j=1}^{24} \Delta C_{p,\sigma}^1(i,j)
 \end{aligned}$$

Parametrelerin Güncellenmesi:

$$\begin{aligned}
 k_{1,p}^1 &\leftarrow k_{1,p}^1 - \alpha \cdot \Delta k_{1,p}^1 \\
 b_p^1 &\leftarrow b_p^1 - \alpha \cdot \Delta b_p^1 \\
 k_{p,q}^2 &\leftarrow k_{p,q}^2 - \alpha \cdot \Delta k_{p,q}^2 \\
 b_q^2 &\leftarrow b_q^2 - \alpha \cdot \Delta b_q^2 \\
 W &\leftarrow W - \alpha \cdot \Delta W \\
 b &\leftarrow b - \alpha \cdot \Delta b
 \end{aligned}$$

3.4. Öğrenme Aktarma (Transfer Learning)

Transfer learning işlemi hazır ağ modellerini kullanarak yeni resimleri sınıflandırma yöntemidir. Sınıf sayısı öğrenme oranı değiştirilebilmektedir. Bu sayede sadece belli parametreler değiştirilerek hazır modelleri kullanıp yeniden sınıflandırma işlemi yapılabilenmektedir. Sıfırdan model oluşturup eğitmek de mümkün olmakla birlikte ön eğitilmiş (önceden eğitilmiş) modeller bir çok çalışmada kullanılmaktadır (Arnavi ve ark., 2015) (Rajkomar ve ark., 2017). Literatürde resim ve obje tanımak için kullanılan birçok model bulunmaktadır. Bazıları VGG16 (Simonyan ve ark., 2014), VGG19 (Russakovsky ve ark., 2015), GoogleNet (Szegedy ve ark., 2015), VGGFace (Parkhi ve ark., 2015), ResNet (He ve

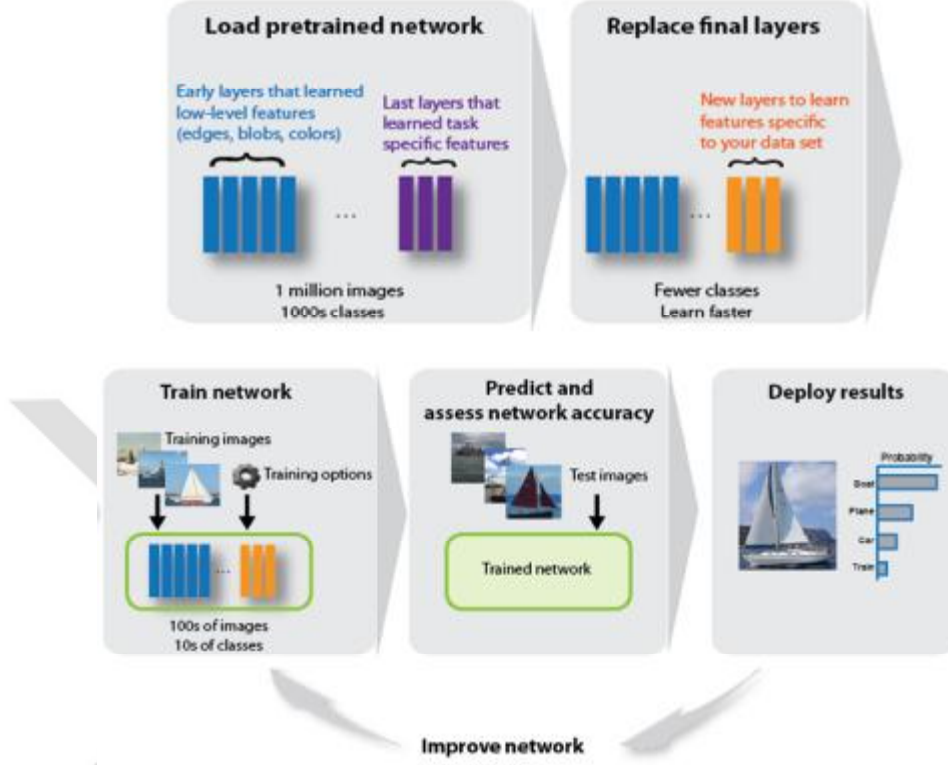
ark.,2016).Bu hazır modellerin kullanıldığı programlama dillerinin arayüz olarak kullanıldığı derin öğrenme kütüphaneleri mevcuttur. Literatürde bulunan ve popüler olanlar şu şekilde sıralanabilir:

- Caffe : Geliştirilmiş C ++ ve Python arayüzleri sağlar. UC Berkeley'deki lisans öğrencileri tarafından geliştirilmiştir(Jia ve ark., 2014).
- Tensorflow : C ++ ve Python ve sağlar. Google tarafından geliştirilen ve Google araştırması tarafından kullanılan arayüz(Abadi ve ark.,2016).
- Theano : Geliştirilmiş bir Python arayüzü sağlar. Montreal'deki MILA lab(Bastien ve ark.,2012).
- Torch : Lua arayüzü sağlar ve diğerleri arasında Facebook AI araştırması tarafından kullanılır(Collobert ve ark.,2011).
- Matconvnet : VLFeat ve C++ tarafından sağlanır. Matlab arayüzü ile kullanılır. Vgg araştırmaları için kullanılır(Vedaldi ve ark.,2015).

Bu kütüphaneler içerisinde popüler modeller hazır halde bulunmaktadır ve transfer learning için kullanılmaktadır. Matlab açık kaynak değildir. Lisanslı programdır. Demo sürümü kullanıldı.

AlexNet bir milyondan fazla görüntü üzerinde eğitilmiştir ve görüntüleri 1000 nesne kategorisine (klavye, kahve kupa, kalem ve birçok hayvan gibi) sınıflandırabilir. Ağ, çok çeşitli görüntüler için zengin özellik temsillerini öğrendi. Ağ, bir görüntüyü giriş olarak alır ve görüntüdeki nesnenin her nesne kategorisi için olasılıklarla birlikte bir etiket çıkarır.

Derin öğrenme uygulamalarında transfer öğrenme yaygın olarak kullanılmaktadır. Önceden hazırlanmış bir ağı alabilir ve yeni bir görev öğrenmek için başlangıç noktası olarak kullanılabilir. Aktarım öğrenmesi ile bir ağı ince ayarlanması genellikle, ağdan rastgele başlatılmış ağırlıkları olan bir ağı eğitmekten çok daha hızlıdır ve daha kolaydır. Öğrenilmiş özellikleri daha az sayıda eğitim görüntüsü kullanarak hızlı bir şekilde yeni bir göreve aktarılabilir.



Şekil.3.9. Transfer Learning(Aktarmalı Öğrenme) Aşamaları

Transfer learning işleminin aşamaları şunlardır:

1. Başla
2. Datayı yükle
3. Ön eğitilmiş ağı yükle
4. Son katmanları değiştir
5. Ağı eğit
6. Validasyon görüntülerini sınıflandır
7. Doğrula
8. Bitir

Son katmanları deęiřtirme: Ön eęitimli aęın son üç katmanı 1000 sınıf için yapılandırılmıştır. Bu üç katman yeni sınıflandırma problemi için ince ayarlanmış olmalıdır. Son üç katman hariç tüm katmanları, ön eęitimli aędan çıkarılır.

Aęın Eęitimi: Ön eęitimli olan aę giriř olarak 227*227*3 boyutlarında resimler ister. Resimler farklı boyutlarda olsa da boyutları 227*227*3 boyutlarına getirilmektedir.

Derin öğrenme çok fazla resim gerektirdięinden data artırma işlemi alexnet modelini kullanırken yapılmaktadır. Yapılan çalışmada 20 pixel düşeyde 20 piksel yatayda kaydırılarak çoęaltılmıştır. Ayrıca resimler yatayda ayna görüntüsü oluşturularak da sayısı artırılmıştır.

Aktarılan ve yeni katmanlardan oluşan aę eęitilir. Varsayılan olarak, trainNetwork varsa bir GPU kullanır (Parallel Computing Toolbox™ ve bilgi işlem kapasitesi 3.0 veya üstü olan bir CUDA özellikli GPU gerektirir). Aksi halde, bir CPU kullanır.

3.5. Alexnet

Alexnet bir derin konvolüsyonel yapay sinir aęı çeřididir.

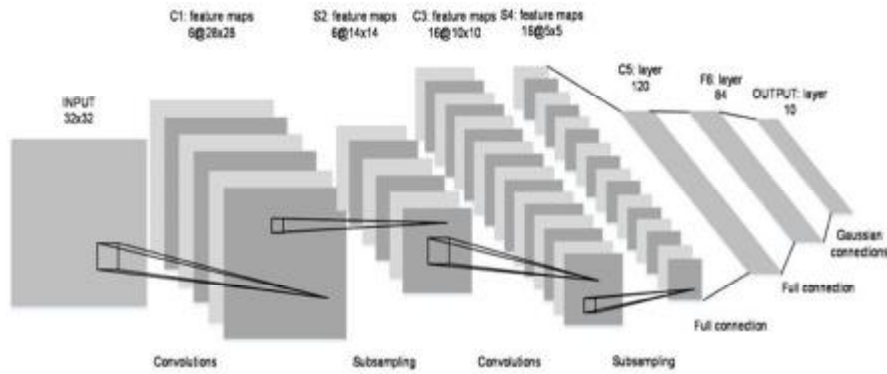
Imagenet LSVRC(ImageNet Large Scale Visual Recognition Challenge)-2010 1.2 milyon görüntüden ve 1000 sınıftan oluşan bir veri seti ile yapılan eęitim sonucunda elde edilen başarı top-1 hatası %37.5 ve top-5 hatası %17'dir. Bu sonuçlar o tarihe kadar elde edilmiş en iyi sonuçlardı. Oluřturdukları aę 5 adet konvolüsyonel 3 adet de tamamen baęlı aęlardan oluşur. 60 milyon parametre, 650.000 nöron ve 1000 vektörlü softmax katmanından oluşur. Eęitimi hızlandırmak amaçlı GPU kullanıldı. Sistemin ezberlemesini azaltmak için dropout denilen kesme yöntemi kullanıldı. Sonuç olarak 2012'de yapılan çalışma 2010'da yapılan çalışmaya rakip oldu ve sonuçlar top-5 test hata oranı %15,3 e düřtü.

Mevcut nesne ve yüz tanıma işlemleri makine öğrenmesi kullanımını mecburi kılmaktadır. Bu da makine öğrenmesi metotlarını geliřtirmeyi

gerektirmektedir. Öğrenmeyi geliştirmek içinse daha büyük data setleri kullanılmalıdır ki literatürde çok sayıda data seti bulunmaktadır. Nesne tanımak için kullanılan etiketlenmiş data setleri ImageNet LSVRC(Krizhevsky ve ark.,2012), NORB(LeCun ve ark., 2004), Caltech-101(Lee ve ark., 2007), Caltech-256(Griffin ve ark.,2007), CIFAR-10/100[9], yüz tanıma içinse LFW(Huang ve ark., 2007), VGG-Face(Parkhi ve ark.,2015) data setleri mevcuttur. Teknolojinin gelişmesi, akıllı telefon ve sosyal medya kullanımının artmasıyla insan yüzlerine ait datasetleri de hızla artmaktadır.

Bu makine öğrenmesi sistemlerinde ezberlemeyi önlemek amacıyla büyük data setleri kullanılır. Mesela mevcut çalışmalar içinde en başarılı sonuçlar veren MNIST data setiyle yapılan çalışmadır.

70.000 adet el yazısı rakamlardan oluşan MNIST data seti 60.000 eğitim, 10.000 tanesi testte kullanılmak üzere ayrılmıştır. İnsan a yakın hata değerinde doğruluk ile tahminde bulunma sonucu elde etmeyi başardılar hata<%0.3(Krizhevsky ve ark., 2012).



Şekil.3.10. Mnist data setinin tabi tutulduğu eğitim modeli Resim (Tra ve ark., 2017)

32*32 boyutlarında el yazısı resimler 5*56adet kernelle konvolüsyona tabi tutulur. Sonrasında 2*2 maxpooling yapılır. Yine 5*5 16 adet kernelle konvolusyon işlemi yapılır, sonrasında 2*2 maxpooling işlemi yapılır. Daha sonra kalanların

hepsi tek boyutlu vektöre çevrilip tam bağlı yapay sinir ağıyla öğrenme sağlanır. En son katmandaki vektörler softmax fonksiyon ile 0'dan 9'a kadar her biriyle hangi olasılıkla eşleştiğini hesaplamaktadır.

Binlerce resmi öğrenebilmesi için büyük öğrenme kapasiteli bir modele ihtiyaç vardır(Krizhevsky ve ark.,2012). Konvolüsyonel sinir ağları bu öğrenme için ideal varsayılan bir seçimdir. CNN(konvolüsyonel sinir ağları) daha az parametre, daha az bağlantı ve dolayısıyla daha kolay eğitim sayesinde daha doğru sonuçlar edebilmektedir.

CNN in avantajlarına rağmen resimlerin boyutlarının büyük olması ve yüksek çözünürlüklü olması yine hesaplamaları yüksek maliyetli hale getirmektedir. Teknoloji gelişimine bağlı olarak GPU(Graphic Proccessing Unit) artık çok rahat ulaşılabilir durumdadır. Derin öğrenme yapılarında görüntünün çok sayıda olması ezberlemenin önüne geçmektedir. Bu da çok sayıdaki etiketli görüntünün daha güçlü ve daha hızlı eğitilmesini sağlamaktadır(Krizhevsky ve ark., 2012).

Alexnet modelinde kendisinden önceki çalışmalarda kullanılmayan birkaç yenilik kullanılmıştır. Derin öğrenme yapısında en büyük problem ezber olduğu için kullanılan yeni yöntemler bunu önlemeye yönelik yöntemlerdir.Alexnet iki adet GTX 580 3GB GPU üzerinde 5-6gün kullanılmıştır.

ImageNet 15 milyonun üzerinde 22.000 kategoride etiketli resimlere sahiptir. Buradaki resimler internetten ve insan eliyle etiketlenen Amazon' MechanicalTurk topluluk kaynağı aracından toplanmıştır. Yani alexnet in eğitimi için totalde1.2 milyon resim eğitim için, 50.000 resim validasyon için, 150.000 resim de test için kullanılmıştır.

ImageNet resimleri farklı çözünürlüktedir. Kullanılan ağ modeli giriş olarak sabit boyutlarda giriş gerektirmektedir. Bu sebeple resimler 256*256 boyutlarına yeniden ölçeklendirmeye getirilmişlerdir.

ReLU (Rectified Linear Unit) bir aktivasyon fonksiyonudur. Genel olarak sinir ağlarında aktivasyon fonksiyonu olarak $f(x)=\tanh(x)$ ya da $f(x)= (1+e^{-x})^{-1}$

kullanılır. Fakat bu fonksiyonlar gradient azalmalı eğitimde doyum nonlineerliğine çok daha yavaş çok daha yavaş ulaşmaktadır. ReLU ile eğitilen derin konvolüsyonel sinir ağı tanh ile eğitilen derin konvolüsyonel sinir ağlarından birkaç kat daha hızlıdır(Krizhevsky ve ark., 2012). Öğrenme işi bu katmanlarda gerçekleşir.

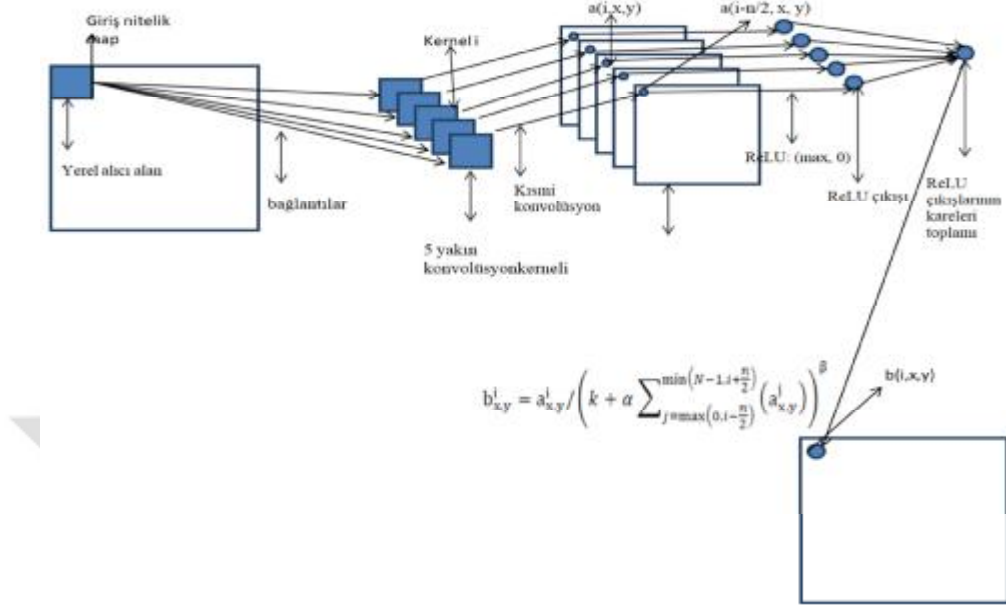
Bu çalışmada iki adet GPU kullanılmıştır. Bir tane GPU 1.2 milyongörüntüyü eğitmek için yetersiz kalmıştır. Bu şekilde yapmak top-1 hatasını %1,7 top-5 hatasını da %1.2 oranlarında indirmiştir(Krizhevsky ve ark., 2012).

Yerel Tepki Normalizasyon(Local Response Normalization(LRN))

$$b_{x,y}^i = a_{x,y}^i / \left(k + \alpha \sum_{j=\max(0, i-\frac{n}{2})}^{\min(N-1, i+\frac{n}{2})} (a_{x,y}^j) \right)^\beta$$

$$k=2, n=5, \alpha=10^{-4}, \beta=0,75$$

ReLU sistemi satürasyondan korumak için normalizasyon gerektirmeyen bir özelliğe sahiptir. Fakat yine de yukarıdaki şema genelleştirilmiş bir normalizasyon çeşididir. Kerneller keyfi ve rastgele seçilir. $a_{x,y}^i$ kernelle işleme tabi tutulup ReLU'dan geçirildikten sonraher bir kernelin (x,y)konumudaki elemanını aynı konumdaki farklı kernellerdeki kareleri toplamının α ile çarpılıp k ile toplanıp β . dereceden üssüne böldüğümüzde her bir elemanı normalize etmiş oluruz. Hesaplama yönteminin nasıl olduğu Şekil.3.11'de gösterilmiştir:



Şekil.3.11. Yerel Tepki Normalizasyon şeması

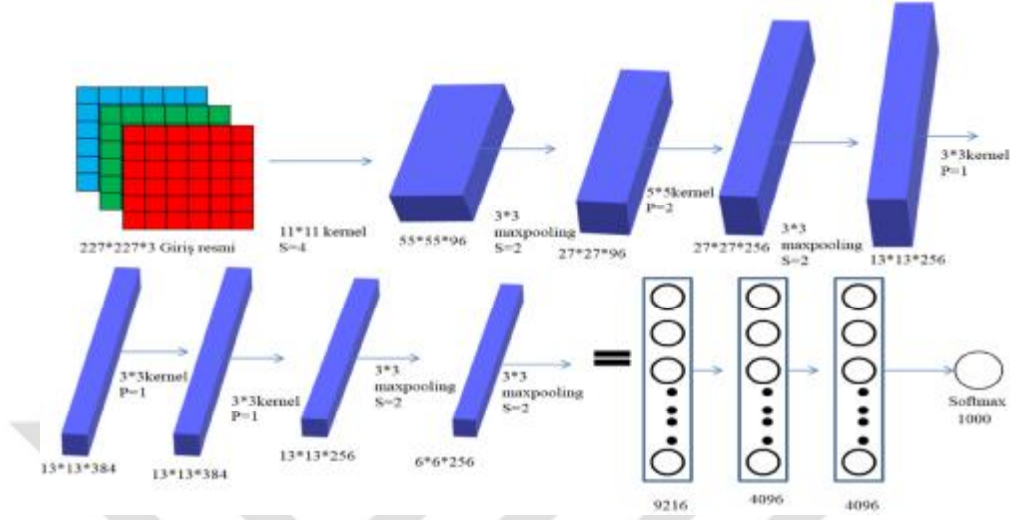
Bu şema kontrast normalizasyon işlemi ile benzerlik göstermektedir (Jarett ve ark. 2009). Fakat burada görüntülerin ortalamasını çıkarma işlemi yapılmadı. Buradaki işleme “aydınlık normalizasyonu” denirse daha doğru olur (Krizhevsky ve ark., 2012). Top-1 ve top-5 hatalarını sırasıyla %1.4 ve %1.2 oranında azaltmıştır.

Anlatılan pooling sisteminden farklı olarak burada çakışan pooling sistemi kullanılmaktadır. Yani 3 adımlı 3*3 pooling değil de 2 adımlı 3*3 pooling kullanılmaktadır. Pooling hesaplama yöntemi Pooling başlığı altında anlatılmıştır. Birçok çalışmada çakışan pooling değil de alışıldık pooling kullanılmaktadır (Jarett ve ark., 2009) (LeCun ve ark., 2010) (Cireşan ve ark., 2012). Bu pooling çeşidi top-1 ve top-5 hatalarını sırasıyla %0.4 ve %0.3 oranlarında azalmasını sağlamıştır.

Ağın yapısı 5 adet konvolüsyon ve 3 adet tamamen bağlı ağlardan katmanlardan oluşmaktadır. Çıkışı ise 1000 sınıf etiketi üzerine dağılım üreten tamamen bağlı son ağdan 1000 yollu softmax ile beslenir. Softmax fonksiyonu çok terimli lojistik regresyona karşılık gelmektedir.

İkinci, dördüncü ve beşinci konvolüsyonel katmanların kernelleri aynı GPU üzerinde bulunan önceki katmandaki sadece o kernel haritalarına bağlıdır. LRN ise ilk ve ikinci konvolüsyonel katmanlardan sonra yapılmaktadır. Aktivasyon fonksiyonu olarak tüm katmanlarda ReLU kullanılmıştır.

Giriş resmi $227 \times 227 \times 3$ boyutunda resimlerden oluşmaktadır. İlk konvolüsyonda 4 adımlı $11 \times 11 \times 3$ boyutlarında 96 kernel kullanıldı. Konvolüsyon, ReLU, LRN ve Max-pooling işleminden sonra ikinci konvolüsyon işlemi yapılır. İkinci konvolüsyonda $5 \times 5 \times 96$ boyutunda 256 kernel kullanılır. Üçüncü, dördüncü ve beşinci katmanlardan sonra herhangi bir pooling ya da normalizasyon işlemi yapılmamaktadır. Üçüncü konvolüsyonda $3 \times 3 \times 256$ boyutlarında 384 adet kernel kullanılmıştır. Dördüncü konvolüsyon katmanı $3 \times 3 \times 384$ boyutlarında 384 kernel kullanılır. Beşinci konvolüsyon katmanı $3 \times 3 \times 384$ boyutlarında 256 kernel kullanılmıştır. Tamamen bağlı katmanlar da 4096×1 boyutlarındadır. Resim sayısının fazla olması da sistemin ezberlemesini önlemeye yetmiyor. AlexNet ağında kullanılan iki yöntem ezberleme ile başa çıkmak için kullanılmıştır. En uygun ve en yaygın ezber azaltma yöntemi etiket korumalı data artırma yöntemidir (Cireşan ve ark., 2012) (Simard ve ark., 2003) (Cireşan ve ark., 2011). Bu çalışmada 2 çeşit data artırma işlemi kullanılıyor. Birinci çeşit data artırma resim öteleme ve yatayda yansıma işlemlerinden oluşur. Bu işlem resimlerden rastgele 227×227 boyutlarında parçalar çıkarılarak yapılır çünkü giriş resmi $227 \times 227 \times 3$ boyutlarındadır. Bu işlem eğitim setini 2048 kat artırır.



Şekil.3.12. AlexNet Şeması

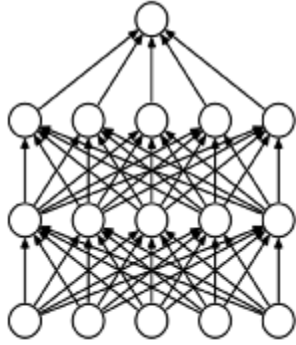
İkinci çeşit data artırma eğitim seti resimlerinde RGB kanal yoğunluklarının değiştirilmesinden oluşur. ImageNet eğitim seti görüntülerinin RGB pixel değerlerinin ayarı üzerinde PCA gerçekleştirilmiştir. Her bir eğitim görüntüsüne bulunan temel bileşen çarpımları eklendi. Yani özdeğerlere karşılık gelen orantısal büyüklüklü kere 0.1 standart sapmalı ve sıfır ortalı Gaussian'dan çekilen rastgele değişken eklenmiş oldu. Böylelikle her bir RGB pikseline $I_{x,y} = [I_{x,y}^R, I_{x,y}^G, I_{x,y}^B]^T$ aşağıdaki miktarda ekleme yapılır:

$$[P_1, P_2, P_3] [\alpha_1 \lambda_1, \alpha_2 \lambda_2, \alpha_3 \lambda_3]^T$$

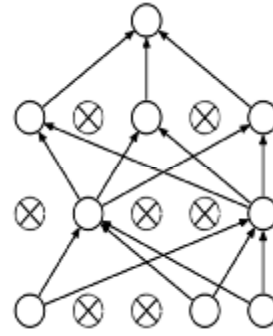
P_i ve λ_i sırasıyla i'ninci özvektör ve RGB piksel değerlerinin 3x3 kovaryans matrisinin özdeğeridir ve α_i yukarıda bahsedilen rastgele değişkendir.

Bu şema yaklaşık olarak doğal resimlerdeki önemli noktaları yakalar. Yani nesnelerin kimlikleri değişimlerden ve renk yoğunluklarının değişiminden bağımsız olur. Bu şema top-1 ve top-5 hata değerlerini yaklaşık olarak %1 oranında düşürmüştür (Krizhevsky ve ark., 2012).

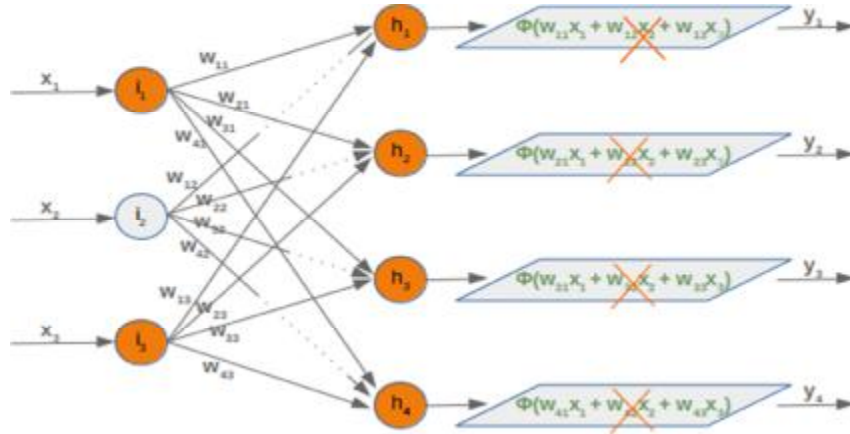
Birçok farklı modelin tahmininin birleşimi test hatalarını azaltmada başarılı bir yoldur(Belland ve ark., 2007)(Breimann,2001). Fakat bu büyük sinir ağları için günler alan maliyetli bir iş olur. Bu maliyeti oldukça azaltan bir yöntem olan “dropout” kullanılmıştır(Hinton ve ark.,2012). İlerleyen kısımlarda anlatılacaktır. Dropout ileri yönde beslemede yapılır sadece. Geri yayılım beslemede yapılmaz. Bu sadece FC katmanlarında yapılır ve burada 0,5 oranında kullanılır. 0.5 oranında rastgele nöronlar kesilir.Dropout olmaksızın yapılan eğitimde ezberleme oldukça fazlaydı(Krizhevsky ve ark.,2012).Aşağıdaki resimde dropout işleminin şekli şematize edilmiştir.



Şekil.3.13. Dropouttan önce



Şekil.3.14. Dropouttan sonra



Şekil.3.15. Dropout Açık Şekil

Stokastik gradyan azalma ile eğitilen sistemde batch boyutu 128 örnekten oluşuyor. Momentum değeri 0.9 ve ağırlık sönümlleme değeri 0.0005'tir. Ağırlık güncelleme kuralı:

$$V_{i+1} = 0.9V_i - 0.0005 * e * w_i - e * \left\langle \frac{\partial L}{\partial w} | w_i \right\rangle D_i$$

$$W_{i+1} = W_i + V_{i+1}$$

Denklemden i iterasyon indisi, V momentum değişkeni, e öğrenme oranı ve $\left\langle \frac{\partial L}{\partial w} | w_i \right\rangle D_i$ de i'ninci batch $D_i W_i$ 'de değerlendirilen w'ye göre objektifin türevidir. Ağırlıklar her katmanda 0.01 standart sapmalı sıfır ortalamalı Gaussian dağılımdan başlatılmıştır. 2,4 ve 5. Konvolüsyonel katmanlar ve tamamen bağlı katmanlarda bias değeri sabit 1 olarak kullanılmıştır. Geri kalan katman ve nöronlarda bias değeri sıfırdır. Bu değerler ReLU ile birlikte öğrenmenin hızlandırır. Ağ 1.2 milyon görüntü ile yaklaşık 90 saykıl eğitilmiştir. 5-6 gün süren eğitimde iki adet GTX 580 3 GB GPU kullanılmıştır.

Top-1 ve top-5 hata oranları sırasıyla %37.5 ve %17 dir.

3.6. Fotoğraf Veri Setinin Oluşturulması

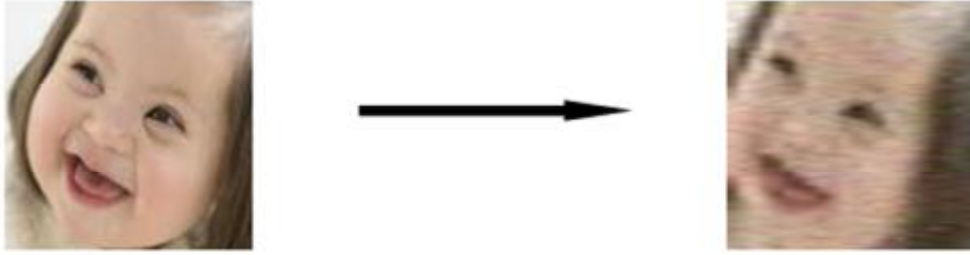
Çalışmadaki temel amaç down sendromlu insanları normal insanlardan ayırt etmek için makine öğrenmesi sistemi kullanılmasıdır. Klinik tanının insan-bilgisayar etkileşimi ile sağlanması amaçlanmaktadır. Kullanılan resimler uluslararası fotoğraf veri setlerinden alınmıştır.

Veri setinde 230 adet normal insan resmi 230 adet down sendromlu insan resimleri kullanılmıştır. 230 adet down sendromlu resim pixabay.com ve pinterest.com sitelerinden alınmıştır. 230 adet normal insan resimleri ise LFW data setinden alınmıştır(Miller ve ark.,2016)(Sanderson&Lovell,2009).

Derin öğrenmenin en önemli kısmı öznetelikleri kendisinin çıkarmasıdır. Fakat bunun için oldukça fazla resim gerekmektedir. LFW data setinde çok fazla

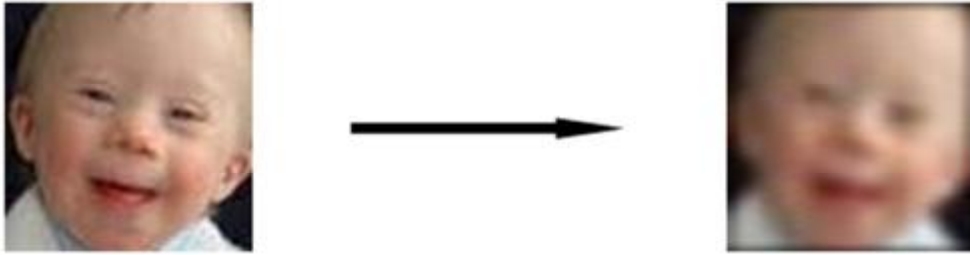
resim olmasına karşın down sendromlu kişiler için yeterli sayıda resim bulunmamaktadır. Bu sebeple resimler data artırma yöntemiyle artırılmıştır.

Down sendromlu kişilerin resimlerine 0.1 oranında gaussian gürültü eklenmiştir. Gürültü eklendikten sonra data setine eklenmiştir. Aşağıdaki şekilde gösterilmektedir.



Şekil.3.16. Gürültü eklenmiş resimler

Normal resimler ve gaussian gürültülü resimler bulandırmak suretiyle veri setine eklendi. Matlab görüntü işleme araç kutusundan yararlanılmıştır.



Şekil.3.17. Bulanıklaştırılmış resimler

3.7. Özniteliklerin Çıkarılması

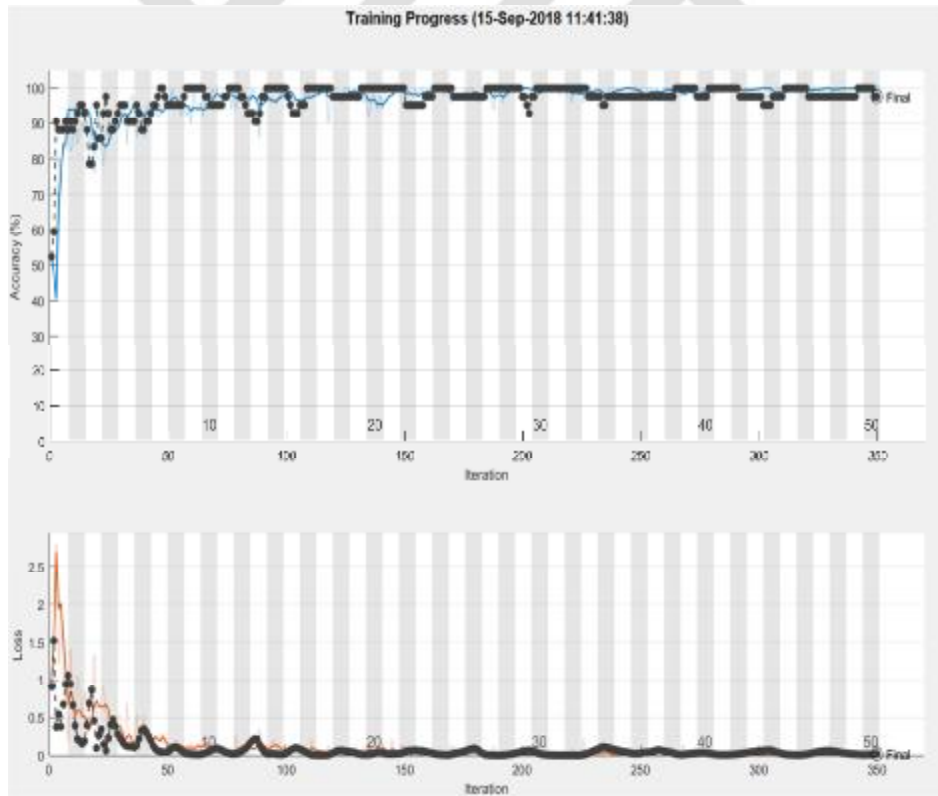
Resimlerin özniteliklerinin çıkarılması için alexnet kullanılmıştır (Krizhevsky ve ark.,2012). Matlab R2018a öğrenci sürümü kullanılmıştır. Mathworks alexnet i hazır bir ağ olarak kullanıcılarına sunuyor. Çalışmamızda bu ağ modelini kullandık. Alexnet modelinin tercih edilmesinin nedeni daha önce literatürde buna benzer bir çalışma yapılmış olmasıdır ve aynı model kullanılarak

%98.80 başarı elde edilmiştir(Shukla ve ark.,2017). Matlab için tasarlanmış olan alexnet modeli matlab paket programı kurduktan sonra indirildi ve kuruldu.Bu modeli indirdikten sonra transfer learning işlemi uygulandı.

3.8. Öğrenme Verileri ve Ön eğitimli Ağ kullanımı

3.8.1. Normal veri seti ile eğitim

210 adet down sendromlu ve 210 adet normal insandan oluşan iki adet temiz fotoğraflarla eğitim yapıldı. Yapılan eğitimin grafiği aşağıda verilmiştir. Bu eğitim esnasında fotoğrafların öznitelikler çıkarılarak sınıflandırılmıştır. Özniteliklerin çıkarılma işlemi konvolüsyon katmanında sınıflandırma işlemi son üç katmandaki tamamıyla bağlı katmanlarda gerçekleşmektedir.



Şekil.3.18. Eğitim sürecinde elde edilen performans başarı grafiği



Şekil.3.19. Optimum değerlerde eğitimin sınıflandırması %97.62



Şekil.3.20. Eğitim setinden seçilen rastgele 16 görüntü

Çizelge 3.1. Optimum eğitimde kullanılan değiştirilebilir hiperparametreler.

Eğitim ve Test seti yüzdesi	0.9
Sınıflandırma katmanındaki ağırlık ve bias öğrenme çarpanı	50
Data artırma kısmında veri setinin kaydırılan piksel sayısı	20
Minibatch boyutu	50
En fazla epok sayısı	50
Başlangıç öğrenme oranı	10^{-4}
Validasyon frekansı	1

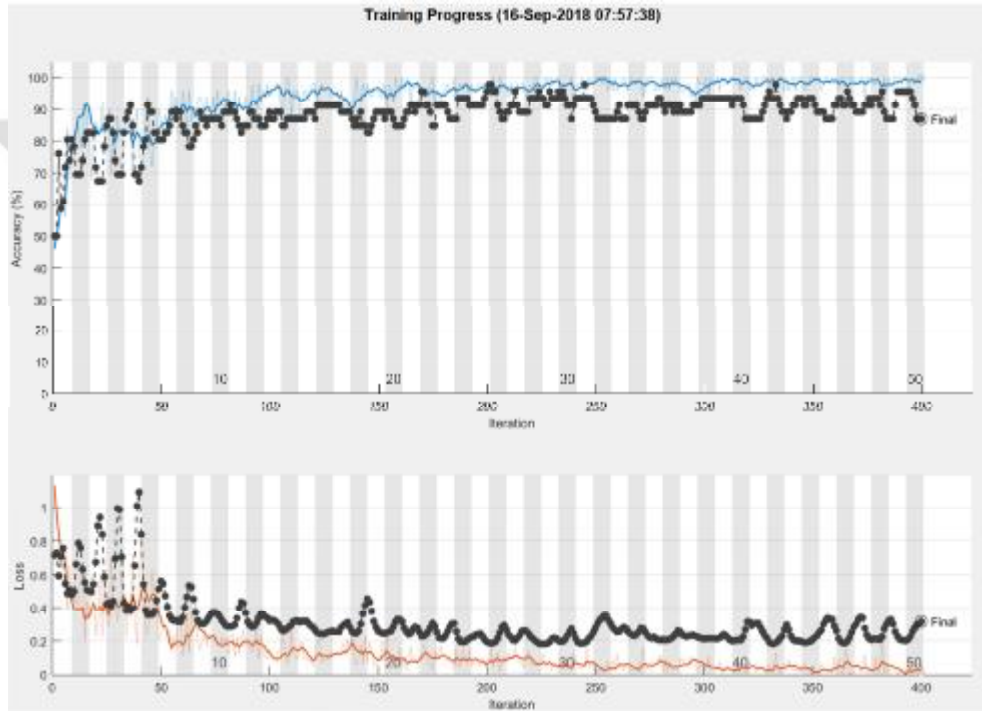
Yukarıdaki tabloda verilen değerlerle eğitim yapılmıştır ve doğruluk değeri grafikte %97.62 olarak gösterilmektedir.



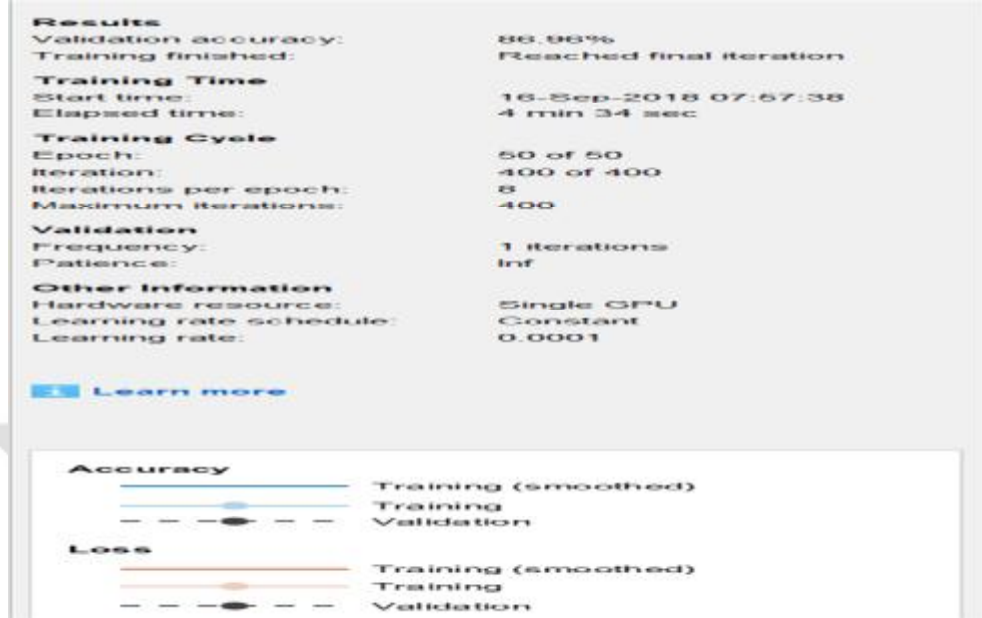
Şekil.3.21. Validasyon setinden seçilen rastgele 4 görüntü

3.8.2. Bozuk veri seti ile eğitim

Resimler 3.12’de anlatıldığı gibi bozulmuş ve bu şekilde eğitime sokulmuştur. Resimler ‘motion’ ve 0.1 oranında ‘salt&pepper’ gürültüler eklenerek resimler bozulmuş ve ayrı ayrı eğitime tabi tutulmuştur. Grafikler ve resimler aşağıdadır.



Şekil.3.22. Eğitim süresince elde edilen grafik(bulanık resimlerle)



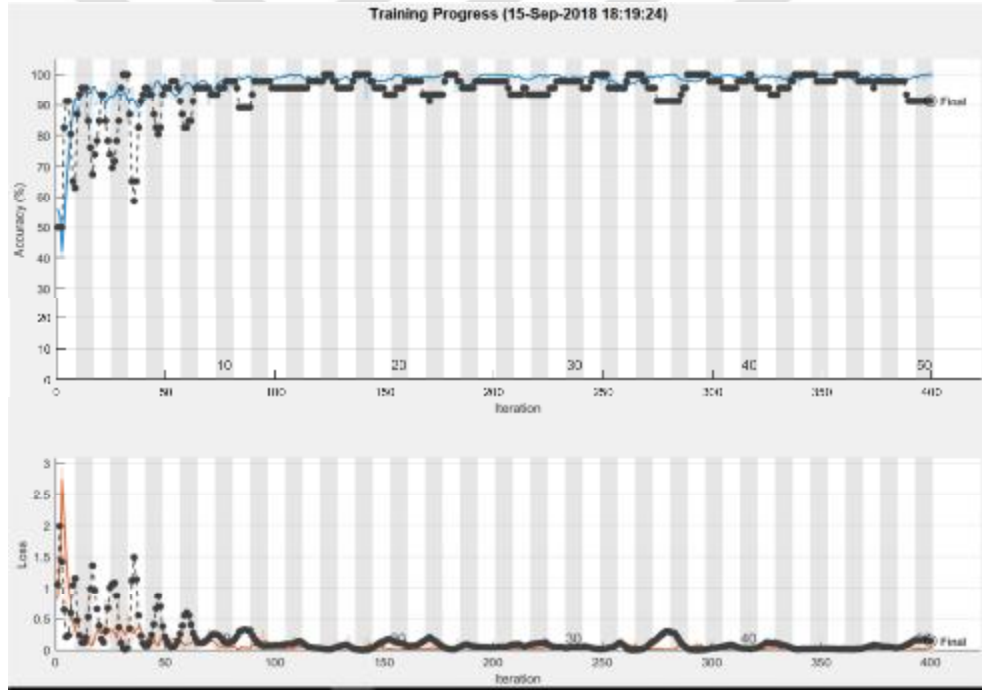
Şekil.3.23. Bulanık resimlerle yapılan eğitim sonucu %86.96 doğruluk oranı.



Şekil.3.24. Eğitim setinden rastgele seçilen bulanık resimler



Şekil.3.25. Validasyon setinden seçilen rastgele 4 görüntü



Şekil.3.26. Eğitim süresince elde edilen grafik(gürültülü resimlerle)



Şekil.3.27. Gürültülü resimle eğitim sonucu sınıflandırma doğruluğu %91.30

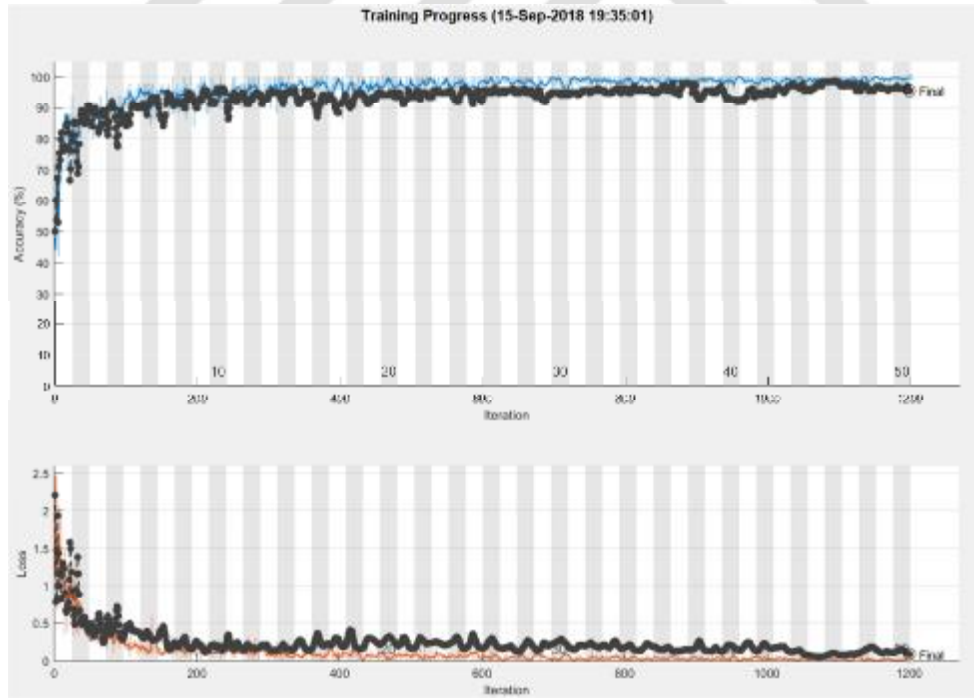


Şekil.3.28. Gürültülü resimlerde eğitimde kullanılan resimlerden rastgele seçilmiş resimler



Şekil.3.29. Sınıflandırma sonucu validasyon resimlerinde rastgele seçilmiş 4 resim

Gürültülü, bulanık ve temiz resimlerden eşit sayıda alınarak oluşturulmuş eğitim seti 690 adet resimden oluşmaktadır. Grafik ve eğitim resimleri aşağıda verilmiştir:



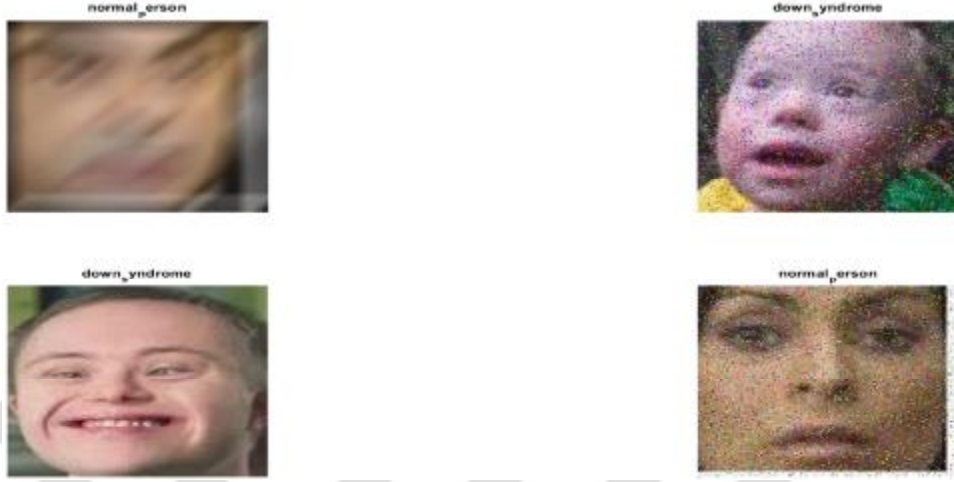
Şekil.3.30. Eğitim süresince elde edilen grafik(karışık resimlerle)



Şekil.3.31. Gürültülü resimle eğitim sonucu sınıflandırma doğruluğu %91.30



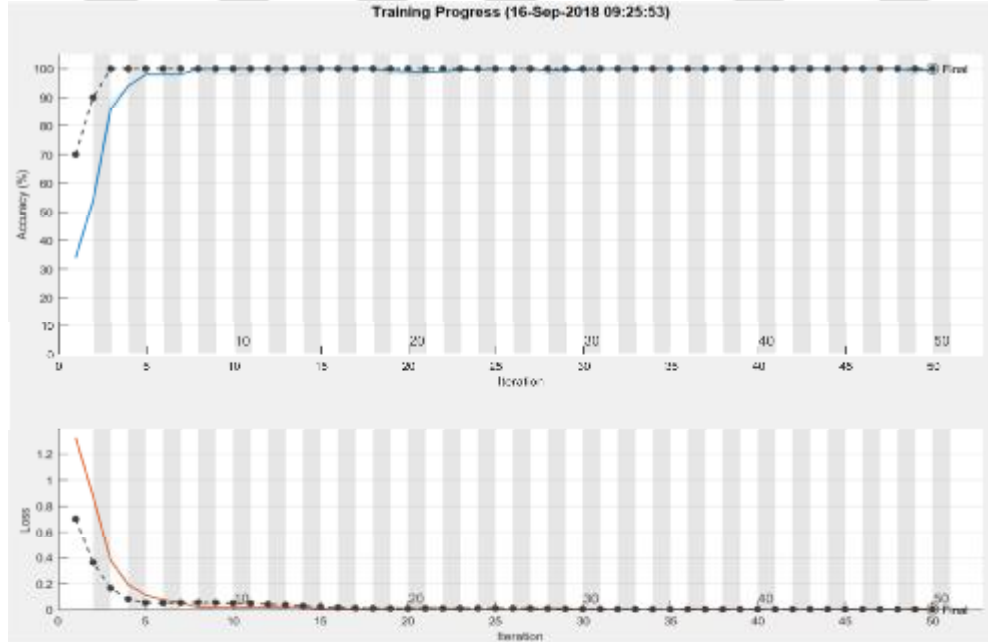
Şekil.3.32. Karışık resimlerde eğitimde kullanılan resimlerden rastgele seçilmiş resimler



Şekil.3.33. Validasyon setinden seçilen rastgele 4 görüntü

3.8.3. Az Sayıda Resimle Yapılan Eğitim

3.8.3.1. 50 Adet Resimle Yapılan Eğitim



Şekil.3.34. Eğitim süresince elde edilen grafik(50 adet resimle yapılan eğitim)



Şekil.3.35. Aynı resimlerden 50 adet seçilip aynı hiperparametrelerle eğitim sonucu

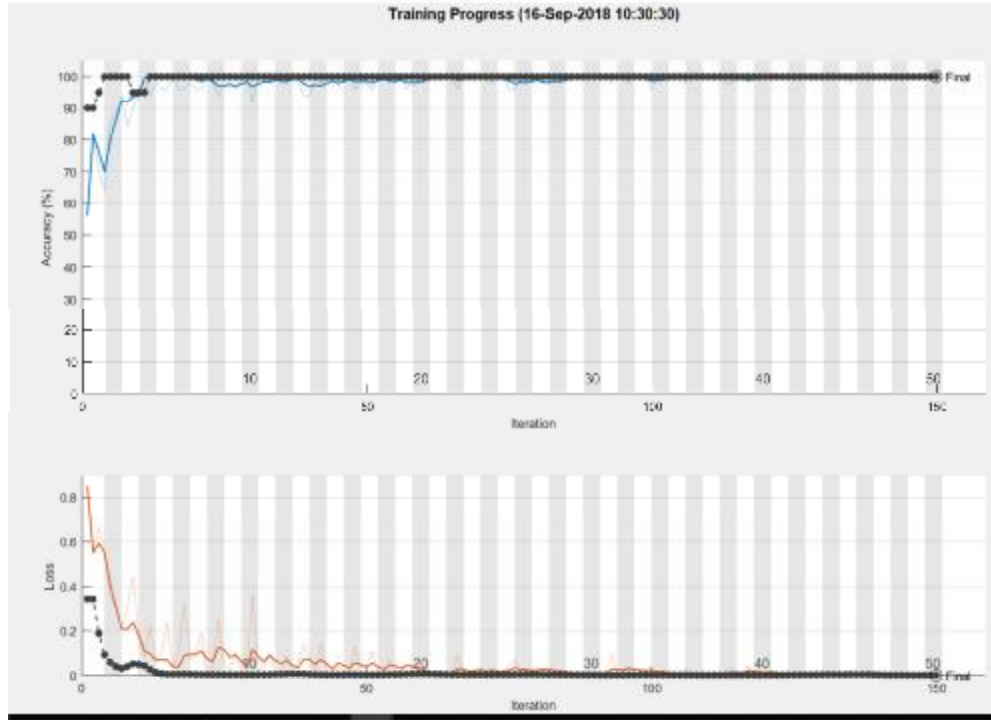


Şekil.3.36. 50 adet resim bulunan eğitim seti için eğitimde kullanılan resimlerden rastgele seçilmiş resimler



Şekil.3.37. Sınıflandırma sonucu Validasyon resimlerinden rastgele seçilen 4 adet resim

3.8.3.2. 100 Adet Resimle Yapılan Eğitim



Şekil.3.38. Eğitim süresince elde edilen grafik(100 adet resimle yapılan eğitim)



Şekil.3.39. Aynı resimlerden 100 adet seçilip aynı hiper parametrelerle eğitim sonucu



Şekil.3.40. 100 adet eğitim veri setinden rastgele seçilmiş resimler



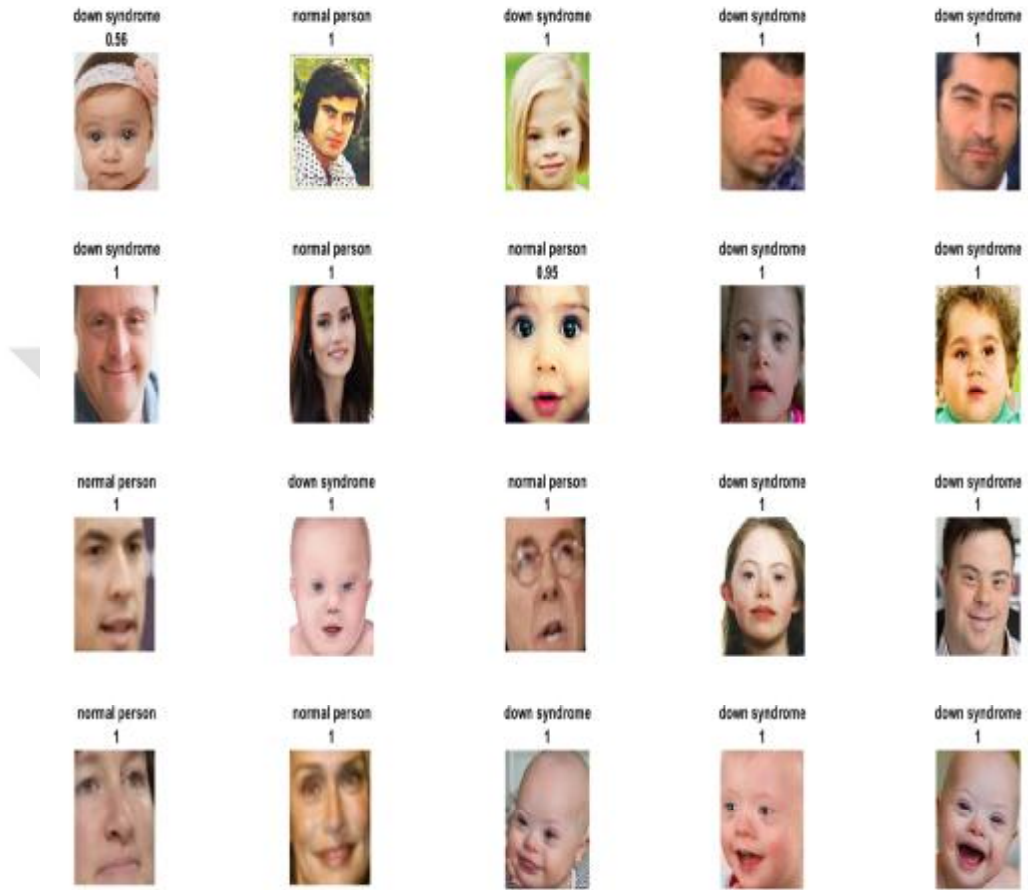
Şekil.3.41. Sınıflandırma sonucu validasyon veri setinden rastgele seçilmiş 4 resim



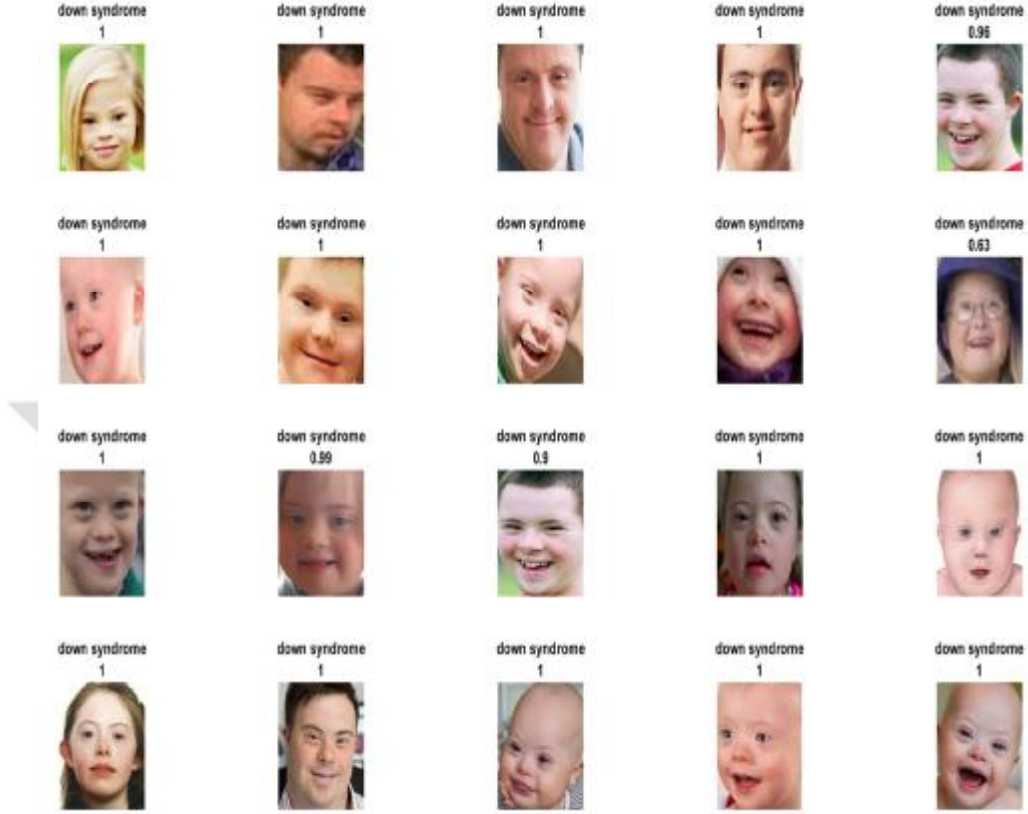
4. SONUÇLAR

Down sendromlu kişilerin yüzlerinden hastalıklarını tespit etmek mümkündür. Çünkü yüz yapıları normal insanların yüzlerinin yapılarından farklıdır ve nitelikleri bellidir. Materyal metot bölümünde anlatıldığı üzere hazır bir model olan Alexnet kullanılmıştır. Alexnet 1000 sınıflı bir nesne tanıma algortimasıdır fakat ağıın yapısında küçük değişiklikler yapıp 2 sınıftan oluşan veri setiyle eğitim gerçekleştirildi. 210 adet Normal insan 210 adet Down sendromlu kişilerin resimleriyle eğitim yapıldı. 230 adet normal insan, 230 adet Down Sendromlu kişilerin resimlerinden 20'şer adet rastgele seçilip performans testi iki adet daha klasör oluşturuldu. 230'ar adet resimler yani oluşturulan ilk klasörlerdeki resimler bulandırılmış ve gürültü eklenmiştir. Bunlar da ayrı ayrı farklı resimlerle performans testine tabi tutulmuştur.

4.1. Normal Eğitim Setiyle Yapılan Eğitim sonuçları



Şekil 4.1. Karışık resimlerle yapılan test sonucu %85 başarı elde edilmiştir.



Şekil 4.2. Down Sendromlu resimlerle yapılan test sonucu %100 başarı elde edilmiştir.

Sadece down sendromlu kişilerin resimlerinden oluşan performans testinde %100 başarı sağlanmıştır. Down sendromlu kişilerin yüzlerinin özneliliklerini sistemin iyi öğrendiği anlamına gelmektedir ve başarı sağlanmıştır.



Şekil 4.3. Normal resimlerle yapılan test sonucu %75 başarı elde edilmiştir.

Sadece normal insanların resimlerinden oluşan performans testinde %75 başarı sağlanmaktadır.

Çizelge 4.1. Data set boyutunun eğitim performansına etkisi

Eğitim setleri	Grafik sonuçları	Performans
		Sonuçları(karışık ve temiz resimler)
50 adet resimle eğitim	%100	%78.94
100 adet resimle eğitim	%100	%84.22
Optimum Eğitim	%97.62	%85

Yukarıdaki tabloyu incelediğimizde resimsayısı arttıkça öğrenme başarısının arttığını ispatladık. Buradan yola çıkarak daha az resim demek öğrenmenin azalması demektir.

Yukarıdaki çizelgede yapılan eğitimde kullanılan parametreler optimum eğitimdeki parametreler kullanılarak yapılmıştır. Resimde bulandırılan ve gürültü eklenmiş nitelikleri tanımada ne kadar başarılı olduğu performans testleriyle belirlenmiştir.

Çizelge4.2’de gürültü eklenmiş ve bulanıklaştırılmış resimlerden oluşmaktadır. Karışık resimler ise 230 adet temiz görüntü aynı resimlerin bulanıklaştırılarak oluşturulmuş 230 adet resim ve aynı resimlere gürültü eklenerek oluşturulmuş 230 adet resim karıştırılıp optimum eğitimde kullanılan parametrelerle eğitim gerçekleştirilmiştir. Optimum parametreler, parametreler değiştirilip eğitilmiş ve karşılaştırma yapılarak tespit edilmiştir.

Bu çalışma sonucunda karışık resimlerle yapılan performans testlerinde %85, down sendromlu kişilerin resimlerinde %100, sadece normal insanların resimlerinden oluşan performans testlerinde %75 başarı elde edilmiştir. Bu oldukça iyi bir başarıdır. Makine öğrenmesi yöntemlerine bakıldığında nitelik çıkarma işleminden sonra sınıflandırma yapılırca da oldukça başarılı sonuçlar elde edilmiştir. Fakat bizim çalışmamızın farklı yanı nitelik çıkarma işlemi yapılmaksızın ham datalarla eğitim ve sınıflandırma gerçekleştirilmiş ve performans testlerine tabi tutulmuştur. Ayrıca Shukla ve arkadaşlarının yaptığı daha kapsamlı çalışmada gürültü eklenip performans testi yapılmamıştır. Bu çalışmanın kısmi yeniliği resimlere gürültü ekleyip performans testi yapılmış olmasıdır. Makine öğrenmesi ile elde edilen sonuçlar ve bu çalışmada elde edilen sonuç Çizelge4.3’de verilmiştir. Derin öğrenme yöntemi kullanarak down sendromlu kişilerin resimlerinin sınıflandırılması çalışması Türkiye’de ilk olmasından ötürü önemlidir.

Çizelge 4.2. Kalitesi bozulmuş ve karıştırılmış resimlerle yapılan eğitimin performansı.

	Grafik sonuçları	Performans Sonuçları (Temiz)	Performans Sonuçları (Gürültülü)	Performans Sonuçları (Bulanık)
Bulanık resimler	%86.96	%8.33	%41.66	%78.95
Gürültülü resimler	%91.30	%50	%75	%78.95
Karışık resimler	%95.65	%91.6	%85	%85

Çizelge 4.3. Makine öğrenmesi yöntemleri kullanılarak yapılan down sendromlu resmi sınıflandırma ve mevcut çalışma.

Çalışma	Yöntem	Sonuç
David et al.,2005	SVM	87%
Saraydemir et al., 2015	GWT+SVM	97,50%
Erogul et al., 2009	YDGM+YSA	68,70%
Saraydemir et al.	GWT+PCA+LDA+KNN	95%
Zhao et al.,2013	Contourlet transform+LBP	97.92%
Bizim Çalışmamız	Alexnet	%85

5. TARTIŞMA

Shukla ve arkadaşlarının çalışması incelendiğinde hem kapsamının geniş hem de performansının yüksek olduğu görülmektedir. Bu çalışmanın zayıf yönü kapsamının az olmasıdır. Güçlü yönü ise resimlere gürültü eklenip eğitime tabi tutulmuş olmasıdır. Bu çalışmanın bir sonraki aşaması sadece down sendromlu kişilerde değil dismorfik bozukluğu olan tüm hastalıklar için yapılmasıdır. Toplum içerisinde yaşaması zor olan bu kişilerin tespiti ve ihtiyaçlarının buna göre belirlenebilmesi için farklı çalışmalarda kullanılabilir. Örneğin, havaalanlarında, otobüs terminallerinde yüz tanıma sistemleri yardımıyla bu insanlar belirlenip gereksinimleri buna göre ağlanabilir. Makine insan etkileşiminde kullanılabilir. Derin öğrenmenin en büyük avantajı öznitelik çıkarma mühendisliği gerektirmemesidir. Özellikle makine-insan etkileşimi düşünüldüğünde insanların yüzlerinden hastalıkları tespit edilip tedavi ve rehabilitasyon yöntemi otomatik olarak seçilebilir. Bu da hem uzmanlar için hem de hasta kişiler için oldukça kolaylık sağlayacaktır. Tedavi ve rehabilitasyon sürecinde başarılı sonuçlar alınabilir.

Bu sistemler gömülü sistemlerde de kullanılmaktadır(Desoli ve ark.,2017). Dolayısıyla robotik tedavi ve rehabilitasyon sistemlerine uyarlanabilir ve kullanılabilir.



KAYNAKLAR

- Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G. S., Davis, A., Dean, J., Devin, M., Ghemawat, S., Goodfellow, I., Harp, A., Irving, G., Isard, M., Jia, Y., Jozefowicz, R., Kaiser, L., Kudlur, M., Levenberg, J., Mane, D., Monga, R., Moore, S., Murray, D., Olah, C., Schuster, M., Shlens, J., Steiner, B., Sutskever, I., Talwar, K., Tucker, P., Vanhoucke, V., Vasudevan, V., Viegas, F., Vinyals, O., Warden, P., Wattenberg, M., Wicke, M., Yu, Y., Zheng, X., 2016. Tensorflow: large-scale machine learning on heterogeneous distributed systems . arxiv:1603.04467
- Alp, M.N., Oral, D. and Budak, T., 2007. Down sendromu ön tanılı 584 olguda sitogenetik çalışma. *Dicle Tıp Dergisi*, 34(4), pp.283-289.-
- Anavi, Y., Kogan, I., Gelbart, E., Geva, O., Greenspan, H., 2015. A comparative study for chest radiograph image retrieval using binary texture and deep learning classification.. In: Proceedings of the IEEE Engineering in Medicine and Biology Society, pp. 2940–2943. doi:10.1109/EMBC.2015.7319008
- Bastien, F., Lamblin, P., Pascanu, R., Bergstra, J., Goodfellow, I., Bergeron, A., Bouchard, N., Warde-Farley, D., Bengio, Y., 2012. Theano: new features and speed improvements. In: Proceedings of the Deep Learning and Unsupervised Feature Learning NIPS 2012 Workshop
- Bell, R.M. and Koren, Y., 2007. Lessons from the Netflix prize challenge. *Acm Sigkdd Explorations Newsletter*, 9(2), pp.75-79..
- Bottou, L., 2010. Large-scale machine learning with stochastic gradient descent. In Proceedings of COMPSTAT'2010 (pp. 177-186). Physica-Verlag HD.
- Breiman, L., 2001. Random Forests Machine Learning. 45: 5–32. *View Article PubMed/NCBI Google Scholar*.

- Broomhead, D. S., & Lowe, D. (1988). *Radial basis functions, multi-variable functional interpolation and adaptive networks* (No. RSRE-MEMO-4148). Royal Signals and Radar Establishment Malvern (United Kingdom)..
- Brunelli, R. and Poggio, T., 1993. Face recognition: Features versus templates. *IEEE transactions on pattern analysis and machine intelligence*, 15(10), pp.1042-1052.
- C. Sanderson, B.C. Lovell. [Multi-Region Probabilistic Histograms for Robust and Scalable Identity Inference](#). ICB 2009, LNCS 5558, pp. 199-208, 2009.
- Ciresan, D.C., Meier, U., Masci, J., Maria Gambardella, L. and Schmidhuber, J., 2011, July. Flexible, high performance convolutional neural networks for image classification. In *IJCAI Proceedings-International Joint Conference on Artificial Intelligence* (Vol. 22, No. 1, p. 1237).
- Cireşan, D., Meier, U. and Schmidhuber, J., 2012. Multi-column deep neural networks for image classification. *arXiv preprint arXiv:1202.2745*.
- Collobert, R., Kavukcuoglu, K., Farabet, C., 2011. Torch7: a matlab-like environment for machine learning. In: *Proceedings of the Advances in Neural Information Processing Systems*.
- David, A. and Lerner, B., 2005. Support vector machine-based image classification for genetic syndrome diagnosis. *Pattern Recognition Letters*, 26(8), pp.1029-1038.
- David, A. and Lerner, B., 2005. Support vector machine-based image classification for genetic syndrome diagnosis. *Pattern Recognition Letters*, 26(8), pp.1029-1038.
- Deng, J., Dong, W., Socher, R., Li, L.J., Li, K. and Fei-Fei, L., 2009, June. Imagenet: A large-scale hierarchical image database. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on* (pp. 248-255). Ieee.

- Erogul, O., Sipahi, M.E., Tunca, Y. and Vurucu, S., 2009, May. Recognition of down syndromes using image analysis. In Biomedical Engineering Meeting, 2009. BIYOMUT 2009. 14th National (pp. 1-4). IEEE.
- Farely, B.G. and Clark, W.A., 1954. Simulation of self-organizing systems by digital computers. *IEEE Transactions of Professional Group of Information Theory, PGIT-4*, pp.76-84.
- Fei-Fei, L., Fergus, R. and Perona, P., 2007. Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories. *Computer vision and Image understanding*, 106(1), pp.59-70.
- Ferry, Q., Steinberg, J., Webber, C., FitzPatrick, D.R., Ponting, C.P., Zisserman, A. and Nellåker, C., 2014. Diagnostically relevant facial gestalt information from ordinary photos. *Elife*, 3, p.e02020.
- Fukushima, K. (1986). A neural network model for selective attention in visual pattern recognition. *Biological Cybernetics*, 55(1), 5-15.
- Fukushima, K., Miyake, S., & Ito, T. (1983). Neocognitron: A neural network model for a mechanism of visual pattern recognition. *IEEE transactions on systems, man, and cybernetics*, (5), 826-834.
- G. Griffin, A. Holub, and P. Perona. Caltech-256 objectcategorydataset. Technical Report 7694, California Institute of Technology, 2007. URL <http://authors.library.caltech.edu/7694>.
- Goodfellow, I., Bengio, Y., Courville, A. and Bengio, Y., 2016. *Deep learning* (Vol. 1). Cambridge: MIT press.
- Guo, G., Li, S.Z. and Chan, K., 2000. Face recognition by support vector machines. In Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on (pp. 196-201). IEEE.
- Haykin, S. (1994). *Neural networks: a comprehensive foundation*. Prentice Hall PTR.

- He, K., Zhang, X., Ren, S. and Sun, J., 2016. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
- Hinton, G.E., Srivastava, N., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R.R., 2012. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580*.
- Huang, G.B., Mattar, M., Berg, T. and Learned-Miller, E., 2008, October. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. In *Workshop on faces in 'Real-Life' Images: detection, alignment, and recognition*. <http://vis-www.cs.umass.edu/lfw/>.
- Huang, G.B., Mattar, M., Berg, T. and Learned-Miller, E., 2008, October. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. In *Workshop on faces in 'Real-Life' Images: detection, alignment, and recognition*.
- Jarrett, K., Kavukcuoglu, K. and LeCun, Y., 2009, September. What is the best multi-stage architecture for object recognition?. In *Computer Vision, 2009 IEEE 12th International Conference on* (pp. 2146-2153). IEEE.
- Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., Guadarrama, S., Darrell, T., 2014. Caffe: convolutional architecture for fast feature embedding. In: Proceedings of the Twenty-Second ACM International Conference on Multimedia, pp. 675–678. doi:10.1145/2647868.2654889.
- Kainz, P., Pfeiffer, M., Urschler, M., 2015. Semantic segm
- Kaynak, O., & Efe, M. Ö. (2000). Yapay Sinir Ağları ve Uygulamaları. *BOĞAZİÇİ ÜNİVERSİTESİ*.
- Keskenler, M.F. and Keskenler, E.F., Geçmişten Günümüze Yapay Sinir Ağları ve Tarihçesi. *Takvim-i Vekayi*, 5(2), pp.8-18.
- Krizhevsky, A. and Hinton, G., 2009. *Learning multiple layers of features from tiny images* (Vol. 1, No. 4, p. 7). Technical report, University of Toronto.

- Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).
- Kurban, O.C., Yildirim, T. and Bilgiç, A., 2017, July. A multi-biometric recognition system based on deep features of face and gesture energy image. In *INnovations in Intelligent SysTems and Applications (INISTA), 2017 IEEE International Conference on* (pp. 361-364). IEEE.
- Lawrence S, Giles CL, Tsoi AC, Back AD. Face recognition: A convolutional neural-network approach. *IEEE transactions on neural networks*. 1997 Jan;8(1):98-113.
- Learned-Miller, E., Huang, G. B., RoyChowdhury, A., Li, H., & Hua, G. (2016). Labeled faces in the wild: A survey. In *Advances in face detection and facial image analysis* (pp. 189-248). Springer, Cham.
- LeCun, Y., Bengio, Y. and Hinton, G., 2015. Deep learning, *nature*, 521(7553), p.436.
- LeCun, Y., Huang, F.J. and Bottou, L., 2004. Learning methods for generic object recognition with invariance to pose and lighting. In *Computer Vision and Pattern Recognition, 2004. CVPR 2004. Proceedings of the 2004 IEEE Computer Society Conference on* (Vol. 2, pp. II-104). IEEE..
- LeCun, Y., Kavukcuoglu, K. and Farabet, C., 2010, May. Convolutional networks and applications in vision. In *ISCAS* (Vol. 2010, pp. 253-256). Rygl M, Kalousova J, Pycha K, Styblova J, Snaidauf J: Current results in treatment of omphalocele and gastroschisis. *Ceska Gynecol*, 69(1):55-59, 2004.
- Liu, D.C. and Nocedal, J., 1989. On the limited memory BFGS method for large scale optimization. *Mathematical programming*, 45(1-3), pp.503-528.
- Loos, H.S., Wiczorek, D., Würtz, R.P., von der Malsburg, C. and Horsthemke, B., 2003. Computer-based recognition of dysmorphic faces. *European Journal of Human Genetics*, 11(8), p.555.

- Lynch, L. and Berkowitz, R.L., 1989. First trimester growth delay in trisomy 18. *American journal of perinatology*, 6(02), pp.237-239.
- Murphy, K., 2012. Machine learning: a probabilistic approach. Massachusetts Institute of Technology, pp.1-21.
- Nakada, M., Wang, H. and Terzopoulos, D., 2017, July. AcFR: active face recognition using convolutional neural networks. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on* (pp. 35-40). IEEE
- Nie, F., Huang, H., Cai, X. and Ding, C.H., 2010. Efficient and robust feature selection via joint ℓ_2 , 1-norms minimization. In *Advances in neural information processing systems* (pp. 1813-1821).
- Nikkawa N, Merrotto E, Kajii T (1977) Origin of acrocentric trisomies in spontaneous abortuses. *Hum Genet* 40:73–78
- Nilson, N. J. (1965). Learning Machines. McGraw-Hill.
- Oztemel, E. (2003). Yapay Sinir Ağları, Papatya Yayıncılık. *Istanbul*. pg48.
- Parkhi, O.M., Vedaldi, A. and Zisserman, A., 2015, September. Deep face recognition. In *BMVC* (Vol. 1, No. 3, p. 6).
- Parkhi, Omkar M., Andrea Vedaldi, and Andrew Zisserman. "Deep face recognition." *BMVC*. Vol. 1. No. 3. 2015.
- Rajkomar, A., Lingam, S., Taylor, A.G., Blum, M., Mongan, J., 2017. High-throughout classification of radiographs using deep convolutional neural networks. *J. Digit. Imaging* 30, 95–101. doi:10.1007/s10278-016-9914-9.
- Rumelhart, D. E., & McClelland, J. L. (1986). Parallel distributed processing: explorations in the microstructure of cognition. volume 1. foundations.
- Rumelhart, D. E., & McClelland, J. L. (1986). Parallel distributed processing: explorations in the microstructure of cognition. volume 1. foundations.
- Rumelhart, D. E., & McClelland, J. L. (1986). Parallel distributed processing: explorations in the microstructure of cognition. volume 1. foundations.

- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M. and Berg, A.C., 2015. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), pp.211-252.
- Saraydemir, Ş., Taşpınar, N., Eroğul, O. and Kayserili, H., 2015. Effects of Training Set Dimension on Recognition of Dysmorphic Faces with Statistical Classifiers. *International Arab Journal of Information Technology (IAJIT)*, 12(2).
- Saraydemir, Ş., Taşpınar, N., Eroğul, O. and Kayserili, H., 2015. Effects of Training Set Dimension on Recognition of Dysmorphic Faces with Statistical Classifiers. *International Arab Journal of Information Technology (IAJIT)*, 12(2).
- Saraydemir, Ş., Taşpınar, N., Eroğul, O., Kayserili, H. and Dinçkan, N., kNN Sınıflandırıcı ile Dismorfik Görüntülerin Ayırıştırılması Discrimination of Dysmorphic Images with kNN Classifier.
- Schmidt, M., 2005. minFunc: unconstrained differentiable multivariate optimization in Matlab. *Software available at [http://www. cs. ubc. ca/~schmidtm/Software/minFunc. htm](http://www.cs.ubc.ca/~schmidtm/Software/minFunc.htm).*
- Shukla, P., Gupta, T., Saini, A., Singh, P., & Balasubramanian, R. (2017, March). A Deep Learning Frame-Work for Recognizing Developmental Disorders. In *Applications of Computer Vision (WACV), 2017 IEEE Winter Conference on* (pp. 705-714). IEEE.
- Simard, P.Y., Steinkraus, D. and Platt, J.C., 2003, August. Best practices for convolutional neural networks applied to visual document analysis. In *null* (p. 958). IEEE.
- Simonyan, K. and Zisserman, A., 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- Sipahi, M. E. (2008). Dismorfik hastalıkların görüntü analizi ile ayırt edilmesi (Yüksek lisans tezi). Başkent Üniversitesi/Fen Bilimleri Enstitüsü, Ankara.

- Specht, D. F. (1988, July). Probabilistic neural networks for classification, mapping, or associative memory. In *IEEE international conference on neural networks* (Vol. 1, No. 24, pp. 525-532).
- Specht, D. F. (1991). A general regression neural network. *IEEE transactions on neural networks*, 2(6), 568-576.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), 1929-1958.
- Sutton, S.G., Holt, M. and Arnold, V., 2016. "The reports of my death are greatly exaggerated"—Artificial intelligence research in accounting. *International Journal of Accounting Information Systems*, 22, pp.60-73.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A., 2015. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*(pp. 1-9).
- Taigman, Y., Yang, M., Ranzato, M.A. and Wolf, L., 2014. Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1701-1708).
- Tra, V., Kim, J., Khan, S.A. and Kim, J.M., 2017. Bearing fault diagnosis under variable speed using convolutional neural networks and the stochastic diagonal levenberg-marquardt algorithm. *Sensors*, 17(12), p.2834..
- Turing, A. (1950). Computing Machinery and Intelligence, in" *Mind*" 59 (236), 433-460. *Cerca con Google*.
- Vedaldi, A., Lenc, K.: MatConvNet: Convolutional Neural Networks for MATLAB. In: 692.qProceedings of the 23rd ACM International Conference on Multimedia. pp. 689 ACM, New York, NY, USA (2015)

- Virtaneva, K., Wright, F.A., Tanner, S.M., Yuan, B., Lemon, W.J., Caligiuri, M.A., Bloomfield, C.D., De La Chapelle, A. and Krahe, R., 2001. Expression profiling reveals fundamental biological differences in acute myeloid leukemia with isolated trisomy 8 and normal cytogenetics. *Proceedings of the National Academy of Sciences*, 98(3), pp.1124-1129.
- Widrow, B. and Hoff, M.E., 1960. Adaptive switching circuits. In 1960 WESCON Convention Record Part IV, pages 96--104. Reprinted in JA Anderson and E. Rosenfeld. *Neurocomputing: Foundations of Research*.
- Zhang, Z., 2016. Derivation of Backpropagation in Convolutional Neural Network (CNN).
- Zhao, Q., Rosenbaum, K., Sze, R., Zand, D., Summar, M. and Linguraru, M.G., 2013, February. Down syndrome detection from facial photographs using machine learning techniques. In Medical Imaging 2013: Computer-Aided Diagnosis (Vol. 8670, p. 867003). International Society for Optics and Photonics.(karşılaştırma makaleleri)
- Zhu, X. and Ramanan, D., 2012, June. Face detection, pose estimation, and landmark localization in the wild. In Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on (pp. 2879-2886). IEEE.
- <https://www.pixabay.com/>
- <https://www.pinterest.com>
- https://tr.wikipedia.org/wiki/Down_sendromu#/media/File:21_trisomy_-_Down_syndrome.png
- https://en.wikipedia.org/wiki/Convolutional_neural_network#/media/File:Typical_cnn.png



ÖZGEÇMİŞ

1991 yılında Denizli’de doğdu. İlk, orta ve lise öğrenimlerini Denizli ’de tamamladı. 2009 yılında Erciyes Üniversitesi Mühendislik Fakültesi Biyomedikal Mühendisliği Bölümü’nde lisans eğitime başladı. 2013 yılında tamamladığı lisans eğitimi sonrasında 2015 yılında Afyon Kocatepe Üniversitesi Fen Bilimleri Enstitüsü Biyomedikal Mühendisliği Anabilim dalında yüksek lisans eğitime başladı. 2017 yılında Çukurova Üniversitesi Fen Bilimleri Enstitüsü Biyomedikal Mühendisliği Anabilim dalına yatay geçişle yüksek lisans eğitime devam etti.