



**T.C.  
GEBZE TEKNİK ÜNİVERSİTESİ**

**Bilgisayar Mühendisliği Bölümü**

# **AGE ESTIMATION WITH RANKING-CNN**

**Burak Furkan AKŞAHİN**

**Danışman  
Doç. Dr. Erchan APTOULA**

**Ocak, 2019  
Gebze, KOCAELİ**





**T.C.  
GEBZE TEKNİK ÜNİVERSİTESİ**

**Bilgisayar Mühendisliği Bölümü**

**AGE ESTIMATION WITH  
RANKING-CNN**

**Burak Furkan AKŞAHİN**

**Danışman  
Doç. Dr. Erchan APTOULA**

**Ocak, 2019  
Gebze, KOCAELİ**



Bu çalışma ....../...../200.. tarihinde aşağıdaki jüri tarafından Bilgisayar Mühendisliği Bölümü’nde Lisans Bitirme Projesi olarak kabul edilmiştir.

Bitirme Projesi Jürisi

Danışman Adı		
Üniversite		
Fakülte		

Jüri Adı		
Üniversite		
Fakülte		

Jüri Adı		
Üniversite		
Fakülte		

## **ÖNSÖZ**

Proje seçim zamanında bana danışmanlık etmeyi kabul eden ve proje süresince bana yol gösterici olan danışman öğretmenim Sayın Doç. Dr. Erchan APTOULA'ya teşekkürlerimi sunarım.

Ayrıca eğitimim süresince bana her konuda tam destek veren aileme ve bana hayatlarıyla örnek olan tüm hocalarıma saygı ve sevgilerimi sunarım.

**Ocak, 2019**

**Burak Furkan Akşahin**

## İçindekiler

ÖNSÖZ.....	vi
ŞEKİL LİSTESİ.....	ix
KISALTMA LİSTESİ .....	xi
SEMBOL LİSTESİ.....	xii
ÖZET .....	xiii
SUMMARY .....	xiv
<b>1. GİRİŞ .....</b>	<b>1</b>
1.1 PROJE TANIMI .....	2
1.2 PROJENİN AMAÇLARI.....	3
<b>2. ÖN BİLGİLER.....</b>	<b>4</b>
2.1 VERİ SETLERİ .....	4
2.1.1 CHALEARN VERİ SETİ[8].....	4
2.1.2 FGNET AGİNG VERİ SETİ[12] .....	6
2.2 YÜZ TESPİTİ .....	7
2.2.1 VİOLA JONES ALGORİTMASI [9] .....	8
2.2.2 OPENCV DNN İLE YÜZ TESPİTİ [10] .....	9
2.3. DERİN ÖĞRENME.....	10
2.3.1. ALEXNET [11].....	11
2.3.2. VGG16 [15] .....	12
2.3.3. VGGFACE [17] .....	13
2.3.4. RANKİNG-CNN [7] .....	13
<b>3. DENEYLER .....</b>	<b>15</b>
3.1 KERAS VE TENSORFLOW KURULUMU.....	15
3.2 VERİ SETİ TEMİZLİĞİ VE YÜZ TESPİTİ .....	16
3.3 ALEXNET İLE EĞİTİMLER.....	17
3.4 VGG16 İLE EĞİTİMLER .....	18
3.4 VGGFACE İLE EĞİTİMLER .....	19

<b>3.5 RANKİNG-CNN IMPLEMENTASYONU .....</b>	<b>20</b>
<b>4. SONUÇ.....</b>	<b>22</b>
<b>KAYNAKLAR .....</b>	<b>23</b>
<b>EKLER.....</b>	<b>25</b>



## ŞEKİL LİSTESİ

Şekil 1.1.1 Ranking-CNN Modeli .....	3
Şekil 2.0.1 FGNet Örnek görüntüler .....	6
Şekil 2.0.2 Yüz tespiti nin önemi .....	7
Şekil 2.0.3 Viola Jones Algoritması .....	8
Şekil 2.0.4 AlexNet Mimarisi .....	11
Şekil 2.0.5 VGG16 Mimarisi [16] .....	12
Şekil 2.0.6 Ranking-CNN Etiketlendirme .....	13
Şekil 2.0.7 Ranking-CNN Kaynaştırma.....	14
Şekil 3.1 Tensorflow MNIST .....	15
Şekil 3.2 Yüz Tespiti Haar ve Res10 .....	16
Şekil 3.3 AlexNet Yapay Olarak Çoğaltılmamış Eğitim .....	17
Şekil 3.4 AlexNet Yapay Olarak Çoğaltılmış Eğitim .....	17
Şekil 3.5 VGG16 Yeni Veri Seti İle .....	18
Şekil 3.6 VGGFace ile Eğitim .....	19
Şekil 3.7 Ranking-CNN veri setleri .....	20
Şekil 3.8 Ranking-CNN Eğitim .....	21

## TABLO LİSTESİ

Tablo 2.1 ChaLearn Veri Seti Yaş Dağılımı.....	5
Tablo 2.2 FGNet Veri Seti Yaş Dağılımı.....	6
Tablo 2.3 Viola Jones İle Yüz Tespiti Sonuçları .....	8
Tablo 2.4 OpenCV DNN ile Yüz Tespiti Sonuçları .....	9

## KISALTMA LİSTESİ

<b>CNN</b>	: Convolutional Neural Network (Evrişimsel Yapay Sinir Ağı)
<b>BIF</b>	: Biologically Inspired Features
<b>SVM</b>	: Support Vector Machines(Destekçi Vectör Makineleri)
<b>SVR</b>	: Support Vector Regression(Destekçi Vectör Regresyonu)
<b>RCNN</b>	:Regional Convolutional Neural Network (Bölgesel Evrişimsel Yapay Sinir Ağı)
<b>CVPR</b>	: The Computer Vision Foundation
<b>LAP</b>	: Looking at People
<b>SLI</b>	: Scan Line Interleave
<b>API</b>	: Application Programming Interface
<b>MNIST</b>	: Mixed National Institute of Standarts and Technology

## SEMBOL LİSTESİ

<i>fk</i>	: Alt modelden elde edilen binary sonuç
<i>rk</i>	: Sistemin kaynaştırılmış sonucu

## ÖZET

İnsan yaşı; reklamcılık, eğlence, güvenlik gibi alanlarda insanları kategorize etmede kullanılan en önemli biyometrik ölçütlerden biridir. Bu yüzden Bilgisayar Bilimi için insanların yüz özelliklerini kullanarak, kişilerin yaş tespitini yapmak, önemli bir amaç haline gelmiştir. Yakın zamandaki araştırmalar göstermiştir ki insan yüzü kullanılarak yapılan yaş tespitlerinde Derin Öğrenmenin kullanılması, elde edilen başarıda büyük bir artış sağlanmıştır.

Bu projede Derin Öğrenme tekniklerinden olan Evrişimsel Yapay Sinir Ağı (CNN) temelli “ranking-CNN” ile yaş tespiti gerçekleştirilmiştir. Yapılan eğitimler ve deneyler sonucu ranking-CNN’in diğer CNN tabanlı modellerden daha iyi bir Derin Öğrenme tekniği olduğu açık ve net bir biçimde kanıtlanamamıştır. Ranking-CNN içerisinde toplam yaş grubu kadar basit ikili CNN modelleri içerdiğinden, bu modellerin ayrı ayrı eğitilmesinin de masrafı düşünüldüğünde; bütün yaş grupları için tek bir model eğitip sonuç almak daha doğru görünmektedir.

## **SUMMARY**

Human age is the one of the most important biometric criteria, used to categorize people for areas such as advertising, entertainment, security. So it has become an important goal for Computer Science, to achieve age estimation with facial features of people. Recent researches have shown that use of Deep Learning in age estimations using human faces has resulted in a significant increase in accuracy of predictions.

In this project, age estimation was made with Convolutional Neural Network (CNN) based ranking-CNN which is one of the Deep Learning techniques. As a result of the trainings and experiments, it has not been clearly proved that ranking-CNN is a better Deep Learning technique than other CNN based models. Since ranking-CNN's training requires training of basic binary CNN's separately; it seems more accurate to train a single model for all age groups.

## 1. GİRİŞ

Güvenlik, reklamcılık, eğlence gibi alanlarda insan kategorizasyonu için kullanılan insan yaşı, çok kritik bir biyometrik özelliktir. Şirketler, hedef kitlelerini belirleyebilmek için reklam ajanslarına ya da kendi iç kaynaklarına yüklü paralar harcayıp, insanları yaş gruplarına ayırıp, o yaş gruplarına özel anketler, çalışmalar hazırlatmaktadırlar.

Çoğu zaman insanların vatandaşlık bilgilerinden ya da bizzat kendisine sorularak elde edilebilen insan yaşı, bazı durumlarda kişiden bizzat temin edilememektedir. İnsanlar bir olay, nesne, kişi hakkındaki görüşlerini sosyal medya üzerinden paylaşırlarken bu görüşlerle birlikte kendi görüntülerini de paylaşmaktadırlar. Bu görüşler ve profil resimleri, şirketler ve devlet kurumları için büyük bir veri kaynağıdır.

İşte bu yüzden, insan yüzünü analiz ederek yaş tahmini yapabilmek Bilgisayar Bilimi için önemli bir hedef haline gelmiştir. 20. yy'da yaş tespiti için insan yüzünden elde edilecek öz nitelik vektörlerinin oluşturulması görüntü işleme tekniklerinden doku analizi ve kenar bulma yöntemleriyle yapılırken[1]; daha sonraları “Biologically Inspired Features” (BIF) [2] ortaya çıkarak, yaş tahmini probleminde kullanılmaya başlamıştır.

Yaş tahmininin, diğer önemli kısmı tahmin etme mekanizmasıdır ve genellikle bu tahmin, regresyon ya da sınıflandırma problemidir. Makine Öğrenmesinde regresyon için yaygın olarak kullanılan yöntem Destekçi Vektör Regresyonu(SVR) [2] iken sınıflandırma için en çok kullanılan yöntemler En Yakın Komşular[3] ile Destekçi Vektör Makineleri(SVM) [2]'dir.

Tahmin etme ile öz nitelik vektörü oluşturma kısımlarını birleştiren Derin Öğrenme yöntemlerinden Evrişimsel Yapay Sinir Ağı (CNN); görüntü sınıflandırma [4], nesne tanıma [4], görüntü anlamlandırma [5] gibi problemlerde makine öğrenmesinin yerini aldığı gibi insan yaşı tahmin etme probleminde de ana araç haline gelmiştir [6,7]. Derin Öğrenme tekniklerinin son on yılda bu kadar çok ön plana çıkmasındaki ana sebep, bilgisayarlara yaptırılan eğitimin CPU tarafından GPU tarafına geçmesi ve elimizde daha önce olmadığı kadar temizlenmiş, doğru etiketlenmiş verinin olması olarak söylenebilir.

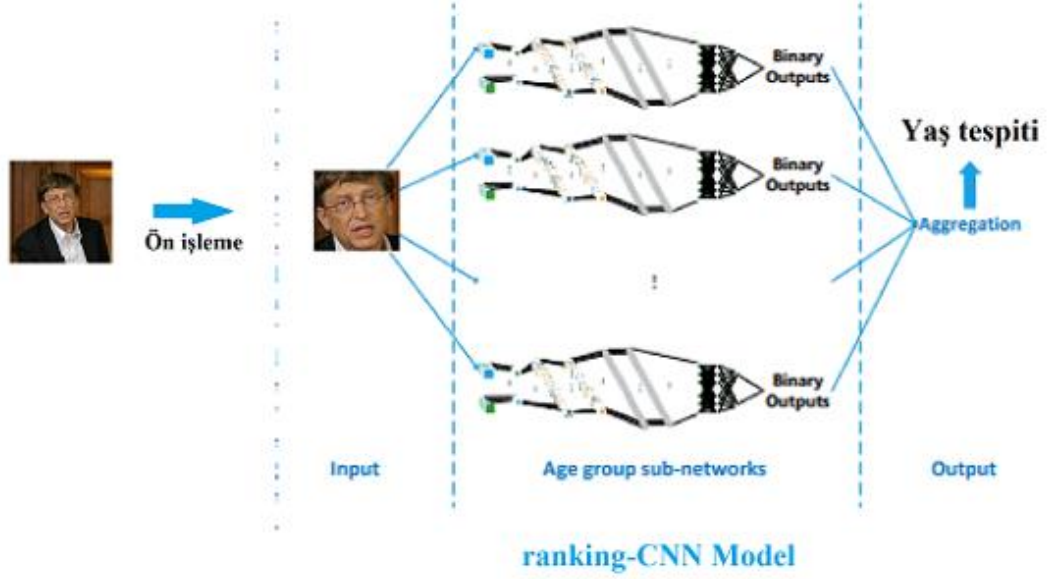
## **1.1 PROJE TANIMI**

İnsanların gerçek yaşının tahmin edilmesine dayalı bir çok proje gerçekleşirken, insanların görünen yaşı ile ilgili bu kadar çalışma yapılmamaktadır. İnsanlar birbirinden farklı bir biçimde yaşlanır, daha net olmak gerekirse insan yaşı problemi lineer bir problem değildir [7]. Bu projede, insanların görüntülerinin, görünen yaşı ile etiketlenmesiyle oluşturulmuş ChaLearn LAP 2016 [8] veri seti kullanarak, bu lineer olmayan problemi çözmek için farklı yaklaşımlarda bulunulmuştur.

Projenin ilk kısmı veri setindeki görüntülerden insan yüzlerinin çıkartılması işlemidir. Bu işlem için öncelikle Viola Jones [9] algoritması kullanılsa da bu algoritmanın verdiği başarı yeterli görülmeyip yüz eşleştirmek için özel eğitilmiş bir ResNet modeli [10] kullanılmıştır.

İkinci ve en önemli kısmı ise; bu tespit edilmiş, kesilmiş ve aynı boyuta getirilmiş insan yüzleri ile eğitilmiş bir Derin Öğrenme modeli oluşturmaktır. Bu model için önce AlexNet [11] düşünülmüş, ancak sonrasında yaşanan problemler nedeni ile önceden insan yüzü görüntüleri ile eğitilmiş ve ağırlıkları tanımlanmış VGG16 [6] modeline geçilmiştir.





Şekil 1.1.1 Ranking-CNN Modeli

Yapılan deneyler sonrasında, projenin asıl hedefi olan ranking-CNN modelinin tasarlanmasına başlanmıştır. Ranking-CNN, içerisinde toplam yaş grubu kadar model barındıran ve toplam sonucunu bu iç modellerin sonucunun kaynaştırılması ile elde eden bir Derin Öğrenme modelidir. Eğitim yapabilmek için mevcut olan veri setini, eğitim yapılacak olan alt modelin yaş grubuna göre iki parçaya ayırmak gereklidir. Bu veri seti varyasyonları, her yaş grubu için özel atanmış şekilde alt modeller eğitilmiş, böylece ranking-CNN modeli oluşturulmuştur.

## 1.2 PROJENİN AMAÇLARI

Bu proje sonucunda;

- Görüntüsü verilen bir kişinin, yüzü tespit edilecek ve ait olduğu yaş grubu sonuç olarak verilecektir.
- Canlı olarak bilgisayar kamerasından alınan görüntüde anlık olarak kişinin yüzü tespit edilecek ve yaş grubu sonuç olarak verilecektir.

## **2. ÖN BİLGİLER**

Bu başlık altında dönem içerisinde yapılan araştırmalar, veri setleri, kullanılan modeller, bu modellerin açıklamaları yapılacaktır.

### **2.1 VERİ SETLERİ**

Derin Öğrenmesi, makine öğrenmesine göre çok daha iyi sonuçlar vermesine rağmen, bunun için bazı fedakarlıklar istemektedir. Bunlardan birisi Derin Öğrenme modelinin tam doygunluğa ulaşması için çokça temiz ve doğru etiketlenmiş veriye ihtiyaç duymasıdır. Bu yüzden bu proje kapsamında birden fazla veri seti bulunup, eğitimde kullanılmıştır.

#### **2.1.1 ChaLearn Veri Seti[8]**

ChaLearn LAP, CVPR '16 etkinliğinin bir yarışması için hazırlanmış bir veri setidir. Diğer yaş tahmini için hazırlanan veri setlerinin aksine, içerisinde bulunan kişilerin gerçek yaşlarıyla değil görünen yaşlarıyla etiketlenmiş veriler içermektedir. Bir kişinin görünen yaşı, Facebook Implementation ve Amazon Mechanical Turk iş birliği ile birden fazla kişinin oylarının ortalaması alınarak etiketlenmiştir. Ayrıca ortalama yaş ile birlikte verilen oyların standart varyasyonu da yanında verilmiştir. İçerisinde toplamda 7591 ayrı insana ait birer görüntü mevcuttur.

**Tablo 0.1 ChaLearn Veri Seti Yaş Dağılımı**

Yaş Grubu	Train	Valid	Test	Toplam
0-6	252	93	153	498
7-13	157	50	108	315
14-20	404	170	109	683
21-27	1107	368	334	1809
28-34	848	332	460	1640
35-41	513	172	271	956
42-48	299	123	164	586
49-55	249	83	135	467
56-62	165	62	77	304
62+	119	47	157	323
Toplam	4113	1500	1978	7591

Bu veri setinin özellikle test için verilen görüntüleri çok zorlu görüntülerdir. Ancak test kümesindeki görüntü sayısını değiştirirsek problem değişeceğinden, test görüntülerine dokunulmamıştır. Eğitimden önce Train ve Valid kısımları gözden geçirilmiş, kişinin yüzünün tam olarak görünmediği veya tam olarak seçilemeyen görüntülerle birlikte çözünürlüğü çok düşük olan görüntüler temizlenmiştir. Bu temizleme sonrası Train 3806 görüntüye, Valid 1414 görüntüye düşmüştür.

### 2.1.2 FGNet Aging Veri Seti[12]



Şekil 2.0.1 FGNet Örnek görüntüler

FGNet Aging, Face And Gesture Recognition Working group üyesi, Andread Lanitis tarafından hazırlanmış bir veri setidir. ChaLearn LAP veri setinin aksine, FGNet Aging veri setinde, bir kişinin çocukluktan belli bir yaşına kadar görüntüleri mevcuttur.

Tablo 0.2 FGNet Veri Seti Yaş Dağılımı

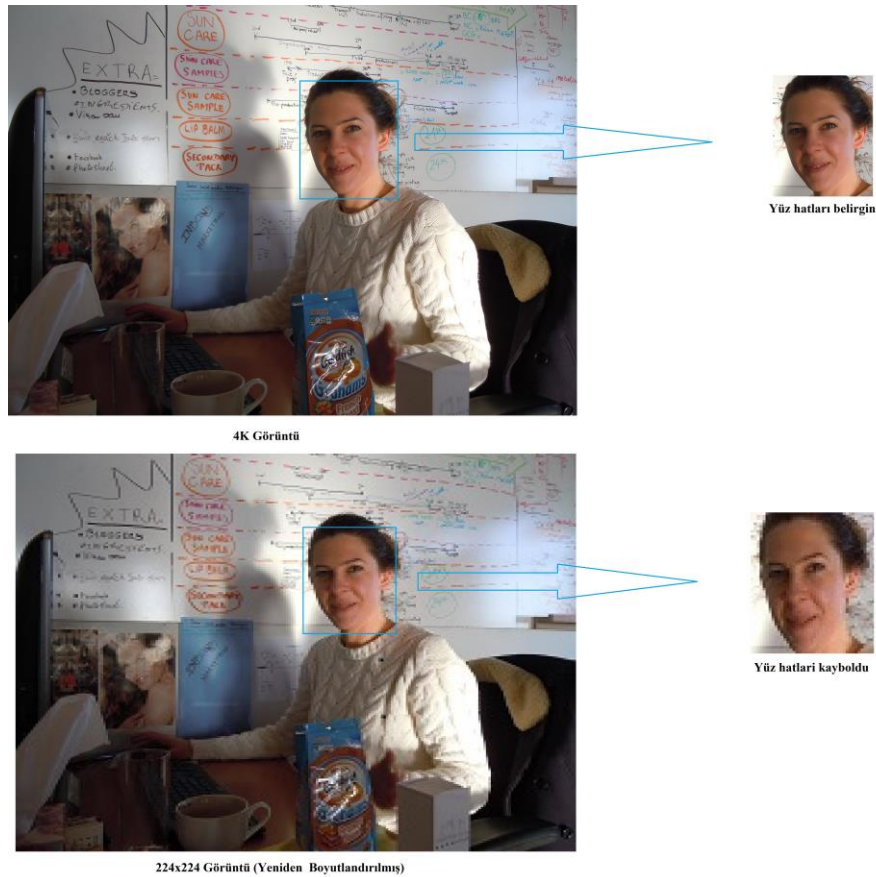
Yaş Grubu	Toplam
0-6	275
7-13	239
14-20	217
21-27	103
28-34	67
35-41	48
42-48	28
49-55	17
56-62	6
62+	3
Toplam	1003

Bu veri setindeki görüntüler, ChaLearn veri setinin train kısmı ile birleştirilerek eğitim için fazla veri elde edilmiştir.

## 2.2 YÜZ TESPİTİ

Eğitime başlamadan önce, elimizdeki verilerden insan yüzlerini bulmak projenin en önemli kısımlarından birisidir. Bunun sebebi, veri setlerindeki görüntülerin aynı çözünürlükte olmaması ve bu görüntülerdeki insan yüzlerinin aynı boyutta olmamasıdır.

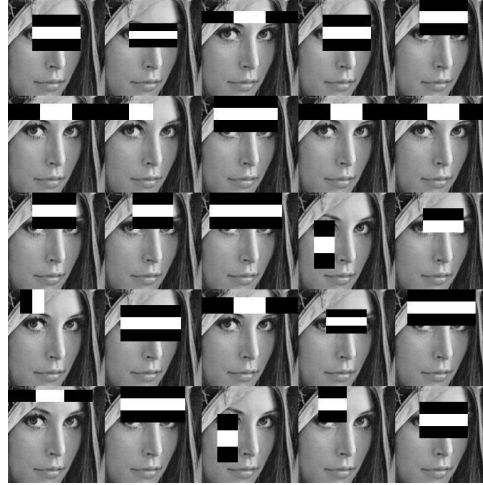
Bu dışarıdan bir problem olarak görülme de bu problemin var olduğunu anlayabilmek için eğitim yapacağımız modellerin aldığı giriş çözünürlüğünü bilmemiz gerekir. Bu proje kapsamında kullanılan modellerin giriş çözünürlüğü  $224 \times 224 \times 3$ 'tür. Örnek vermek gerekirse, 4k verilen bir görüntüde, çok uzaktaki bir kişinin yaşı tespit edilecekse; bütün görüntünün  $224 \times 224$  çözünürlüğüne getirilmesi, bu görüntüdeki insan yüzünün eğitim için kullanılamayacağı anlamına gelmektedir. Çünkü görüntüde ayrıntı boyutuna kadar küçültülmüş olacaktır.



Şekil 2.0.2 Yüz tespitinin önemi

### 2.2.1 Viola Jones Algoritması [9]

OpenCV içerisindeki Haar Cascade sınıflandırıcıların altındaki mantık, Viola Jones Algoritmasıdır. Bu sınıflandırıcılar, algoritmanın adını alan Paul Viola ve Micheal Jones tarafından oluşturulmuştur. Bu algoritmanın çalışma prensibi şu şekildedir: Öncelikle görüntünün integrali alınmaktadır. Böylece “Feature” denilen çerçeveler görüntü üzerinde tekrar tekrar gezdirilmek yerine, önceki aşamalardaki karşılaştırmaları tarayarak çok daha hızlı bir biçimde bu sınıflandırmayı yapabilir. Çerçeveler aydınlık ve karanlık bölgeler görüntüyle karşılaştırılır. Eğer uygun bir bölge yoksa, bu verilen görüntüde yüz yok denir. Eleme sistemi ile çalıştığından, çok hızlı bir yüz tespit algoritmasıdır.



Şekil 2.0.3 Viola Jones Algoritması

OpenCV'nin CascadeClassifierini kullanarak yapılan yüz tespiti, çok hızlı olmasına rağmen çok yanlış sonuçlar vermiştir. Özellikle veri setinin test kısmında, yüz bulunduğu yerlerde yüz bulunmaması, görüntülerde hiç yüz bulmaması gibi sorunlarla karşılaşmıştır. Bu yüzden Viola Jones algoritmasından vazgeçilip, farklı bir yüz tespiti sistemine geçilmiştir.

Tablo 0.3 Viola Jones İle Yüz Tespiti Sonuçları

	Train	Valid	Test	Total
Bulunan Yüz / Toplam Görüntü	2561/4113	1247/1500	1143 / 1978	%65

### 2.2.2 OpenCV DNN ile Yüz Tespiti [10]

OpenCV haar cascader ile birlikte Derin Öğrenmeye dayalı sınıflandırıcılar da bulundurmaktadır. Bu sınıflandırıcılar Caffé modelleri ile çalışmaktadırlar. Bunun için 2 ayrı dosyaya ihtiyaç vardır: Caffé .prototxt dosyası (Modelin mimarisini belirtir.) ve .caffemodel dosyası (Model içerisindeki ağırlıkları belirtir). Bu iki dosya da OpenCV'nin GitHub linkinden[13] temin edilebilmektedir.

Çalışma prensibi ise şu şekildedir: Görüntüden “blob” denilen parçalar alınır, bu parçalar Caffé dosyalarından çekilen modele verilir. Bütün yüz tespit edilmiş bloblara ne kadar emin olduğu sorulur, eğer verdiğimiz eşik değerinden daha yüksek bir kararlılığı varsa burada insan yüzü var denilebilir.

**Tablo 0.4 OpenCV DNN ile Yüz Tespiti Sonuçları**

	Train	Valid	Test	Total
Bulunan Yüz / Toplam Görüntü	3806/4113	1414/1500	1978 / 1978	%94

Yöntemi değiştirmek, yüz tespiti oranını %65'ten %94'e yükseltti.

### 2.3. DERİN ÖĞRENME

Proje kapsamındaki görüntülerin eğitimi Derin Öğrenme ile yapılmıştır. Derin Öğrenmenin temelini “Perceptron” adı verilen yapay sinir ağları oluşturur. Bu ağların birbiri arkasına eklenmesi, derinlik kavramını oluşturur. Derin Öğrenme konsepti aslında çok yeni bir konsept değildir. Bundan 76 sene önce nörobilimin ilk adımları[14] atılmasına rağmen, Derin Öğrenme son 20 yıla kadar fazla popüler değildi. Bunun sebepleri, ağları eğitmennin çok uzun sürmesi, çok miktarda veriye ihtiyaç duyması ve en önemli problemlerden biri olan “Back propagation”ın çözülmemesi idi.

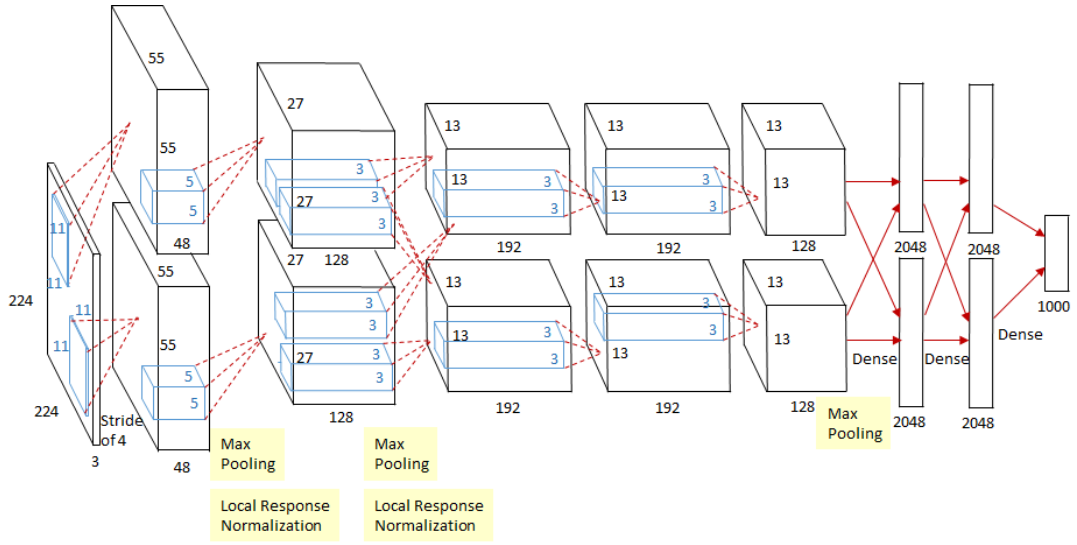
Grafik kartlarının model eğitiminde kullanılmaya başlanması, Google gibi büyük şirketlerin Büyük Veri’ye önem vermesi ile bir çok veri setinin ulaşılabilir olması ve bir çok şirketin Derin Öğrenme kütüphaneleri açık kaynak kod olarak paylaşması ile Derin Öğrenme çok yaygın olarak kullanılmaya ve konuşulmaya başlamıştır.

Bu başlık altında kullanılan Derin Öğrenme modellerinin açıklamalarını, Ranking-CNN mimarisi gibi konular anlatılacaktır.



### 2.3.1. AlexNet [11]

ImageNet yarışmasında ortaya çıkan AlexNet, yarışmadaki sonucundan öte kullanılan grafik kartının eğitim sırasındaki etkisini arttırması ile ünlenmiştir. İçerisinde birden fazla paralel ilerleyiş bulundural bu mimari, özellikle SLI grafik kartlarıyla eğitimi çok efektif hale getirmiştir.

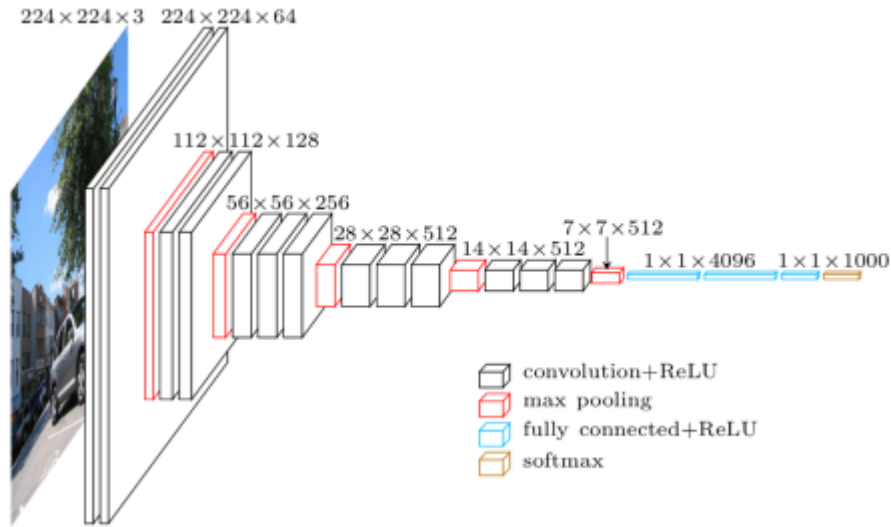


Şekil 2.0.4 AlexNet Mimarisi

Projenin Derin Öğrenme ile eğitim kısmına AlexNet ile başlanılmıştır ancak Keras kütüphanesi üzerinde AlexNet mimarisi implemente edilmediğinden çok iyi sonuçlar alınamamıştır. Bu alınan sonuçlar raporun tartışma ve sonuçlar kısmında yer almaktadır. Bu sebeple başka bir mimari kullanılması gerekliliği anlaşılmıştır.

### 2.3.2. VGG16 [15]

VGG16 ya da bir diğer ismi ile VGGNet, AlexNet gibi ImageNet yarışması için hazırlanmış bir Derin Öğrenme modelidir. VGGNet AlexNet'in aksine tek bir akış içerisinde ilerler ve toplam 16 adet evrimsel katmanı bulunur. Sınıflandırma için kullanılmasının yanında, en son katmanları çıkarılarak ön nitelik vektörleri oluşturmak için de kullanılabilir. Keras kütüphanesi içerisinde VGGNet'in ImageNet yarışmasında eğitilmiş halinin ağırlıkları ile bulundurmaktadır.



Şekil 2.0.5 VGG16 Mimarisi [16]

VGG16 ile yapılan eğitimler, AlexNet'e göre çok daha iyi sonuç vermesine rağmen, problemimiz insan yüzü olduğu için birkaç eğitim tekrarımda model “overfit” olmuştur. Bunu engelleyebilmek için ya VGG16 üzerinde yüzbinlerce veri ile insan yüzünü tanıtılması ya da bu şekilde eğitim yapılmış bir modelin ağırlıklarının bulunması gerekmektedir.

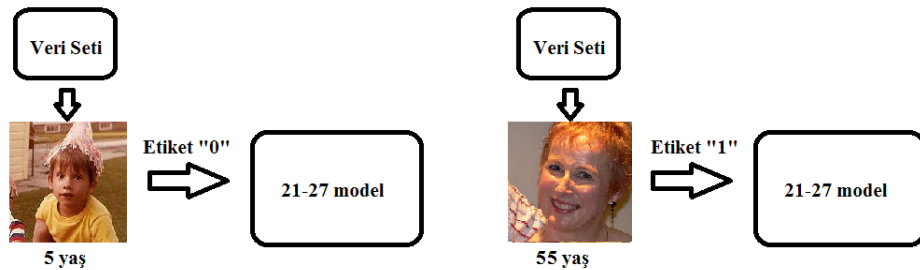
### 2.3.3. VGGFace [17]

Temeli VGG16 olan bir modelin, Labeled Faces in the Wild[18] ve Youtube Faces[19] veri setleri kullanılarak eğitilmesiyle oluşturulan VGGFace, Oxford Üniversitesi tarafından araştırma için kullanılmak üzere yayınlanmıştır. Ancak bu yayınlanan model Caffe modeli şeklinde olması sebebi ile Keras tarafından sağlanmamaktadır. GitHub[20] üzerinden elde edilen Keras implementasyonu kullanıldığında VGG16'ya göre küçükte olsa bir başarı artışı sağlanmıştır. Bu başarı sonrası, ranking-CNN içerisinde kullanılacak temel model VGGFace olarak seçilmiş olmaktadır.

### 2.3.4. Ranking-CNN [7]

Ranking-CNN, literatürde yaş tahmini probleminde “the state of art” olarak tanımlanmaktadır. Bu yüzden bu projede hali hazırda denenmiş olan AlexNet, VGG16 gibi mimariler ile Ranking-CNN kıyaslanacaktır.

Ranking-CNN içerisinde toplam (yaş grubu-1) kadar alt model barındırır. Bu modeller her yaş grubu için özel olarak ayrılmış veri setleriyle eğitilir. Örnek vermek gerekirse; (21-27) yaş grubu için olan alt model, 27 yaşından küçük kişilerin görüntüleri için “0” etiketi ile, 27 yaşından büyük kişilerin görüntülerini “1” etiketi ile eğitilir. Toplamda 10 yaş grubu için 9 adet alt modeli de eğitmemiz gerekmektedir. Bu yüzden elimizdeki eğitim veri setinin 9 adet farklı varyasyonu bulunacaktır.



Şekil 2.0.6 Ranking-CNN Etiketlendirme

Bu mimariden sonuç alabilmek için ise, alt modelleri nasıl tek tek eğittiysek bu sefer de her modelden tek tek sonuç almak gerekir. Bu sonuçlar ikili sistemde olacaktır. Bu sonuçlar doğruluk operatöründen geçirilir. Doğruluk operatörü, eğer sonuç 0'dan büyük ise 1, sıfıra eşit ya da küçük ise 0 döndürür. Doğruluk operatörlerinden geçirilen sonuçların toplamı, o görüntü için sistemin verdiği sonucu döndürür.

$$r(x_i) = 1 + \sum_{k=1}^{K-1} [f_k(x_i) > 0].$$

**Şekil 2.0.7 Ranking-CNN Kaynaştırma**

Görünebileceği üzere, toplam mimarinin hata payı alt modellerin hata payına bağlıdır. Alt modellerin hata payı ne kadar düşük olursa, toplam mimari o kadar doğru sonuç verir. Bu yüzden her bir alt modelin tam doygunluğa gelene kadar eğitilmesi gerekmektedir.

### 3. DENEYLER

#### 3.1 KERAS VE TENSORFLOW KURULUMU

Derin Öğrenme araçlarını kullanabilmek için öncelikle Tensorflow ve Keras kütüphanelerinin kurulması gerekmektedir. Bu kütüphanelerin Nvidia kartlarıyla uyumlu versiyonları (Tensorflow-gpu ve Keras-gpu) ile eğitimler işlemci üzerinde değil grafik kartı üzerinde yapılabilir. Windows işletim sistemi üzerinde Anaconda sanal lokasyonunda kütüphanelerin kurulumu sırasında dışarıdan anlaşılamayacak bir sorunla karşılaşılmıştır. Bu sorun Tensorflow-gpu kurulduktan sonra Keras-gpu kurulurken, Tensorflow-gpu kütüphanesini Tensorflow-base ile değiştirmesi kaynaklı bir sorundur. Bu yüzden Keras-gpu kurulduktan sonra, Tensorflow-gpu kütüphanesinin tekrar kurulması ile bu sorun çözülmüştür.

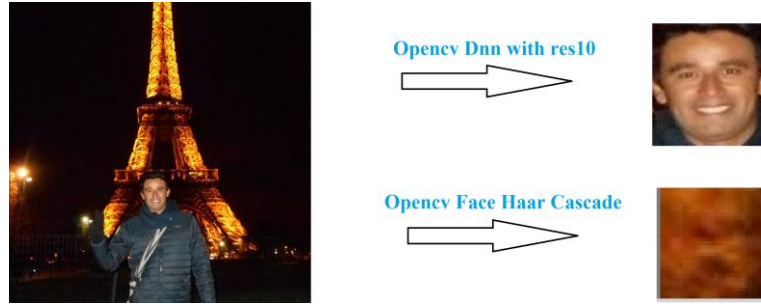
Tensorflow burada backend işlemlerini yürütürken, Keras frontend olarak kullanılır. Keras API'na alışabilmek için en güzel örneklerden olan MNIST veri seti ile eğitim gerçekleştirilmiş, böylece kütüphanelerin düzgün kurulduğundan emin olunmuştur.

```
Iter 194000, Minibatch Loss= 1109.702000, Training Accuracy= 0.96875
Iter 195840, Minibatch Loss= 627.169189, Training Accuracy= 0.96875
Iter 197120, Minibatch Loss= 500.741211, Training Accuracy= 0.95312
Iter 198400, Minibatch Loss= 226.016846, Training Accuracy= 0.98438
Iter 199680, Minibatch Loss= 1598.335938, Training Accuracy= 0.89062
Optimization Finished!
Testing Accuracy: 0.96875
```

Şekil 3.1 Tensorflow MNIST

### 3.2 VERİ SETİ TEMİZLİĞİ VE YÜZ TESPİTİ

Kurulumun doğruluğundan emin olduktan sonra, insan yüzlerinden yeni bir veri seti oluşturulması gerekmektedir. Bunun sebebi Ön Bilgiler başlığı altında açıklanmaktadır. Bunun için önce OpenCV Haar Cascadeleri kullanılmıştır. Veri setinin büyüklüğü düşünüldüğünde hızlı çalışması bir artı olarak gözükse de, çok fazla veri kaybı olması bu yaklaşımdan vazgeçilmesi gerekliliği doğurmuştur.

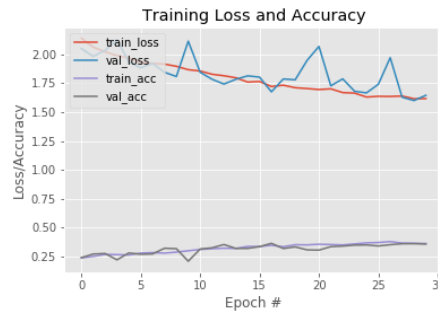


Şekil 3.2 Yüz Tespiti Haar ve Res10

Ardından araştırmalar sonucu OpenCV içerisinde yüz tespiti için özel eğitilmiş bir Caffe Res10 modeli bulunduğu keşfedilmiştir. Bu model OpenCV DNN modülü ile açılarak, veri集中的 insan yüzleri bulunmuş kesilmiş ve ayrılmıştır. Ne kadar Haar Cascade yönteminden daha iyi sonuç verse de insan yüzü bulamadığı ya da yanlış yerlerde insan yüzü bulunduğu durumlarla da karşılaşmıştır. Tekrar bir gözle temizlik yapıp veri setinin son hali ortaya çıkmıştır. (Bkz Tablo 2.4)

### 3.3 ALEXNET İLE EĞİTİMLER

AlexNet ile yapılan eğitimlerde, OpenCV Haar Cascade bulunmuş ve elle temizlenmiş veriler kullanılmıştır. AlexNet mimarisi Keras kütüphanesinde mevcut olmadığından önce Tensorflow “frontend”i ile daha sonra Keras kullanıcıların kendi imkanlarıyla oluşturduğu AlexNet Keras “frontend”leriyle denemeler yapılmış, AlexNet’in orijinaline en yakın olan implementasyonda testler gerçekleştirilmiştir. Bu testler proje aşamasının ilk kısımlarında geçtiğinden, toplam yaş grubunun o zaman 6 olduğunu belirtmek gerekir. Altta verilen isabet oranları bu 6 yaş grubuna göre verilmiştir.



Şekil 3.3 AlexNet Yapay Olarak Çoğaltılmamış Eğitim

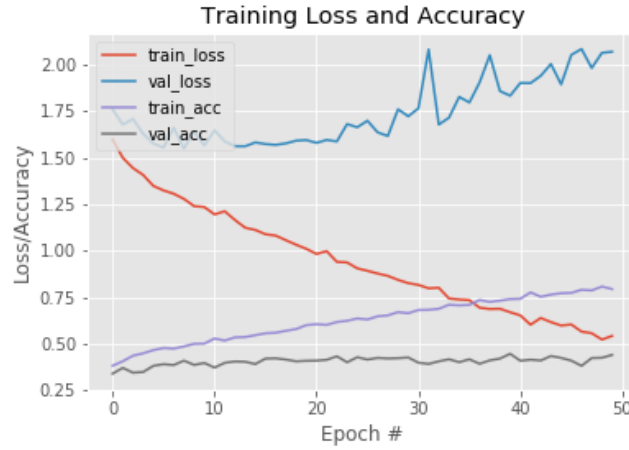
Validasyon başarı oranı %35'lere varsa da maalesef bu haliyle ancak %27 test başarı oranı alınabilmektedir. Bu da elimizdeki veriyle bu modelin “overfit” olduğu anlamına gelmesidir. Bu yüzden elimizdeki seçeneklerden birisi olan verinin yapay olarak çoğaltılması uygulanmıştır ve %27'lik test başarı oranı %35'e çıkarılmıştır.



Şekil 3.4 AlexNet Yapay Olarak Çoğaltılmış Eğitim

### 3.4 VGG16 İLE EĞİTİMLER

OpenCV Haar Cascade kullanılarak yapılan yüz tespiti ile elde edilmiş veriyi AlexNet'te eğitildiğinde alınan kötü sonuçlar, başka bir yüz tespiti yöntemi ile daha iyi bir mimari kullanmaya itmiştir. Araştırmalar sonucunda Derin Öğrenme mimarisi AlexNet'ten VGG16'ya geçmiştir. Önceden 6 olan yaş grubu sayısı, 10'a çıkararak, projede alınabilecek test isabet yüzdesinde çok büyük bir düşüş beklenmekteyken, hem mimarinin daha iyi olması hem de kullanılan verilerin sayısının ve kalitesinin artması test isabet yüzdesinde bir artış sağlamıştır.



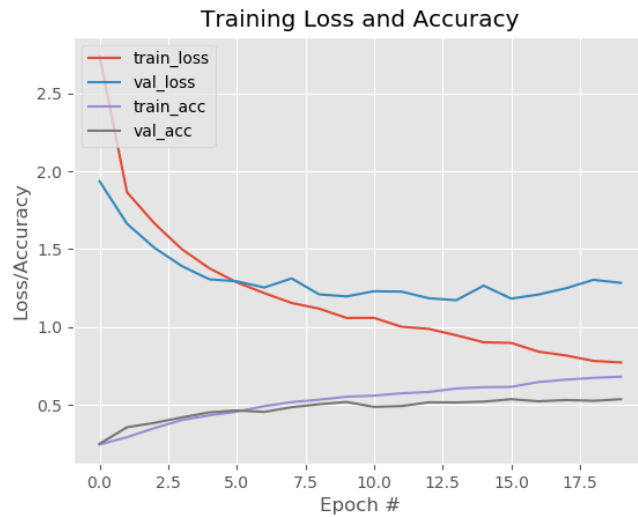
Şekil 3.5 VGG16 Yeni Veri Seti İle

Tablo 2.1'deki yaş grupları ile yapılan 50 tekrarlı bir eğitim sonrası, %53 validasyon isabet oranı ve %44 test isabet oranı elde edilmiştir. Ancak validasyon loss fonksiyonunun 25. Tekrardan sonra yükselişe geçmesi, bu modelin “overfit” olduğunu göstermektedir. Her ne kadar ImageNet yarışmasındaki eğitim ağırlıkları olan bir model kullansak da, insan yüzünü daha iyi tanıyan bir modele geçilmesi gerekliliği ortaya çıkmıştır.



### 3.4 VGGFACE İLE EĞİTİMLER

VGG16 mimarisinin ImageNet ağırlıklarıyla yapılan eğitimlerinde “overfit” olması, insan yüzüyle ön eğitime sokulmuş başka modellere yöneltmiştir. Araştırmalar sonucu Facebook firmasının DeepFace, Google firmasının FaceNet gibi mimarilerinin yanında akademik bir çalışma olan VGGFace bulunmuştur. VGGFace iki farklı büyük insan yüzü veri seti ile eğitilmiş VGG16 modelinin ağırlıklarının kullanıcıyla paylaşılmış halidir. Bu ağırlıklar Oxford Üniversitesi tarafından Caffé uzantısında verilse de, İstanbul Teknik Üniversitesi Öğretim Görevlisi Refik Can Mallı bu Caffé modellerini Keras arayüzü için implement etmiştir.



Şekil 3.6 VGGFace ile Eğitim

Alınan %53 validasyon isabet oranı ve %44 test isabet oranı VGGFace’in ranking-CNN temel modeli olarak seçilmesini sağlamıştır.

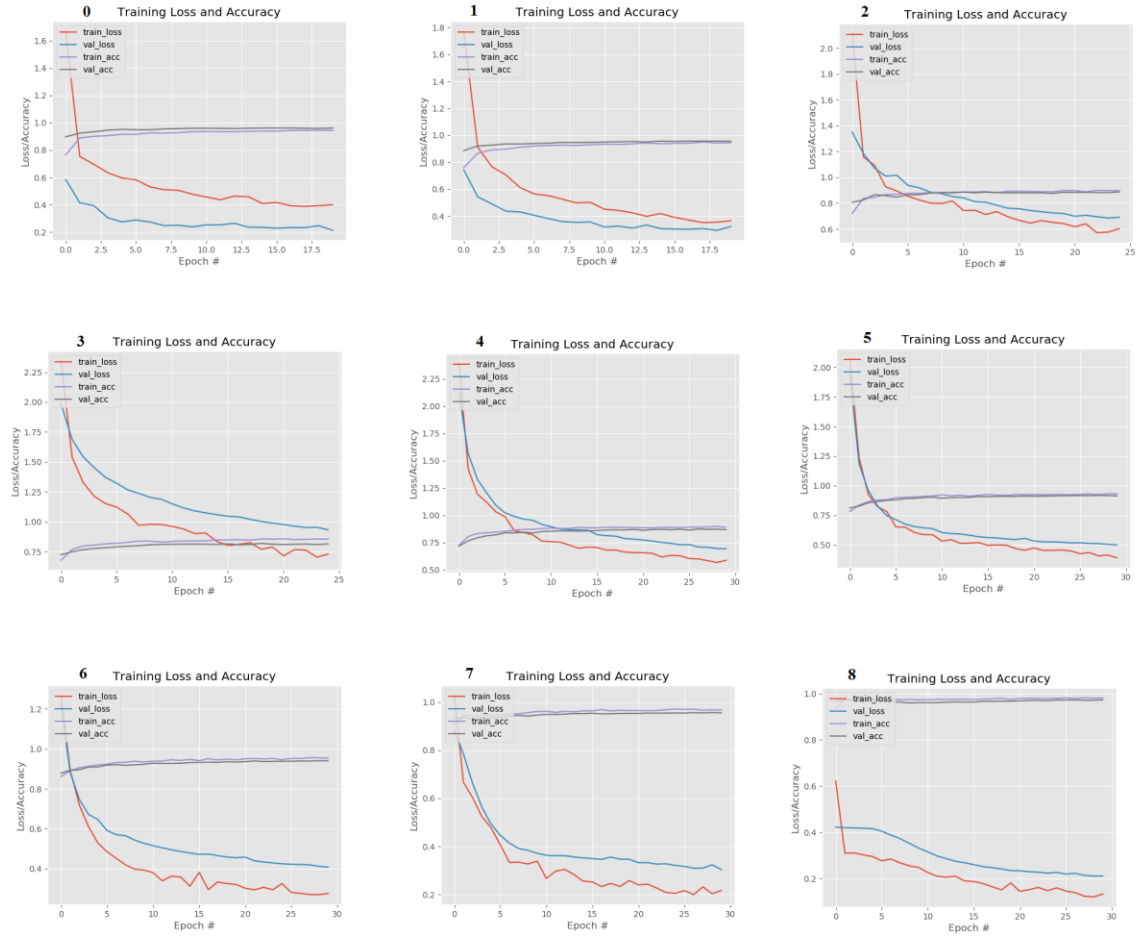
### 3.5 RANKİNG-CNN IMPLEMENTASYONU

Ranking-CNN için temel model seçildikten sonra, yapılması gereken ilk iş veri setinin ranking-CNN için uyumlu hale getirilmesi olmuştur. Bunun için toplam 10 yaş grubu için olan 9 alt model için 9 ayrı veri seti varyasyonu oluşturulmuştur.

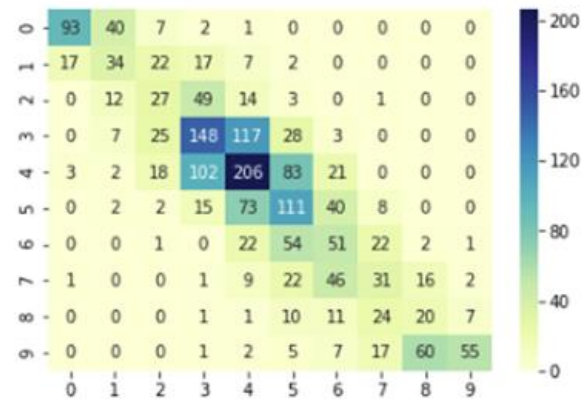


Şekil 3.7 Ranking-CNN veri setleri

Her bir rakam hangi yaş grubunun veri seti olduğunu temsil eder. 9. yaş grubu için bir veri setinin olmamasının nedeni, 8. yaş grubu için hazırlanan modelin şu soruyu sormasından kaynaklıdır: Bu gelen görüntü 8. yaş grubundan büyük bir kişiye mi ait? Eğer bu sorunun cevabı evet ise bu verilen görüntü 9. yaş grubuna ait bir görüntüdür. Bu yüzden toplam 10 yaş grubu için 9 adet model eğitilmiştir.



	Accuracy
0	95%
1	94%
2	90%
3	80%
4	82%
5	87%
6	90%
7	94%
8	94%



Şekil 3.8 Ranking-CNN Eğitim

#### 4. SONUÇ

Yapılan eğitimler ve denemeler sonucu, ranking-CNN alt modelleri doayunluğa ulaşmadığı ve bu alt modellerin kaynaştırması geliştirilmediği için, ranking-CNN birden çok sınıfı tek bir modelde eğiten modellere bir üstünlük kuramamıştır. Belki veri seti farklılığından belki de yukarda bahsettiğim durumdan ötürü, isabet oranında bir yükseliş görünmemiştir. Her bir alt modelin en az bir buçuk saat eğitilmesi ve toplamda 9 alt model olmasından kaynaklı eğitimin yaklaşık 14 saat sürmesi sebebi ile, tek bir VGGFace modelinde yaş tahmini yapmak, ranking-CNN kullanmaktan daha doğru görünmektedir.

Ancak ranking-CNN'in isabet oranı benim kaldığım yerden daha yukarıya taşınabilir görünüyor. Alt modeller tam doayunluğa getirilip, hata oranları düşürülürse, belki daha çok ve kaliteli görüntüye sahip bir veri seti ile birlikte daha yüksek isabet oranlarına ulaşılabilir. ChaLearn veri seti görünen yaş etiketli bir veri seti olduğu için herhangi bir yaş veri seti ile bu veri setini birleştirmek isabet oranında düşüşe sebep olmaktadır. Bir başka sebep de bu olabilir. Çünkü ranking-CNN eğitimi için daha çok veriye ihtiyaç olunduğu bilindiğinden FGNet veri seti ile ChaLearn veri seti birleştirilip ranking-CNN öyle eğitilmiştir.

## KAYNAKLAR

- [1] Young. H. KWON ve Niels da Vitoria LOBOY, "Age Classification from Facial Images", Computer Vision and Image Understanding, Vol. 74, No. 1, pp 1 -24, April 1999.
- [2] G. Guo, G. Mu, Y. Fu, and T. S. Huang. "Human age estimation using bio-inspired features". In IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 112–119, 2009
- [3] A. Gunay and V. V. Nabiyev. "Automatic age classification with lbp. In Computer and Information Sciences". ISCIS'08. 23rd International Symposium on, pages 1–4. IEEE, 2008.
- [4] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, Z. Huang, A. Karpathy, A. Khosla, M. Bernstein, A. C. Berg, and L. Fei-Fei. "ImageNet Large Scale Visual Recognition Challenge". International Journal of Computer Vision (IJCV), 115(3):211–252, 2015.
- [5] S. Zheng, S. Jayasumana, B. Romera-Paredes, V. Vineet, Z. Su, D. Du, C. Huang, and P. Torr. "Conditional random fields as recurrent neural networks". In International Conference on Computer Vision (ICCV), 2015.
- [6] Refik Can Malli, Mehmet Aygun, Hazim Kemal Ekenel; The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops, 2016, pp. 9-16
- [7] Chen, Shixing, et al. "Using ranking-cnn for age estimation." The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2017.
- [8] <http://chalearnlap.cvc.uab.es/dataset/19/description/> [Ziyaret Tarihi: 2 Ocak 2019]
- [9] Viola, Paul, and Michael Jones. "Rapid object detection using a boosted cascade of simple features." Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on. Vol. 1. IEEE, 2001.
- [10] <https://www.pyimagesearch.com/2018/02/26/face-detection-with-opencv-and-deep-learning/> [Ziyaret Tarihi: 2 Ocak 2019]
- [11] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems. 2012.
- [12] <http://ktisis.cut.ac.cy/handle/10488/4045> [Ziyaret Tarihi: 3 Ocak 2019]
- [13] [https://github.com/opencv/opencv/tree/master/samples/dnn/face\\_detector](https://github.com/opencv/opencv/tree/master/samples/dnn/face_detector) [Ziyaret Tarihi: 4 Ocak 2019]
- [14] McCulloch, Warren S., and Walter Pitts. "A logical calculus of the ideas immanent in nervous activity." *The bulletin of mathematical biophysics* 5.4 (1943): 115-133.
- [15] <https://keras.io/applications/#vgg16> [Ziyaret Tarihi: 3 Ocak 2019]
- [16] <https://www.cs.toronto.edu/~frossard/post/vgg16/> [Ziyaret Tarihi: 3 Ocak 2019]

- [17] Parkhi, Omkar M., Andrea Vedaldi, and Andrew Zisserman. "Deep face recognition." *BMVC*. Vol. 1. No. 3. 2015.
- [18] G. B. Huang, M. Ramesh, T. Berg, E. Learned-Miller. "Labeled faces in the wild: A database for studying face recognition in unconstrained environments". Technical Report 07-49, University of Massachusetts, Amherst, 2007.
- [19] L. Wolf, T. Hassner, I. Maoz. "Face Recognition in Unconstrained Videos with Matched Background Similarity". *Computer Vision and Pattern Recognition (CVPR)*, 2011.
- [20] <https://github.com/rcmalli/keras-vggface> [Ziyaret Tarihi: 3 Ocak 2019]

## **EKLER**

### **A. Lisans Bitirme Projesi Konusu Bildirme Formu**

<b>T.C. GEBZE TEKNİK ÜNİVERSİTESİ MÜHENDİSLİK FAKÜLTESİ BİLGİSAYAR MÜHENDİSLİĞİ LİSANS BİTİRME PROJESİ KONUSU BİLDİRME FORMU</b>	
<b>ÖĞRENCİ ADI SOYADI :</b>	
<b>ÖĞRENCİ NO :</b>	<b>İMZA:</b>
<b>PROJE KONU BAŞLIĞI:</b>	
<b>PROJENİN AMACI :</b> Projenin amacı ile yapılacak çalışmaların kısa özeti bu kısma yazılacaktır.	
<b>FAYDALANILACAK KAYNAKLAR :</b> Bitirme projesi önerisi hazırlanırken yararlanılan ve proje çalışmasına teşkil eden kaynaklar yazılacaktır.	
<b>PROJE DANIŞMANI:</b>	<b>İMZA:</b>
<b>BÖLÜM BAŞKANI:</b>	<b>İMZA:</b>

- Bu form bilgisayar ortamında 2 nüsha olarak düzenlenecek, bir nüsha bölüm başkanlığına ve bir nüsha proje danışmanına verilecektir.