

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

**SCORENET: DERİN KASKAT SKOR SEVİYESİ FÜZYONU İLE
KISITLANMAMIŞ GÖRÜNTÜLER ÜZERİNDEN KULAK TANIMA**

DOKTORA TEZİ

Ümit KAÇAR

Elektronik ve Haberleşme Mühendisliği Anabilim Dalı

Elektronik Mühendisliği Programı

ŞUBAT 2019

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

**SCORENET: DERİN KASKAT SKOR SEVİYESİ FÜZYONU İLE
KISITLANMAMIŞ GÖRÜNTÜLER ÜZERİNDEN KULAK TANIMA**

DOKTORA TEZİ

**Ümit KAÇAR
(504132206)**

Elektronik ve Haberleşme Mühendisliği Anabilim Dalı

Elektronik Mühendisliği Programı

Tez Danışmanı: Doç. Dr. Mürvet KIRCI

ŞUBAT 2019

İTÜ, Fen Bilimleri Enstitüsü'nün 504132206 numaralı Doktora Öğrencisi Ümit KAÇAR, ilgili yönetmeliklerin belirlediği gerekli tüm şartları yerine getirdikten sonra hazırladığı “SCORENET: DERİN KASKAT SKOR SEVİYESİ FÜZYONU İLE KISITLANMAMIŞ GÖRÜNTÜLER ÜZERİNDEN KULAK TANIMA” başlıklı tezini aşağıda imzaları olan jüri önünde başarı ile sunmuştur.

Tez Danışmanı : **Doç. Dr. Mürvet KIRCI**
İstanbul Teknik Üniversitesi

Jüri Üyeleri : **Prof. Dr. Ece Olcay GÜNEŞ**
İstanbul Teknik Üniversitesi

Doç. Dr. Ender Mete EKŞİOĞLU
İstanbul Teknik Üniversitesi

Dr. Öğr. Üyesi Bülent BOLAT
Yıldız Teknik Üniversitesi

Dr. Öğr. Üyesi Tolga İNAN
Çankaya Üniversitesi

Teslim Tarihi : 28 Aralık 2018
Savunma Tarihi : 22 Şubat 2019





Çok Değerli Eşime,



ÖNSÖZ

Doktora eğitimim süresince; sabrı, disiplini, kişiliği, tecrübesi, titizliği ile bu çalışmanın var olmasını sağlayan ve her türlü olanağı sunan, her daim saygı ile hatırlayacağım sevgili hocam Doç. Dr. Mürvet Kırıcı'ya teşekkür ederim.

Beni her daim destekleyen, zor günlerimde motivasyon kaynağım sevgili eşim Birgül Kaçar'a en içten dileklerle teşekkür ederim.

Doktora çalışmalarım boyunca görüş ve önerileri ile desteklerini esirgemeyen hocalarım Prof. Dr. Ece Olcay Güneş, Prof. Dr. Bilge Günsel ve Dr. Öğr. Üyesi Tolga İnan'a teşekkür ederim.

Birlikte çalıştığım ve silah arkadaşlarım Yük. Müh. Kd. Yzb. Cemal Gürpınar, Yük. Müh. Kd. Yzb. Deniz Kumlu ve Yük. Müh. Kd. Ütgm. Mahmut Burak Karadeniz'e desteklerinden dolayı teşekkür ederim.

Akademik çalışmalarım süresince fikir alışverişinde bulunduğum Araştırma Görevlisi Yüksek Mühendis arkadaşlarım Halil Durmuş, Muhammet Sait Altuner, Mesut Atasoyu, Emre Göncü, Emrah Abtioğlu, Bilgi Görkem Yazgaç, Sercan Aygün, Muhammet Şahin ile Dr. Ersin Alaybeyoğlu ve Dr. Musa Nurullah Yazar'a desteklerinden dolayı teşekkür ederim.

Bilim insanının yanında olan TÜBİTAK'a dört yıl boyunca 2211-E Doğrudan Yurt İçi Doktora Burs Programı desteğinden dolayı teşekkür ederim.

Yenilikçi yaklaşımları ile ufkumu açan başta Bilişim Teknoloji Yönetici Ekberjan Derman olmak üzere tüm Cutesafe Teknoloji A.Ş. ailesine teşekkür ederim.

Şubat 2019

Ümit KAÇAR
(Elektronik Mühendisi)



İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	vii
İÇİNDEKİLER	ix
KISALTMALAR	xi
SEMBOLLER	xiii
ÇİZELGE LİSTESİ.....	xv
ŞEKİL LİSTESİ.....	xvii
ÖZET.....	xix
SUMMARY	xxiii
1. GİRİŞ	1
1.1 Motivasyon.....	3
1.2 Literatür Katkısı	5
1.3 Tezin Bölümleri.....	6
2. KULAK BİYOMETRİSİ	9
2.1 Kulağın Embriyolojik Gelişimi ve Yapısı.....	9
2.2 İlgili Çalışmalar	10
3. KULAK TANIMA YÖNTEMLERİ.....	17
3.1 Skor Seviyesi Füzyon Mekanizmaları.....	18
3.1.1 Klasik skor seviyesi füzyon	18
3.1.2 Hiyerarşik skor seviyesi füzyon	19
3.1.3 Kaskat skor seviyesi füzyon.....	19
3.1.4 Hibrit skor seviyesi füzyon	20
3.2 Önerilen Mimari – “ScoreNet”	21
3.2.1 Modalite havuzunun oluşturulması.....	23
3.2.1.1 Renk haritası.....	24
3.2.1.2 Görüntü boyutu	25
3.2.1.3 Ön işleme ve görüntü iyileştirme	25
3.2.1.4 Öznetelik çıkarımı ve parametre setleri	26
3.2.1.5 Öznetelik azaltılması.....	29
3.2.1.6 Sınıflandırıcı.....	29
3.2.1.7 Metrik uzayı	31
3.2.2 Önerilen derin kaskat skor seviyesi füzyon algoritması	32
4. DENEY KURULUMU.....	39
4.1 AWE Kulak Veri Seti.....	39
4.2 WPUTE Kulak Veri Seti	40
4.3 UERC Kulak Veri Seti	40
4.4 Çalışma Ortamı	41
4.5 Performans Değerlendirilmesi.....	41
5. DENEY SONUÇLARI.....	43
6. SONUÇLAR VE ÖNERİLER	53
KAYNAKLAR	57

EKLER.....	65
ÖZGEÇMİŞ.....	75



KISALTMALAR

AOVY	: Ayırt Edici Ortak Vektör Yaklaşımı (Discriminative Common Vector Approach)
AutoFL	: Automated Fusion Learning (Otomatik Füzyon Öğrenme)
AWE	: Annotated Web Ears
BSIF	: Binarised Statistical Image Features (İkili İstatistiksel Görüntü Özellikleri)
BYDAA	: Bütünsel Yerel Doğrusal Ayırtaç Analizi (Complete Global-Local Linear Discriminant Analysis)
CCTV	: Close Circuit Television (Kapalı Devre Televizyonu)
CLAHE	: Contrast Limited Adaptive Histogram Equalization (Kontrast Sınırlı Uyarlamalı Histogram Eşitleme)
CMC	: Cumulative Matching Characteristic (Kümülatif Eşleştirme Karakteristiği)
CNN	: Convolutional Neural Network (Evrişimsel Sinir Ağları)
CPU	: Graphics Processing Unit (Grafik İşlem Birimi)
DAA	: Doğrusal Ayırtaç Analizi (Linear Discriminant Analysis)
DCSLF	: Deep Cascade Score Level Fusion (Derin Kaskat Skor Seviyesi Füzyon)
DVM	: Destek Vektör Makineleri
GAA	: Genelleştirilmiş Ayırtaç Analizi (Generalized Discriminant Analysis)
GABOR	: Gabor Filtreleri
GPU	: Central Processing Unit (Merkezi İşlem Birimi)
HOG	: Histogram of Gradients (Yönlü Gradyan Histogramı)
IJCB	: International Joint Conference on Biometrics (Uluslararası Biyometrik Ortak Konferansı)
ILSVRC	: ImageNet Large-Scale Visual Recognition Challenge
LBP	: Local Binary Pattern (Yerel İkili Örüntü)
LPQ	: Local Phase Quantisation (Yerel Faz Nicelemesi)
LWIR	: Long-Wave Infrared (Uzun Dalga Kızılötesi Işıma)
MWIR	: Medium-Wave Infrared Radiation (Orta Dalga Kızılötesi Işıma)
PTZ	: Pan-Tilt-Zoom (Kamerayı sırasıyla sağa-sola, yukarı-aşağı, yakınlaştırma-uzaklaştırma hareket ettirme özelliği)
ScoreNet	: Modalite havuzu ve derin kaskat skor seviyesi füzyon öğrenme algoritmasını birlikte içeren önerilen mimarinin adı
SM	: Score Matching (Skor Eşleştirme)
TBA	: Temel Bileşen Analizi (Principal Component Analysis)
UERC	: Unconstrained Ear Recognition Challenge (Kısıtlanmamış Görüntüler Üzerinde Kulak Tanıma Yarışması)
ULDP	: Unconstrained and Limited Data Problem (Kısıtlanmamış ve Sınırlı Veri Problemi)
WPU	: West Pomeranian University of Technology Ear



SEMBOLLER

D_s	: Modalite havuzu için kullanılan eğitim ve validasyon setinin birleşimi
D_{sx}	: Modalite havuzu için kullanılan eğitim seti
D_{sy}	: Modalite havuzu için kullanılan validasyon seti
D_T	: Füzyon öğrenme için test seti
i	: İndeks (1, 2, 3, ..., n)
$M_i (D_{sx}, D_{sy})$	: Skor eşleştirme matrisinin elde edilmesi için kullanılan i.inci modalite fonksiyonu
SM_i	: i'nci modalitenin skor eşleştirme matrisi
$\widehat{SM}_i$	: i'nci modalitenin normalize edilmiş skor eşleştirme matrisi
μ	: Ortalama
σ	: Varyans
SLF_i	: i'nci skor seviyesi füzyon matrisi
$\widehat{SLF}_i$	: i'nci normalize edilmiş skor seviyesi füzyon matrisi
β_i	: i'nci füzyon öğrenme oranı (başlangıç değeri $\beta_1 = 1$)
$normalise(X)$	: X matrisinin normalize işlemi için kullanılan fonksiyon



ÇİZELGE LİSTESİ

Sayfa

Çizelge 1.1 : CCTV kamera çeşitleri.	2
Çizelge 1.2 : Kısıtlanmamış yüz veri setleri ve özellikleri.	2
Çizelge 1.3 : Kısıtlanmamış kulak veri setleri ve özellikleri.	3
Çizelge 2.1 : Kulak biyometrisi içeren çoklu biyometrik sistemler.	14
Çizelge 3.1 : DCSLF algoritmasının sözde kodu.	38
Çizelge 5.1 : Kısıtlanmamış kulak tanıma için genel performans değerlendirmesi. .	44
Çizelge 5.2 : Kısıtlanmamış kulak tanıma için ölçeklenebilirlik değerlendirmesi. ...	44
Çizelge 5.3 : Oryantasyon ve normalizasyon işlemlerinden sonra kısıtlanmamış kulak tanıma için genel performans değerlendirmesi.	47
Çizelge 5.4 : Oryantasyon ve normalizasyon işlemlerinden sonra kısıtlanmamış kulak tanıma için ölçeklenebilirlik değerlendirmesi.	48
Çizelge 5.5 : Önerilen yöntemlerin görüntü başına ortalama test süresi.	50
Çizelge A.1 : “ScoreNet-1” in içeriği ile detaylı validasyon ve test sonuçları.	68
Çizelge A.2 : “ScoreNet-2” nin içeriği ile detaylı validasyon ve test sonuçları.	70
Çizelge A.3 : “ScoreNet-3” ün içeriği ile detaylı validasyon ve test sonuçları.	72
Çizelge A.4 : “ScoreNet-4” ün içeriği ile detaylı validasyon ve test sonuçları.	74



ŞEKİL LİSTESİ

Sayfa

Şekil 2.1 : Kulağın embriyolojik gelişimi.....	9
Şekil 2.2 : Kulağın yapısı.	10
Şekil 2.3 : Çoklu biyometrik füzyon stratejileri.	12
Şekil 2.4 : Farklı füzyon seviyelerinin kategorizasyonu.	14
Şekil 3.1 : Klasik skor seviyesi füzyon mekanizması.....	18
Şekil 3.2 : Hiyerarşik skor seviyesi füzyon mekanizması.	19
Şekil 3.3 : Kaskat skor seviyesi füzyon mekanizması.	20
Şekil 3.4 : Hibrit skor seviyesi füzyon mekanizması.	20
Şekil 3.5 : Önerilen mimari – “ScoreNet”.	22
Şekil 3.6 : Renk haritaları.	24
Şekil 3.7 : Popüler CNN mimarileri.	28
Şekil 3.8 : SM matrisi kombinasyon çeşitleri.	33
Şekil 3.9 : DCSLF algoritmasının akış şeması.	37
Şekil 4.1 : AWE kulak veri setinden örnek görüntüler.....	40
Şekil 4.2 : WPUTE kulak veri setinden örnek görüntüler.	40
Şekil 4.3 : UERC kulak veri setinden örnek görüntüler.	41
Şekil 5.1 : “ScoreNet-1” in tanıma oranı değişim grafiği.	45
Şekil 5.2 : “ScoreNet-2” nin tanıma oranı değişim grafiği.	46
Şekil 5.3 : Yerel işaretçiler ile kulak oryantasyon ve normalizasyon işlemleri.....	46
Şekil 5.4 : “ScoreNet-3” ün tanıma oranı değişim grafiği.....	48
Şekil 5.5 : “ScoreNet-4” ün tanıma oranı değişim grafiği.....	49
Şekil 5.6 : “ScoreNet” ler ile tüm diğer yöntemlerin CMC eğrileri.....	51
Şekil A.1 : Validasyon ve test seti için “ScoreNet-1” in CMC eğrileri.	67
Şekil A.2 : Validasyon ve test seti için “ScoreNet-2” nin CMC eğrileri.....	69
Şekil A.3 : Validasyon ve test seti için “ScoreNet-3” ün CMC eğrileri.....	71
Şekil A.4 : Validasyon ve test seti için “ScoreNet-4” ün CMC eğrileri.	73



SCORENET: DERİN KASKAT SKOR SEVİYESİ FÜZYONU İLE KISITLANMAMIŞ GÖRÜNTÜLER ÜZERİNDEN KULAK TANIMA

ÖZET

Günümüzün en büyük problemlerinden biri kişisel verilerin saklanması, doğrulanması ve yapay bilgilerden doğru bir biçimde ayrıştırılmasıdır. Kişisel kimlik bilgisi söz konusu olduğunda da biyometrik veriler ilk akla gelen veriler olmaktadır. Biyometrik verilerin elde edilmesinde yaygın olarak kullanılan kapalı devre televizyon (CCTV) kameralar şüpheli kişi tespit ve tanıma için oldukça büyük bir potansiyele sahiptir. CCTV kameralar ile şüpheli kişilerin yüz ve kulak biyometrik verileri elde edilebilmektedir. Ancak, gerçek dünya senaryolarında biyometrik verinin poz, aydınlatma, arka plan, düşük çözünürlük, oklüzyon vb. nedenlerden dolayı kaliteli olarak elde edilmesi oldukça zordur. Ayrıca güvenlik ve emniyet birimlerinin veri tabanında, bireylerin biyometrik verisi sınırlı sayıda mevcuttur. Bununla birlikte, şüpheli kişilerin doğal olarak işbirlikçi olmamaları nedeni ile şüpheli kişilerden elde edilecek biyometrik verinin kalitesi oldukça düşük olabilmektedir.

Tüm bu bahsedilen sorunlar “*Kısıtlanmamış ve Sınırlı Veri Problemi (ULDP)*” olarak tanımlanmıştır. Bu problemin üstesinden gelebilmek için “*ScoreNet*” adı verilen yeni bir mimari önerilmiştir. “*ScoreNet*” mimarisi hem modalite havuzunu hem de füzyon öğrenme algoritmasını içermektedir.

Füzyon öğrenme algoritması için “*Derin Kaskat Skor Seviyesi Füzyon (DCSLF)*” algoritması geliştirilmiş; modalite havuzu için de çeşitlilik bakımından zengin ve modüler yapıya sahip özgün bir modalite havuzu oluşturulmuştur.

“*ScoreNet*”, bugüne kadar geliştirilen kulak tanımlarını değerlendirmek için en zor veri tabanı olarak kabul edilen “*Kısıtlanmamış Görüntüler Üzerinde Kulak Tanıma Yarışması (UERC)*” kulak veri seti kullanılarak değerlendirilmiştir.

Genel performans ve ölçeklenebilirlik olmak üzere iki deney gerçekleştirilmiştir. Ayrıca, yüz tanıma literatüründe standart hale gelen ve kulak tanıma performansının iyileştirilmesini sağlayan yerel işaretçilere dayalı oryantasyon ve normalizasyon işlemleri uygulandıktan sonra söz konusu iki deney tekrarlanmıştır.

Genel performans için yapılan birinci deney, bireyin biyometrik özelliklerinin veri tabanında mevcut olduğu kapalı set kimliklendirme senaryosunu temsil etmektedir. Genel performans değerlendirme sonucunda, önerilen “ScoreNet-1”, UERC kulak tanıma sonuçlarına göre Sıra-1’de en az %9 daha iyi performans göstermiştir.

Ölçeklenebilirlik değerlendirmesi için yapılan ikinci deney, bireyin veri tabanında olup olmadığını belirleyen ve veri tabanındaki kişinin kimliğini teşhis eden açık set kimliklendirme senaryosuna uygundur. Ölçeklenebilirlik değerlendirmesi sonuçlarına göre “ScoreNet-2”, UERC kulak tanıma yöntemlerinden Sıra-1’de en az %12,5’luk bir iyileşme sağlamıştır.

Yerel işaretçilere dayalı oryantasyon ve normalizasyon işlemleri uygulandıktan sonra iki deney tekrar gerçekleştirilmiş olup “ScoreNet-3” ve “ScoreNet-4”, UERC sonuçlarına göre sırasıyla en az %23 ve %19’luk bir iyileştirme sağlarken; oryantasyon ve normalizasyon işlemlerinden sonra el yapımı ve evrimsel sinir ağları (CNN) tabanlı derin öğrenme algoritmaları içeren daha yeni bir çalışmadan da sırasıyla %12 ve %3,5’luk bir iyileştirme gerçekleştirmiştir.

Böylece “ScoreNet” mimarisi kısıtlanmamış görüntülerden oluşan ve bilinen en zorlu kulak veri seti olan UERC kulak veri setinde en son teknoloji doğruluğunu elde etmiştir. Ancak, “ScoreNet” mimarisinin oluşturulması ve eğitilmesi önemli bir iş yükü de getirmektedir. “ScoreNet” mimarisi için modalite havuzunun oluşturulması ve eğitim sürecinin uzun sürmesi bu mimarinin dezavantajlarıdır.

“ScoreNet” mimarisi kısıtlanmamış görüntüler üzerinde kulak tanıma için ilk “Otomatik Füzyon Öğrenme (AutoFL)” yaklaşımını temsil etmektedir. Aynı zamanda bu mimarinin gelecekteki “AutoFL” uygulamalarına öncülük edeceği değerlendirilmektedir.

“ScoreNet” mimarisi modüler bir yapıda olmasından dolayı geliştirilmeye yatkındır. Füzyon çalışmalarının performansının iyileştirilmesi için gerekli en önemli ihtiyaç çeşitliliktir. Daha iyi sonuçlar için “ScoreNet” mimarisinin modalite havuzunun çeşitliliği genişletilebilir. Bununla birlikte, eğitim ve validasyon (doğrulama) setinde içerik olarak karmaşıklık artırılarak daha iyi füzyon öğrenme gerçekleştirilebilir.

Sonuç olarak, “*ScoreNet*”in özgün mimarisi ile kısıtlanmamış görüntüler üzerinde kulak tanıma alanında önemli bir iyileştirme sağladığı UERC deney sonuçları ile kanıtlanmıştır.





SCORENET: UNCONSTRAINED EAR RECOGNITION WITH DEEP CASCADE SCORE LEVEL FUSION

SUMMARY

Nowadays, closed circuit television (CCTV) cameras, which are widely used in the electronic security industry, have a tremendous potential for detecting and recognizing suspicious people.

Face and ear biometric data of suspicious persons can be obtained with CCTV cameras. However, in real-world scenarios, biometric data is exposed to exposure, lighting, background, low resolution, occlusion, and so on. It is very difficult to obtain quality as a result.

There is also a limited number of biometric data of individuals in the database of security and security units. However, the quality of biometric data from suspicious persons has the potential to be very low due to the fact that suspicious persons are not naturally cooperative.

All of these problems are defined as Unconstrained and Limited Data Problem (ULDP). To overcome this problem, a new architecture called “*ScoreNet*” was proposed. The “*ScoreNet*” architecture includes both the modality pool and the fusion learning algorithm.

Proposed “*Deep Cascade Score Level Fusion (DCSLF)*” algorithm for fusion learning algorithm; for the modality pool, a unique pool of modality (rich in diversity) and modular structure was created.

The “*ScoreNet*” was evaluated using the Unconstrained Ear Recognition Challenge (UERC) ear data set on images, which is considered to be the most difficult database to evaluate ear recognition ever developed.

Two experiments were performed, namely general performance and scalability. In addition, these two experiments were repeated after applying orientation and

normalization procedures based on landmarks which became standard in facial recognition literature and improved ear recognition performance.

The first experiment for general performance represents the closed set identification scenario in which the individual's biometric properties are present in the database. The “*ScoreNet-1*”, which was proposed as a result of the overall performance evaluation, performed at least 9% better in the Rank-1 compared to the UERC ear recognition results.

The second experiment for scalability assessment is in accordance with the open set identification scenario, which determines whether the individual is in the database and identifies the person in the database. According to the results of the scalability assessment, the “*ScoreNet-2*” provided an improvement of at least 12.5% in Rank-1 from UERC ear recognition methods.

After performing orientation and normalization procedures based on landmarks, two experiments were repeated and the “*ScoreNet-3*” and the “*ScoreNet-4*” provided an improvement of at least 23% and 19% respectively according to the UERC results. After the orientation and normalization procedures, a 12% and 3.5% improvement was achieved in a new study which included deep learning algorithms based on hand-crafted and convolutional neural networks (CNN).

Thus, the “*ScoreNet*” architecture has achieved the state-of-the-art accuracy in the UERC ear data set, which is the most challenging ear data set of unconstrained images. However, the creation and training of “*ScoreNet*” architecture brings with it an important workload. The creation of a modality pool for the “*ScoreNet*” architecture and the long time of the training process are the disadvantages of this architecture.

The “*ScoreNet*” architecture represents the first “*Automated Fusion Learning (AutoFL)*” approach for unconstrained ear recognition. It is also considered that this architecture will lead the future of AutoFL applications.

The “*ScoreNet*” architecture is modular in nature and is prone to development. The most important need for improving the performance of fusion studies is diversity. For better results, the diversity of “*ScoreNet*” architecture can be extended. However, better fusion learning can be achieved by increasing complexity as content in the training and validation set.

As a result, the “*ScoreNet*” has been proven with UERC experiment results that provide a significant improvement in the field of unconstrained ear recognition with its original architecture.





1. GİRİŞ

Günümüzde teknolojinin gelişimine bağlı olarak yeni güvenlik ve gözetleme sistemleri ortaya çıkmaktadır. Bireylerin davranışsal ve biyolojik özelliklerine göre otomatik olarak tanınmasını sağlayan biyometrik sistemler de bu yeni güvenlik ve gözetleme sistemleri içerisinde yer almaktadır. Biyometrik sistemler ile kişi tanıma ve kişi doğrulama işlemi ihtiyaç duyulan her alanda kullanılmaya başlanmıştır. Günlük yaşamda hastanelerde avuç içi ve parmak damar izi tanıma biyometrik sistemleri ile, bazı bankalarda kullanılan iris ve yüz tanıma biyometrik sistemleri, ön plana çıkmaktadır. Bu biyometrik sistemler, genellikle kişi doğrulama amaçlı olarak kullanılmakta olup hastanelerde başkasının yerine tedavi olmanın veya bankalarda sahte kimlik kartları ile kredi çekilmesinin önüne geçmektedir. Ancak bu senaryolarda kontrollü bir ortam söz konusu olup biyometrik verinin en iyi şekilde alınması sağlanmaktadır. Bunun sonucunda yüksek doğruluklu biyometrik sistemler oluşturulabilmektedir. Ancak, kontrolsüz ortamlarda biyometrik verinin poz, aydınlatma, arka plan, düşük çözünürlük, oklüzyon vb. nedenlerden dolayı kaliteli olarak elde edilmesi oldukça zordur. Bu tip senaryolar, özellikle güvenlik ve suçla mücadele kapsamında şüpheli kişi tanıma işi, güvenlik ve emniyet birimlerinin önemli bir ihtiyacıdır.

Günümüzde elektronik güvenlik endüstrisinde yaygın olarak kullanılan kapalı devre televizyon (Close Circuit Television (CCTV)) kameraları şüpheli kişileri tespit ve tanıma için oldukça büyük bir potansiyele sahiptir. Sabit, sabit dom, 360 derece ve Pan-Tilt-Zoom (PTZ) tiplerine göre sınıflandırılmış çok çeşitli CCTV kameraları mevcut olup Çizelge 1.1’de gösterilmiştir.

CCTV kameralar ile şüpheli kişilerin yüz ve kulak biyometrik verileri elde edilebilmektedir. Bununla birlikte şüpheli kişilerin doğal olarak işbirlikçi olmaması nedeni ile şüpheli kişilerden elde edilecek biyometrik verinin kalitesi oldukça düşük olabilmektedir. Ayrıca güvenlik ve emniyet birimlerinin veri tabanında bireylerin biyometrik verisi sınırlı sayıda mevcuttur. Bu nedenle yüz ve kulak biyometrikleri ile kişi tanıma için kontrolsüz ortamlardan oluşturulan veri setlerine gereksinim vardır.

Çizelge 1.1 : CCTV¹ kamera çeşitleri.

Kamera tipi	Sabit	Sabit dom	360 derece	PTZ
Kapalı alan				
Açık alan				

Literatürde açık kaynaklardan erişilebilen çok çeşitli kısıtlanmamış yüz veri setleri bulunmakta olup Çizelge 1.2’de gösterilmiştir [1-13] . Bu kısıtlanmamış yüz veri setleri hem kişi sayısı hem de görüntü sayısı bakımından oldukça zengindir. Ancak kısıtlanmamış kulak veri setlerinde aynı seviyeye henüz ulaşamamıştır. Literatürde açık kaynaklardan erişilebilen kısıtlanmamış kulak veri setleri Çizelge 1.3’te sunulmuştur [14-16].

Çizelge 1.2 : Kısıtlanmamış yüz veri setleri ve özellikleri.

Kısıtlanmamış yüz veri setleri	Kişi sayısı	Görüntü sayısı	Her kişiden görüntü sayısı min/ort./mak.	Yıl
LFW [5]	5749	13233	1/2,3/530	2007
YTF [12]	1595	3425 video	-	2011
CelebFaces [10]	10177	202599	19,9	2014
CASIA-WebFace [13]	10575	494414	2/46,8/804	2014
IJB-A [7]	500	5712 görüntü + 2085 video	11,4	2015
IJB-B [11]	1845	11754 görüntü + 7011 video	36,2	2017
IJB-C [8]	3531	31334 görüntü + 11779 video	36,3	2018
VGGFace [9]	2622	2,6 milyon	1000/1000/1000	2015
MegaFace [6]	690572	4,7 milyon	3/7/2469	2016
MS-Celeb-1M [4]	100000	10 milyon	100	2016
UMDFaces [2]	8501	367920	43,3	2016
UMDFaces-Video [1]	3107	22075 video	-	2017
VGGFace2 [3]	9131	3,31 milyon	80/362,6/843	2018

¹ <https://www.security.us.panasonic.com/white-papers/paper/panasonic-i-pro-video-surveillance>

Çizelge 1.2 ve Çizelge 1.3 karşılaştırıldığında literatürde kulak veri set gelişimi yüz veri setlerine göre oldukça yavaştır. Özellikle kısıtlanmamış kulak veri seti azlığı sebebi ile literatürde kısıtlanmamış görüntüler üzerinde kulak tanıma çalışmaları yeterli seviyede bulunmamaktadır.

Çizelge 1.3 : Kısıtlanmamış kulak veri setleri ve özellikleri.

Kısıtlanmamış veri setleri	kulak	Kişi sayısı	Görüntü sayısı	Her görüntü sayısı min/ort./mak.	Yıl
WPUTE [14]		474	3348	4/7/10	2010
AWE [15]		100	1000	10/10/10	2016
UERC [16]		3540	9500	1/2,7/10	2017

Bu literatür boşluğunu gidermek ve kısıtlanmamış kulak görüntülerinden kulak tanıma problemini incelemek için, 2017 tarihinde 3. Uluslararası Biyometrik Ortak Konferansı (3rd International Joint Conference on Biometrics (IJCBI)) kapsamında Kısıtlanmamış Görüntüler Üzerinde Kulak Tanıma Yarışması (Unconstrained Ear Recognition Challenge (UERC)) düzenlenmiştir.

Yarışmanın amacı, çoklu araştırma gruplarının birleşik uzmanlığını ve tanıma teknolojilerini kapsamlı bir şekilde değerlendirmek; araştırmaları pekiştirmek, ortak bir veri kümesi ve iyi tanımlanmış bir deney protokolü ile açık zorlukları tanımlamaktır. Kısıtlanmamış ortamlardan elde edilen büyük ve zorlu kulak görüntülerinin veri kümesi, UERC için özel olarak toplanmış olup algoritma geliştirme ve test için katılımcılara sunulmuştur. Bu tez çalışmasında UERC için sağlanan kulak veri seti ve deney protokolü kullanılmıştır.

1.1 Motivasyon

Literatürde kulak tanıma alanında kontrollü ortamlardan elde edilen kulak veri setleri kullanılarak birçok el yapımı yöntem geliştirilmiştir. Örneğin; Yönlü Gradyan Histogramı (Histogram of Gradients (HOG)) [17], Yerel İkili Örüntü (Local Binary Pattern (LBP)) [18], Yerel Faz Nicelemesi (Local Phase Quantisation (LPQ)) [19], İkili İstatistiksel Görüntü Özellikleri (Binarised Statistical Image Features (BSIF)) [20], Gabor filtreleri (GABOR) [21] vb.

Bu çalışmalar:

- Geometrik,

- Bütünsel,
- Yerel,
- Hibrit,

olmak üzere dört genel yaklaşımda toplanmıştır. El yapımı yöntemler kontrollü kulak veri setlerinde tanıma oranı açısından başarılı olsa da, kısıtlanmamış kulak veri setlerinde oldukça düşük performans sergilemiştir [15].

Son zamanlarda, derin öğrenme - özellikle, evrimsel sinir ağlarının (Convolutional Neural Network (CNN)) kullanımı - sınıflandırma uygulamalarında çekici bir yaklaşım haline gelmiştir [22]. Yüksek performans doğruluğu ile zengin öznitelikler elde etmek için CNN'ler uygulanabilmesine rağmen, CNN'lerin kullanılması için yapılandırılması zor olabilen büyük ölçekli eğitim veri setlerine [23] ihtiyaç vardır. Başlıca popüler CNN tabanlı derin öğrenme ağları; AlexNet [22], VGG [24], Inception (GoogLeNet) [25], ResNet [26], DenseNet [27] olarak sıralanabilir. Mevcut kulak veri setleri nispeten küçük olduğu için yalnızca CNN'lerin kullanımı kısıtlanmamış görüntüler üzerinde kulak tanıma henüz yeterli seviyeye ulaşmamıştır [16].

Bu tez çalışmasında, kulak tanıma için yukarıda söz edilen problem “Kısıtlanmamış ve Sınırlı Veri Problemi (Unconstrained and Limited Data Problem (ULDP)) olarak tanımlanmıştır. ULDP'nin üstesinden gelmek için füzyon yaklaşımlarının uygulanmasına gereksinim vardır. Bu tür füzyon sistemleri genel olarak iki sınıfa ayrılmıştır [28, 29]:

- Seviye içi (intra) füzyon sistemi,
- Seviyeler arası (inter) füzyon sistemi.

Seviyeler arası füzyon sistemi kulak, yüz, ses, parmak izi ve iris vb. farklı biyometrik sistemlerden bilgi alırken [30-32], seviye içi füzyon sistemi tek bir biyometrik sistemden çıkarılan çeşitli tiplerdeki verileri kullanmaktadır [33].

Literatürde, seviyeler arası ve seviye içi füzyon sistemlerinden, sırasıyla çoklu model (multi-modal) ve çoklu algoritma füzyon stratejileri ile birlikte bahsedilmektedir [34]. Bu çözümler ile diğer çoklu biyometrik füzyon stratejileri ile ilgili daha ayrıntılı bilgiler Bölüm-2'de anlatılmıştır.

Kulak tanıma literatüründe çoklu algoritma ile ilgili çalışmalar düşük karmaşıklık türündeki kombinasyon (iki [30] veya üç algoritma [35] birleşimi) ile sınırlıdır.

Bu tez çalışmasındaki amaç; kulak tanıma alanında ULDP'nin üstesinden gelebilmek için karmaşıklığı yükseltebilecek bir modalite havuzu oluşturmak ve bu karmaşıklık ile başa çıkabilecek “*otomatik füzyon öğrenme (Automated Fusion Learning (AutoFL))*” algoritması geliştirmektir.

1.2 Literatür Katkısı

Kulak tanıma alanında ULDP'nin üstesinden gelebilmek için “*ScoreNet*” adı verilen yeni bir mimari önerilmiştir. “*ScoreNet*” mimarisi hem modalite havuzunu hem de füzyon öğrenme algoritmasını içermektedir. Füzyon öğrenme algoritması için Derin Kaskat Skor Seviyesi Füzyon (Deep Cascade Score Level Fusion (DCSLF)) algoritması geliştirilmiş; modalite havuzu için de çeşitlilik bakımından zengin ve modüler yapıya sahip özgün bir modalite havuzu oluşturulmuştur.

“*ScoreNet*”, bugüne kadar geliştirilen kulak tanımlarını değerlendirmek için en zor veri tabanı olarak kabul edilen UERC kulak veri seti [16] kullanılarak değerlendirilmiş, daha önce bildirilen tüm diğer yöntemlerden daha iyi performans göstermiş ve en son teknoloji doğruluğunu (state-of-the-art accuracy) elde etmiştir.

Bilinen kadarı ile “*ScoreNet*”, kısıtlanmamış görüntüler üzerinde kulak tanıma için ilk “*AutoFL*” yaklaşımını temsil etmekte olup “*ScoreNet*” mimarisinin avantajları aşağıda sunulmuştur:

- Bağımsız olarak oluşturulan yöntemler (derin öğrenme ve el yapımı yöntemler dahil) birlikte kullanılabilir.
- En iyi sonuç birden fazla çözüm ile elde edilebilir.
- Füzyon işleminden elde edilen veriler, sonraki füzyon için tekrar kullanılabilir.
- Kullanıcı müdahalesi ve rastgele başlatma mümkündür.
- Mimari, tüm uçtan uca operasyonları otomatikleştirmek için bir öğrenme sürecinin uygulanmasını ifade eden “*AutoFL*” için uygundur.

- Her modalite birbirinden bağımsız olduğundan paralel işlem gerçekleştirilebilir.

Ancak, “ScoreNet” mimarisinin bazı dezavantajları vardır:

- Eğitim süreci uzun olabilmektedir.
- Uygun çeşitliliğe ulaşmak için bir modalite havuzuna gereksinimi vardır.
- Modalite sayısına bağlı olarak karmaşıklık artmaktadır.

Bu tez çalışmanın ana katkıları aşağıdaki gibi özetlenmiştir:

- Kısıtlanmamış görüntüler üzerinde kulak tanıma için “ScoreNet” adı verilen yeni bir mimari önerilmiştir.
- Bildiğimiz kadarıyla, önerilen mimari kulak tanımada hatta daha geniş spektrumda biyometrik alanında “AutoFL” için ilk uygulamadır ve bu mimarinin gelecekteki uygulamalara öncülük edeceği değerlendirilmektedir.
- Çeşitlilik bakımından zengin ve modüler yapıya sahip özgün bir modalite havuzu oluşturulmuştur.
- Füzyon öğrenmenin gerçekleştirilmesi için DCSLF adı verilen yeni bir algoritma geliştirilmiştir.
- Yukarıda listelenen yeniliklere dayanarak, CNN'lerin kullanımına dayanan son teknoloji ürünü kulak tanıma yaklaşımlarından daha iyi performans gösteren hem CNN hem de el yapımı yöntemleri içeren ve derin kaskat skor seviyesi füzyona dayalı bir kulak tanıma mimarisi tasarlanmıştır.

1.3 Tezin Bölümleri

Tezin diğer bölümleri aşağıdaki gibi düzenlenmiştir:

- Bölüm 2'de, kulak biyometrisi ile ilgili çalışmalar kısaca gözden geçirilmiştir.
- Bölüm 3'te önerilen mimari, “ScoreNet” ayrıntılı bir şekilde anlatılmıştır.
- Bölüm 4'te, önerilen yaklaşımı doğrulamak için kullanılan deney setleri ve deney düzeneği açıklanmıştır.
- Bölüm 5'te deney sonuçları sunulmuştur.

- Bölüm 6'da sonuçlar yorumlanmıştır.

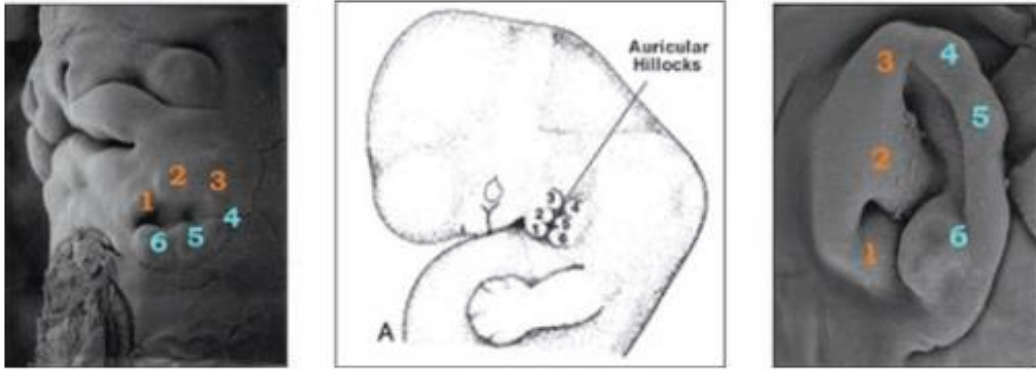




2. KULAK BİYOMETRİSİ

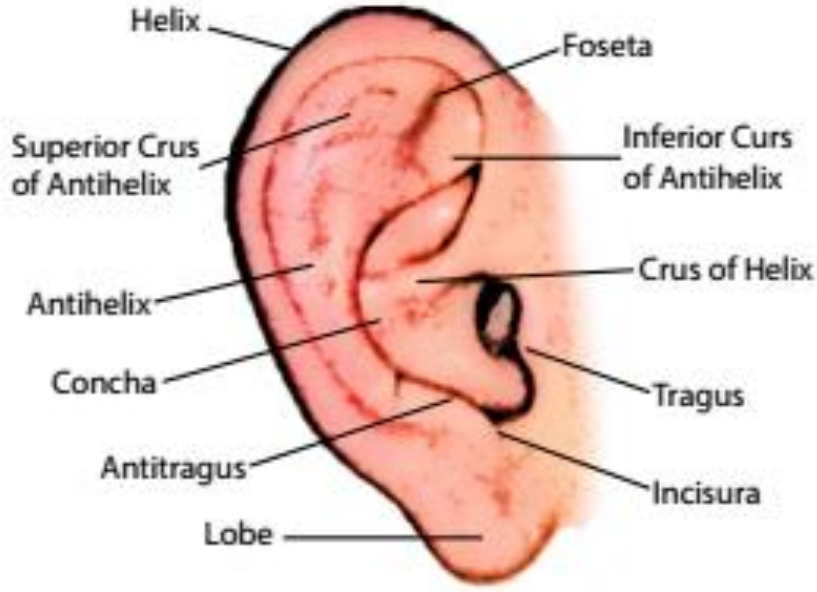
2.1 Kulağın Embriyolojik Gelişimi ve Yapısı

İnsan kulağı, hamileliğin beşinci ve yedinci haftaları arasında görünmeye başlamaktadır. Embriyolojik gelişimin beşinci haftasında ortaya çıkan “*auricular hillocks*” adı verilen yapı, embriyolojik gelişimin altıncı ve dokuzuncu haftaları arasında genişler, farklılaşır ve en son kaynaşarak dış kulağın yapısını oluşturur. Dokuzuncu haftada oluşan morfoloji bir insan kulağı olarak seçilebilmekte olup Şekil 2.1’de 1-6 numara ile işaretlenen bölgeler kulağın “*tragus, helix, cymba concha, antitragus, antihelix ve concha*” kısımlarını oluşturmaktadır. Adli tıp literatürü, doğumun ilk dört ayından sonra kulak büyümesinin son derece doğrusal olduğunu bildirmektedir. 4 aylık ile 8 yaş arası kulağın genişleme oranı normalden yaklaşık beş kat daha fazla olabilmekte olup 8 ile 70 yaş arası kulak büyüklüğü değişmemektedir. [30].



Şekil 2.1 : Kulağın embriyolojik gelişimi [30].

Kulağın yapısı Şekil 2.2’de sunulmuştur [15]. Literatürde insan kulağının simetri özelliği de araştırılmıştır [36]. İncelemeler sonucunda sağ ve sol kulaklar birbirine benzemekte olup bir dereceye kadar simetri veya asimetri özelliği taşıyabileceği ortaya çıkmıştır [36].



Şekil 2.2 : Kulağın yapısı [15].

2.2 İlgili Çalışmalar

Bu bölümde, el yapımı ve derin öğrenme yöntemlerine dayanan, kulak tanıma yaklaşımları üzerine yapılan çalışmaların yanı sıra kulak tanıma alanındaki füzyon teknikleri ile ilgili literatür de kısaca özetlenmiştir.

Kulak tanıma için el yapımı yöntemler bir önceki bölümde bahsedildiği gibi geometrik, bütünsel, yerel veya hibrit teknikler olarak sınıflandırılabilir.

Geometrik teknikler, kulak üzerinde önceden belirlenmiş noktalar arasındaki metriklere dayanmaktadır. Geometrik yaklaşımların bir sınırlaması, aydınlatma varyasyonlarına ve gürültüye duyarlı olduğu bilinen kenar dedektörlerine olan bağımlılıklarıdır. Birçok teknik, kulak noktalarının kesin konumlarına ihtiyaç duymakta olup düşük kaliteli görüntülerde kulak noktalarının tespitini yapmak oldukça zorlaşmaktadır [15, 33, 37].

Bütünsel yaklaşımlar, kulağı temsil edecek öznitelik vektörlerini çıkarmak için bir bütün olarak kulağın tüm görüntüsünü kullanmaktadır. Kuvvet alanı dönüşümü [38], kulak görüntüsündeki her pikselin yoğunluğuna ve konumuna göre bir kuvvet alanı ile ilişkili olduğu popüler bir bütünsel yöntemdir. Diğer bütünsel yöntemler, bir öznitelik vektörü ile kulağı temsil etmek için, doğrusal veya doğrusal olmayan alt uzay izdüşüm tekniklerini kullanmaktadır [39, 40].

Yerel yöntemler, kulağın belirli bölgelerinden öznitelik vektörleri çıkarmakta ve bunları alt dizilim yolu ile tek bir vektöre dönüştürmektedir. Çoğu yerel yöntem, küçük görüntü döndürme ve hizalama için daha dirençlidir. Yerel tekniklerin tipik örnekleri arasında HOG [17], LBP [18], LPQ [19], BSIF [20] ve GABOR [21] bulunmaktadır.

Hibrit teknikler, geometrik, yerel ve bütünsel özellikleri birleştiren ve tanıma oranını arttırmak için birden fazla yöntem kullanan yaklaşımları içermektedir [41].

CNN'ler nesne tespiti, görüntü tanıma alanları için yakın zamanda geliştirilmiş olan derin öğrenme yöntemleridir. AlexNet [22], 1.2 milyon görüntü ile 1000 sınıf içeren ImageNet Large-Scale Visual Recognition Challenge (ILSVRC)-2010 yarışmasında el yapımı yöntemlerden daha fazla doğruluğa sahip ilk CNN tabanlı derin öğrenme yöntemi olmuştur.

VGG, oldukça küçük (3×3) konvolüsyon filtrelerine sahip daha derin bir ağ uygulayan bir diğer CNN tabanlı yöntemdir [24, 42]. Sırasıyla 16 ve 19 katman içeren VGG-16 ve VGG-19 ağları, araştırmacıların ConvNet modelleri olarak yararlanabilmesi için kamuya açık hale getirilmiştir. VGG modeli, diğer tanıma problemleri (örneğin; yüz tanıma [3, 9] vb.) ile birlikte kulak tanıma alanına da uygulanmış olup zorlu kulak veri setlerinde el yapımı yöntemlerine göre daha yüksek tanıma oranı sağlamıştır. [16, 23, 43-46].

Inception (GoogLeNet) derin öğrenme yöntemi [25], daha büyük veri kümelerinin basitleştirilmiş eğitimini sağlayarak oldukça düşük hesaplama maliyeti ile VGG'den daha iyi performans göstermektedir. Bununla birlikte, Inception, genel çerçeve içinde ek bir küçük ağ oluşturan bir başlangıç modülünün kullanılması durumunda VGG'den daha karmaşık bir model haline gelmektedir.

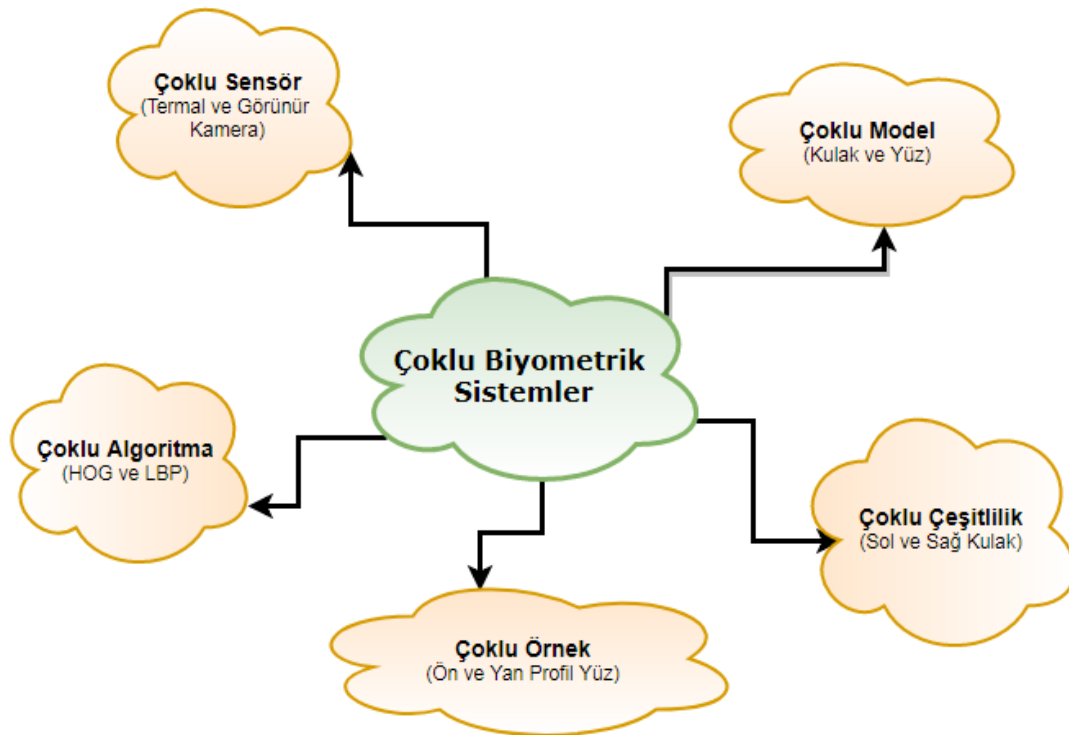
ResNet [26], bitişik katmanlar arasındaki girişleri ilerletmek için artık bir bağlantı kullanmaktadır. Böylece daha derin ağlarda görülen bozulma sorunları azalmaktadır. Başka bir ifade ile, ResNet, CNN modellerine kısayol veya atlama bağlantıları sağlamaktadır. Bu bağlantılar, bazı konvolüsyon katmanlarını atlayan ve bilgiyi aşağıdan yüksek model katmanlara ileten kimlik kısayollarını temsil etmektedir. Bu durum, ağ boyunca bilginin kaybolmasını engellemektedir. Aynı zamanda, daha fazla sayıda katmanı olan derin modellerin eğitimini de kolaylaştırmaktadır. Model eğitimi sırasında kısayol bağlantılarının katma değeri, model parametrelerini öğrenmek için

kullanılan geri yayılım algoritması için kısayollar olarak işlev görmeleridir ve böylece gradyanların ağda kaybolmasının önüne geçilebilmektedir.

DenseNet, bir CNN'nin her katmanını diğer tüm katmanlara bağlayarak, ağdaki tüm katmanların önceki tüm katmanların öz niteliklerini girdi olarak kullanmasına izin vermektedir. DenseNet modeli, gradyanların yok olma probleminin hafifletilmesi, öz nitelik yayılımının güçlendirilmesi, öz niteliklerin yeniden kullanımının desteklenmesi ve parametrelerin sayısında önemli bir azalma sağlanması vb. gibi çeşitli avantajlar sağlamaktadır.

Bu tez çalışmasında el yapımı yerel tekniklerden HOG, LBP, LPQ, BSIF ve GABOR ile DenseNet modeli hariç yukarıda bahsedilen CNN modelleri, öz nitelik çıkarımı için modalite havuzuna eklenmiştir. CNN modeller tekrar eğitilmemiş olup orijinal halleri ile kullanılmıştır.

Bir önceki bölümde bahsedilen çoklu biyometrik füzyon stratejileri, literatürde biyometrik bilginin kaynağının niteliğine bağlı olarak 5 sınıfa ayrılmış olup Şekil 2.3'te sunulmuştur [47, 48].



Şekil 2.3 : Çoklu biyometrik füzyon stratejileri.

- **Çoklu sensör:** Bir bireyin tek bir biyometrik özelliğini yakalamak için birden fazla sensör kullanarak biyometrik verinin çeşitliliğini arttıran bir sistemdir. Örneğin, termal ve görünür kamera.
- **Çoklu algoritma:** Tanıma oranını yükseltmek için birden çok algoritma kullanan ve hesaplama karmaşıklığı yüksek olan bir sistemdir. Örneğin, HOG ve LBP algoritmaları.
- **Çoklu örnek (multi-sample):** Tek bir sensör ile aynı biyometrik özellikten türetilmiş çoklu örnek kullanan bir sistemdir. Örneğin, ön ve yan profil yüz.
- **Çoklu çeşitlilik (multi-instance):** Bu sistemde, biyometrik bilgiler aynı vücut özelliklerinin çoklu örneklerinden çıkarılmıştır. Örneğin, bir bireyin sağ ve sol kulakları.
- **Çoklu model (multi-modal):** Birden fazla biyometrik kaynaktan elde edilen sonuçları birleştiren bir sistemdir. Örneğin, kulak ve yüz biyometrileri.

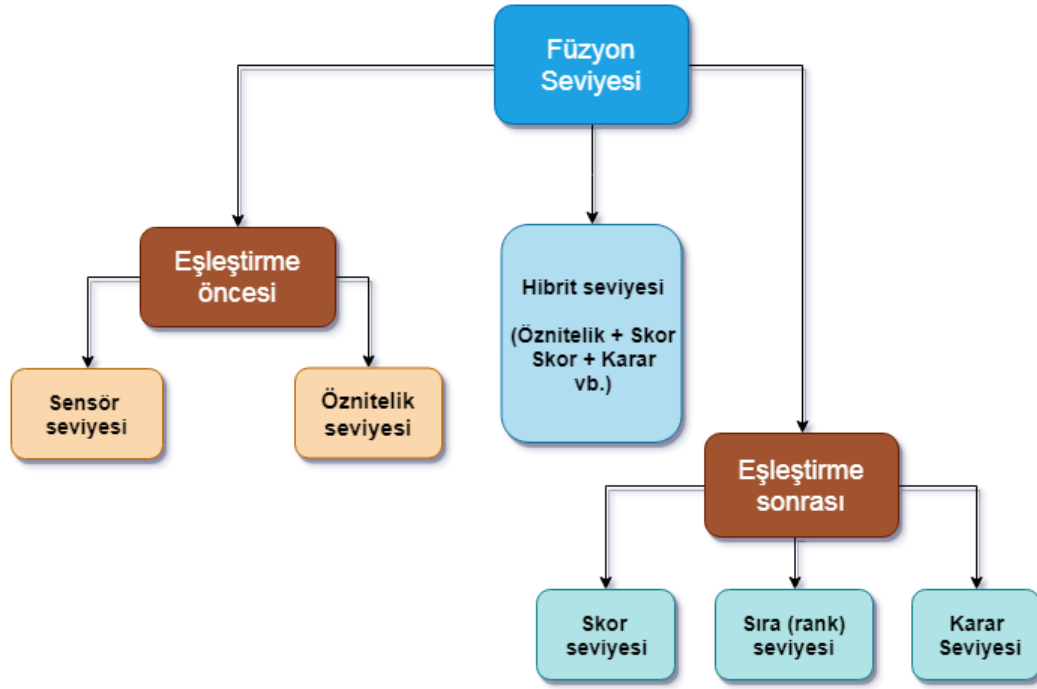
Ayrıca, literatürde füzyon sürecinde kullanılan beş çeşit füzyon seviyesi yöntemi bulunmaktadır [47, 48]:

- Sensör seviyesi füzyon,
- Öznitelik seviyesi füzyon,
- Skor seviyesi füzyon,
- Sıra (rank) seviyesi füzyon,
- Karar seviyesi füzyon.

Bu füzyon seviyeleri de bir üst kategoride üç ana sınıfa bölünmüştür [48]:

- Eşleştirme öncesi füzyon,
- Eşleştirme sonrası füzyon,
- Hibrit seviyesi füzyon.

Sensör ve öznitelik seviyesi füzyonlar, eşleştirme işleminden önce uygulanmakta olup skor eşleştirme, sıra (rank) ve karar seviyesi füzyonlar ise, eşleştirmeden sonra yapılmaktadır. Ayrıca, bu füzyon seviyelerine ek olarak iki veya daha fazla füzyon seviyesinin birleştirilmesi ile oluşan hibrit seviyesi füzyon da mevcuttur. Öznitelik ve skor eşleştirme füzyonlarının birlikte kullanılması hibrit seviyesi füzyona örnek olarak verilebilir. Tüm bu bahsedilen füzyon seviyelerinin sınıflandırılmış durumu Şekil 2.4'te gösterilmiştir [48].



Şekil 2.4 : Farklı füzyon seviyelerinin sınıflandırılması.

Literatürde kulak biyometrisini içerecek şekilde çoklu model (multi-modal) ve çoklu algoritma füzyon stratejilerine yönelik oldukça çeşitli çalışmalar yapılmış olup Çizelge 2.1’de sunulmuştur [31, 32, 39, 41, 47, 49-65].

Çizelge 2.1 : Kulak biyometrisi içeren çoklu biyometrik sistemler.

Füzyon Sistem Tipleri	Çoklu biyometrik sistemler	Füzyon Seviyesi	Spektrum
Seviyeler arası	Kulak & ön profil yüz [54-56]	Skor	Görünür
	Kulak & ön profil yüz [61, 65]	Öznitelik	Görünür
	Kulak & yan profil yüz [52]	Öznitelik	Görünür
	Kulak & yüz & parmak yüzeyi [51]	Skor	Görünür
	Kulak & yüz & imza [57]	Sıra (rank)	Görünür
	Kulak & iris [31]	Skor	Görünür
	Kulak & iris [63]	Öznitelik	Görünür
	Kulak & iris [59]	Karar	Görünür
	Kulak & iris & yüz [62]	Sıra (rank)	Görünür
	Kulak & ses [49]	Skor	Görünür
Seviye içi	Kulak & çoklu algoritma [50, 58]	Skor	Görünür
	Kulak & çoklu algoritma [53]	Öznitelik	Görünür
	Kulak & çoklu algoritma [60]	Skor	Termal-MWIR
	Kulak & çoklu algoritma [35]	Skor	Termal-LWIR

Çizelge 2.1'deki arařtırmalar incelendiğinde çoklu model (multi-modal) ve çoklu algoritma füzyon stratejileri düşük karmařıklık türündeki ikili ve üçlü kombinasyonları içermektedir. Kulak biyometrisinde gerçek dünya senaryolarında bulunan zorluklar ile başa çıkılması için karmařıklığı yüksek çözümlere ihtiyaç duyulmaktadır [66-68].

Bu tez çalışmasında kulak tanıma alanında iki boyutlu, görünür bant spektrumunda, çoklu algoritma, skor seviyesi füzyon ve yüksek karmařıklık içeren özgün bir çalışma gerçekleştirilmiştir.





3. KULAK TANIMA YÖNTEMLERİ

Herhangi bir füzyon işleminin temel amacı, ilgili problemi daha iyi bir yere taşıyabilmek için uzmanlar tarafından üretilen kararı, optimum bir şekilde birleştirebilecek uygun bir işlevi gerçekleştirebilmesidir [47]. Bu temel amacı çeşitli başlıklar altında inceleyen kulak biyometrisi alanındaki füzyon çalışmalarından Bölüm-1 ve 2’de ayrıntılı bir şekilde bahsedilmiş ve bu çalışmalar aşağıda özet halinde sunulmuştur:

- Füzyon sistem tipleri (seviyeler arası (inter) ve seviye içi (intra)),
- Çoklu biyometrik füzyon stratejileri (çoklu sensör, çoklu algoritma, çoklu örnek (multi-sample), çoklu çeşitlilik (multi-instance), çoklu model (multi-modal)),
- Füzyon seviyesi yöntemleri (sensör, öznelilik, skor eşleştirme, sıra (rank) ve karar),
- Füzyon seviyesi üst sınıf yöntemleri (eşleştirme öncesi ve sonrası, hibrit).

Çoklu biyometrik bir sistemde füzyon stratejisinin kullanımı sistemin genel performansını olumlu veya olumsuz yönde etkileyebilmektedir. Bu bağlamda, senaryoya özgü optimum çözüm için, maliyet-etkinlik analizi de yapılarak, füzyon seviyesi yöntemleri ile füzyon stratejilerinin birlikte değerlendirilmesine gereksinim vardır. Burada gerçekleştirilecek senaryonun kazanımı da önemlidir. Örneğin, biyometrik sistemlerde hız ve doğruluk arasında doğrusal olmayan tersine bir ilişki bulunmaktadır.

Hız daha baskın olduğunda doğruluktan ödün verilerek daha az karmaşık yapı kullanma ihtiyacı oluşurken; doğruluk veya tanıma oranının yükseltilmesi söz konusu olunca daha karmaşık füzyon yaklaşımlarının uygulanmasına gereksinim duyulmakta ve buna bağlı olarak hız da azalmaktadır.

Kulak tanıma için geleneksel füzyon yöntemleri, doğruluğu en iyi olanlardan iki veya üç modaliteyi; öznitelik, skor, sıra (rank) veya karar seviyesi füzyon ile birleştirerek tanıma oranında kısıtlı bir iyileştirme gerçekleştirmektedir [35, 39, 60]. Bu geleneksel füzyon yöntemleri daha çok kontrollü kulak veri setlerine uygulanmış olup kısıtlanmamış görüntüler üzerinde gerçekleştirildiğinde ise bu kısıtlı iyileştirme daha da azalabilmektedir [15, 16]. Daha iyi sonuçlar için karmaşıklığı yüksek füzyon ağlarına ihtiyaç duyulmaktadır. Bu bağlamda, ilk olarak çeşitli skor seviyesi füzyon mekanizmaları ele alınmıştır.

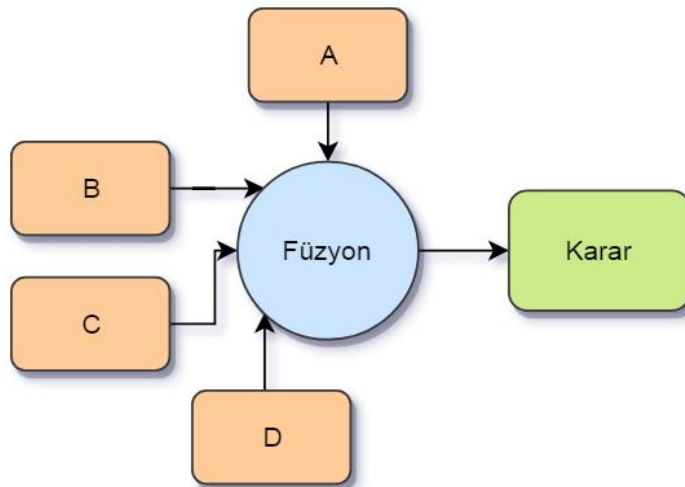
3.1 Skor Seviyesi Füzyon Mekanizmaları

Literatürde genel olarak dört farklı tipte skor seviyesi füzyon mekanizması bulunmaktadır [66-69]:

- Klasik skor seviyesi füzyon,
- Hiyerarşik skor seviyesi füzyon,
- Kaskat skor seviyesi füzyon,
- Hibrit skor seviyesi füzyon.

3.1.1 Klasik skor seviyesi füzyon

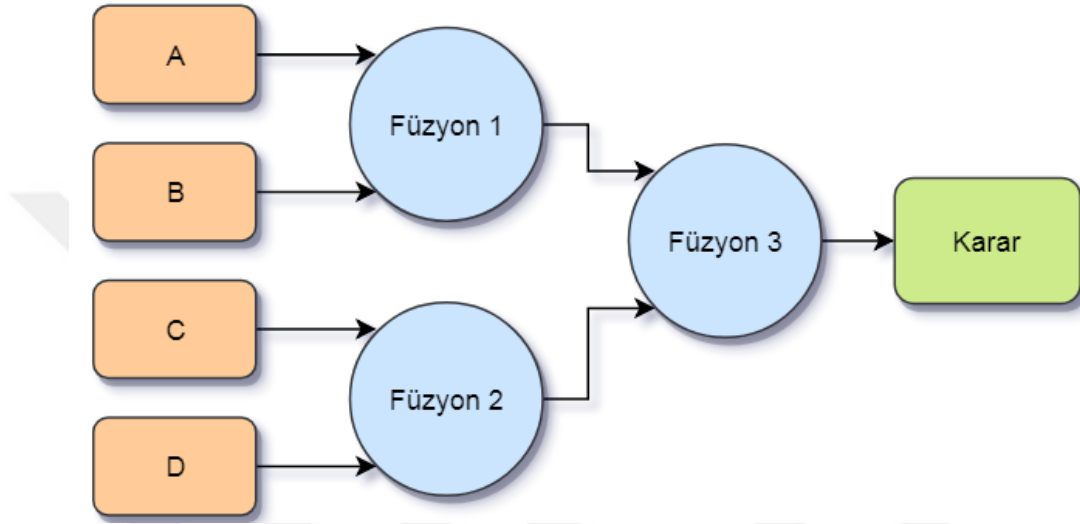
Klasik skor seviyesi füzyon mekanizması, tek bir denklem kullanarak optimal ağırlık bulmayı amaçlamaktadır. Ancak, modalite sayısı arttıkça, arama alanında genişlemenin bir sonucu olarak, bir ağ için en uygun çözümü bulmak giderek zorlaşmaktadır. Klasik skor seviyesi füzyon mekanizması Şekil 3.1’de gösterilmiştir.



Şekil 3.1 : Klasik skor seviyesi füzyon mekanizması.

3.1.2 Hiyerarşik skor seviyesi füzyon

Hiyerarşik skor seviyesi füzyon mekanizması, klasik skor seviyesi füzyon mekanizmasından farklı olarak çoklu ağ şubelerinden oluşmaktadır. Ancak, ağ şubelerinin sayısı arttıkça, çeşitlilik açısından bileşenlerden bağımsız olarak ikili kombinasyonların oluşturulmasında artan zorluklar ile karşılaşilmektedir. Ayrıca, bu durum tek bir modalitenin ağa doğrudan katkısını gözlemlemeyi imkansız kılmaktadır. Hiyerarşik skor seviyesi füzyon mekanizması Şekil 3.2’de gösterilmiştir.



Şekil 3.2 : Hiyerarşik skor seviyesi füzyon mekanizması.

3.1.3 Kaskat skor seviyesi füzyon

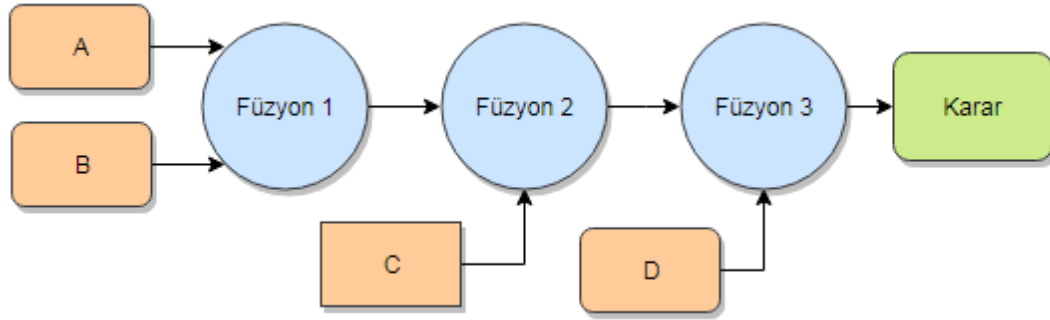
Kaskat skor seviyesi füzyon mekanizmasında, modaliteler sıra ile denenmektedir. Modalitelere belirli bir eşik değerine göre kabul veya ret işlemi uygulanabilmektedir.

Kaskat skor seviyesi füzyon mekanizmasında tek bir modalitenin doğrudan etkisi gözlenebilmektedir. Bu avantaj, modalite havuzu içinde sistemin genel performansını daha iyi yapacak modalitenin tespitini kolaylaştırmaktadır. Bununla birlikte, kaskat skor seviyesi füzyon mekanizmasının yapısı gereği sürekli iki modalitenin ağırlık parametresinin bulunmasına ihtiyaç vardır. Bu sayede, modalite etraflı veya tam arama (exhaustive search) işlemine yatkın hale gelmektedir. Etraflı veya tam arama (exhaustive search), parametrenin optimum değerinin tüm değerlerinin denenerek bulunmasıdır.

Ayrıca, kaskat skor seviyesi füzyon mekanizmasının modalite karmaşıklığını kontrol edebilecek bir üstünlüğü de bulunmaktadır. Çünkü, ağırlık parametreleri her bir katmandan bağımsız olarak bulunmaktadır. Hiyerarşik skor seviyesi füzyon

mekanizmasında da bu durum ağ şubeleri için geçerlidir. Ancak, ağ şubelerinin iç karmaşıklığı da mevcuttur.

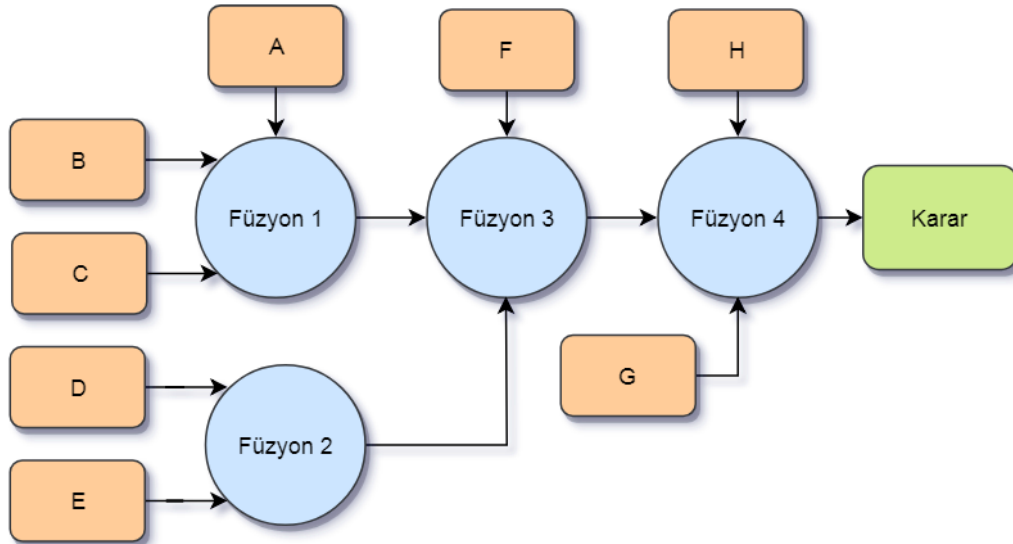
Bu tez çalışmasında, söz konusu avantajlar nedeni ile kaskat skor seviyesi füzyon mekanizması kullanılmıştır. Kaskat skor seviyesi füzyon mekanizması Şekil 3.3'te gösterilmiştir.



Şekil 3.3 : Kaskat skor seviyesi füzyon mekanizması.

3.1.4 Hibrit skor seviyesi füzyon

Hibrit skor seviyesi füzyon mekanizması; klasik, hiyerarşik ve kaskat skor seviyesi mekanizmalarının birleşiminden oluşmaktadır ve daha karmaşık bir yapıya sahiptir. Hibrit skor seviyesi füzyon mekanizması Şekil 3.4'te gösterilmiştir.



Şekil 3.4 : Hibrit skor seviyesi füzyon mekanizması.

3.2 Önerilen Mimari – “ScoreNet”

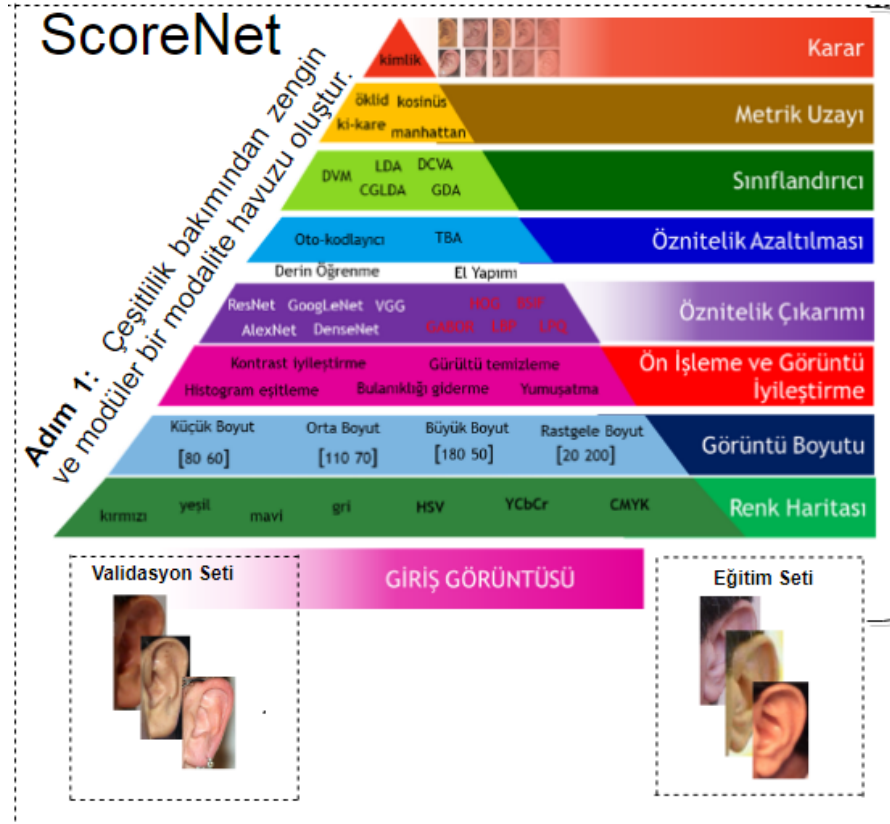
Diğer bölümde bahsedilen tüm skor seviyesi füzyon mekanizmalarının işletilmesi için, çeşitlilik bakımından zengin ve füzyon öğrenimi için kaynak olarak kullanılabilecek bir modalite havuzuna gereksinim vardır.

Bu tez çalışmasında, bir modalite; bir giriş görüntüsü seti, ara operasyonları içeren bir dizi işlem, bir çıktı ve skor eşleştirme matrisinden oluşmaktadır. Modalite havuzu ise, birden çok modalitenin oluşturduğu topluluktan meydana gelmiştir.

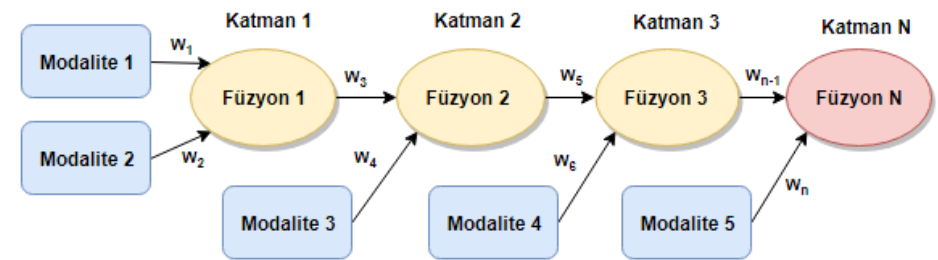
Bu tez çalışmasında hem modalite havuzunu hem de füzyon öğrenme algoritmasını içeren “ScoreNet” adı verilen yeni bir mimari sunulmuştur. Füzyon öğrenimini gerçekleştirmek için de “*Derin Kaskat Skor Seviyesi Füzyon (DCSLF)*” algoritması geliştirilmiştir. Ayrıca, çeşitlilik bakımından zengin ve modüler yapıya sahip özgün bir modalite havuzu oluşturulmuştur.

Şekil 3.5’te sunulan “ScoreNet” mimarisi üç temel adımdan oluşmaktadır:

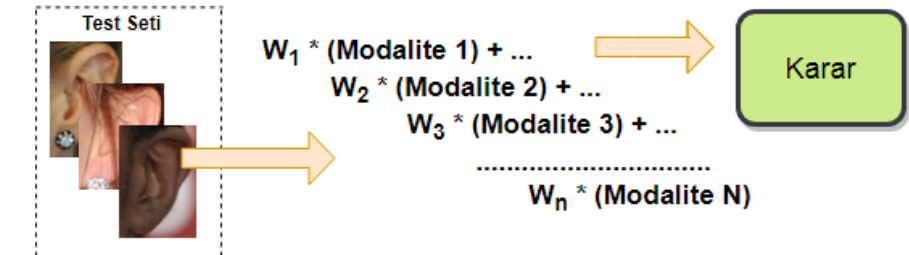
- 1) Çeşitlilik bakımından zengin ve modüler bir modalite havuzu oluşturun.
- 2) Eğitim ve validasyon (doğrulama) setini giriş olarak kullanarak rastgele modaliteler oluşturun. DCSLF algoritması ile hem en iyi modaliteyi seç hem de seçilmiş modalitelerin füzyon ağırlıklarını bulun.
- 3) Kaskat ağ yapısında bulunan seçilmiş tüm modaliteleri ve hesaplanan füzyon ağırlıklarını sabitleyerek test veri setine uygulayın.



Adım 2: Eğitim ve validasyon setini giriş olarak kullanarak rastgele modaliteler oluşturun. DCSLF algoritması ile hem en iyi modaliteyi seçin hem de seçilmiş modalitelerin füzyon ağırlıklarını bulun.



Adım 3: Kaskat ağ yapısında bulunan seçilmiş tüm modaliteleri ve hesaplanan füzyon ağırlıklarını sabitleyerek test veri setine uygulayın.



Şekil 3.5 : Önerilen mimari – “ScoreNet”.

3.2.1 Modalite havuzunun oluşturulması

Çeşitlilik bakımından zengin bir modalite havuzu oluşturmak için, yedi katmandan oluşan modüler bir yapı önerilmiş olup her katman aşağıda sırasıyla sunulmuştur:

- Renk haritası,
- Görüntü boyutu,
- Ön işleme ve görüntü iyileştirme,
- Öznitelik çıkarımı ve parametre setleri,
- Öznitelik azaltılması,
- Sınıflandırıcı,
- Metrik uzayı.

Modalite havuzunun yapısı modüler olduğundan, her katman kolayca genişletilebilmektedir. Ancak, bazı katmanlarda seçilecek aday işlemler birbiri ile uyumsuz veya tek başına gereksiz olabileceğinden ihmal edilebilmektedir. Örneğin; öznitelik vektörünün boyutu küçük ise öznitelik azaltma işlemi göz ardı edilebilmektedir. Başka bir örnek olarak da öznitelik çıkarımı kullanılan metrik ile uyumsuz olduğunda diğer metrikler arasından seçim yapılmaktadır.

Modalite havuzundan modalite oluşturma işlemi rastgele (otomatik) veya kullanıcı (manuel) tarafından yapılabilir. Doğal olarak modalite havuzu, literatürde yeni yöntemler geliştirildikçe devamlı genişleyecek modüler bir yapıya sahiptir. Bu tez çalışmasındaki amaç, yüksek karmaşıklıkla bir modalite havuz kaynağını kullanarak füzyon öğrenme algoritması geliştirmektir. Bu nedenle modalite havuzu, geliştirilen füzyon algoritmasının literatürdeki yöntemlerden üstünlüğünün gösterilmesine kadar genişletilmiştir. Ancak, modalite havuzunun her katmanında bulunan mevcut yöntemlere ek olarak gelecek uygulamalarda eklenecek potansiyel çalışmalardan da aşağıda sırasıyla bahsedilmiştir:

3.2.1.1 Renk haritası

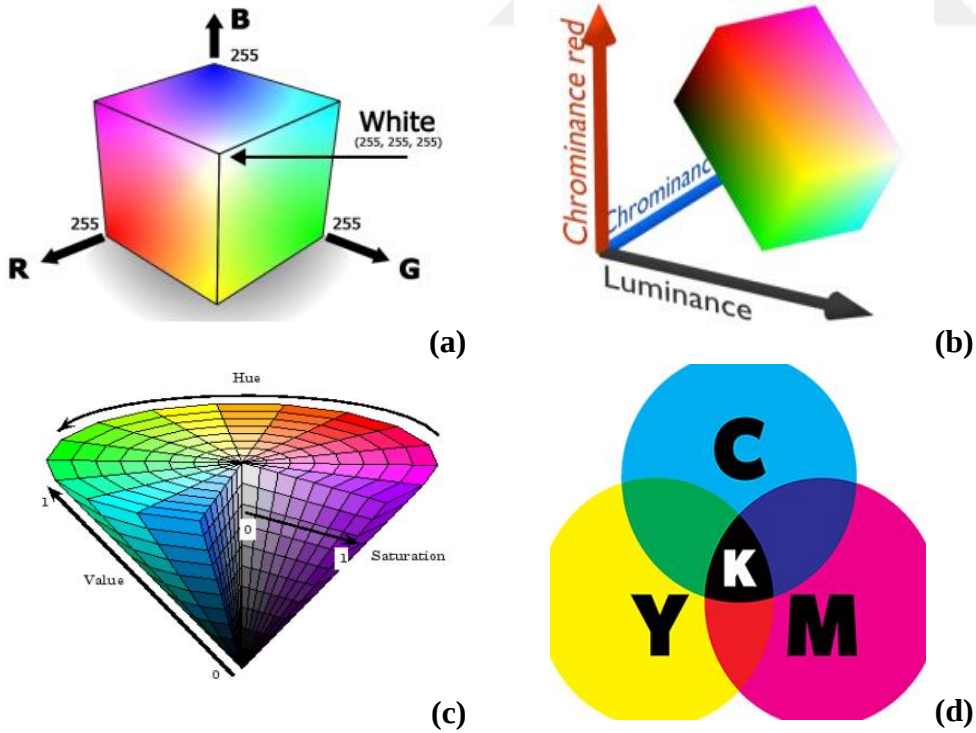
İlk adımda, giriş görüntüsünün kırmızı, yeşil ve mavi renkleri veya bunların ağırlıklı dönüştürülmüş hali kullanılmakta olup örnek renk haritaları Şekil 3.6’da gösterilmiştir. Bu ilk katman, sensör çeşitliliğine bağlı olarak, görünür ve kızılötesi hiperspektral görüntüleme ile genişletilebilir [70].

Bu tez çalışmasında aşağıdaki renk haritaları kullanılmıştır:

- Gri tonlama,
- Kırmızı, yeşil, mavi (RGB),
- Parlaklık, krominans (YCbCr),
- Ton, doygunluk, değer (HSV) [71, 72].

Diğer potansiyel eklenebilecek çalışmalar aşağıda sunulmuştur:

- Camgöbeği, fuşya, sarı ve siyah (CMYK) [72],
- Rastgele ağırlıklı renk haritaları,
- 3B görüntüleme.



Şekil 3.6 : Renk haritaları: (a) RGB (b) YCbCr (c) HSV (d) CMYK [71, 72].

3.2.1.2 Görüntü boyutu

İlk adımda renk haritası seçimi yapıldıktan sonra ikinci adımda görüntü boyutu,

- Küçük (80×60 piksel),
- Orta (110×70 piksel),
- Büyük (180×50 piksel),
- Rastgele ($20\text{--}200 \times 20\text{--}200$ piksel arası),

olmak üzere dört seçenek arasından seçilir. Bu tez çalışmasında rastgele görüntü boyutu kullanılmıştır. Görüntü çözünürlüğüne göre de görüntü boyutu ayarlanabilir.

3.2.1.3 Ön işleme ve görüntü iyileştirme

Üçüncü adımda, daha iyi öznitelikler elde etmek için görüntüye ön işleme ve görüntü iyileştirme işlemleri uygulanır. Literatürde çok fazla sayıda ön işleme ve görüntü iyileştirme işlemleri mevcut olup yaygın olarak kullanılanlar aşağıda sunulmuştur:

- Kontrast iyileştirme,
- Histogram eşitleme,
- Yumuşatma (smoothing),
- Gürültü temizleme (denoising),
- Bulanıklığı giderme (deblurring).

Bu tez çalışmasında, üç farklı ön işleme ve görüntü iyileştirme yöntemi kullanılmıştır:

- Kontrast Sınırlı Uyarlamalı Histogram Eşitleme (Contrast Limited Adaptive Histogram Equalization (CLAHE)) [73],
- 2B Gaussian filtresi ile yumuşatma,
- Görüntü yoğunluğu değerlerini ayarlama.

Diğer potansiyel eklenebilecek çalışmalar aşağıda sunulmuştur:

- Derin öğrenmeye dayalı görüntü iyileştirme [74, 75], gürültü temizleme (denoising) [76] ve bulanıklığı giderme (deblurring) [77] ,
- Düşük çözünürlüklü görüntüleri iyileştirmek için derin öğrenmeye dayalı süper çözünürlük yöntemi [78].

3.2.1.4 Öznitelik çıkarımı ve parametre setleri

Dördüncü aşamada, öznitelik çıkarımı için hem el yapımı hem de derin öğrenme yöntemleri kullanılır. Literatürde en çok kullanılan ve Bölüm-1’de de bahsedilen el yapımı yöntemler aşağıdadır:

- HOG, görüntüdeki gradyan yönelimleri ve büyüklükleri kullanarak öznitelik çıkarımı gerçekleştirir [17].
- LBP, görüntü piksellerinin komşuluklarına bakarak öznitelik çıkarır [18].
- LPQ, görüntünün frekans spektrumu elde edildikten sonra LBP yöntemine benzer bir şekilde öznitelik çıkarımı yapar [19].
- BSIF, giriş görüntüsünü önceden öğrenilmiş bir dizi filtreden geçirir ve daha sonra filtre yanıtlarını birleştirerek öznitelikleri elde eder [20].
- Gabor filtreleri, harmonik bir fonksiyon ile Gaussian bir fonksiyonun çarpımından oluşur. Bu Gabor filtre bankası kullanılarak görüntüden öznitelik çıkarımı gerçekleştirilir [21].

Her öznitelik çıkarma yönteminde iç parametre değerleri de değişiklik göstermektedir. Ayrıca, bu aşamada her algoritma için 10 farklı iç parametre değeri içeren parametre seti hazırlanır. İç parametre değerleri algoritmanın optimum çalışabildiği aralık içinden belirlenir.

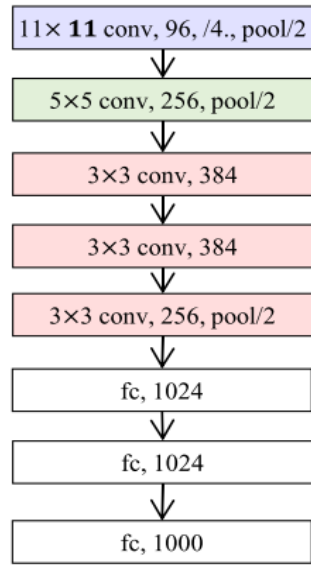
Benzer şekilde, literatürde yaygın olarak kullanılan ve Bölüm-2’de de bahsedilen CNN tabanlı derin öğrenme yöntemleri aşağıda sunulmuş olup mimari yapıları Şekil 3.7’de gösterilmiştir:

- AlexNet, 8 katmandan oluşan seri ağ yapısında ilk CNN mimarisidir [22].
- VGG, daha derin seri ağlardan oluşmakta olup sırasıyla 16 ve 19 katman olmak üzere iki farklı modeli bulunmaktadır [24].
- Inception (GoogLenet), paralel modüller kullanarak VGG’den daha düşük hesaplama maliyetleri sağlamıştır ve 22 katmandan oluşmaktadır [25].
- ResNet, artık bir bağlantı kullanarak daha derin ağlarda gradyanların ağda kaybolmasını engellemiştir ve 152 katman bulunmaktadır [26].
- DenseNet, artık bir bağlantı yerine her katmandan diğer tüm katmanlara bağlantı sağlayan bir CNN mimarisidir [27].

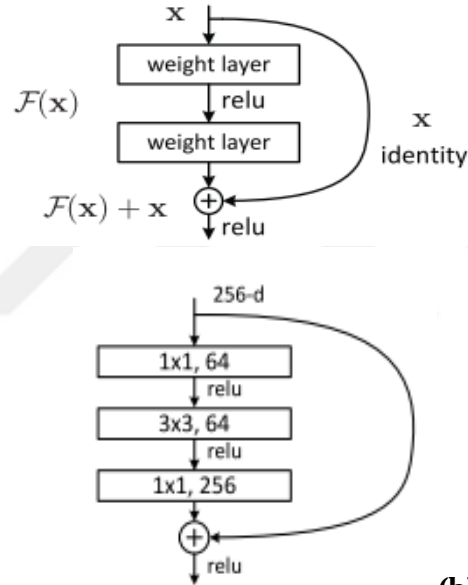
Her CNN tabanlı derin öğrenme algoritmaları çoklu katmanlardan oluşmaktadır. Her katmanın çıktısı birbirinden farklı bir öznetelik çıkarıcı olarak kullanılabilir. Bir görüntü, CNN tabanlı bir ağı giriş olarak verilip ileri yayılım (forward propagation) işlemi yapıldıktan sonra ağı her konvolüsyon katmanında bulunan, önceden eğitilmiş ve sabitlenmiş filtrelerden öznetelik vektörü çıkarılabilir. Örneğin, AlexNet mimarisinde öznetelik çıkarmak için kullanılacak yedi katman mevcuttur.

Bu tez çalışmasında, DenseNet hariç diğer tüm CNN tabanlı derin öğrenme yöntemlerinden öznetelik çıkarımı yapılmıştır.

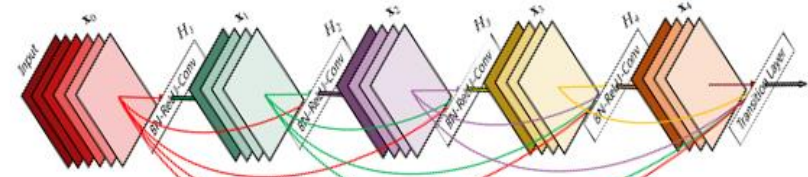




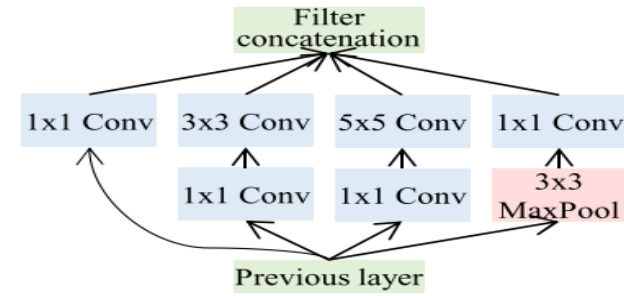
(a)



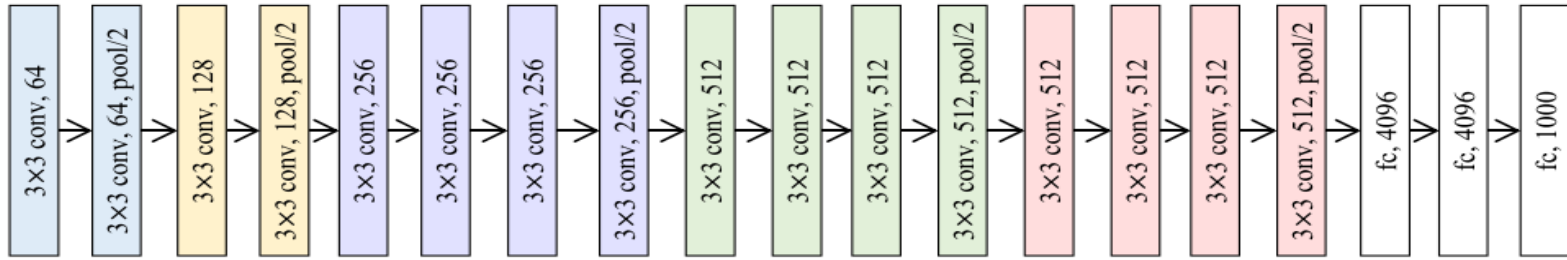
(b)



(c)



(d)



(e)

Şekil 3.7 : Popüler CNN mimarileri: (a) AlexNet (b) Resnet (c) DenseNet (d) Inception (GoogLenet) (e) VGG [22, 24-27, 79].

3.2.1.5 Öznitelik azaltılması

Beşinci adımda, hesaplama maliyetini düşürmek için öznitelik boyut azaltma yöntemleri uygulanır ve literatürde popüler olan iki yöntem aşağıda sunulmuştur:

- Temel Bileşen Analizi (TBA) (Principal Component Analysis (PCA)) [39].
- Oto-kodlayıcılar [80, 81].

Bu çalışmada, yüksek boyutlu öznitelikleri azaltmak için TBA kullanılmış olup ayrıntıları aşağıda anlatılmıştır:

- TBA, bir veri kümesinin, maksimum varyansı temsil eden daha az sayıda değişken ile ifade edilmesini sağlayan bir veri dönüştürme yöntemidir. TBA yönteminin amacı, yüksek boyutlu veriden daha az sayıda veri kullanarak en iyi dönüşümü sağlamayı hedefler. Bu dönüşümü verinin maksimum varyansına göre düzenler. TBA yöntemi denklem (3.1)'i sağlamayı hedefler:

$$J_{PCA}(W_{opt}) = \arg \max_W |W^T S_T W| \quad (3.1)$$

Burada S_T , verinin toplam dağılım matrisi ve W , sütunları ortonormal izdüşüm vektörlerinden oluşan bir izdüşüm matrisidir. TBA, maksimum varyans doğrultusunda ortogonal temel vektör setini bularak verinin boyut küçültme işlemini gerçekleştirir.

3.2.1.6 Sınıflandırıcı

Altıncı adımda, çeşitli sınıflandırıcı yöntemleri kullanılmıştır:

- *Doğrusal Ayırtaç Analizi (DAA) (Linear Discriminant Analysis (LDA))*: DAA sınıflandırıcının amacı, sınıfların birbirinden ayırt ediciliğini en üst düzeye çıkaran doğrusal ayırtaç izdüşüm vektörlerini bulmaktır [40]. DAA yöntemi denklem (3.2)'de gösterilen oranı en üst düzeye çıkarmaya çalışır:

$$J_{DAA}(W_{opt}) = \arg \max_W \left| \frac{W^T S_B W}{W^T S_W W} \right| \quad (3.2)$$

S_B , S_W ve W sırasıyla sınıflar arası dağılım matrisi, sınıf içi dağılım matrisi ve izdüşüm matrisini temsil etmektedir. Böylece, DAA sınıflar içi dağılım matrisini en aza indirirken sınıflar arası dağılım matrisini büyütme çalışır.

- *Ayırt Edici Ortak Vektör Yaklaşımı (AOVY) (Discriminative Common Vector Approach (DCVA))*: AOVY sınıflandırıcının amacı, tüm sınıfların sınıf içi dağılım matrisini kullanarak ortak vektörler elde etmektir. Ortak vektörler elde edildikten sonra, bunların toplam dağılımı optimal izdüşüm vektörlerini elde etmek için maksimize edilir [82, 83]. AOVY yöntemi denklem (3.3)'ü sağlamayı hedefler:

$$\begin{aligned}
 J_{AOVY}(W_{opt}) &= \arg \max_{|W^T S_w W|=0} |W^T S_B W| = \arg \max_{|W^T S_w W|=0} |W^T S_T W| \\
 &= \arg \max_W |W^T S_{com} W|
 \end{aligned} \tag{3.3}$$

Burada W , sütunları ortonormal optimal izdüşüm vektörleri olan bir izdüşüm matrisini ve S_{com} ise ortak vektörlerin dağılım matrisini temsil etmektedir. AOVY yöntemi ortak vektörlerin dağılım matrisini en üst düzeye çıkarmaya çalışır.

- *Bütünsel Yerel Doğrusal Ayırtaç Analizi (BYDAA) (Complete Global-Local Linear Discriminant Analysis (CGLDA))*: BYDAA sınıflandırıcının amacı, hem yerel hem de bütünsel bilgiyi yakalamaktır [84, 85]. BYDAA yöntemi denklem (3.4)'deki oranı en üst düzeye çıkarmayı hedefler:

$$J_{CGLDA}(W_{opt}) = \arg \max_W \left| \frac{W^T (\alpha S_B + (1 - \alpha) S_{TL}) W}{W^T (\varepsilon S_w + (1 - \varepsilon) S_{LW}) W} \right| \tag{3.4}$$

Denklem (3.4)'daki parametrelerin açıklamaları aşağıdadır:

- S_B : Sınıflar arası dağılım matrisi,
- S_{TL} : Toplam yerel örüntü varyasyon dağılım matrisi,
- S_w : Sınıf içi dağılım matrisi,
- S_{LW} : Yerel sınıf içi dağılım matrisi,
- W : İzdüşüm matrisi,
- α, ε : İnce ayar parametreleri.

α parametresi, bütünsel ayırt edici bilgi ile yerel örüntü varyasyonu arasındaki dengeyi kontrol ederken; ε parametresi ise bütünsel benzerlik bilgisi ile yerel benzerlik varyasyonu arasındaki dengeyi sağlar.

- *Genelleştirilmiş Ayırtaç Analizi (GAA) (Generalized Discriminant Analysis (GDA))*: GAA sınıflandırıcının amacı, eşlenmiş bir öznitelik uzayı içinde sınıflar arası dağılım matrisinin, sınıf içi dağılım matrisine oranını maksimize etmektir [86]. Örneğin; X uzayının doğrusal olmayan bir eşleme fonksiyonu Φ , ile bir Hilbert uzayına F , eşlendiği varsayılın.

$$\Phi: X \rightarrow F, x \rightarrow \Phi(x) \quad (3.5)$$

Bu şartlarda GAA yöntemi denklem (3.6)'daki kernel fonksiyonunu kullanarak girdi setini doğrusal olmayan bir yapıya dönüştürür:

$$(k_{ij})_{pq} = \varphi^t(x_{pi}) \cdot \varphi(x_{qj}) = k(x_{pi}, x_{qj}) = e^{-\frac{|x_{pi}-x_{qj}|^2}{\sigma}} \quad (3.6)$$

Burada x_{pi} parametresi p sınıfının i . örneğini, x_{qj} parametresi q sınıfının j . örneğini ve σ ise varyansı temsil etmektedir.

Bu çalışmada, yukarıdaki bahsedilen sınıflandırıcı yöntemleri kullanılmış olup sınıflandırıcı yöntemi artırılmasına yönelik literatürde yaygın olarak kullanılan Destek Vektör Makinesi (DVM) (Support Vector Machine (SVM)) [87-89] eklenebilir.

3.2.1.7 Metrik uzayı

Bir metrik, bir kümedeki her bir eleman çifti arasındaki uzaklığı ölçen ve böylece benzerlik ölçüsünü hesaplayan bir fonksiyondur. Literatürde metrikler, önceden tanımlanmış ve öğrenilmiş olmak üzere iki kategoriye ayrılmaktadır [90, 91].

Önceden tanımlanmış metrikler veri bilgisini kullanmamakta olup “*Manhattan, Öklid, Kosinüs, Ki-kare, Hamming ve Korelasyon*” örnek olarak verilebilir.

Öğrenilmiş metrikler ise veri bilgisini kullanmakta olup örnek olarak “*Mahalanobis*” ve derin öğrenme tabanlı öğrenilmiş metrik yöntemleri verilebilir [92-94].

Bu tez çalışmasında, “*Öklid, Kosinüs ve Manhattan*” olmak üzere önceden tanımlanmış üç farklı metrik kullanılmıştır.

3.2.2 Önerilen derin kaskat skor seviyesi füzyon algoritması

Çeşitlilik bakımından zengin ve modüler yapıya sahip bir modalite havuzu oluşturulduktan sonra, eğitim ve test seti için tamamen birbirinden farklı en az iki kulak veri seti kullanılmasına ihtiyaç vardır. Buradaki amaç, çalışmanın koşullarını zorlaştırarak önerilen yöntemin kazanımlarını daha iyi analiz edebilmektir.

Bölüm-1’de kısıtlanmış ve sınırlı veri problemi olarak tanımlanmış ULDP’nin üstesinden gelebilmek için derin kaskat skor seviyesi füzyon (DCSLF) algoritması önerilmiştir [34]. DCSLF algoritmasının kullanılması için bir modalite havuzu oluşturulması gereksinimi mevcuttur. Söz konusu modalite havuzu Bölüm-3.2.1’de ayrıntılı bir şekilde anlatılmış olup genişletilmesi veya küçültülmesi de mümkündür. DCSLF algoritmasının ayrıntıları aşağıdaki sunulmuştur:

$$D_S = D_{SX} + D_{SY} \quad (3.7)$$

$$D_S \cap D_T = \emptyset \quad (3.8)$$

Burada, D_S , eğitim ve validasyon (doğrulama) seti; D_{SX} , sadece eğitim seti; D_{SY} , sadece validasyon (doğrulama) seti ve D_T , test setini temsil etmektedir. Denklem (3.7) ve (3.8) ise bu veri setleri arasındaki ilişkiyi açıklamaktadır. Buradaki en önemli noktalardan biri, eğitim ve test fazında birbirinden tamamen bağımsız veri setleri kullanılmasıdır.

D_{SX} ve D_{SY} modalite havuzuna girdi olarak verilir. Modalite havuzundan bir modalitenin iç yapısı rastgele (otomatik) veya kullanıcı (manuel) tarafından oluşturulabilir. Bir modalitenin iç yapısı oluşturma örneği, aşağıda gösterilmiştir:

- | | |
|---|------------------------|
| 1) Renk haritası | : Kırmızı, yeşil, mavi |
| 2) Görüntü boyutu | : 80x60 |
| 3) Ön işleme ve görüntü iyileştirme | : Histogram eşitleme |
| 4) Öznitelik çıkarımı ve parametre seti | : HOG, [8,2,1,12,1] |
| 5) Öznitelik azaltılması | : TBA |
| 6) Sınıflandırıcı | : DAA |
| 7) Metrik uzayı | : Kosinüs |

Yukarıdaki dördüncü adım sonuna kadar eğitim ve validasyon seti için yapılan işlemler tamamen birbirinden bağımsızdır. Ancak, beşinci adımda eğitim setinin izdüşüm matrisi validasyon seti için kullanılır. Benzer şekilde altıncı adımda da eğitim setindeki sınıflandırıcı izdüşüm matrisi validasyon seti için kullanılır. Yedinci adımda ise eğitim ve validasyon seti için bir metrik ölçümü kullanılarak skor eşleştirme matrisi elde edilir.

Bir modaliteden skor eşleştirme (score matching (SM)) matrisi elde edilmesi için kullanılan modalite fonksiyonu denklem (3.9)'da tanımlanmıştır:

$$SM_i = M_i(D_{SX}, D_{SY}) \quad i \in (1, 2, 3, \dots, n) \quad (3.9)$$

Burada, D_{SX} ve D_{SY} modalite fonksiyonlarının girdisi, SM_i , i.nci modalite fonksiyonunun çıktısıdır. SM matrisi, çeşitli kombinasyonlarda elde edilebilmekte olup SM matrisi kombinasyon çeşitleri Şekil 3.8'de sunulmuştur:



Şekil 3.8 : SM matrisi kombinasyon çeşitleri.

SM matrisi kombinasyonlarının içinde en zor olanı eğitim setini içinde barındırmayan iki numaralı kombinasyondur. Bu tez çalışmasında iki numaralı kombinasyon uygulanmıştır.

Yukarıdaki işlemler özetlenirse, D_{SX} ve D_{SY} modalite havuzuna girdi olarak verilir. Modalite havuzundan rastgele (otomatik) veya kullanıcı (manuel) tarafından bir modalitenin iç yapısı üretilir. Ardından, modalite fonksiyonu kullanılarak SM matrisi elde edilir. Buraya kadar yapılan işlemler eğitim fazı bitine kadar sürekli tekrarlanacaktır. Bundan sonraki amaç, modalite havuzundan füzyon için en verimli modaliteleri seçmek ve modalitelere ait SM matrislerinin füzyon için gerekli ağırlık parametrelerini bulmaktır. Denklem (3.9) birinci kez uygulandığında,

$$SM_1 = M_1(D_{SX}, D_{SY}) \quad (3.10)$$

elde edilir. Burada SM_1 , ilk skor eşleştirme matrisidir. Skor eşleştirme matrislerinin değerleri farklı aralıklarda olduğu için, skor eşleştirme matrisleri füzyon işlemi için doğrudan kullanılamaz veya birleştirilemez. Bu nedenle her modalite için elde edilen skor eşleştirme matrisini normalleştirmek için “ $\tanh$ ” operatörü kullanılır. “ $Min-max$ ” ve “ $z-score$ ” normalizasyonu gibi diğer normalizasyon teknikleri, verilerdeki aykırı değerlere karşı duyarlı olsa da, “ $\tanh$ ” normalizasyonu, skor normalleştirme için gürbüz (robust) ve etkili bir işlemdir [95]. İlk skor eşleştirme matrisi, “ $\tanh$ ” normalizasyonu kullanılarak aşağıdaki gibi normalize edilir:

$$\widehat{SM}_1 = \frac{1}{2} * \left\{ \tanh \left(0.01 * \left(\frac{SM_1 - \mu}{\sigma} \right) \right) + 1 \right\} \quad (3.11)$$

Burada, $\widehat{SM}_1$, normalize edilmiş skor eşleştirme matrisi olup μ ve σ , sırasıyla, eşleştirme matrisinin ortalama ve varyans değerleridir. Başlangıçta, ilk normalize edilmiş skor seviyesi füzyon matrisi, $\widehat{SLF}_1$, $\widehat{SM}_1$ matrisinin değerini alır:

$$\widehat{SLF}_1 = \widehat{SM}_1, \quad (\beta_1 = 1) \quad (3.12)$$

Her döngüde ($i = 1, 2, 3, \dots, n$), SM_i matrisi aşağıdaki döngülere [denklem (3.13), (3.14) ve (3.15)] aşağıdaki gibi eklenir²:

² İlk döngü denklem (3.12) ile başladığından indeks $i+1$ ile başlamıştır.

$$\widehat{SM}_{i+1} = \text{normalize}(SM_{i+1}) \quad (3.13)$$

$$SLF_{i+1} = (1 - \beta_{i+1}) * \widehat{SLF}_i + \beta_{i+1} * \widehat{SM}_{i+1} \quad (3.14)$$

$$\widehat{SLF}_{i+1} = \text{normalize}(SLF_{i+1}) \quad (3.15)$$

Burada, $\widehat{SLF}_i$, normalize edilmiş i.nci skor seviyesi füzyon matrisidir. β parametresi, füzyon öğrenme oranı olarak tanımlanmış olup başlangıç değeri ilk döngü olan denklem (3.12)'de $\beta_1=1$ olarak atanmıştır. β parametresinin değer aralığı 0 ile 1 arasındadır. β parametresi, aynı zamanda yeni modalitenin ağırlık füzyonunu temsil etmektedir. Herhangi bir katmanda, β parametresinin değerinin 0,5'in üzerinde olması, yeni modalitenin tek başına diğer önceki skor seviyesi füzyon edilmiş modalitelerden daha çok tanıma oranı sağladığını göstermektedir.

İkinci döngüde, modalite havuzundan seçilen ikinci modalitenin skor eşleştirme matrisi normalize edildikten sonra denklem (3.14) ile füzyon edilir. Ardından, füzyon sonucunda elde edilen skor seviyesi füzyon matrisi normalize edilir.

Füzyon öğrenimi, validasyon seti için elde edilen kulak tanıma oranına göre gerçekleştirilir. Yeni seçilen modalite, skor seviyesi füzyon işlemi sonucu kulak tanıma oranını yükseltmesi durumunda sistem tarafından kabul edilip aksi takdirde reddedilir.

Yeni seçilen modalitenin kabul edilmesi durumunda β değeri, etraflı veya tam arama (exhaustive search) kullanılarak bulunur. β parametresi kullanılarak elde edilen skor seviyesi füzyon matrisi mevcut skor seviyesi füzyon matrisinin yerine atanır. Bu işlemler her döngü için tekrarlanır.

Döngüye kabul edilen her yeni modalite diğer modalitelerin ağırlığını azaltmaktadır. Bir önceki döngüde kullanılan denklem (3.14) bir sonraki döngüye eklendiğinde aşağıdaki sonuç bulunur:

$$SLF_{i+2} = (1 - \beta_{i+2}) * [(1 - \beta_{i+1}) * \widehat{SLF}_i + \beta_{i+1} * \widehat{SM}_{i+1}] + \beta_{i+2} * \widehat{SM}_{i+2} \quad (3.16)$$

Elde edilen bu skor seviyesi füzyon matrisi aşağıdaki forma göre düzenlenebilir:

$$SLF_{i+2} = (1 - \beta_{i+2}) * (1 - \beta_{i+1}) * \widehat{SLF}_i + (1 - \beta_{i+2}) * \beta_{i+1} * \widehat{SM}_{i+1} + \beta_{i+2} * \widehat{SM}_{i+2} \quad (3.17)$$

Birinci ve ikinci modalitelerin ağırlıklarının azaldığı denklem (3.17)'de açıkça görülmektedir. Dolayısıyla son durumda sıfır veya sıfıra yakın ağırlığa sahip olan modaliteler füzyondan otomatik olarak uzaklaştırılır, yani füzyon katkıları ihmal edilebilir.

0 ve 1 arasında değişen optimum β değerini etraflı veya tam arama (exhaustive search) ile bulmak için bir adım sayısı belirlenir. Bu adım sayısı β çözünürlüğü olarak tanımlanmıştır. Bu tez çalışmasında, β çözünürlüğünün başlangıç değeri 0,05 olarak atanmıştır. Füzyon edilen modalite sayısı arttıkça veya başka bir ifade ile füzyon öğrenme gerçekleştikçe β çözünürlüğünün azaltılmasına ihtiyaç duyulmaktadır. Bu tez çalışmasında, 15 modaliteyi içeren füzyon işlemi uygulandıktan sonra β çözünürlüğü 0,01 olarak değiştirilmiştir. İşlem hızının arttırılması için de β değer aralığı 0 ve 0,2 arasına çekilmiştir.

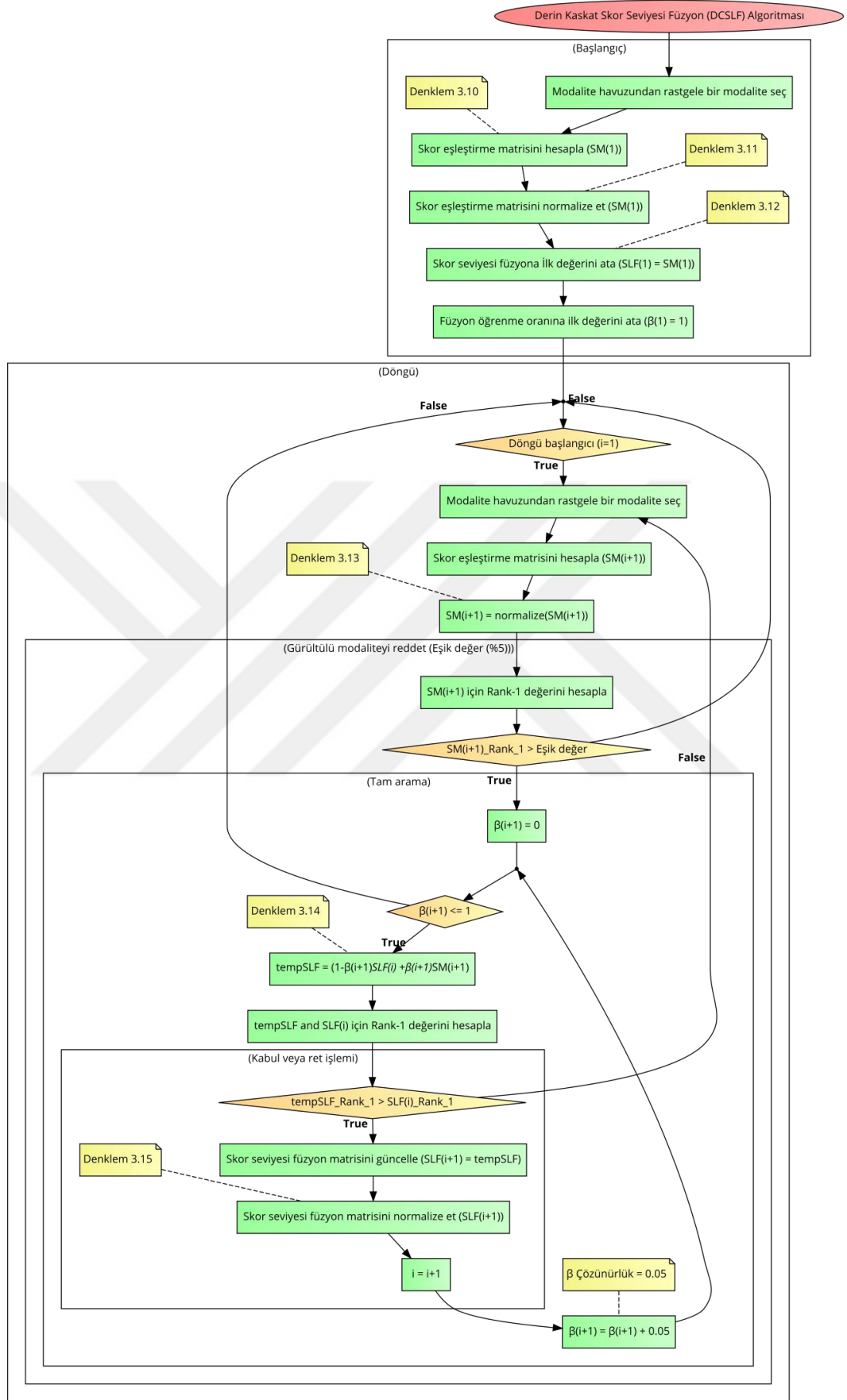
Füzyon öğrenimi sürecinde, gürültü sonucu doğruluk açısından önemli ölçüde zayıf performansa sahip modaliteler kaskat ağ yapısına dahil olabilmektedir. Bu nedenle istenmeyen modalitelerin sisteme dahil olmasını engellemek için bir eşik değeri belirlenir. Bu tez çalışmasında tanıma oranı eşik değeri %5 olarak belirlenmiştir.

Füzyon öğrenme işleminde, validasyon seti için elde edilen kulak tanıma oranında değişim sıfıra yaklaşılmaya başladığında füzyon eğitimi durdurulur. Ardından, kaskat füzyon ağındaki tüm füzyon ağırlıkları ve kabul edilmiş modaliteler sabitlenerek D_T 'ye karşı test edilir.

Yukarıda açıklanan DCSLF algoritmasının akış şeması³ Şekil 3.9'da, sözde kodu⁴ Çizelge 3.1'de sunulmuştur.

³ Akış şeması <https://code2flow.com/app> web sitesi kullanılarak çizilmiştir.

⁴ Sözde kod <https://code2flow.com/app> web sitesi kullanılarak yazılmıştır.



Şekil 3.9 : DCSLF algoritmasının akış şeması.

Çizelge 3.1 : DCSLF algoritmasının sözde kodu.

Derin Kaskat Skor Seviyesi Füzyon (DCSLF) Algoritması;

block (Başlangıç)

```
{
Modalite havuzundan rastgele bir modalite seç;
Skor eşleştirme matrisini hesapla ( $SM_1$ ); //Denklem 3.10
Skor eşleştirme matrisini normalize et ( $\widehat{SM}_1$ ); //Denklem 3.11
Skor seviyesi füzyona ilk değerini ata ( $\widehat{SLF}_1 = \widehat{SM}_1$ ); //Denklem 3.12
Füzyon öğrenme oranına ilk değerini ata ( $\beta_1 = 1$ );
}
```

block (Döngü)

```
{
while (Döngü başlangıcı (i=1))
{
XX:
Modalite havuzundan rastgele bir modalite seç;
Skor eşleştirme matrisini hesapla ( $SM_{i+1}$ );
 $\widehat{SM}_{i+1} = \text{normalize}(SM_{i+1})$ ; //Denklem 3.13
block (Gürültülü modaliteyi reddet (Eşik değer (%5)))
{
 $SM_{i+1}$  için Rank-1 değerini hesapla;
if ( $SM_{i+1\_Rank\_1} > \text{Eşik değeri}$ )
{
block (Tam arama)
{
for ( $\beta_{(i+1)} = 0$ ;  $\beta_{(i+1)} \leq 1$ ;  $\beta_{(i+1)} = \beta_{(i+1)} + 0.05$ ) // $\beta$  çözünürlük = 0.05
{
geçici_SLF =  $(1 - \beta_{i+1}) * \widehat{SLF}_i + \beta_{i+1} * \widehat{SM}_{i+1}$ ; //Denklem 3.14
geçici_SLF and  $\widehat{SLF}_i$  için Rank-1 değerlerini hesapla;
block (Kabul veya ret işlemi)
{
if ( $\text{geçici\_SLF\_Rank\_1} > \widehat{SLF}_i\_Rank\_1$ )
{
Skor seviyesi füzyon matrisini güncelle ( $SLF_{(i+1)} = \text{geçici\_SLF}$ );
Skor seviyesi füzyon matrisini normalize et ( $\widehat{SLF}_{i+1}$ ); //Denklem 3.15
i = i+1;
}
}
}
}
}
}
}
}
```

4. DENEY KURULUMU

Daha önceki kulak tanıma çalışmalarında, tek bir kulak veri seti, eğitim ve test seti olmak üzere iki kısma ayrılmıştır. Ancak, bu şekilde test edilen yöntemler veri setine özgü sonuçlar üretebilmekte ve farklı bir kulak veri setine karşı zayıf bir performans sergileyebilmektedir. Literatürde, test setindeki kimlikler eğitim setinde bulunması durumunda kişiye bağımlı; mevcut olmaması durumunda ise kişiye bağımsız senaryo olarak adlandırılmıştır [79].

Bu tez çalışmasında, eğitim ve validasyon (doğrulama) setinde bulunan kimlikler test setinde bulunmayacak şekilde ayarlanmıştır. Bu bağlamda, üç farklı kulak veri seti kullanılmıştır:

- Annotated Web Ears (AWE) [15],
- West Pomeranian University of Technology Ear (WPUTE) [14],
- Unconstrained Ear Recognition Challenge (UERC) [16].

AWE ve WPUTE kulak veri setleri birleşiminden %50 eğitim ve %50 validasyon (doğrulama) olmak üzere iki veri seti hazırlanmıştır. Test seti ise bu iki kulak veri setinden tamamen bağımsız UERC veri setinden oluşmaktadır. Kullanılan tüm veri setleri kısıtlanmamış kulak veri setleri statüsündedir.

4.1 AWE Kulak Veri Seti

AWE kulak veri seti [15], çoğunlukla aktör, müzisyen, politikacı vb. kişilerden oluşmakta olup internet ortamından toplanmıştır. AWE kulak veri seti 100 kişinin toplam 1000 kulak görüntüsüne sahiptir. AWE kulak veri setinde en küçük kulak görüntüsü 15x29 piksel, en büyük kulak görüntüsü 473x1022 piksel ve ortalama kulak görüntü boyutu 83x160 pikseldir. AWE kulak veri setinden örnek görüntüler Şekil 4.1’de sunulmuştur.



Şekil 4.1 : AWE kulak veri setinden örnek görüntüler.

4.2 WPUTE Kulak Veri Seti

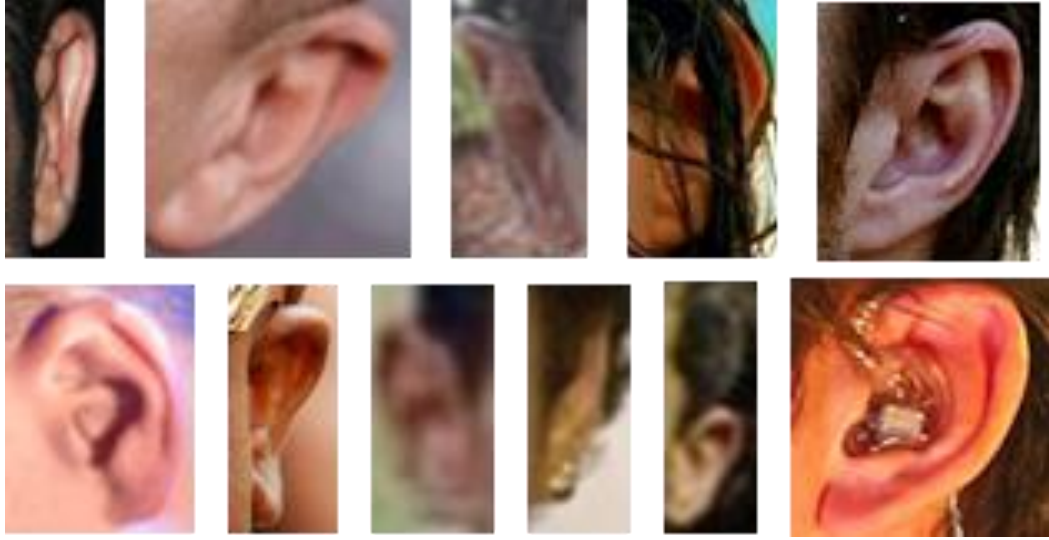
WPUTE kulak veri seti [14], 474 kişinin 3348 kulak görüntüsünü içermektedir. Ancak, WPUTE kulak veri setinde 1388 kulak görüntüsü birebir aynı olduğu için toplamda 1960 kulak görüntüsü mevcuttur. Ayrıca, her kişiden en az dört adet kulak görüntüsü bulunmakta olup kulak görüntü boyutu 380x500 pikseldir. WPUTE kulak veri setinden örnek görüntüler Şekil 4.2’de gösterilmiştir.



Şekil 4.2 : WPUTE kulak veri setinden örnek görüntüler.

4.3 UERC Kulak Veri Seti

UERC kulak veri seti [16], katılımcıların 3540 kişiye ait 9500 galeri görüntüsünün 1482 kişiye ait 7742 prob görüntüsü ile karşılaştırması için bir kulak tanıma yarışması kapsamında geliştirilmiştir. Katılımcılardan kulak tanıma performans değerlendirilmesi için 9500x7742 boyutunda benzerlik matrisinin üretilmesi istenmiştir. UERC kulak veri setinden örnek görüntüler Şekil 4.3’te sunulmuştur.



Şekil 4.3 : UERC kulak veri setinden örnek görüntüler.

4.4 Çalışma Ortamı

Tüm uygulamalar *HP Z440* iş istasyonu *Intel® Xeon® CPU E5-1650 v4 @ 3.60 Ghz* (6 Çekirdek), *16 GB RAM* özellikleri bulunan bilgisayarda, *Linux* işletim sisteminde *MATLAB 2017b öğrenci versiyonu* ile gerçekleştirilmiştir.

4.5 Performans Değerlendirilmesi

Kulak tanıma performans değerlendirilmesi kapsamında deney sonuçlarının görselleştirilmesi için Kümülatif Eşleştirme Karakteristiği (Cumulative Matching Characteristic (CMC)) eğrisi kullanılmıştır. Ayrıca, kulak tanıma performansı aşağıdaki nicel metrikler ile de ölçülmüştür:

- Sıra-1 (Rank-1) : Prob kulak görüntüsünü birinci sırada tanıma oranı,
- Sıra-5 (Rank-5) : Prob kulak görüntüsünü beşinci sırada tanıma oranı,
- AUC : CMC eğrisinin altında kalan alan.



5. DENEY SONUÇLARI

Bölüm-2’de bahsedilen, 2017 yılında IJCB kapsamında düzenlenen “Kısıtlanmamış Görüntüler Üzerinde Kulak Tanıma Yarışması” deney protokolü ve performans ölçümlerine uygun olarak, önerilen ve daha önce bildirilmiş diğer tüm kulak tanıma yöntemlerinin genel performansını ve ölçeklenebilirliğini (scalability) değerlendirmek ve karşılaştırmak için iki deney gerçekleştirilmiştir.

Genel performans değerlendirmesi için gerçekleştirilen ilk deneyde, 180 kişiden oluşan 1800 test kulak görüntüsü kullanılmıştır. Bu deneyde, tümüne karşı tümü (all-versus-all) senaryosuna istinaden 1800×1800 skor eşleştirme matrisi hesaplanmıştır. Literatürde bu tip senaryo karşılaştırılması, kapalı set kimliklendirme (closed set identification) olarak adlandırılmıştır. Kapalı set kimliklendirme, kimliği belirlenmemiş bir bireyin biyometrik özelliklerinin veri tabanında olduğu ve sistemin bu kişiyi belirlemeye çalıştığı biyometrik bir görev olarak tanımlanmıştır [79].

Ölçeklenebilirlik (scalability) değerlendirmesi için yapılan ikinci deneyde, 3540 kişiden oluşan 9500 kulak görüntüsü, her bir kişiden en az iki görüntü içeren 1482 kişinin 7742 kulak görüntüsü ile karşılaştırılmıştır. Bu durumda, eşleşen çiftlerden 7742×9500 boyutunda skor eşleştirme matrisi oluşmuştur. Literatürde bu tip senaryo karşılaştırılması, açık set kimliklendirme (open set identification) veya izleme listesi (watchlist) olarak isimlendirilmiştir. Açık set kimliklendirme veya izleme listesi, bir kişinin veri tabanında olup olmadığını belirlemek ve veri tabanındaki kişinin kaydını bulmak için operasyonel biyometrik sistem koşullarını daha yakından takip eden bir görev olarak tanımlanmıştır [79].

Önerilen “ScoreNet-1”, genel performans değerlendirmesi için gerçekleştirilen ilk deneyde Çizelge 5.1’de listelenen deney sonuçlarından en iyi kulak tanıma oranını elde etmiştir. “ScoreNet-1”, WPUTE ve AWE kulak veri setleri kullanılarak Bölüm-3’te bahsedilen “Derin Kaskat Skor Seviyesi Füzyonu (DCSLF)” ile eğitilmiştir. Füzyon eğitimden sonra, “ScoreNet-1” in kaskat ağ yapısı, 39 modalite ve öğrenilmiş füzyon ağırlıklarını içermektedir.

Çizelge 5.1 : Kısıtlanmamış kulak tanıma için genel performans değerlendirmesi.

Yöntem	Rank-1 (%)	Rank-5 (%)	AUC
UCCS [16]	90,4*	100*	0,994*
ICL [16]	5,3	14,8	0,717
IAU [16]	38,5	63,2	0,940
IITK [16]	22,7	43,6	0,861
ITU-I [16]	24,0	46,0	0,890
ITU-II [16]	27,3	48,3	0,877
<i>“ScoreNet-1”</i>	47,8	72,1	0,958

*Bu sonuçların doğrulanmasına ihtiyaç vardır.

Colorado Colorado Springs Üniversitesi (UCCS) tarafından üretilen sonuçlar hariç, *“ScoreNet-1”* diğer tüm çalışmalardan Sıra-1 (Rank-1)’de en az %9 daha iyi tanıma oranı performansı göstermiştir. Ancak, UCCS sonuçlarının daha fazla doğrulanması gerektiği ve eğitim için test görüntülerinin kullanıp kullanılmadığının belirlenmesine ihtiyaç olduğu bildirilmiştir [44].

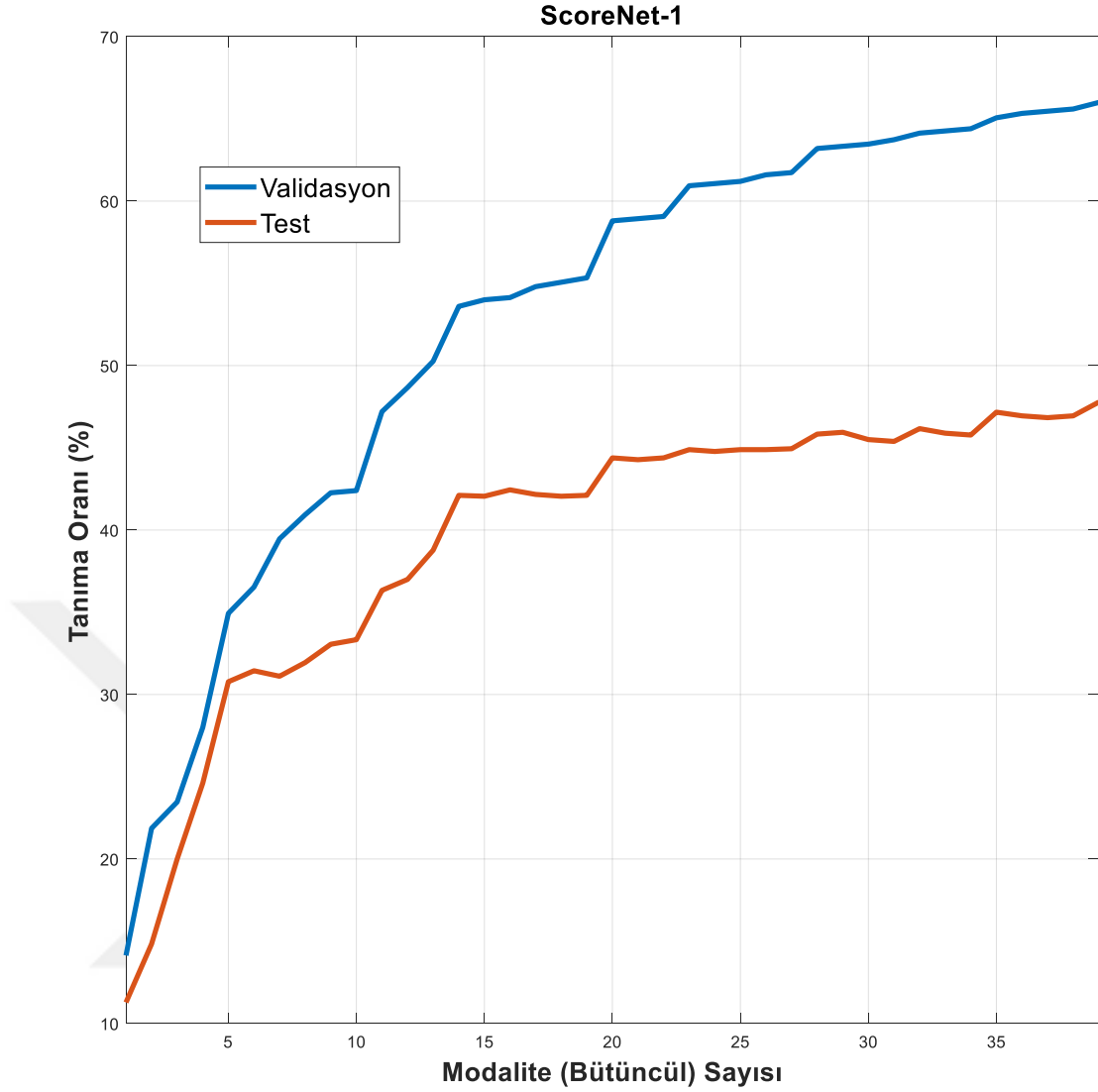
Validasyon (doğrulama) ve test deneylerinde *“ScoreNet-1”*’in Sıra-1 (Rank-1)’de tanıma oranı değişim grafiği Şekil 5.1’de sunulmuş olup içeriği, CMC eğrileri, detaylı validasyon (doğrulama) ve test sonuçları Ek’tte listelenmiştir.

Ölçeklenebilirlik (scalability) değerlendirmesi için yapılan ikinci deneyde ise *“ScoreNet-2”*’nin diğer yaklaşımlar ile karşılaştırılması Çizelge 5.2’de listelenmiştir. İkinci deney sonuçlarına göre önerilen yöntem Sıra-1 (Rank-1)’de en az %12,5 bir iyileşme sağlamıştır. *“ScoreNet-2”*’nin eğitim süreci *“ScoreNet-1”* ile aynı olup eğitim sonunda *“ScoreNet-2”*’nin kaskat ağ yapısı, 14 modalite ve öğrenilmiş füzyon ağırlıklarını içermektedir.

Çizelge 5.2 : Kısıtlanmamış kulak tanıma için ölçeklenebilirlik değerlendirmesi.

Yöntem	Rank-1 (%)	Rank-5 (%)	AUC
UCCS [16]	22,3*	26,3*	0,858*
ICL [16]	2,9	7,08	0,823
IAU [16]	8,4	14,2	0,810
IITK [16]	5,4	10,2	0,814
ITU-I [16]	6,1	12,6	0,842
ITU-II [16]	6,9	12,8	0,844
<i>“ScoreNet-2”</i>	20,9	33,42	0,883

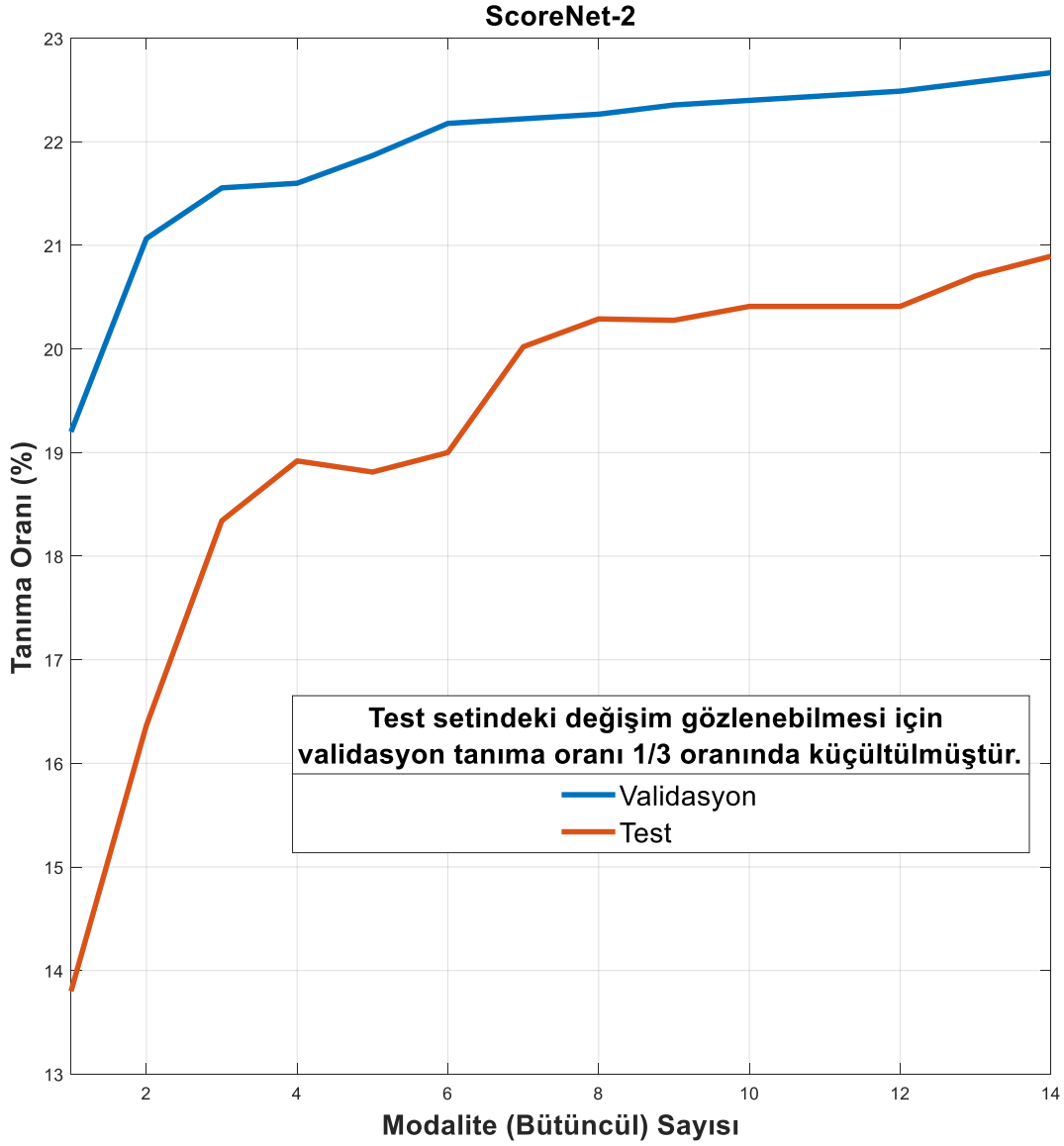
*Bu sonuçların doğrulanmasına ihtiyaç vardır.



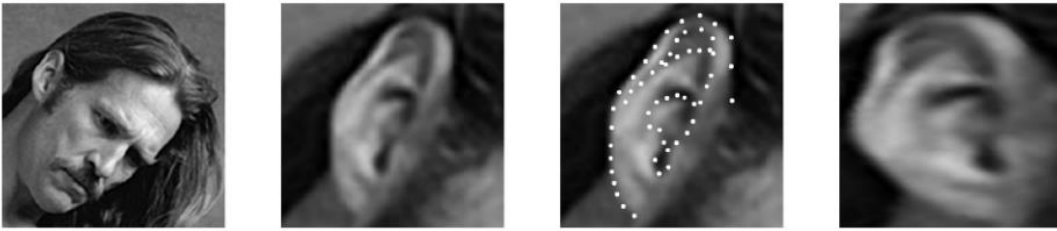
Şekil 5.1 : “ScoreNet-1 ”in tanıma oranı değişim grafiği.

Validasyon (doğrulama) ve test deneylerinde “ScoreNet-2”nin Sıra-1 (Rank-1)’de tanıma oranı değişim grafiği Şekil 5.2’de sunulmuş olup içeriği, CMC eğrileri, detaylı validasyon (doğrulama) ve test sonuçları Ek’te listelenmiştir.

Literatürde kulak tanıma performansının iyileştirilmesinde yerel işaretçilere (landmarks) dayalı kulak oryantasyonu ve normalizasyon işlemleri kullanılmakta olup gerçekleştirilen örnek bir çalışma Şekil 5.3’te sunulmuştur [44]. Ayrıca, yerel işaretçiler (landmarks) tabanlı oryantasyon ve normalizasyon işlemlerinin yüz tanıma için standart hale gelmesi de dikkat çekicidir [79].



Şekil 5.2 : “ScoreNet-2”nin tanıma oranı değişim grafiği.



Şekil 5.3 : Yerel işaretçiler ile kulak oryantasyon ve normalizasyon işlemleri [44].

Bu kapsamda, yukarıdaki kulak oryantasyon ve normalizasyon prosedürleri uygulandıktan sonra “ScoreNet” mimarisi aynı eğitim sürecine ve aynı testlere tabi tutulmuştur. Çizelge 5.3’te listelenen sonuçlar, “ScoreNet-3”ün kısıtlanmamış görüntüler üzerinde kulak tanıma için genel performans açısından en iyi sonuçları aldığını göstermekte olup önerilen yöntem Sıra-1 (Rank-1)’de en az %12’lik bir

iyileşme daha sağlamıştır. Bununla birlikte, eğitim sonunda “ScoreNet-3”ün kaskat ağ yapısı, 37 modalite ve öğrenilmiş füzyon ağırlıklarını içermektedir.

Çizelge 5.3 : Oryantasyon ve normalizasyon işlemlerinden sonra kısıtlanmamış kulak tanıma için genel performans değerlendirmesi.

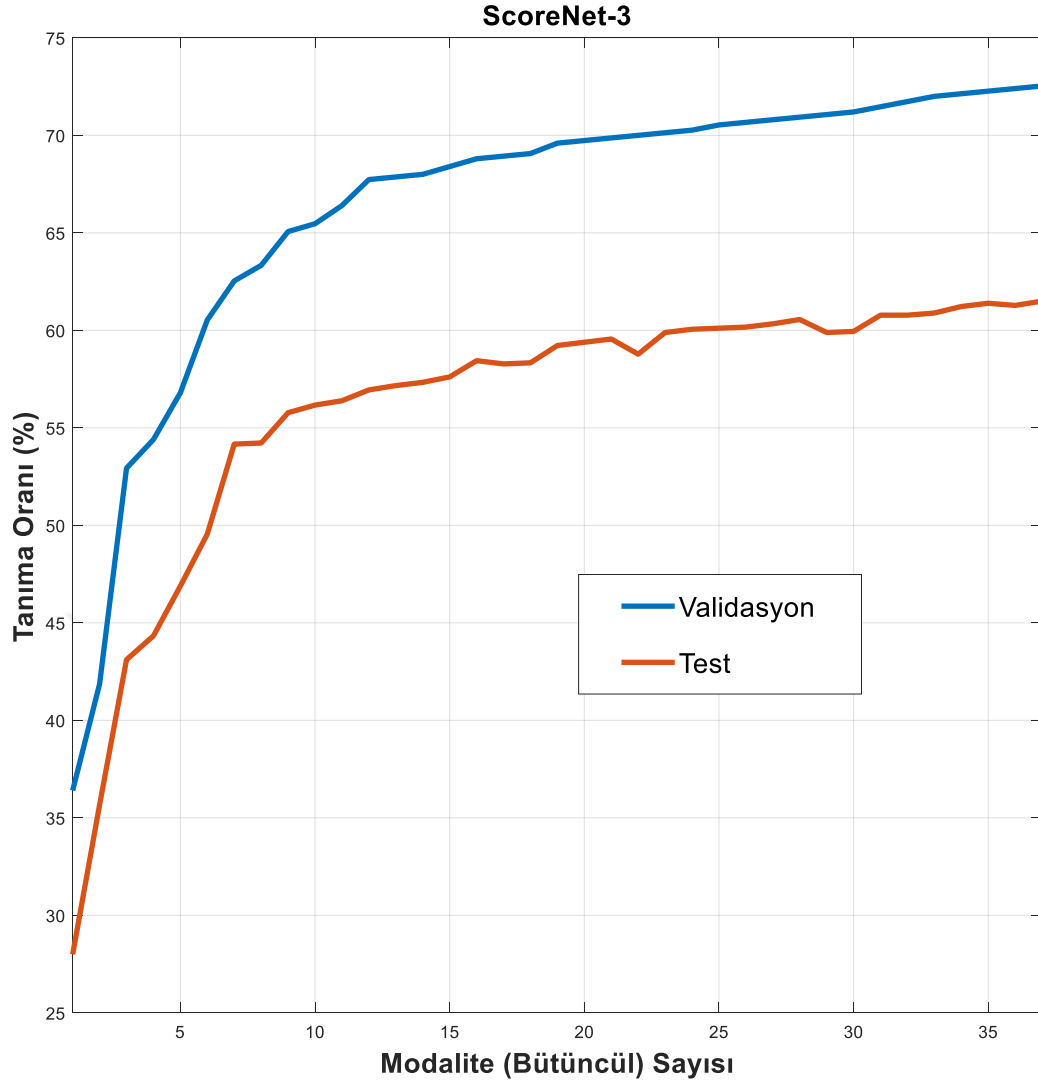
Yöntem	Rank-1 (%)	Rank-5 (%)	AUC
UCCS [16]	90,4*	100*	0,994*
ICL [16]	5,3	14,8	0,717
IAU [16]	38,5	63,2	0,940
IITK [16]	22,7	43,6	0,861
ITU-I [16]	24,0	46,0	0,890
ITU-II [16]	27,3	48,3	0,877
“ScoreNet-1 ”	47,8	72,1	0,958
CNN+HOG [44]	49,06	69,94	0,951
CNN+POEM [44]	47,28	70,00	0,948
“ScoreNet-3 ”	61,50	80,89	0.971

*Bu sonuçların doğrulanmasına ihtiyaç vardır.

Validasyon (doğrulama) ve test deneylerinde “ScoreNet-3”ün Sıra-1 (Rank-1)’de tanıma oranı değişim grafiği Şekil 5.4’te sunulmuş olup içeriği, CMC eğrileri, detaylı validasyon (doğrulama) ve test sonuçları Ek’te listelenmiştir.

Kulak oryantasyon ve normalizasyon işlemleri uygulandıktan sonra ölçeklenebilirlik (scalability) değerlendirmesi için yapılan ikinci deney tekrar değerlendirmeye tabi tutulmuştur. WPUTE ve AWE kulak veri setleri kullanılarak yeniden eğitilen “ScoreNet-4”, Çizelge 5.4’te listelenen deney sonuçlarından en iyi tanıma oranı performansı elde etmiş olup önerilen yöntem literatürde elde edilen en iyi sonuçtan Sıra-1 (Rank-1)’de en az %3,5 bir iyileşme daha sağlamıştır. Bununla birlikte, “ScoreNet-4” kaskat ağ yapısı, 39 modalite ve öğrenilmiş füzyon ağırlıklarını içermektedir.

Validasyon (doğrulama) ve test deneylerinde “ScoreNet-4”ün Sıra-1 (Rank-1)’de tanıma oranı değişim grafiği Şekil 5.5’te sunulmuş olup içeriği, CMC eğrileri, detaylı validasyon (doğrulama) ve test sonuçları Ek’te listelenmiştir.

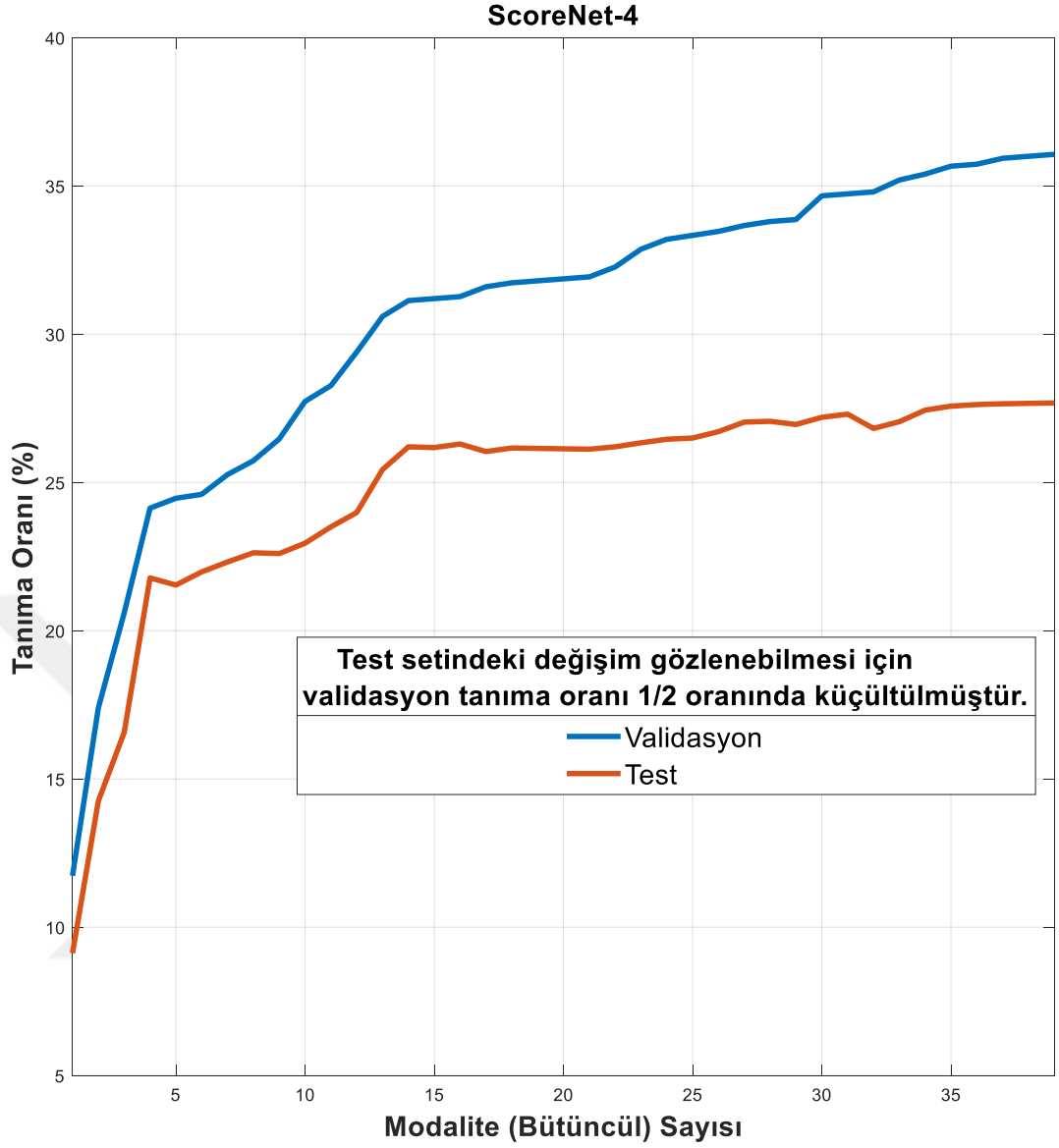


Şekil 5.4 : “ScoreNet-3”ün tanıma oranı değişim grafiği.

Çizelge 5.4 : Oryantasyon ve normalizasyon işlemlerinden sonra kısıtlanmamış kulak tanıma için ölçeklenebilirlik değerlendirme.

Yöntem	Rank-1 (%)	Rank-5 (%)	AUC
UCCS [16]	22,3*	26,3*	0,858*
ICL [16]	2,9	7,08	0,823
IAU [16]	8,4	14,2	0,810
IITK [16]	5,4	10,2	0,814
ITU-I [16]	6,1	12,6	0,842
ITU-II [16]	6,9	12,8	0,844
“ScoreNet-2”	20,9	33,42	0,883
CNN+HOG [44]	24,17	36,43	0,881
CNN+POEM [44]	23,02	35,70	0,882
“ScoreNet-4”	27,68	38,61	0,886

*Bu sonuçların doğrulanmasına ihtiyaç vardır.



Şekil 5.5 : “ScoreNet-4”ün tanıma oranı değişim grafiği.

Özet olarak hem genel performans hem de ölçeklenebilirlik (scalability) değerlendirmesi için yapılan deney sonuçlarına istinaden “ScoreNet”ler ile önceden bildirilmiş diğer tüm yöntemlerin CMC eğrileri Şekil 5.6’da sunulmuştur.

Son olarak, her bir “ScoreNet” için görüntü başına ortalama test süreleri modalitelerin seri ve paralel çalışmasına göre hesaplanmış olup Çizelge 5.5’te gösterilmiştir. Ancak, bu test süreleri kullanılan donanımın CPU ve GPU işlemcilerine göre farklılık gösterecektir.

Çizelge 5.5 : Önerilen yöntemlerin görüntü başına ortalama test süresi.

Yöntem	Modalite sayısı	Modaliteler paralel çalışmakta (ms)	Modaliteler seri çalışmakta (ms)
“ScoreNet-1”	39	85	1387
“ScoreNet-2”	14	58	581
“ScoreNet-3”	37	67	1020
“ScoreNet-4”	39	89	1479

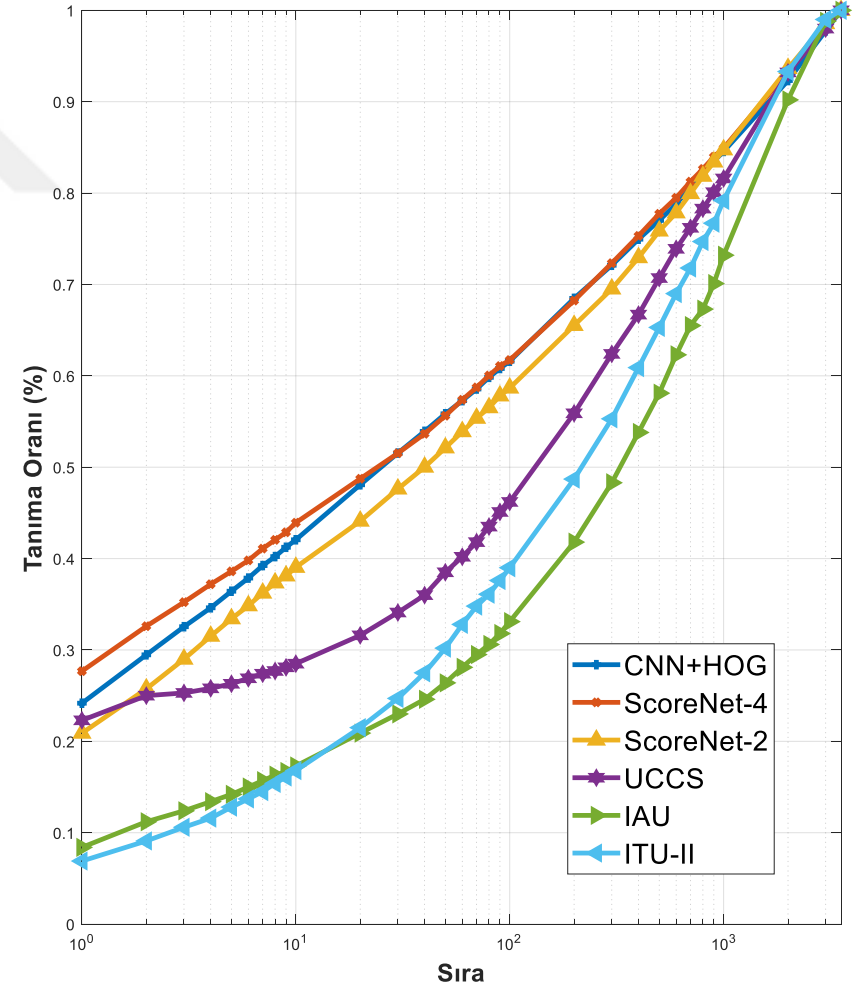
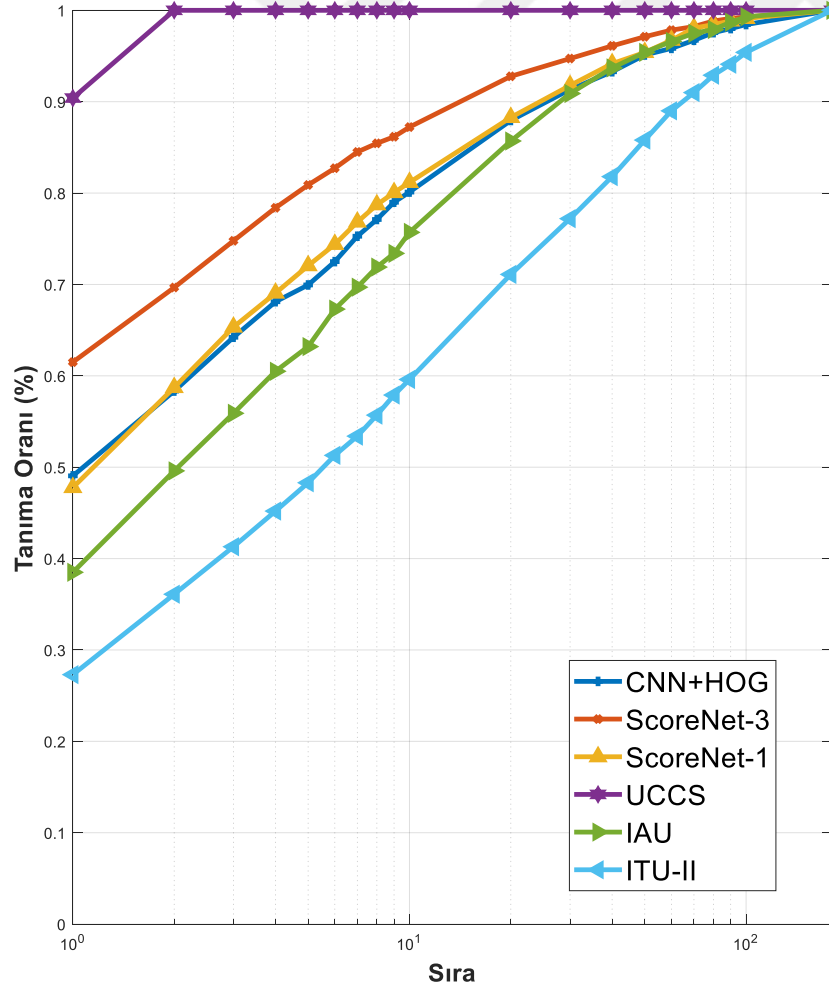
“ScoreNet” kaskat ağ yapısında bulunan modaliteler birbirinden bağımsız olarak çalışabildikleri için paralel işleme uygundur. Ancak, modalitelerin çıktısından elde edilen skor eşleştirme matrislerinin füzyon ağırlıkları ile çarpılmasını ve “*tanh*” normalizasyon hesaplamalarını içeren füzyon işlemleri seri olarak çalışmaktadır. Bundan dolayı modalitelerin sayısının fazla olması hız açısından dezavantaj oluşturmaktadır. Bu nedenle, gerçek dünya senaryolarına göre hız ve tanıma oranı arasında bir denge kurulmasına ihtiyaç duyulacaktır.

“ScoreNet” mimarisi “AutoFL” kapsamında eğitilmesi durumunda modalite havuzundan rastgele modaliteler üretilecek olup bu işlem birçok düşük doğruluklu modalitelerin oluşmasına neden olacaktır. Bunun sonucunda toplam füzyon öğrenme eğitim süresi, yaklaşık iki güne kadar sürebilmektedir. Eğitim süresinin kısaltılması için aşağıdaki kullanıcı müdahaleleri kullanılabilir:

- Önceden bilinen yüksek doğruluklu modaliteler ile başlamak,
- Modalitelerin her katmanında bulunan algoritmaların parametre değerlerini daha iyi sonuç verecekleri aralıklara kısıtlamak,
- Bir ön adım olarak daha az sayıda eğitim ve validasyon (doğrulama) seti kullanarak daha hızlı bir şekilde modalite adayı aramaktır.

Bununla birlikte daha gürbüz (robust) kaskat ağ yapısı oluşturmak için aşağıdaki yöntemler uygulanabilir:

- Düşük modalitelerin kaskat ağ yapısına dahil olmasını engellemek için bir eşik değer kullanmak,
- Validasyon set sayısını birden fazla yaparak daha kararlı modalite adayı seçmektir.



Şekil 5.6 : “ScoreNet”ler ile tüm diğer yöntemlerin CMC eğrileri: (a) Genel performans (b) Ölçeklenebilirlik.



6. SONUÇLAR VE ÖNERİLER

Kulak tanıma alanında “*Kısıtlanmamış ve Sınırlı Veri Problemi (ULDP)*”nin üstesinden gelebilmek için “*ScoreNet*” adı verilen yeni bir mimari önerilmiştir. “*ScoreNet*” mimarisi hem modalite havuzunu hem de füzyon öğrenme algoritmasını içermektedir. Füzyon öğrenme algoritması için “*Derin Kaskat Skor Seviyesi Füzyon (DCSLF)*” algoritması geliştirilmiş; modalite havuzu için de çeşitlilik bakımından zengin ve modüler yapıya sahip özgün bir modalite havuzu oluşturulmuştur.

“*ScoreNet*”, bugüne kadar geliştirilen kulak tanıma yöntemlerini değerlendirmek için en zor veri tabanı olarak kabul edilen UERC kulak veri seti kullanılarak değerlendirilmiştir. Genel performans ve ölçeklenebilirlik (scalability) olmak üzere iki deney gerçekleştirilmiştir. Ayrıca, yüz tanıma literatüründe standart hale gelen ve kulak tanıma performansının iyileştirilmesini sağlayan yer işaretçilere (landmarks) dayalı oryantasyon ve normalizasyon işlemleri uygulandıktan sonra söz konusu iki deney tekrarlanmıştır.

Genel performans için yapılan birinci deney, bireyin biyometrik özelliklerinin veri tabanında mevcut olduğu kapalı set kimliklendirme senaryosunu temsil etmektedir. Genel performans değerlendirme sonucunda önerilen “*ScoreNet-1*”, UERC kulak tanıma sonuçlarına göre Sıra-1 (Rank-1)’de en az %9 daha iyi performans göstermiştir.

Ölçeklenebilirlik (scalability) değerlendirmesi için yapılan ikinci deney, bireyin veri tabanında olup olmadığını belirleyen ve veri tabanındaki kişinin kimliğini teşhis eden açık set kimliklendirme senaryosuna uygundur. Ölçeklenebilirlik (scalability) değerlendirmesi sonuçlarına göre “*ScoreNet-2*”, UERC kulak tanıma yöntemlerinden Sıra-1 (Rank-1)’de en az %12,5’luk bir iyileşme sağlamıştır.

Yerel işaretçilere (landmarks) dayalı oryantasyon ve normalizasyon işlemleri uygulandıktan sonra iki deney tekrar gerçekleştirilmiş olup “*ScoreNet-3*” ve “*ScoreNet-4*”, UERC sonuçlarına göre sırasıyla en az %23 ve %19’luk bir iyileştirme sağlarken; oryantasyon ve normalizasyon işlemlerinden sonra el yapımı ve CNN

tabanlı derin öğrenme algoritmaları içeren daha yeni bir çalışmadan da sırasıyla %12 ve %3,5'luk bir iyileştirme gerçekleştirmiştir.

Böylece “*ScoreNet*” mimarisi kısıtlanmamış görüntülerden oluşan ve bilinen en zorlu kulak veri seti olan UERC kulak veri setinde en son teknoloji doğruluğunu (state-of-the-art accuracy) elde etmiştir. Ancak, “*ScoreNet*” mimarisinin oluşturulması ve eğitilmesi önemli bir iş yükü de getirmektedir. “*ScoreNet*” mimarisi için modalite havuzunun oluşturulması ve eğitim sürecinin uzun sürmesi bu mimarinin dezavantajlarıdır.

“*ScoreNet*” mimarisi kısıtlanmamış görüntüler üzerinde kulak tanıma için ilk “*otomatik füzyon öğrenme (AutoFL)*” yaklaşımını temsil etmektedir. Aynı zamanda bu mimarinin gelecekteki “*AutoFL*” uygulamalarına öncülük edeceği değerlendirilmektedir.

“*ScoreNet*” mimarisinin en iyi sonuçları incelendiğinde kulak tanıma oranı kapalı set kimliklendirme için %61,5 ve açık set kimliklendirme için ise %27,68'dir. Kulak tanıma oranlarının bu kadar düşük olması şöyle özetlenebilir:

- Eğitim ve validasyon (doğrulama) setinde bulunan kişi sayısının az olması,
- Eğitim ve validasyon (doğrulama) setinde bulunan kişiye ait örnek sayısının az olması,
- Test setinde çok düşük çözünürlüklü kulak görüntülerin fazla olması,
- Validasyon (doğrulama) setinin test seti kadar büyük olmaması,
- Eğitim ve validasyon (doğrulama) setinin test seti ile aynı içerikte karmaşıklığa sahip olmaması (Örneğin, çok düşük çözünürlüklü kulak görüntülerin eğitim ve validasyon (doğrulama) setinde test setine göre daha az bulunması),
- Önerilen mimaride düşük çözünürlüklü görüntüleri iyileştirmek için derin öğrenmeye dayalı süper çözünürlük yöntemlerinin henüz yer almamasıdır.

“*ScoreNet*” mimarisi modüler bir yapıda olmasından dolayı geliştirilmeye yatkındır. Füzyon çalışmalarının performansının iyileştirilmesi için gerekli en önemli ihtiyaç çeşitlilik (diversity)'tir. Daha iyi sonuçlar için “*ScoreNet*” mimarisinin modalite havuzunun çeşitliliği genişletilebilir. Bununla birlikte, eğitim ve validasyon (doğrulama) setinde içerik olarak karmaşıklık artırılarak daha iyi füzyon öğrenme gerçekleştirilebilir.

Sonuç olarak, “*ScoreNet*”in özgün mimarisi ile kısıtlanmamış görüntüler üzerinde kulak tanıma alanında önemli bir iyileştirme sağladığı UERC deney sonuçları ile kanıtlanmıştır.





KAYNAKLAR

- [1] **Bansal, A., Castillo, C. D., Ranjan, R., & Chellappa, R.** (2017). The Do's and Don'ts for CNN-based Face Verification. In ICCV Workshops (Vol. 1, No. 2, p. 7).
- [2] **Bansal, A., Nanduri, A., Castillo, C. D., Ranjan, R., & Chellappa, R.** (2017). Umdfaces: An annotated face dataset for training deep networks. In Biometrics (IJCB), 2017 IEEE International Joint Conference on (pp. 464-473). IEEE.
- [3] **Cao, Q., Shen, L., Xie, W., Parkhi, O. M., & Zisserman, A.** (2018). Vggface2: A dataset for recognising faces across pose and age. In Automatic Face & Gesture Recognition (FG 2018), 2018 13th IEEE International Conference on (pp. 67-74). IEEE.
- [4] **Guo, Y., Zhang, L., Hu, Y., He, X., & Gao, J.** (2016). Ms-celeb-1m: A dataset and benchmark for large-scale face recognition. In European Conference on Computer Vision (pp. 87-102). Springer, Cham.
- [5] **Huang, G. B., Mattar, M., Berg, T., & Learned-Miller, E.** (2008). 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.
- [6] **Kemelmacher-Shlizerman, I., Seitz, S. M., Miller, D., & Brossard, E.** (2016). The megaface benchmark: 1 million faces for recognition at scale. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 4873-4882).
- [7] **Klare, B. F., Klein, B., Taborsky, E., Blanton, A., Cheney, J., Allen, K., & Jain, A. K.** (2015). Pushing the frontiers of unconstrained face detection and recognition: Iarpa janus benchmark a. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1931-1939).
- [8] **Maze, B., Adams, J., Duncan, J. A., Kalka, N., Miller, T., Otto, C., & Grother, P.** (2018). IARPA Janus Benchmark-C: Face Dataset and Protocol. In 11th IAPR International Conference on Biometrics.
- [9] **Parkhi, O. M., Vedaldi, A., & Zisserman, A.** (2015). Deep face recognition. In BMVC (Vol. 1, No. 3, p. 6).
- [10] **Sun, Y., Wang, X., & Tang, X.** (2014). Deep learning face representation from predicting 10,000 classes. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1891-1898).
- [11] **Whitelam, C., Taborsky, E., Blanton, A., Maze, B., Adams, J. C., Miller, T., & Cheney, J.** (2017). IARPA Janus Benchmark-B Face Dataset. In CVPR Workshops (Vol. 2, No. 5, p. 6).

- [12] **Wolf, L., Hassner, T., & Maoz, I.** (2011). Face recognition in unconstrained videos with matched background similarity. In *Computer Vision and Pattern Recognition (CVPR), 2011 IEEE Conference on* (pp. 529-534). IEEE.
- [13] **Yi, D., Lei, Z., Liao, S., & Li, S. Z.** (2014). Learning face representation from scratch. *arXiv preprint arXiv:1411.7923*.
- [14] **Frejlichowski, D., & Tyszkiewicz, N.** (2010). The west pomeranian university of technology ear database—a tool for testing biometric algorithms. In *International Conference Image Analysis and Recognition* (pp. 227-234). Springer, Berlin, Heidelberg.
- [15] **Emeršič, Ž., Štruc, V., & Peer, P.** (2017). Ear recognition: More than a survey. *Neurocomputing*, 255, 26-39.
- [16] **Emeršič, Ž., Štepec, D., Štruc, V., Peer, P., George, A., Ahmad, A., & Zafeiriou, S.** (2017, October). The unconstrained ear recognition challenge. In *Biometrics (IJCB), 2017 IEEE International Joint Conference on* (pp. 715-724). IEEE.
- [17] **Dalal, N., & Triggs, B.** (2005). Histograms of oriented gradients for human detection. In *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on* (Vol. 1, pp. 886-893). IEEE.
- [18] **Ojala, T., Pietikainen, M., & Maenpaa, T.** (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on pattern analysis and machine intelligence*, 24(7), 971-987.
- [19] **Ahonen, T., Rahtu, E., Ojansivu, V., & Heikkila, J.** (2008). Recognition of blurred faces using local phase quantization. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on* (pp. 1-4). IEEE.
- [20] **Kannala, J., & Rahtu, E.** (2012). Bsif: Binarized statistical image features. In *Pattern Recognition (ICPR), 2012 21st International Conference on* (pp. 1363-1366). IEEE.
- [21] **Haghighat, M., Zonouz, S., & Abdel-Mottaleb, M.** (2015). CloudID: Trustworthy cloud-based and cross-enterprise biometric identification. *Expert Systems with Applications*, 42(21), 7905-7916.
- [22] **Krizhevsky, A., Sutskever, I., & Hinton, G. E.** (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).
- [23] **Emeršič, Ž., Štepec, D., Štruc, V., & Peer, P.** (2017). Training convolutional neural networks with limited training data for ear recognition in the wild. *arXiv preprint arXiv:1711.09952*.
- [24] **Simonyan, K., & Zisserman, A.** (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.
- [25] **Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z.** (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2818-2826).

- [26] **He, K., Zhang, X., Ren, S., & Sun, J.** (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).
- [27] **Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q.** (2017). Densely connected convolutional networks. In CVPR (Vol. 1, No. 2, p. 3).
- [28] **Susmitha, P. N., Reddy, S. S. S., & Pradesh, A.** (2014). A Review on Biometrics and Analysis of Fusion Methods for Ear Biometrics.
- [29] **Liu, K. H., Yan, S., & Kuo, C. C. J.** (2015). Age estimation via grouping and decision fusion. *IEEE Transactions on Information Forensics and Security*, 10(11), 2408-2423.
- [30] **Abaza, A., Ross, A., Hebert, C., Harrison, M. A. F., & Nixon, M. S.** (2013). A survey on ear biometrics. *ACM computing surveys (CSUR)*, 45(2), 22.
- [31] **Nadheen, M. F., & Poornima, S.** (2013). Fusion in multimodal biometric using iris and ear. In *Information & Communication Technologies (ICT), 2013 IEEE Conference on* (pp. 83-87). IEEE.
- [32] **Akin, C., Kacar, U., & Kirci, M.** (2018). A Multi-Biometrics for Twins Identification Based Speech and Ear. *arXiv preprint arXiv:1801.09056*.
- [33] **Pflug, A., & Busch, C.** (2012). Ear biometrics: a survey of detection, feature extraction and recognition methods. *IET biometrics*, 1(2), 114-129.
- [34] **Kacar, U., & Kirci, M.** (2019). ScoreNet: Deep Cascade Score Level Fusion for Unconstrained Ear Recognition. *IET Biometrics*, 8(2), 109-120.
- [35] **Kacar, U., & Kirci, M.** (2018). Ear Recognition With Score-Level Fusion Based On CMC In Long-Wave Infrared Spectrum. *arXiv preprint arXiv:1801.09054*.
- [36] **Abaza, A., & Ross, A.** (2010). Towards understanding the symmetry of human ears: A biometric perspective. In *BTAS* (pp. 1-7).
- [37] **Yan, P., & Bowyer, K. W.** (2007). Biometric recognition using 3D ear shape. *IEEE Transactions on pattern analysis and machine intelligence*, 29(8), 1297-1308.
- [38] **Hurley, D. J., Nixon, M. S., & Carter, J. N.** (2000). A new force field transform for ear and face recognition. In *Image Processing, 2000. Proceedings. 2000 International Conference on* (Vol. 1, pp. 25-28). IEEE.
- [39] **Victor, B., Bowyer, K., & Sarkar, S.** (2002). An evaluation of face and ear biometrics. In *Pattern Recognition, 2002. Proceedings. 16th International Conference on* (Vol. 1, pp. 429-432). IEEE.
- [40] **Xiaoxun, Z., & Yunde, J.** (2007). Symmetrical null space LDA for face and ear recognition. *Neurocomputing*, 70(4-6), 842-848.
- [41] **Wang, Y., Mu, Z. C., & Zeng, H.** (2008). Block-based and multi-resolution methods for ear recognition using wavelet transform and uniform local binary patterns. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on* (pp. 1-4). IEEE.

- [42] **Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., & Berg, A. C.** (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), 211-252.
- [43] **Eyiokur, F. I., Yaman, D., & Ekenel, H. K.** (2017). Domain adaptation for ear recognition using deep convolutional neural networks. *IET Biometrics*, 7(3), 199-206.
- [44] **Hansley, E. E., Segundo, M. P., & Sarkar, S.** (2018). Employing fusion of learned and handcrafted features for unconstrained ear recognition. *IET Biometrics*, 7(3), 215-223.
- [45] **Emeršič, Ž., Meden, B., Peer, P., & Štruc, V.** Evaluation and analysis of ear recognition models: performance, complexity and resource requirements. *Neural Computing and Applications*, 1-16.
- [46] **Dodge, S., Mounsef, J., & Karam, L.** (2018). Unconstrained ear recognition using deep neural networks. *IET Biometrics*, 7(3), 207-214.
- [47] **Modak, S. K. S., & Jha, V. K.** (2018). Multibiometric Fusion strategy and its Applications: A Review. *Information Fusion*.
- [48] **El-sayed, A.** (2015). Multi-biometric systems: a state of the art survey and research directions. *IJACSA) International Journal of Advanced Computer Science and Applications*, 6.
- [49] **Iwano, K., Hirose, T., Kamibayashi, E., & Furui, S.** (2003). Audio-visual person authentication using speech and ear images. In *Workshop on (Multimodal-User Authenticaion (MMUA)*, pages 85–90.
- [50] **Yan, P., & Bowyer, K. W.** (2005). Multi-biometrics 2D and 3D ear recognition. In *International Conference on Audio-and Video-Based Biometric Person Authentication* (pp. 503-512). Springer, Berlin, Heidelberg.
- [51] **Woodard, D. L., Faltemier, T. C., Yan, P., Flynn, P. J., & Bowyer, K. W.** (2006, June). A comparison of 3D biometric modalities. In *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW'06. Conference on* (pp. 57-57). IEEE.
- [52] **Xu, X., & Mu, Z.** (2007). Feature fusion method based on KCCA for ear and profile face based multimodal recognition. In *Automation and Logistics, 2007 IEEE International Conference on* (pp. 620-623). IEEE.
- [53] **Badrinath, G. S., & Gupta, P.** (2009). Feature level fused ear biometric system. In *2009 Seventh International Conference on Advances in Pattern Recognition* (pp. 197-200). IEEE.
- [54] **Islam, S. M., Bennamoun, M., Mian, A. S., & Davies, R.** (2009). Score level fusion of ear and face local 3D features for fast and expression-invariant human recognition. In *International conference image analysis and recognition* (pp. 387-396). Springer, Berlin, Heidelberg.
- [55] **Kisku, D. R., Sing, J. K., & Gupta, P.** (2009). Multibiometrics belief fusion. In *Machine Vision, 2009. ICMV'09. Second International Conference on* (pp. 37-40). IEEE.

- [56] **Mahoor, M. H., Cadavid, S., & Abdel-Mottaleb, M.** (2009). Multi-modal ear and face modeling and recognition. In *Image Processing (ICIP), 2009 16th IEEE International Conference on* (pp. 4137-4140). IEEE.
- [57] **Monwar, M. M., & Gavrilova, M. L.** (2009). Multimodal biometric system using rank-level fusion approach. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39(4), 867-878.
- [58] **Arbab-Zavar, B., & Nixon, M. S.** (2011). On guided model-based analysis for ear biometrics. *Computer Vision and Image Understanding*, 115(4), 487-502.
- [59] **Kumar, A., Hanmandlu, M., Sharma, V., & Gupta, H. M.** (2011). Rank based hybrid multimodal fusion using PSO. In *International Conference on Swarm, Evolutionary, and Memetic Computing* (pp. 217-224). Springer, Berlin, Heidelberg.
- [60] **Abaza, A., & Boursai, T.** (2013). On ear-based human identification in the mid-wave infrared spectrum. *Image and Vision Computing*, 31(9), 640-648.
- [61] **Huang, Z., Liu, Y., Li, C., Yang, M., & Chen, L.** (2013). A robust face and ear based multimodal biometric system using sparse representation. *Pattern Recognition*, 46(8), 2156-2168.
- [62] **Monwar, M. M., & Gavrilova, M.** (2013). Markov chain model for multimodal biometric rank fusion. *Signal, Image and Video Processing*, 7(1), 137-149.
- [63] **Ghoualmi, L., Chikhi, S., & Draa, A.** (2014). A SIFT-based feature level fusion of iris and ear biometrics. In *IAPR Workshop on Multimodal Pattern Recognition of Social Signals in Human-Computer Interaction* (pp. 102-112). Springer, Cham.
- [64] **Annapurani, K., Sadiq, M. A. K., & Malathy, C.** (2015). Fusion of shape of the ear and tragus—a unique feature extraction method for ear authentication system. *Expert Systems with Applications*, 42(1), 649-656.
- [65] **Haghighat, M., Abdel-Mottaleb, M., & Alhalabi, W.** (2016). Discriminant correlation analysis: Real-time feature level fusion for multimodal biometric recognition. *IEEE Transactions on Information Forensics and Security*, 11(9), 1984-1996.
- [66] **Tulyakov, S., Jaeger, S., Govindaraju, V., & Doermann, D.** (2008). Review of classifier combination methods. In *Machine learning in document analysis and recognition* (pp. 361-386). Springer, Berlin, Heidelberg.
- [67] **Mohandes, M., Deriche, M., & Aliyu, S. O.** (2018). Classifiers Combination Techniques: A Comprehensive Review. *IEEE Access*, 6, 19626-19639.
- [68] **Chitroub, S.** (2010). Classifier combination and score level fusion: concepts and practical aspects. *International Journal of Image and Data Fusion*, 1(2), 113-135.
- [69] **Giot, R., & Rosenberger, C.** (2012). Genetic programming for multibiometrics. *Expert Systems with Applications*, 39(2), 1837-1847.

- [70] **Ghiass, R. S., Arandjelović, O., Bendada, A., & Maldague, X.** (2014). Infrared face recognition: A comprehensive review of methodologies and databases. *Pattern Recognition*, 47(9), 2807-2824.
- [71] **Kolkur, S., Kalbande, D., Shimpi, P., Bapat, C., & Jatakia, J.** (2017). Human skin detection using RGB, HSV and YCbCr color models. *arXiv preprint arXiv:1708.02694*.
- [72] **Tkalcic, M., & Tasic, J. F.** (2003). Colour spaces: perceptual, historical and applicational background (Vol. 1, pp. 304-308). *IEEE*.
- [73] **Zuiderveld, K.** (1994). Contrast limited adaptive histogram equalization. *Graphics gems*, 474-485.
- [74] **Lore, K. G., Akintayo, A., & Sarkar, S.** (2017). LLNet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition*, 61, 650-662.
- [75] **Shen, L., Yue, Z., Feng, F., Chen, Q., Liu, S., & Ma, J.** (2017). Msr-net: Low-light image enhancement using deep convolutional network. *arXiv preprint arXiv:1711.02488*.
- [76] **Xie, J., Xu, L., & Chen, E.** (2012). Image denoising and inpainting with deep neural networks. In *Advances in neural information processing systems* (pp. 341-349).
- [77] **Nimisha, T. M., Singh, A. K., & Rajagopalan, A. N.** (2017). Blur-Invariant Deep Learning for Blind-Deblurring. In *ICCV* (pp. 4762-4770).
- [78] **Kim, J., Kwon Lee, J., & Mu Lee, K.** (2016). Accurate image super-resolution using very deep convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1646-1654).
- [79] **Wang, M., & Deng, W.** (2018). Deep Face Recognition: A Survey. *arXiv preprint arXiv:1804.06655*.
- [80] **Møller, M. F.** (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural networks*, 6(4), 525-533.
- [81] **Olshausen, B. A., & Field, D. J.** (1997). Sparse coding with an overcomplete basis set: A strategy employed by V1?. *Vision research*, 37(23), 3311-3325.
- [82] **Cevikalp, H., Neamtu, M., & Wilkes, M.** (2006). Discriminative common vector method with kernels. *IEEE Transactions on Neural Networks*, 17(6), 1550-1565.
- [83] **Cevikalp, H., Neamtu, M., Wilkes, M., & Barkana, A.** (2005). Discriminative common vectors for face recognition. *IEEE Transactions on pattern analysis and machine intelligence*, 27(1), 4-13.
- [84] **Zhang, D., He, J., Zhao, Y., Luo, Z., & Du, M.** (2014). Global plus local: a complete framework for feature extraction and recognition. *Pattern Recognition*, 47(3), 1433-1442.
- [85] **Sakarya, U.** (2016). Dimension reduction using global and local pattern information-based maximum margin criterion. *Signal, Image and Video Processing*, 10(5), 903-909.

- [86] **Baudat, G., & Anouar, F.** (2000). Generalized discriminant analysis using a kernel approach. *Neural computation*, 12(10), 2385-2404.
- [87] **Schölkopf, B., Smola, A. J., & Bach, F.** (2002). *Learning with kernels: support vector machines, regularization, optimization, and beyond*. MIT press.
- [88] **Liu, F., Fry, C. S., Mula, J., Jackson, J. R., Lee, J. D., Peterson, C. A., & Yang, L.** (2013). Automated fiber-type-specific cross-sectional area assessment and myonuclei counting in skeletal muscle. *Journal of Applied Physiology*, 115(11), 1714-1724.
- [89] **Kacar, U., Kirci, M., Gunes, E. O., & Inan, T.** (2015). A comparison of PCA, LDA and DCVA in ear biometrics classification using SVM. In *Signal Processing and Communications Applications Conference (SIU)*, 2015 23th (pp. 1260-1263). IEEE.
- [90] **Url-1** <http://slazebni.cs.illinois.edu/spring17/lec09_similarity.pdf>, erişim tarihi 22.02.2019
- [91] **Bellet, A., Habrard, A., & Sebban, M.** (2013). A survey on metric learning for feature vectors and structured data. *arXiv preprint arXiv:1306.6709*.
- [92] **Hu, J., Lu, J., & Tan, Y. P.** (2014). Discriminative deep metric learning for face verification in the wild. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1875-1882).
- [93] **Yi, D., Lei, Z., Liao, S., & Li, S. Z.** (2014, August). Deep metric learning for person re-identification. In *Pattern Recognition (ICPR), 2014 22nd International Conference on* (pp. 34-39). IEEE.
- [94] **Hoffer, E., & Ailon, N.** (2015). Deep metric learning using triplet network. In *International Workshop on Similarity-Based Pattern Recognition* (pp. 84-92). Springer, Cham.
- [95] **Jain, A., Nandakumar, K., & Ross, A.** (2005). Score normalization in multimodal biometric systems. *Pattern recognition*, 38(12), 2270-2285.



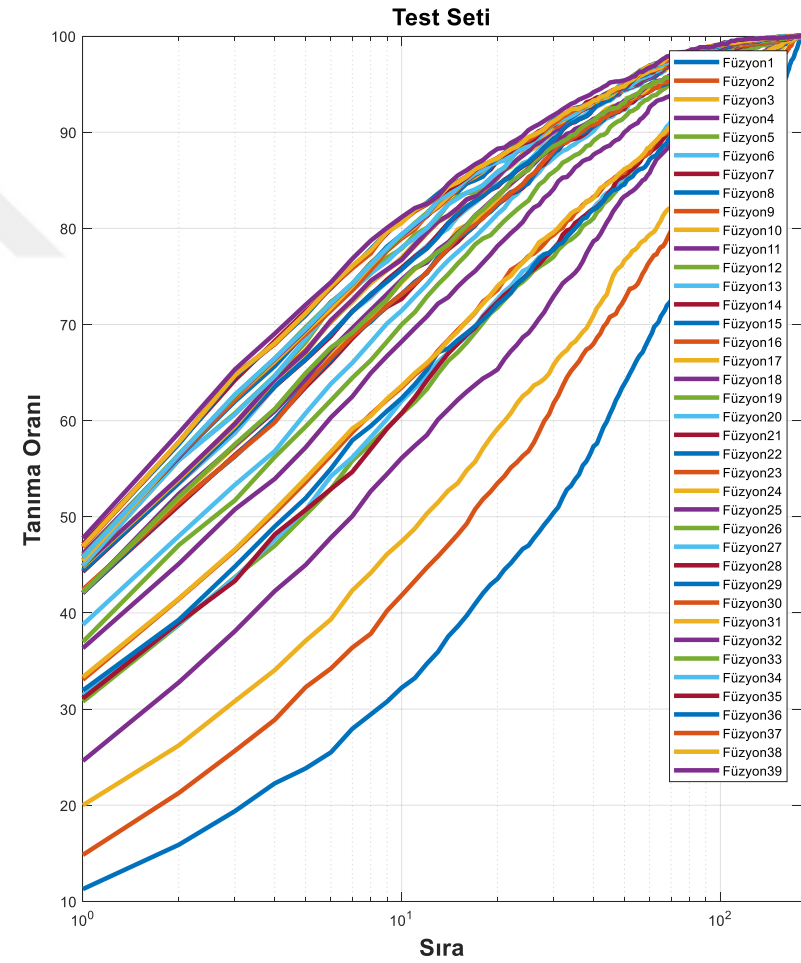
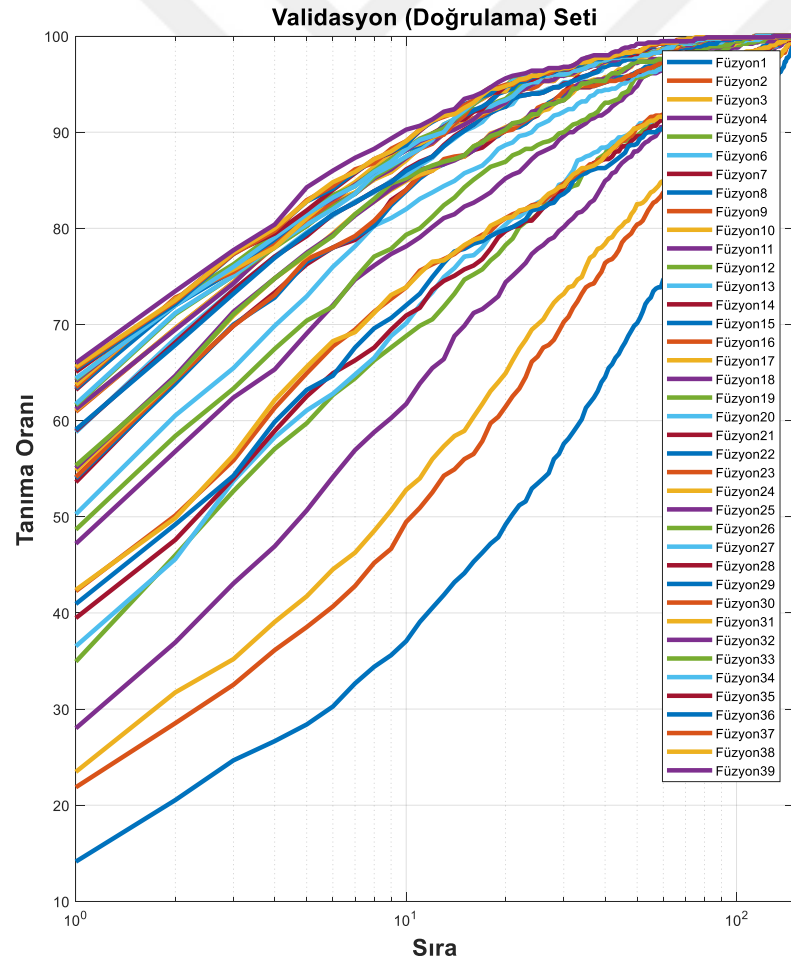
EKLER

EK A: Önerilen mimariden elde edilen detaylı sonuçlar





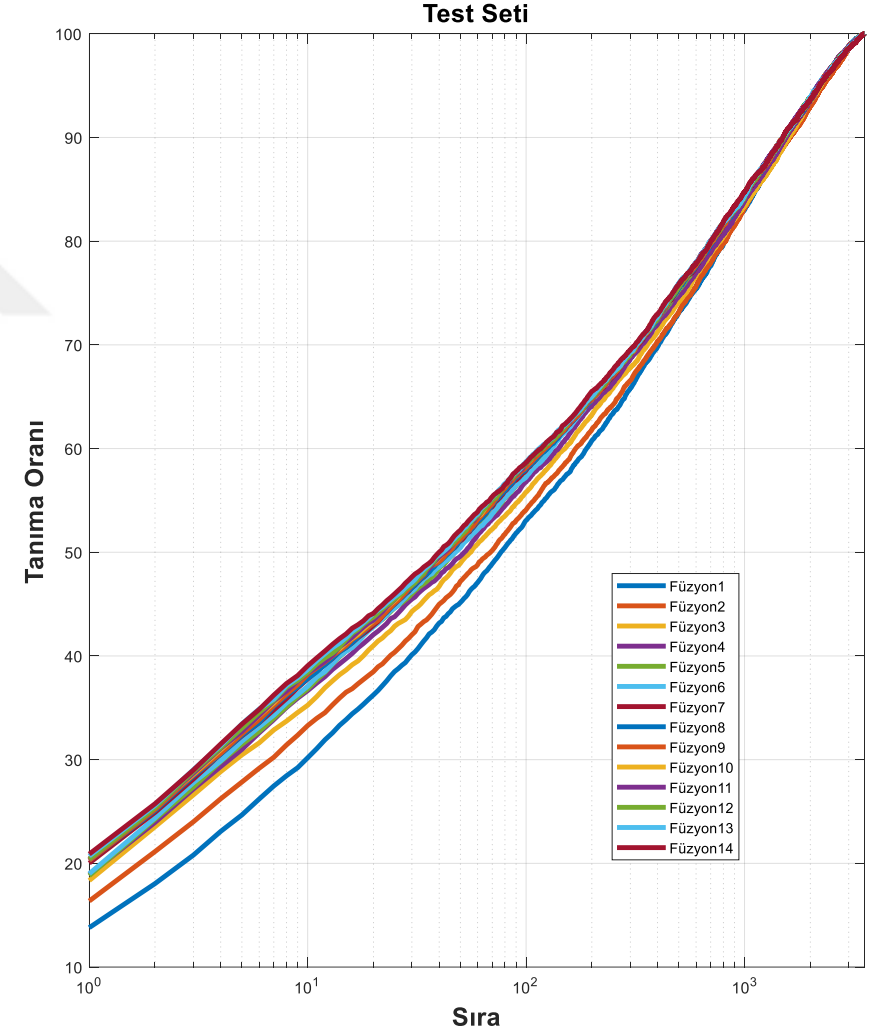
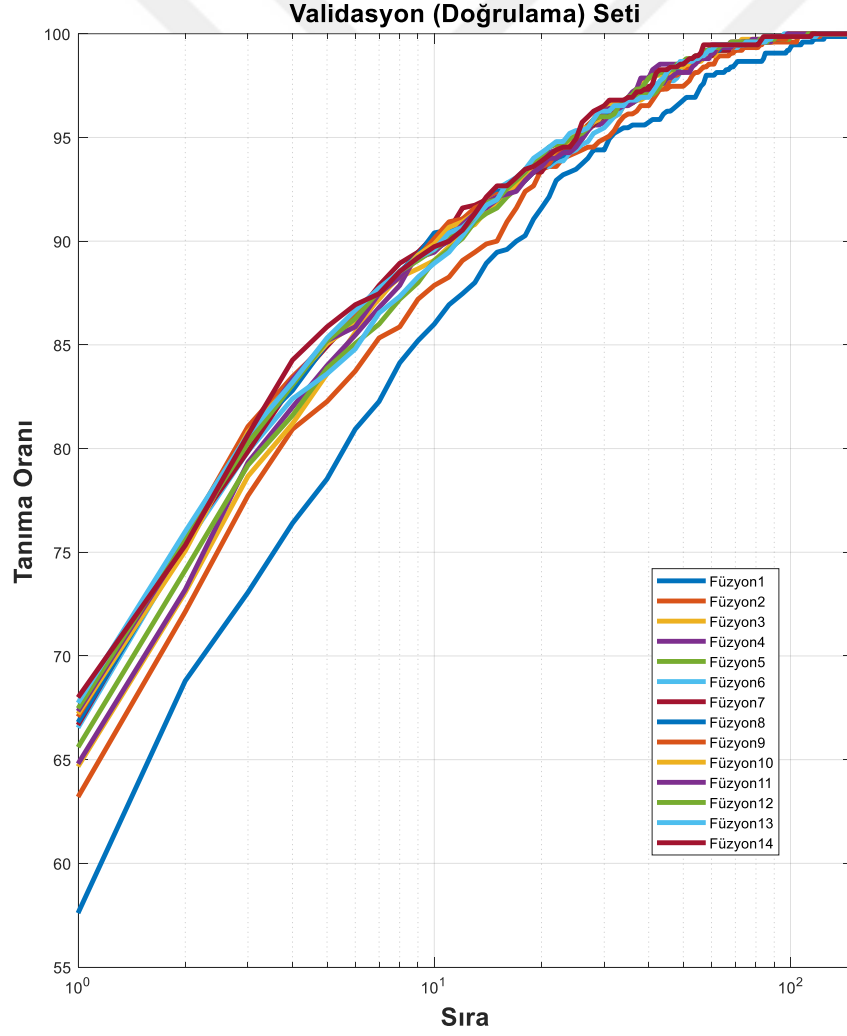
EK A: Önerilen mimariden elde edilen detaylı sonuçlar



Şekil A.1 : Validasyon ve test seti için “ScoreNet-1”in CMC eğrileri.

Çizelge A.1 : “ScoreNet-1 ”in içeriği ile detaylı validasyon ve test sonuçları.

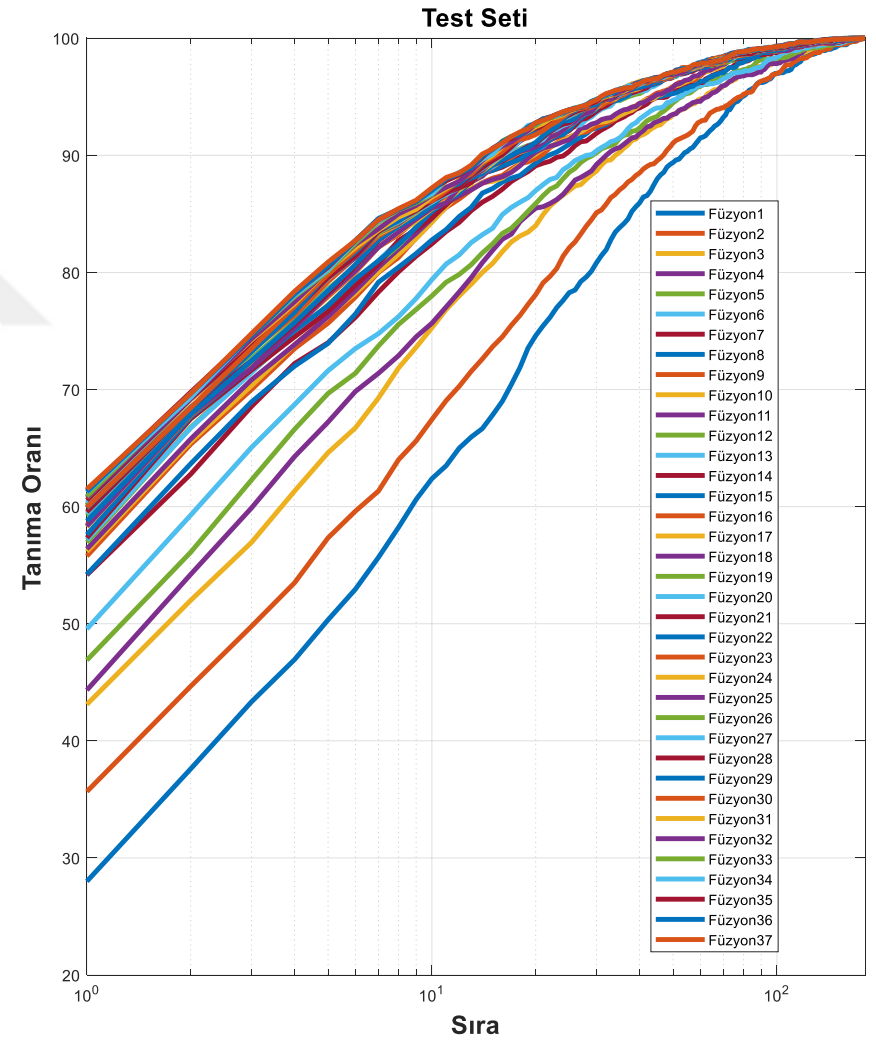
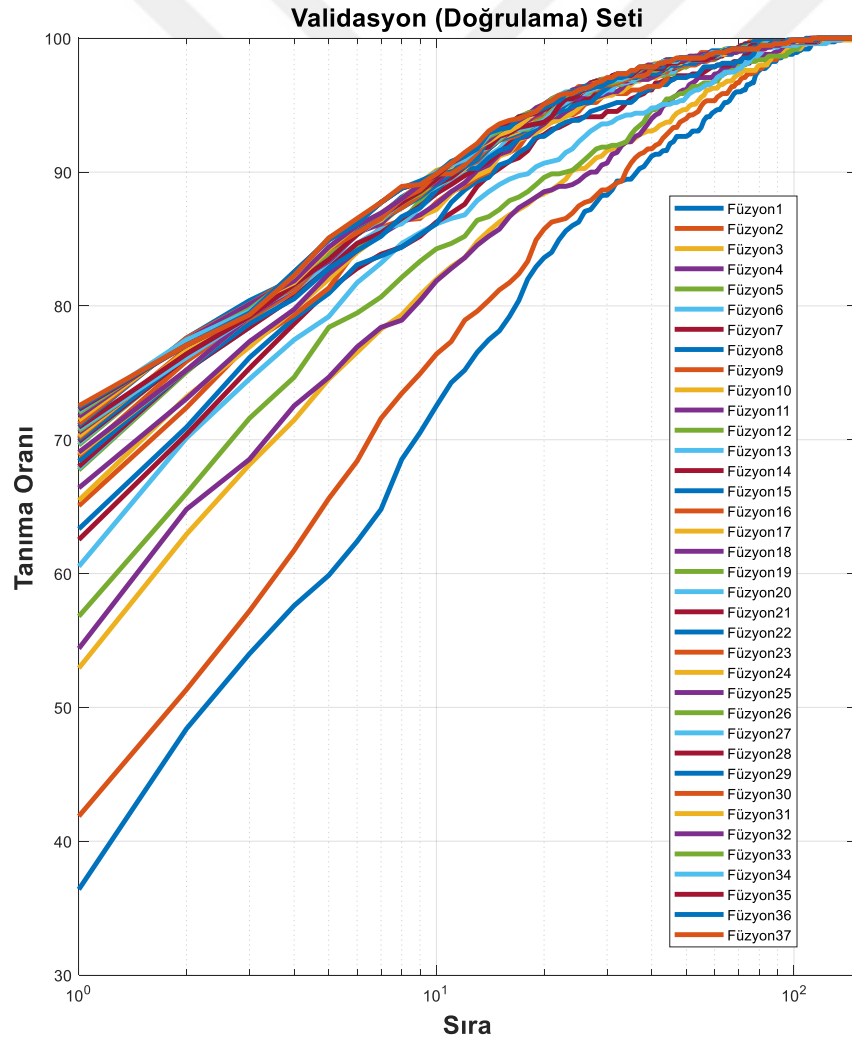
Füzyon sayısı	Renk haritası			Görüntü boyutu [satır-sütun]		Ön işleme ve görüntü iyileştirme	Öznitelik çıkarımı	Parametre set indeksi	Öznitelik azaltılması	Sınıflandırıcı	Metrik uzayı	Füzyon öğrenme oranı	Her kişiden örnek sayısı	Validasyon tek modalite sıra-1	Validasyon füzyon modalite sıra-1	Test tek modalite sıra-1	Test füzyon modalite sıra-1
1	V	yeşil	kırmızı	107	120	imadjust	BSIF	10	PCA	DCVA	manhattan	1	3	14.133	14.133	11.278	11.278
2	yeşil	V	Y	85	109	imadjust	GABOR	2	PCA	CGLDA	manhattan	0.55	4	13.333	21.867	10.389	14.833
3	mavi	Y	V	48	58	imgaussfilt	vgg16	25	PCA	CGLDA	öklid	0.25	5	9.600	23.467	11.278	20.000
4	Y	gri	V	115	117	imadjust	GABOR	2	PCA	CGLDA	kosinüs	0.35	3	22.800	28.000	18.556	24.611
5	kırmızı	mavi	Y	114	60	normal	HOG	6	PCA	DCVA	öklid	0.6	2	28.800	34.933	22.444	30.778
6	yeşil	kırmızı	Y	112	57	adapthisteq	resnet50	60	PCA	DCVA	öklid	0.15	3	10.533	36.533	6.389	31.444
7	yeşil	mavi	V	76	77	imgaussfilt	vgg19	34	PCA	GDA	kosinüs	0.3	4	6.133	39.467	4.167	31.111
8	mavi	yeşil	Y	58	99	normal	vgg19	30	PCA	GDA	kosinüs	0.25	5	6.800	40.933	4.778	31.944
9	Y	mavi	yeşil	79	70	adapthisteq	inceptionv3	107	PCA	CGLDA	öklid	0.15	2	16.533	42.267	9.722	33.056
10	Y	yeşil	mavi	89	120	imadjust	LBP	7	PCA	DCVA	öklid	0.05	3	19.867	42.400	14.222	33.333
11	gri	kırmızı	V	88	42	adapthisteq	resnet50	132	PCA	CGLDA	öklid	0.45	4	32.000	47.200	18.333	36.333
12	gri	mavi	Y	83	46	adapthisteq	vgg16	32	PCA	CGLDA	kosinüs	0.15	4	24.800	48.667	18.611	37.000
13	gri	Y	kırmızı	106	71	imadjust	googlenet	62	PCA	DCVA	kosinüs	0.3	4	28.000	50.267	19.611	38.778
14	gri	yeşil	V	116	59	imgaussfilt	HOG	1	PCA	LDA	kosinüs	0.4	5	36.267	53.600	24.778	42.111
15	Y	kırmızı	mavi	82	119	adapthisteq	vgg16	36	PCA	LDA	kosinüs	0.02	3	7.200	54.000	6.444	42.056
16	Y	mavi	kırmızı	77	45	imgaussfilt	resnet50	94	PCA	DCVA	öklid	0.02	2	27.733	54.133	21.278	42.444
17	kırmızı	Y	yeşil	89	101	adapthisteq	resnet101	316	PCA	LDA	manhattan	0.2	5	16.933	54.800	8.944	42.167
18	mavi	Y	V	59	51	adapthisteq	HOG	4	PCA	LDA	kosinüs	0.03	4	23.200	55.067	20.444	42.056
19	Y	V	mavi	86	54	imadjust	BSIF	9	PCA	CGLDA	manhattan	0.04	3	15.467	55.333	10.500	42.111
20	yeşil	Y	gri	57	99	normal	vgg19	33	PCA	DCVA	kosinüs	0.19	4	51.333	58.800	30.500	44.389
21	kırmızı	Y	yeşil	51	120	imgaussfilt	alexnet	5	PCA	LDA	öklid	0.03	5	19.733	58.933	11.333	44.278
22	yeşil	kırmızı	V	106	65	imgaussfilt	googlenet	91	PCA	DCVA	öklid	0.01	5	10.267	59.067	6.056	44.389
23	mavi	yeşil	kırmızı	99	88	imgaussfilt	vgg19	31	PCA	LDA	kosinüs	0.17	5	57.600	60.933	34.722	44.889
24	mavi	gri	Y	112	64	imgaussfilt	LPQ	10	PCA	DCVA	manhattan	0.01	2	18.800	61.067	14.611	44.778
25	yeşil	Y	kırmızı	43	50	imadjust	vgg16	29	PCA	LDA	öklid	0.01	4	33.733	61.200	22.556	44.889
26	gri	V	yeşil	88	101	normal	inceptionv3	257	PCA	DCVA	kosinüs	0.18	4	28.667	61.600	16.611	44.889
27	gri	V	yeşil	101	102	adapthisteq	HOG	4	PCA	DCVA	kosinüs	0.02	2	32.267	61.733	22.722	44.944
28	kırmızı	mavi	Y	69	91	imadjust	resnet101	232	PCA	LDA	kosinüs	0.17	5	50.400	63.200	29.667	45.833
29	yeşil	Y	kırmızı	71	75	adapthisteq	vgg16	25	PCA	GDA	öklid	0.04	4	6.933	63.333	3.389	45.944
30	V	kırmızı	Y	40	112	imadjust	vgg16	33	PCA	CGLDA	manhattan	0.09	4	11.067	63.467	8.000	45.500
31	yeşil	Y	V	68	89	imgaussfilt	resnet101	84	PCA	DCVA	kosinüs	0.02	5	44.667	63.733	26.444	45.389
32	V	kırmızı	Y	42	59	normal	resnet101	303	PCA	DCVA	manhattan	0.06	5	38.800	64.133	22.333	46.167
33	gri	kırmızı	Y	112	79	adapthisteq	alexnet	10	PCA	LDA	kosinüs	0.02	3	32.933	64.267	25.056	45.889
34	V	kırmızı	gri	46	119	imgaussfilt	vgg16	30	PCA	LDA	manhattan	0.02	3	20.533	64.400	15.667	45.778
35	mavi	gri	V	118	90	imadjust	inceptionv3	131	PCA	LDA	kosinüs	0.17	5	52.933	65.067	34.333	47.167
36	yeşil	kırmızı	mavi	91	66	normal	googlenet	110	PCA	DCVA	öklid	0.05	5	23.467	65.333	14.389	46.944
37	V	kırmızı	mavi	82	71	normal	LBP	6	PCA	LDA	öklid	0.03	4	13.333	65.467	6.889	46.833
38	Y	V	yeşil	92	61	normal	alexnet	12	PCA	CGLDA	kosinüs	0.02	4	38.000	65.600	28.444	46.944
39	gri	mavi	yeşil	112	41	normal	inceptionv3	125	PCA	DCVA	kosinüs	0.19	4	50.933	66.000	32.722	47.778



Şekil A.2 : Validasyon ve test seti için “ScoreNet-2”nin CMC eğrileri.

Çizelge A.2 : “ScoreNet-2”nin içeriği ile detaylı validasyon ve test sonuçları.

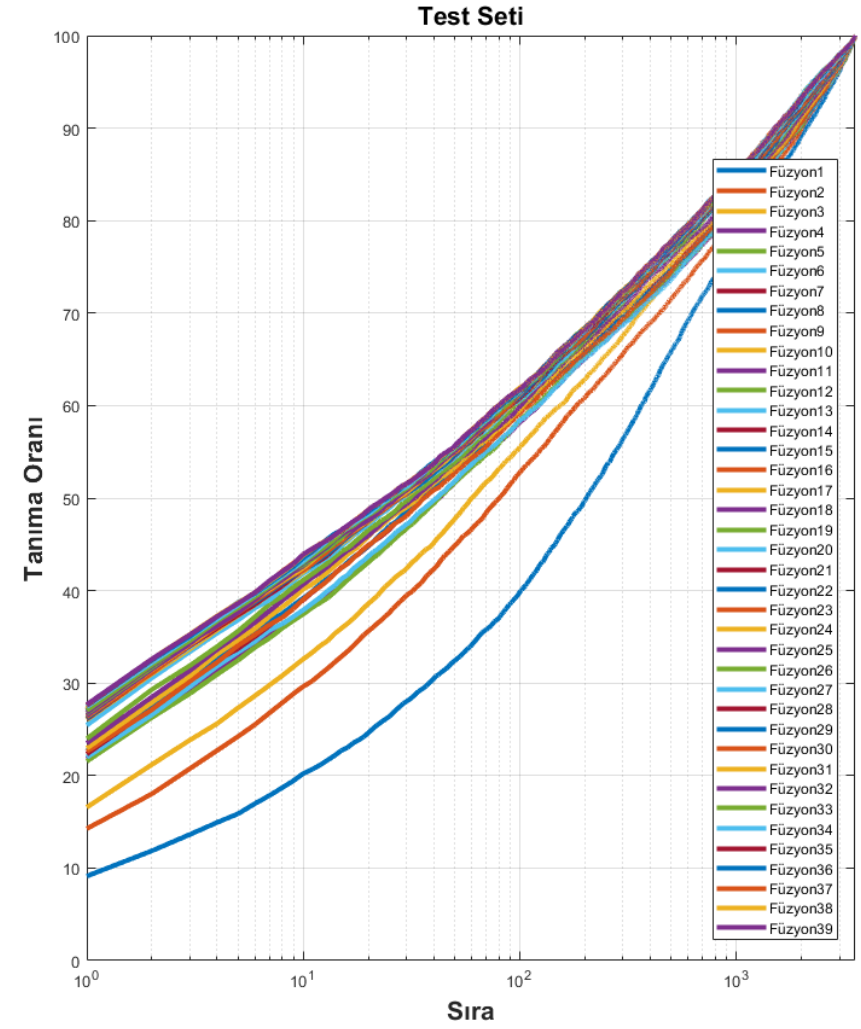
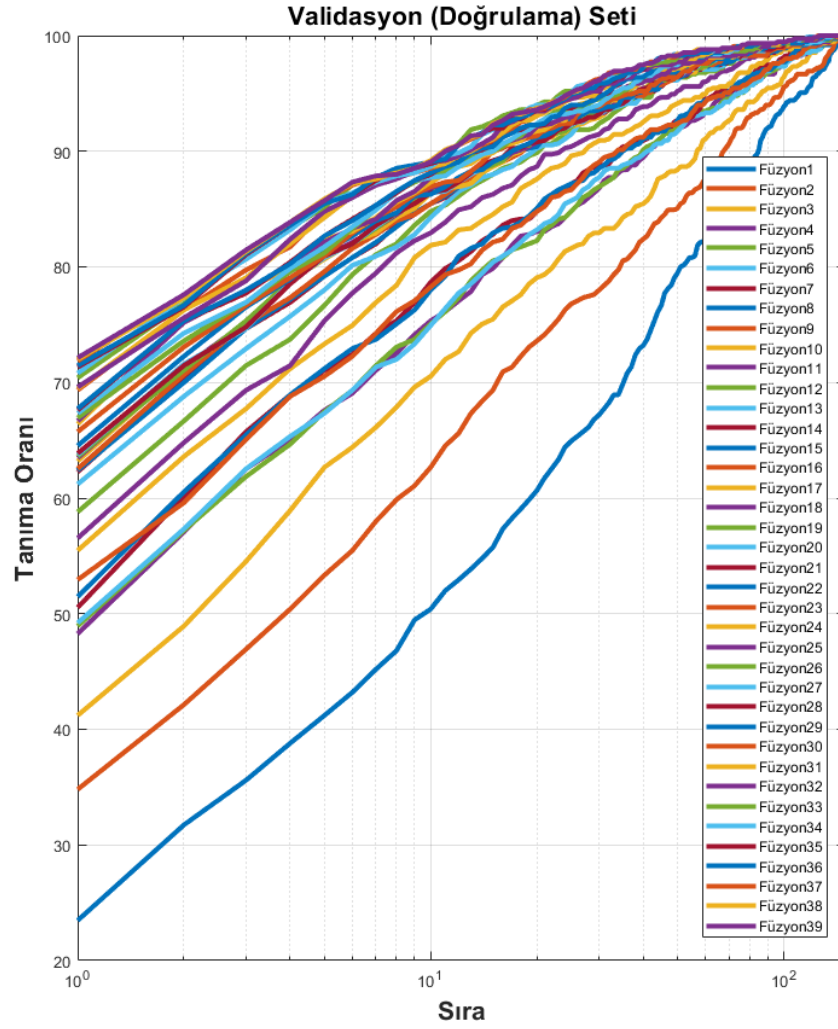
Füzyon sayısı	Renk haritası			Görüntü boyutu [satır-sütun]		Ön işleme ve görüntü iyileştirme	Öznitelik çıkarımı	Parametre set indeksi	Öznitelik azaltılması	Sınıflandırıcı	Metrik uzayı	Füzyon öğrenme oranı	Her kişiden örnek sayısı	Validasyon tek modalite sıra-1	Validasyon füzyon modalite sıra-1	Test tek modalite sıra-1	Test füzyon modalite sıra-1
1	mavi	yeşil	kırmızı	99	88	imgaussfilt	vgg19	31	PCA	LDA	kosinüs	1	5	57.600	57.600	13.800	13.800
2	mavi	gri	V	118	90	imadjust	inceptionv3	131	PCA	LDA	kosinüs	0.35	5	52.933	63.200	10.965	16.367
3	gri	mavi	yeşil	112	41	normal	inceptionv3	125	PCA	DCVA	kosinüs	0.5	4	50.933	64.667	13.142	18.342
4	kırmızı	mavi	Y	69	91	imadjust	resnet101	232	PCA	LDA	kosinüs	0.25	5	50.400	64.800	10.777	18.920
5	V	kırmızı	Y	42	59	normal	resnet101	303	PCA	DCVA	manhattan	0.3	5	38.800	65.600	7.955	18.812
6	Y	V	yeşil	92	61	normal	alexnet	12	PCA	CGLDA	kosinüs	0.2	4	38.000	66.533	10.642	19.000
7	gri	yeşil	V	116	59	imgaussfilt	HOG	1	PCA	LDA	kosinüs	0.2	5	36.267	66.667	8.049	20.021
8	mavi	Y	V	59	51	adapthisteq	HOG	4	PCA	LDA	kosinüs	0.05	4	23.200	66.800	6.289	20.290
9	kırmızı	Y	yeşil	89	101	adapthisteq	resnet101	316	PCA	LDA	manhattan	0.05	5	16.933	67.067	3.064	20.277
10	V	yeşil	kırmızı	107	120	imadjust	BSIF	10	PCA	DCVA	manhattan	0.05	3	14.133	67.200	4.206	20.411
11	mavi	yeşil	kırmızı	58	95	imadjust	GABOR	1	PCA	LDA	manhattan	0.05	4	8.400	67.333	2.876	20.411
12	Y	V	kırmızı	118	86	normal	resnet101	152	PCA	CGLDA	manhattan	0.1	2	39.600	67.467	8.009	20.411
13	kırmızı	mavi	yeşil	66	45	imadjust	inceptionv3	206	PCA	DCVA	öklid	0.05	5	34.933	67.733	8.264	20.707
14	kırmızı	Y	gri	103	53	adapthisteq	LPQ	10	PCA	DCVA	manhattan	0.1	4	19.467	68.000	3.843	20.895



Şekil A.3 : Validasyon ve test seti için “ScoreNet-3”ün CMC eğrileri.

Çizelge A.3 : “ScoreNet-3”ün içeriği ile detaylı validasyon ve test sonuçları.

Füzyon sayısı	Renk haritası			Görüntü boyutu [satır-sütun]		Ön işleme ve görüntü iyileştirme	Öznitelik çıkarımı	Parametre set indeksi	Öznitelik azaltılması	Sınıflandırıcı	Metrik uzayı	Füzyon öğrenme oranı	Her kişiden örnek sayısı	Validasyon tek modalite sıra-1	Validasyon füzyon modalite sıra-1	Test tek modalite sıra-1	Test füzyon modalite sıra-1
1	gri	kırmızı	yeşil	89	115	imgaussfilt	resnet50	136	PCA	DCVA	kosinüs	1	3	36.400	36.400	28.000	28.000
2	V	kırmızı	Y	71	64	normal	LPQ	1	PCA	LDA	manhattan	0.3	2	20.800	41.867	22.722	35.667
3	kırmızı	mavi	Y	89	118	normal	resnet50	116	PCA	DCVA	manhattan	0.45	5	36.133	52.933	27.111	43.111
4	gri	Y	V	109	77	normal	GABOR	10	PCA	CGLDA	kosinüs	0.3	4	19.467	54.400	17.833	44.333
5	Y	kırmızı	V	67	71	adapthisteq	inceptionv3	155	PCA	DCVA	manhattan	0.35	3	32.667	56.800	24.833	46.889
6	V	Y	kırmızı	94	85	normal	BSIF	2	PCA	CGLDA	manhattan	0.3	5	26.400	60.533	18.278	49.556
7	gri	yeşil	kırmızı	44	106	normal	googlenet	25	PCA	DCVA	kosinüs	0.4	3	50.133	62.533	42.222	54.167
8	mavi	gri	V	100	115	adapthisteq	resnet101	242	PCA	LDA	manhattan	0.1	3	27.733	63.333	22.444	54.222
9	kırmızı	Y	yeşil	65	75	imadjust	vgg19	29	PCA	DCVA	kosinüs	0.3	3	52.533	65.067	44.500	55.778
10	V	kırmızı	yeşil	72	69	adapthisteq	alexnet	7	PCA	LDA	kosinüs	0.1	4	42.533	65.467	35.778	56.167
11	gri	V	mavi	107	57	imgaussfilt	googlenet	17	PCA	DCVA	kosinüs	0.05	5	54.667	66.400	41.000	56.389
12	gri	kırmızı	V	65	119	adapthisteq	vgg19	30	PCA	DCVA	kosinüs	0.35	4	56.800	67.733	42.722	56.944
13	V	gri	yeşil	103	50	adapthisteq	inceptionv3	292	PCA	CGLDA	kosinüs	0.05	2	6.933	67.867	9.056	57.167
14	gri	V	Y	106	91	adapthisteq	LBP	10	PCA	CGLDA	öklid	0.1	2	23.867	68.000	22.722	57.333
15	V	Y	yeşil	66	115	imadjust	LBP	6	PCA	CGLDA	manhattan	0.06	5	19.333	68.400	15.500	57.611
16	kırmızı	V	gri	95	90	imgaussfilt	resnet50	143	PCA	CGLDA	manhattan	0.17	5	44.933	68.800	28.111	58.444
17	V	Y	yeşil	104	49	adapthisteq	alexnet	16	PCA	CGLDA	kosinüs	0.03	4	41.200	68.933	31.778	58.278
18	kırmızı	Y	mavi	85	87	imadjust	BSIF	5	PCA	CGLDA	öklid	0.01	4	24.000	69.067	19.611	58.333
19	mavi	V	gri	100	109	normal	LBP	5	PCA	DCVA	kosinüs	0.08	4	39.467	69.600	31.389	59.222
20	V	gri	kırmızı	114	110	imadjust	vgg19	31	PCA	GDA	kosinüs	0.01	5	16.267	69.733	6.611	59.389
21	kırmızı	gri	yeşil	113	77	adapthisteq	LPQ	4	PCA	DCVA	kosinüs	0.02	2	35.600	69.867	28.500	59.556
22	Y	kırmızı	mavi	70	74	imadjust	resnet101	240	PCA	GDA	manhattan	0.04	3	6.800	70.000	4.556	58.778
23	mavi	kırmızı	V	107	90	imadjust	LBP	7	PCA	LDA	kosinüs	0.09	4	45.467	70.133	35.444	59.889
24	kırmızı	V	gri	89	52	adapthisteq	LBP	6	PCA	CGLDA	manhattan	0.04	4	15.067	70.267	14.111	60.056
25	yeşil	mavi	Y	54	68	imgaussfilt	inceptionv3	174	PCA	LDA	öklid	0.02	4	29.467	70.533	22.056	60.111
26	gri	V	kırmızı	90	65	imgaussfilt	resnet101	160	PCA	CGLDA	kosinüs	0.02	3	61.067	70.667	47.056	60.167
27	V	kırmızı	mavi	119	87	adapthisteq	BSIF	3	PCA	DCVA	öklid	0.01	4	24.800	70.800	19.389	60.333
28	V	Y	gri	109	94	imgaussfilt	LPQ	10	PCA	DCVA	kosinüs	0.01	4	51.867	70.933	40.333	60.556
29	yeşil	mavi	kırmızı	65	88	imgaussfilt	vgg19	34	PCA	DCVA	kosinüs	0.16	4	50.133	71.067	37.167	59.889
30	kırmızı	mavi	gri	104	119	imgaussfilt	resnet50	84	PCA	LDA	manhattan	0.04	3	37.600	71.200	32.944	59.944
31	V	mavi	Y	77	84	imadjust	HOG	4	PCA	DCVA	öklid	0.06	5	46.267	71.467	35.944	60.778
32	gri	mavi	kırmızı	108	96	adapthisteq	LPQ	10	PCA	CGLDA	kosinüs	0.02	5	52.400	71.733	33.944	60.778
33	Y	kırmızı	V	97	64	imgaussfilt	HOG	10	PCA	GDA	kosinüs	0.06	4	5.467	72.000	5.833	60.889
34	kırmızı	V	gri	49	87	imadjust	HOG	8	PCA	CGLDA	kosinüs	0.01	4	21.333	72.133	21.556	61.222
35	mavi	gri	yeşil	92	83	normal	HOG	4	PCA	DCVA	manhattan	0.02	5	42.800	72.267	34.333	61.389
36	Y	mavi	kırmızı	93	110	normal	googlenet	122	PCA	DCVA	kosinüs	0.01	4	29.867	72.400	18.944	61.278
37	kırmızı	V	mavi	111	94	imgaussfilt	HOG	3	PCA	LDA	kosinüs	0.02	4	51.600	72.533	44.722	61.500



Şekil A.4 : Validasyon ve test seti için “ScoreNet-4”ün CMC eğrileri.

Çizelge A.4 : “ScoreNet-4”ün içeriği ile detaylı validasyon ve test sonuçları.

Füzyon sayısı	Renk haritası			Görüntü boyutu [satır-sütun]		Ön işleme ve görüntü iyileştirme	Öznitelik çıkarımı	Parametre set indeksi	Öznitelik azaltılması	Sınıflandırıcı	Metrik uzayı	Füzyon öğrenme oranı	Her kişiden örnek sayısı	Validasyon tek modalite sıra-1	Validasyon füzyon modalite sıra-1	Test tek modalite sıra-1	Test füzyon modalite sıra-1
1	V	yeşil	kırmızı	107	120	imadjust	BSIF	10	PCA	DCVA	manhattan	1	3	23.467	23.467	9.124	9.124
2	yeşil	V	Y	85	109	imadjust	GABOR	2	PCA	CGLDA	manhattan	0.45	4	17.200	34.800	6.342	14.257
3	Y	gri	V	115	117	imadjust	GABOR	2	PCA	CGLDA	kosinüs	0.3	3	29.600	41.200	10.763	16.568
4	kırmızı	mavi	Y	114	60	normal	HOG	6	PCA	DCVA	öklid	0.55	2	39.867	48.267	18.087	21.782
5	V	yeşil	mavi	64	61	imadjust	GABOR	4	PCA	CGLDA	öklid	0.1	3	10.000	48.933	2.929	21.540
6	kırmızı	V	mavi	110	97	imgaussfilt	LBP	10	PCA	LDA	öklid	0.25	3	24.667	49.200	11.583	21.983
7	Y	mavi	yeşil	79	70	adapthisteq	inceptionv3	107	PCA	CGLDA	öklid	0.2	2	19.867	50.533	6.517	22.319
8	Y	yeşil	mavi	89	120	imadjust	LBP	7	PCA	DCVA	öklid	0.15	3	28.400	51.467	10.320	22.628
9	V	yeşil	kırmızı	118	107	imgaussfilt	LBP	10	PCA	GDA	kosinüs	0.15	3	6.133	52.933	1.720	22.601
10	gri	kırmızı	V	88	42	adapthisteq	resnet50	132	PCA	CGLDA	öklid	0.16	4	41.733	55.467	9.325	22.951
11	gri	mavi	Y	83	46	adapthisteq	vgg16	32	PCA	CGLDA	kosinüs	0.11	4	33.600	56.533	7.135	23.502
12	gri	Y	kırmızı	106	71	imadjust	googlenet	62	PCA	DCVA	kosinüs	0.14	4	37.200	58.800	9.258	23.985
13	gri	yeşil	V	116	59	imgaussfilt	HOG	1	PCA	LDA	kosinüs	0.16	5	42.133	61.200	15.439	25.437
14	yeşil	mavi	gri	111	77	imadjust	HOG	6	PCA	DCVA	kosinüs	0.14	5	45.600	62.267	15.641	26.203
15	mavi	gri	kırmızı	117	111	normal	LPQ	1	PCA	CGLDA	öklid	0.01	5	37.733	62.400	10.602	26.176
16	Y	mavi	kırmızı	77	45	imgaussfilt	resnet50	94	PCA	DCVA	öklid	0.08	2	37.200	62.533	13.854	26.297
17	kırmızı	Y	yeşil	89	101	adapthisteq	resnet101	316	PCA	LDA	manhattan	0.18	5	13.067	63.200	3.709	26.041
18	mavi	Y	V	59	51	adapthisteq	HOG	4	PCA	LDA	kosinüs	0.04	4	32.667	63.467	11.744	26.162
19	kırmızı	Y	yeşil	51	120	imgaussfilt	alexnet	5	PCA	LDA	öklid	0.02	5	21.467	63.600	9.057	26.149
20	Y	yeşil	V	42	87	imadjust	GABOR	7	PCA	LDA	manhattan	0.02	2	13.200	63.733	7.364	26.135
21	mavi	Y	V	61	115	imgaussfilt	HOG	8	PCA	CGLDA	kosinüs	0.01	3	33.867	63.867	13.840	26.122
22	gri	Y	V	97	100	imgaussfilt	resnet101	133	PCA	DCVA	manhattan	0.11	5	43.600	64.533	11.032	26.203
23	yeşil	mavi	V	83	83	imgaussfilt	BSIF	2	PCA	CGLDA	kosinüs	0.11	5	43.067	65.733	11.556	26.337
24	mavi	V	gri	71	87	imgaussfilt	resnet50	106	PCA	DCVA	öklid	0.15	3	46.000	66.400	14.055	26.458
25	gri	mavi	yeşil	102	56	imgaussfilt	BSIF	1	PCA	LDA	kosinüs	0.05	5	29.467	66.667	10.589	26.498
26	V	yeşil	mavi	94	52	imgaussfilt	alexnet	5	PCA	DCVA	kosinüs	0.2	5	42.533	66.933	14.956	26.713
27	kırmızı	yeşil	V	97	57	imgaussfilt	HOG	4	PCA	DCVA	manhattan	0.11	2	43.067	67.333	17.872	27.036
28	gri	V	yeşil	88	101	normal	inceptionv3	257	PCA	DCVA	kosinüs	0.11	4	27.733	67.600	7.310	27.063
29	kırmızı	mavi	gri	41	57	adapthisteq	inceptionv3	37	PCA	CGLDA	manhattan	0.01	2	22.933	67.733	9.850	26.955
30	kırmızı	mavi	Y	69	91	imadjust	resnet101	232	PCA	LDA	kosinüs	0.16	5	59.867	69.333	13.410	27.197
31	V	yeşil	Y	89	95	imadjust	HOG	1	PCA	DCVA	öklid	0.05	5	42.267	69.467	12.644	27.304
32	gri	kırmızı	V	97	109	adapthisteq	vgg19	39	PCA	CGLDA	manhattan	0.09	5	5.467	69.600	1.854	26.821
33	yeşil	Y	V	68	89	imgaussfilt	resnet101	84	PCA	DCVA	kosinüs	0.15	5	54.667	70.400	15.318	27.049
34	kırmızı	Y	yeşil	72	53	normal	LPQ	3	PCA	CGLDA	kosinüs	0.09	4	40.667	70.800	13.894	27.439
35	Y	V	kırmızı	72	78	imadjust	LBP	4	PCA	DCVA	öklid	0.07	4	18.533	71.333	7.525	27.573
36	yeşil	mavi	gri	113	108	imgaussfilt	resnet50	107	PCA	DCVA	kosinüs	0.01	4	55.200	71.467	14.647	27.627
37	gri	mavi	kırmızı	108	96	adapthisteq	LPQ	10	PCA	CGLDA	kosinüs	0.09	5	52.400	71.867	12.739	27.654
38	kırmızı	V	mavi	111	94	imgaussfilt	HOG	3	PCA	LDA	kosinüs	0.01	4	51.600	72.000	17.925	27.667
39	yeşil	mavi	V	83	83	imgaussfilt	BSIF	2	PCA	CGLDA	kosinüs	0.01	5	43.067	72.133	11.556	27.681

ÖZGEÇMİŞ

Ad-Soyad : Ümit KAÇAR
Doğum Tarihi ve Yeri : Şişli, 1985
E-posta : umitkacar@itu.edu.tr

ÖĞRENİM DURUMU:

- **Lisans** : 2007, Deniz Harp Okulu, Elektrik-Elektronik Mühendisliği
- **Yüksek Lisans** : 2013, İstanbul Teknik Üniversitesi, Elektronik Mühendisliği

MESLEKİ DENEYİM VE ÖDÜLLER:

- 2003 yılında Nişantaşı Anadolu Lisesi'nden birincilik ile mezun olmuştur.
- 2007 yılında Deniz Harp Okulu'ndan birincilik ile mezun olmuştur.
- 2013 yılında İstanbul Teknik Üniversitesi Elektronik Mühendisliği Yüksek Lisans eğitimini 4.00 ortalama ile tamamlamıştır.
- 2011-2013 yılları arasında Tübitak Yüksek Lisans Başarı Bursu almıştır.
- 2013-2017 yılları arasında Tübitak Doktora Başarı Bursu almıştır.

DOKTORA TEZİNDEN TÜRETİLEN YAYINLAR:

- **Kacar, U., & Kirci, M.** (2019). ScoreNet: Deep Cascade Score Level Fusion for Unconstrained Ear Recognition. IET Biometrics, 8(2), 109-120.
- **Kacar, U., Kirci, M., Güneş, E. O., & İnan, T.** (2015). A comparison of PCA, LDA and DCVA in ear biometrics classification using SVM. In Signal Processing and Communications Applications Conference (SIU), 2015 23th (pp. 1260-1263). IEEE.
- **Kacar, U., & Kirci, M.** (2017). Ear Recognition With Score-Level Fusion Based On CMC In Long-Wave Infrared Spectrum. 40th Anniversary of International Conference on Telecommunications and Signal Processing (TSP).
- Akin, C., **Kacar, U., & Kirci, M.** (2016). A Multi-Biometrics for Twins Identification Based Speech and Ear, 3rd International Conference on Electrical and Electronics Engineering (ICEEE).

DİĞER YAYINLAR:

- **Kacar, U.,** Kirci, M., Kus, M., & Gunes, E. O. (2013). An embedded biometric system. In Proceedings of the 16th International Conference on Information Fusion (pp. 736-742). IEEE.
- Kus, M., **Kacar, U.,** Kirci, M., & Gunes, E. O. (2013). ARM based ear recognition embedded system. In EUROCON, 2013 IEEE (pp. 2021-2028). IEEE.
- **Kacar, U.,** Kumlu, D., & Kirci, M. (2015). A Novel approach for automatic ship type classification. In Signal Processing and Communications Applications Conference (SIU), 2015 23th (pp. 2153-2156). IEEE.
- **Kaçar, U.** (2013). ARM Tabanlı Gömülü Sistemlerde Kulak Tanıma Sisteminin Gerçeklenmesi, (Yüksek Lisans Tezi). İstanbul Teknik Üniversitesi, Fen Bilimleri Enstitüsü, İstanbul.

