



**DERİN METRİK ÖĞRENME İLE HİSTOPATOLOJİK GÖRÜNTÜLERİN
SINIFLANDIRILMASI**

Mahmut KAYA

DOKTORA TEZİ

BİLGİSAYAR MÜHENDİSLİĞİ ANA BİLİM DALI

**GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

ARALIK 2019

Mahmut KAYA tarafından hazırlanan “DERİN METRİK ÖĞRENME İLE HİSTOPATOLOJİK GÖRÜNTÜLERİN SINIFLANDIRILMASI” adlı tez çalışması aşağıdaki jüri tarafından OY BİRLİĞİ ile Gazi Üniversitesi Bilgisayar Mühendisliği Ana Bilim Dalında DOKTORA TEZİ olarak kabul edilmiştir.

Danışman: Doç. Dr. Hasan Şakir BİLGE

Elektrik Elektronik Mühendisliği Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum.

.....

Başkan: Prof. Dr. Halil Tanyer EYYUBOĞLU

Telekomünikasyon Ana Bilim Dalı, Çankırı Karatekin Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum.

.....

Üye: Prof. Dr. Suat ÖZDEMİR

Bilgisayar Mühendisliği Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum.

.....

Üye: Prof. Dr. Hasan OĞUL

Bilgisayar Mühendisliği Ana Bilim Dalı, Başkent Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum.

.....

Üye: Prof. Dr. Fırat HARDALAÇ

Elektrik Elektronik Mühendisliği Ana Bilim Dalı, Gazi Üniversitesi

Bu tezin, kapsam ve kalite olarak Doktora Tezi olduğunu onaylıyorum.

.....

Tez Savunma Tarihi: 19/12/2019

Jüri tarafından kabul edilen bu tezin Doktora Tezi olması için gerekli şartları yerine getirdiğini onaylıyorum.

.....

Prof. Dr. Sena YAŞYERLİ

Fen Bilimleri Enstitüsü Müdürü

ETİK BEYAN

Gazi Üniversitesi Fen Bilimleri Enstitüsü Tez Yazım Kurallarına uygun olarak hazırladığım bu tez çalışmada;

- Tez içinde sunduğum verileri, bilgileri ve dokümanları akademik ve etik kurallar çerçevesinde elde ettiğimi,
- Tüm bilgi, belge, değerlendirme ve sonuçları bilimsel etik ve ahlak kurallarına uygun olarak sunduğumu,
- Tez çalışmada yararlandığım eserlerin tümüne uygun atıfta bulunarak kaynak gösterdiğimi,
- Kullanılan verilerde herhangi bir değişiklik yapmadığımı,
- Bu tezde sunduğum çalışmanın özgün olduğunu,

bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

Mahmut KAYA

19/12/2019

DERİN METRİK ÖĞRENME İLE HİSTOPATOLOJİK GÖRÜNTÜLERİN
SINIFLANDIRILMASI
(Doktora Tezi)

Mahmut KAYA

GAZİ ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

Aralık 2019

ÖZET

Metrik öğrenme, örnekler arasındaki benzerliği ölçmek için bir benzerlik metriğinden yararlanır. Genellikle doğrusal bir izdüşüm kullanan metrik öğrenme yöntemleri, doğrusal olmayan özellikleri gösteren gerçek dünyadaki problemleri çözmede sınırlıdır. Doğrusal olmayan veriler için metrik öğrenmede çekirdek yaklaşımları kullanılır. Son yıllarda popülerliği artan derin metrik öğrenme konusunda dikkat çeken çalışmalar bulunmaktadır. Derin metrik öğrenme (DML) sahip olduğu aktivasyon fonksiyonları ile doğrusal olmayan veriler için de en uygun çözümü elde etmekte oldukça başarılıdır. Bu tez çalışmasında derin metrik öğrenmenin literatürdeki konumu, önemi ve katkıları hakkında detaylıca bilgi verilmiştir. DML benzer sınıf örnekleri arasındaki uzaklığı azaltmaya çalışırken, farklı sınıf örnekleri arasındaki uzaklığı ise arttırmaya çalışmaktadır. Derin metrik öğrenmede bunu gerçekleştirmek için, örnekleme stratejisi, ağ mimarisi ve kayıp fonksiyonu ağı başarısında temel faktörlerdir. Genel itibarıyla literatürü incelediğimizde çalışmaların birçoğunun Siamese ve Triplet ağlarından esinlendiği görülmüştür. Bu tez çalışmasında metrik öğrenme bakış açısıyla üç farklı uygulama geliştirilmiştir. Öncelikle geleneksel metrik öğrenme yaklaşımı kullanılarak Parkinson çoklu ses veri kümesi üzerinde deneyler yapılmıştır. Bu deneylerde sınıflandırma başarımını arttırmak için en uygun W ağırlık matrisi genetik algoritma kullanılarak elde edilmiştir. Daha sonra histopatolojik veri kümesinde derin metrik öğrenme modelleri kullanılarak çalışmalar yapılmıştır. Son olarak ise, histopatolojik veri kümesi üzerinde bölütleme yaklaşımı ile hücre çekirdekler tespit edilmiştir. Tespit edilen hücre çekirdek görüntülerinin derin metrik öğrenme yaklaşımıyla sınıflandırma işlemi gerçekleştirilmiştir. Elde edilen sonuçlar ile derin metrik öğrenmenin genellikle örnek sayısı sınırlı olan histopatolojik veriler için başarıyla uygulanabileceği görülmüştür.

Bilim Kodu : 92431
Anahtar kelimeler : Metrik öğrenme, derin metrik öğrenme, sınıflandırma, benzerlik
Sayfa Sayısı : 112
Danışman : Doç. Dr. Hasan Şakir BİLGE

CLASSIFICATION OF HISTOPATHOLOGICAL IMAGES WITH DEEP METRIC LEARNING

(Ph. D. Thesis)

Mahmut KAYA

GAZİ UNIVERSITY
GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES

December 2019

ABSTRACT

Metric learning uses a similarity metric to measure similarity between samples. Metric learning methods, which generally use a linear projection, are limited in solving real-world problems that exhibit non-linear properties. Kernel approaches are used for nonlinear data in metric learning. There are remarkable studies on deep metric learning with increasing popularity in recent years. Deep metric learning (DML) is very successful in obtaining optimal solution for nonlinear data with its activation functions. In this thesis, detailed information is given about the position, importance and contribution of deep metric learning in the literature. DML tries to minimize the distance between similar class instances and maximizes the distance between different class instances. In order to achieve this in deep metric learning, sampling strategy, network architecture and loss function are key factors in network success. In general, when we examined the literature, it was seen that most of the studies were inspired by Siamese and Triplet networks. In this thesis, three different applications have been developed from a metric learning perspective. First, experiments were performed on a Parkinson's multiple sound dataset using the traditional metric learning approach. In these experiments, the most suitable W weight matrix was obtained by using genetic algorithm to increase the classification performance. Afterwards, studies were conducted using deep metric learning models in histopathological data set. Finally, nuclei were determined by segmentation approach on histopathological data set. The detected nuclei images were classified using a deep metric learning approach. The results showed that deep metric learning can be applied successfully for histopathological data, which usually has limited sample size.

Science Code : 92431
Key Words : Metric learning, deep metric learning, classification, similarity
Page Number : 112
Supervisor : Assoc. Prof. Dr. Hasan Şakir BİLGE

TEŞEKKÜR

Bu tez çalışmamda büyük emekleri olan kıymetli danışmanım Doç. Dr. Hasan Şakir BİLGE'ye, tez izleme komitemde bulunup değerli katkılar sunan Prof. Dr. Suat ÖZDEMİR ve Prof. Dr. Hasan OĞUL hocalarıma, tıbbi bilgilerin edinilmesi, histopatoloji görüntülerinin tez çalışması kapsamında kullanımına izin verdiğinden dolayı Prof. Dr. Mehmet GÜL'e; hayatımın her anında benden desteğini esirgemeyen eşim Kübra KAYA'ya ve mutluluk kaynağım olan oğlum Yusuf Selim KAYA'ya, bu günlere gelmemde büyük emeği olan anneme, babama ve kardeşlerime çok teşekkür ederim.



İÇİNDEKİLER

	Sayfa
ÖZET	iv
ABSTRACT.....	v
TEŞEKKÜR.....	vi
İÇİNDEKİLER	vii
ÇİZELGELERİN LİSTESİ.....	ix
ŞEKİLLERİN LİSTESİ.....	x
SİMGELER VE KISALTMALAR.....	xii
1. GİRİŞ	1
2. METRİK ÖĞRENME	7
2.1. Uzaklık Metriğine Dayalı Öğrenme Yaklaşımı	9
2.2. Uzaklık Metriğine Dayalı Yöntemler	11
2.3. Uzaklık Metriğine Dayalı Öğrenme Algoritmaları	12
2.3.1. Gözetimli uzaklık metriği öğrenme	13
2.3.2. Gözetimsiz uzaklık metriği öğrenme	14
2.3.3. Yarı gözetimli uzaklık metriği öğrenme	16
2.4. Uzaklık Metriğinde Ağırlık Hesaplama	17
3. DERİN ÖĞRENME	21
3.1. Derin Öğrenme	21
3.2. Evrişimli Sinir Ağları	23
3.2.1. Evrişim katmanı	23
3.2.2. Örnekleme katmanı	24
3.2.3. Tam bağlı katman	25
3.2.4. Aktivasyon fonksiyonları	25

	Sayfa
3.3. Transfer Öğrenme	26
4. DERİN METRİK ÖĞRENME	31
4.1. Derin Metrik Öğrenme Uygulama Alanları	35
4.2. Derin Metrik Öğrenmede Örnek Seçimi	45
4.3. Derin Metrik Öğrenmede Kayıp Fonksiyonları	48
5. METRİK ÖĞRENME YAKLAŞIMIYLA SINIFLANDIRMA	57
5.1. Materyal Ve Metot	58
5.1.1. Veri kümesi	58
5.1.2. Önerilen yöntem	58
5.2. Sentetik Veri Üzerinde Metrik Öğrenme Yaklaşımı	61
5.3. Parkinson Çoklu Ses Kaydetme Veri Kümesinde Uygulama	62
6. HİSTOPATOLOJİK GÖRÜNTÜLERİN DERİN METRİK ÖĞRENME İLE SINIFLANDIRILMASI	65
6.1. H & E Patolojik Görüntü	65
6.2. H & E Patolojik Görüntü Veri Kümesi	66
6.3. Derin Metrik Öğrenme ile Histopatolojik Görüntüleri Sınıflandırma	67
6.3.1. Çeşitli derin metrik öğrenme modelleri ile sınıflandırma	71
6.3.2. Rekabetçi ağ modellerini kullanarak derin metrik öğrenme modelleri ile sınıflandırma	76
6.3.3. Deneysel sonuçlar	78
7. HİSTOPATOLOJİK GÖRÜNTÜLERDE ÇEKİRDEK TESPİTİ VE SINIFLANDIRMA	89
8. SONUÇ VE DEĞERLENDİRMELER	93
KAYNAKLAR	97
ÖZGEÇMİŞ	111

ÇİZELGELERİN LİSTESİ

Çizelge	Sayfa
Çizelge 2.1. Gözetimli uzaklık metriği öğrenme algoritmaları	13
Çizelge 2.2. Gözetimsiz uzaklık metriği öğrenme algoritmaları	15
Çizelge 2.3. Yarı gözetimli uzaklık metriği öğrenme algoritmaları	16
Çizelge 2.4. W ağırlık hesaplama	18
Çizelge 4.1. Derin metrik öğrenme problemlerinin karşılaştırmalı sonuçları	40
Çizelge 4.2. Kayıp fonksiyonları	55
Çizelge 5.1. Deneysel sonuçlar	64
Çizelge 5.2. Karşılaştırmalı sonuçlar	64
Çizelge 6.1. Parametreler	72
Çizelge 6.2. Siamese eğitim ağı örnekleri	75
Çizelge 6.3. Siamese test ağı örnekleri	75
Çizelge 6.4. Triplet eğitim ağı örnekleri	76
Çizelge 6.5. Triplet test ağı örnekleri	76
Çizelge 6.6. Siamese sonuçları	86
Çizelge 6.7. Triplet sonuçları	87
Çizelge 7.1. Çekirdek sınıflandırma sonuçları	91

ŞEKİLLERİN LİSTESİ

Şekil	Sayfa
Şekil 2.1. Metrik öğrenme	8
Şekil 2.2. Uzaklık metrikleri (a) öklid uzaklığı (b) mahalanobis uzaklığı	9
Şekil 3.1. Derin öğrenme seviyeleri	22
Şekil 3.2. Evrişimli sinir ağı modeli	23
Şekil 3.3. Aktivasyon fonksiyonları	26
Şekil 3.4. Transfer öğrenme ile makine öğrenmesi farkı	27
Şekil 3.5. AlexNet	28
Şekil 4.1. Derin metrik öğrenme	32
Şekil 4.2. Siamese ağı benzer ve benzemeyen görüntüler arası uzaklık ilişkisi	34
Şekil 4.3. Literatürde derin metrik öğrenme ile ilgili yayın sayıları	36
Şekil 4.4. Negatif madencilik	47
Şekil 4.5. Siamese ve triplet ağı	49
Şekil 4.6. Kayıp fonksiyonları	50
Şekil 5.1. Önerilen yöntem	60
Şekil 5.2. Sentetik veri kümesi örnekleri	61
Şekil 5.3. Sentetik veri kümeleri sınıflandırma sonuçları	62
Şekil 5.4. Genetik kodlama	63
Şekil 6.1. Patolojik görüntü elde etme süreci	66
Şekil 6.2. Bir patolojik görüntü	67
Şekil 6.3. Patolojik veri kümesi örnekleri	68
Şekil 6.4. Evrişimsel sinir ağı modeli	69
Şekil 6.5. Siamese + Evrişimsel sinir ağı	69
Şekil 6.6. Triplet + Evrişimsel sinir ağı modeli	70

Şekil	Sayfa
Şekil 6.7. Budanmış ve rasgele dağılım ağırlıklar	71
Şekil 6.8. Denenmiş ağ modelleri	73
Şekil 6.9. VGG-19 Ağ modeli	77
Şekil 6.10. Siamese VGG-19 ağ modeli	77
Şekil 6.11. Triplet VGG -19 ağ modeli	78
Şekil 6.12. Destek vektör makinesi marjın ayrımı	80
Şekil 6.13. Siamese eğitim sonuçları	81
Şekil 6.14. Siamese test sonuçları	82
Şekil 6.15. Triplet eğitim sonuçları	83
Şekil 6.16. Triplet test sonuçları	84
Şekil 6.17. Siamese eğitim kayıp performansı	85
Şekil 6.18. Triplet eğitim kayıp performansı	87
Şekil 7.1. Çekirdek bölütleme adımları	89
Şekil 7.2. Histopatolojik görüntü bölütleme	91
Şekil 7.3. Çekirdek bölütleme adımları	92

SİMGELER VE KISALTMALAR

Bu çalışmada kullanılmış simgeler ve kısaltmalar, açıklamaları ile birlikte aşağıda sunulmuştur.

Kısaltmalar	Açıklamalar
3B	3 Boyutlu
ANMM	Average neighborhood margin maximization
CNN	Convolutional neural network
DCG	İndirimli kümülatif kazanç
DML	Deep metric learning
DWS	Distance weighted sampling
E	Emeasure
EER	Eşit hata oranı
FP	False-Positive
FT	First tier
GDA	Gradient descent algorithm
GPU	Graphics processing unit
GSS	Greedy search strategy
H & E	Hematoxylin & Eosin
ISOMAP	Isometric feature mapping
ITML	Information-theoretic metric learning
KANMM	Kernel average neighborhood margin maximization
KLDA	Kernel linear discriminant analysis
KLRML	Kernel laplacian regularized metric learning
kNN	k en yakın komşu
KPCA	Kernel principal component analysis
KRCA	Kernel relevant component analysis
LDA	Linear discriminant analysis
LE	Laplacian embedding
LLE	Locally linear embedding
LMNN	Large margin nearest neighbor

Kısaltmalar**Açıklamalar**

LPP	Locality preserving projections
LRML	Laplacian regularized metric learning
LSTM	Long short-term memory
mAP	Mean average precision
MDCF	Minimum karar maliyet fonksiyonu
MSE	Ortalama kare hatası
NCA	Neighborhood component analysis
NMI	Normalized mutual information
PCA	Principal component analysis
R	Recall
RCA	Relevant component analysis
ResNet	Residual network
SSDR	Semi-supervised dimensionality reduction
ST	Second tier
SVM	Support vector machines
VGG	Visual Geometry Group

1. GİRİŞ

Bu bölümde tezde ele alınan mevcut problem hakkında bilgi verilmiş, tezin amacı, kapsamı, literatüre katkısı, önemi ve yapısı hakkında her bir bölüm hakkında değerlendirmelerde bulunularak tezin katkılarının özet şekilde verilmesi hedeflenmiştir.

Günümüzde bilgisayar teknolojisinin gelişmesiyle insan hayatını kolaylaştırma ve teknoloji çağına ayak uydurma amacıyla problemlerin her biri olabilirliklerine bağlı olarak bilgisayar ortamına aktarılmaktadır. İnsanın öğrenme kabiliyeti gibi bilgisayarlara da nesneleri öğretme kabiliyeti kazandırma makine öğrenmesi kavramını ortaya çıkarmıştır [1]. Bugün, hayatımızı daha kolay hale getirmek için birçok alanda makine öğrenmesinden faydalanabiliriz. Makine öğrenmesi yüz tanıma, sağlık verilerinde, saldırı tespit sistemlerinde, konuşma tanıma, ses tanıma, metin madenciliğinde, örüntü tanıma vb. amaçlarla farklı disiplinlerde çözümler sunmaktadır. Makine öğrenmede bu tür problemler amacına yönelik olarak sınıflandırma, kümeleme gibi yaklaşımlarla çözümler sunmaktadır. Gerçek dünya problemlerinde veri boyutunun artması ve problemlerinde farklılaşmasıyla birlikte makine öğrenmesi teknikleriyle en iyi çözümü elde etmek başlı başına bir problemdir [2].

Makine öğrenme algoritmaları ile mevcut verilere dayanan çok başarılı sınıflandırma modelleri üretebilir. Önerilen modelin hem verilerin eğitim sürecinde hem de eğitime katılmayan veriler üzerinde başarılı sonuçlar vermesi hedeflenmektedir. Bununla birlikte, önerilen modelin herhangi bir problemde istenen sonucu üretmesini beklemek her zaman mümkün olmayabilir, çünkü her verinin çözülmesi gereken kendi temel problemleri vardır. Örneğin, bir yüz tanıma problemiyle uğraşırken, her biri verinin sınıflandırılmasında çeşitli sorunlara neden olan poz farklılıkları, aydınlatma farklılıkları, ölçeklendirme, arka plan ve ifade gibi faktörlerle uğraşabilirsiniz. Bu nedenle, bu faktörlerin her birinin üstesinden gelmek için; yani verilerin doğru bir şekilde sınıflandırılması için verilerin ayırt edici özellikleri maksimum seviyeye çıkarılmalıdır.

k en yakın komşu, destek vektör makineleri ve naive bayes sınıflandırıcıları makine öğrenme uygulamalarında birer sınıflandırıcı olarak kullanılmaktadır. Bu algoritmalar belirli bir sınıflandırma performansına sahip olsa da verilerin daha iyi bir şekilde temsil

edilmesi mümkündür. Bu algoritmalar orijinal bir veri kümesini yeni bir dönüşüm uzayına taşımazlar. Doğrudan veri üzerinden sınıflandırma yaparlar. Ancak, bir veri kümesindeki her öz niteliğin etkisi sınıflandırma açısından eşit değildir. Bu nedenle, sınıflandırma öncesi öz niteliklerin ağırlıklandırılması daha anlamlı sonuçların üretilmesini sağlar. Ham veriler bulunduğu uzaydan yeni bir temsili uzaya da taşınabilirler. Bunu gerçekleştirmek için Temel Bileşen Analizi ve Doğrusal Ayırıştırma Analizi gibi veri dönüştürme algoritmalarından yararlanılmaktadır. Bu algoritmaları kullanılarak verinin dönüştürülmüş uzaydaki temsili üzerinden sınıflandırma işlemini gerçekleştirmekle daha sağlıklı sonuçların elde edilmesi beklenir.

Metrik öğrenme, nesneler arasında benzerlik veya benzersizlik oluşturmayı amaçlayan doğrudan bir metriğe dayanan bir yaklaşımdır. Metrik öğrenme, benzer nesneler arasındaki mesafeyi azaltmayı amaçlarken, farklı nesneler arasındaki mesafeyi de arttırmayı amaçlar. Basit ama veri sınıflandırmada etkili olup nesneler arası uzaklık bilgisini hesaplayan k-en yakın komşular gibi yaklaşımlar ve verilerin yeni bir gösterime dönüştürüldüğü benzerlik tabanlı yaklaşımlar bulunmaktadır.

Metrik öğrenme yaklaşımları mesafe bilgisiyle dönüşüm alanına taşınırken, yöntem temel olarak bir W projeksiyon matrisine dayanır. Mevcut çalışmalar genel olarak Mahalanobis uzaklık metriği ile doğrudan ilişkilidir [3-5]. Mahalanobis mesafesi Öklid mesafesine dönüştürüldüğünde, metrik öğrenme yaklaşımı, kovaryans matrisinin ayrışmasına ve bu işlemleri yaparken simetrik pozitif tanımlı matrislerin kullanımına dayanarak sunulur.

Veri hacmi gün geçtikçe artmakta ve bu da daha doğru sınıflandırma için önemli avantajlar sağlamaktadır. Ancak, bu durum aynı zamanda çok sayıda verinin işlenmesi ve yüksek hesaplama gücünü ortaya çıkarmaktadır. Günümüzde çok fazla hesaplama gereksinimi nedeniyle işlemlerin parçalar halinde ve birlikte yapılması önemli hesaplama avantajları sağlayacaktır. Bu nedenle, paralel hesaplamanın gücü sayesinde makine öğreniminde daha hızlı ve başarılı çözümler elde etmek mümkündür. Çok katmanlı yapı ile derin öğrenme, son yıllarda GPU teknolojisinin gelişmesiyle birlikte bilgisayar bilimlerinde zamanımızın en popüler konularından biri haline gelmiştir [6].

Verilerin ham veriler üzerinden yeni bir temsiliyi sağlayan derin öğrenme, verilerin yeni temsili uzaya dönüştürülmesi sırasında, hedefinde daha yüksek soyutlama seviyeleri elde

etmek olduđu ayırt edici öz niteliklerin otomatik olarak çıkarılmasını sağlamaktadır [7]. Derin öğrenme kendi başına kompakt bir yapı sunar ve mimarisinde sınıflandırmayı da bünyesinde bulundurmaktadır. Ayrıca, aktivasyon fonksiyonlarının gücü ile gerçek dünyadaki problemleri tespit etmemiz için bize daha gerçekçi yaklaşımlar sağlayabilecek doğrusal olmayan bir yapıya sahiptir.

Son yıllarda derin öğrenmenin metrik öğrenmeyle kesişimiyle derin metrik öğrenme kavramı ortaya çıkmıştır [8, 9]. Derin metrik öğrenme örneklerin benzerliği noktasına odaklanmaktadır. Öğrenilecek olan derin metrik benzer ya da aynı sınıftan olan nesneleri birbirine yaklaştırmaya odaklanırken, benzemeyen ya da farklı sınıftan olan nesneleri ise birbirinden uzaklaştırmayı amaç edinmektedir. DML günümüzde çok farklı alanlarda problemlere uygulanmış ve oldukça başarılı sonuçlar elde edilmiştir.

Literatürde derin metrik öğrenme yaklaşımlarını incelediğimizde genel olarak Siamese [10] ve Triplet [11] ağlarından esinlenildiği görülmektedir. Bu ağ modellerinde veriler ağa ikili ya da üçlü örnekler halinde sunulurlar. Ağ sunulan örnekler aynı mimarideki ağırlıkları paylaşımlı bir şekilde kullanırlar. Kullanılan bu paylaşımlı ağlarda verinin daha yüksek bir temsilini elde etmek için bir metrik öğrenmeye odaklanılmaktadır. Öğrenilecek olan metrikte ağa sunulacak olan örneklerin benzerliğine odaklanmaktadır.

Derin metrik öğrenme (DML) yüz doğrulama, tanıma, kişinin yeniden tanımlanması ve 3D şekil elde etme problemlerine başarıyla uygulanmaktadır. Bu görevler çok sayıda kategoriden ve her bir kategoride mevcut olan sınırlı eğitim örneklerinden oluşur. Herhangi bir kategori için yetersiz sayıda örnek, başarılı bir eğitim sürecini zorlaştırabilir. Ancak DML bu tür problemleri çözmede oldukça başarılıdır. DML yüksek ayırım gücüyle bu tür problemlerde başarıyla uygulansa da bu bir sınırlama değildir.

Bu tez çalışmasında DML yaklaşımlarının histopatolojik veriler üzerinde de başarıyla uygulanabileceğine odaklanılmıştır. Bu amaçla bu tez çalışmasında özgün bir patolojik veri kümesinde DML yaklaşımları ile sınıflandırma konusuna odaklanılmıştır. Tez çalışmasında kullanılan veri kümesi dengeli bir veri kümesi değildir. 2 sınıflı bir probleme sahip olunan veri kümesinde verinin 2/3'ü bir sınıftayken, 1/3'ü diğer sınıfın örneklerini barındırmaktadır. Veri kümesinin bir diğer problemi verilerin örnek sayısının az olmasıdır. Ancak DML yapısı gereğince bu tür problemleri aşmak mümkündür. Öncelikle veri

dengelemesi için aynı anda ağa sunulacak örnek ikili ya da üçlülere ile herhangi bir veri artırma tekniğine ihtiyaç duyulmadan veri artışı yapmak mümkündür. Uygulanacak örnek seçimiyle hem eğitim veri kümesi dengelenebilir hem de ağa sunulacak örnek sayısı arttırılabilir. Bu da DML'in histopatolojik veri problemlerine uygunluğunu da ortaya koymaktadır. Böylece tez çalışmasında DML problemleri farklı bir alana uygulanmış ve bu alana özgü çözüm geliştirilerek literatüre katkı sağlanmıştır.

Bu tez çalışması kapsamında yapılan incelemeler ve gerçekleştirilen deneysel çalışmalar sonucunda literatüre sunulan katkılar maddeler halinde aşağıda verilmiştir;

- Literatürde metrik öğrenme, Mahalanobis uzaklık metriğinin Öklid uzaklık cinsinden ifade edilmesine dayanmaktadır. Literatürdeki metrik öğrenme kavramının W dönüşüm matrisiyle doğrudan ilgili olduğu ve tez çalışmasında bu durum net bir şekilde ifade edilmiştir. Bu yaklaşım yapılan deneysel uygulamalarla desteklenmiş ve en uygun W matrisiyle çok başarılı sonuçlar elde edilmiştir.
- Literatürde bir boşluk olarak görülen kapsamlı bir derin metrik öğrenme incemelesi yapılmış, bunun sonucunda derin metrik öğrenme modellerinin avantajları, dezavantajları, problemi ele alma şekilleri, yaklaşımsal farklılıkları detaylıca verilmiştir. Aynı zamanda farklı alanlarda uygulama sonuçları karşılaştırmalı olarak detaylı bir şekilde verilmiştir.
- Deneysel çalışmalarda transfer öğrenmenin neden önemli olduğu deneysel olarak ispat edilmiş ve yapılan deneyler ile başarılı sonuçlar edilmiştir.
- Metrik öğrenmede genetik algoritma yaklaşımı kullanılarak en uygun W matrisiyle başarılı bir sınıflandırma işlemi yapılmıştır.
- Derin metrik öğrenme histopatolojik veri kümesi üzerinde uygulanmış ve başarılı sonuçlar elde edilmiştir. Ayrıca derin metrik öğrenme ile bir model önerisi sunulmuştur.
- Histopatolojik veriler üzerinde öncelikle çekirdek bölütleme işlemi yapılmış, daha sonrasında derin metrik öğrenme uygulaması yapılarak bu alanda da uygulama geliştirilmiştir.

Bu tez çalışmasının her bir bölümünde yapılanlara aşağıda maddeler halinde yer verilmektedir:

İkinci bölümde metrik öğrenme ile neyin kastedildiği ve metrik öğrenmenin literatürde yerine odaklanılmıştır. Bu bölümde metrik öğrenmenin matematiksel alt yapısı ve ait olması gereken özellikleri sıralanmıştır. Genel metrik öğrenme yaklaşımında hedeflenenin en iyi W dönüşüm matrisini elde etmek olduğu ve bu sebeple bu tez çalışmasında yapılan uygulamalarda da en uygun W ağırlıkları tespit edilmeye çalışılmıştır.

Üçüncü bölümde derin öğrenme konusunda bazı temel bilgiler verilmiştir. Daha sonra bu tez çalışmasında kullanılan veri kümesi görüntüler içerdiği için derin öğrenmede bulunan evrişimli sinir ağlarından temel olarak bahsedilmiştir. Son olarak transfer öğrenmenin derin sinir ağlarına olan pozitif etkisinden hareketle; transfer öğrenmede görüntü verileri için sıklıkla başvurulmuş bazı derin öğrenme modelleri hakkında bilgiler verilmiştir.

Dördüncü bölümde tezin temel konusu olan derin metrik öğrenmenin literatürdeki yeri ve katkılarına detaylıca yer verilmiştir. Derin metrik öğrenme üç temel yapıdan oluşmaktadır. Derin metrik öğrenmenin temel problemleri ağ modeli, DML kayıp fonksiyonları ve ağa sunulacak örneklerin seçiminden oluşmaktadır. Her bir problem detaylıca ele alınmış olup veriye bağımlılığın ve veriye özgü çözümlerin üzerinde durulmuştur.

Beşinci bölümde metrik öğrenme yaklaşımından esinlenilerek bir uygulama gerçekleştirilmiştir. Literatürde genetik algoritma optimizasyon amacıyla kullanılırken yapılan bu çalışmada özgün bir yaklaşım sunularak genetik algoritma ile en uygun W dönüşüm ağırlıkları tespit edilerek sınıflandırma başarısının artırılması hedeflenmiştir.

Altıncı bölümde derin metrik öğrenmede sınıflandırma amacıyla kullanılacak olan histopatolojik veri kümesi tanıtılmış ve bu veri kümesinin ne gibi problemleri olduğu hususuna odaklanılarak veriye özgü çözüme odaklanılmıştır. Bu sebeple özgün veri kümesi üzerinde DML algoritmaları uygulanmış ve bu ağ modelleri ile veri kümesi üzerinde oldukça yüksek sonuçlar elde edilmiştir. Literatürde DML benzerlik yaklaşımıyla özellikle bilgi edinim görevleri için kullanılırken; bu tez çalışmasında ise DML'in aynı zamanda bir sınıflandırma yaklaşımı olarak da kullanılabileceği ve bu amaçla DML'in kullanılmadığı bir alan olan histopatolojik görüntüler üzerinde başarılı bir şekilde sınıflandırma model yaklaşımları sunulmuştur.

Yedinci bölümde histopatholojik görüntülerin literatürde ki bir diğer problemi olan hücre çekirdeklerinin sınıflandırılması çözümü için histopatolojik görüntüler üzerinde öncelikle çekirdeklerin tespiti başarılı bir şekilde gerçekleştirilmiştir. Daha sonra tespit edilen bu çekirdekler bir sınıflandırma işlemine tabi tutulmuştur. Sadece çekirdekler kullanılarak yapılan sınıflandırma ile oldukça memnun edici sonuçlar elde edilmiştir.

Sekizinci bölümde tezin literatüre katkısı, problemlerin çözümü ve DML'in histopatolojik veriler üzerinde başarıyla uygulanabildiği konusunda genel değerlendirmelerde bulunulmuştur.



2. METRİK ÖĞRENME

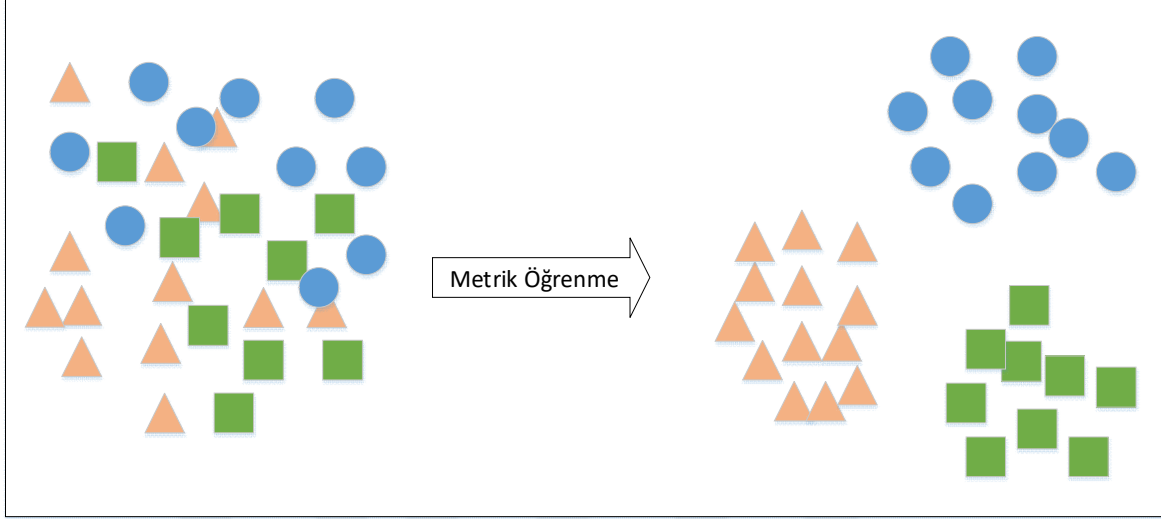
Metrik öğrenme benzerlik esaslı bir yaklaşım olup, örnekler arasındaki benzerliği tespit etmek için bir uzaklık metriği öğrenmeyi amaçlamaktadır [9]. Görüntü, metin, ses gibi gerçek dünya verilerinin her birinin kendisine özgü çeşitli problemleri bulunmaktadır. İnsan doğası gereği bu tür problemleri çözmede başarılı olmasına karşın günümüz teknolojisinde Gauss dağılımı gibi bir dağılıma sahip olmayan veriler üzerinde işlemleri gerçekleştirmek ve öğrenme sürecini maksimum başarı ile elde etmek güçtür. Metrik öğrenme bu problemi aşmak için veriye dayalı bir uzaklık metriğine odaklanmaktadır. Burada veriden metrik öğrenmeyi amaç edinerek; belirlenecek uzaklık metriği ile verinin kendine özgü problemlerini çözmede oldukça etkin çözümler sunulmaktadır. Öğrenilecek uzaklık metriği ile sınıflandırma, kümeleme gibi makine öğrenmesi konularında doğrudan veriden öğrenen yaklaşımlar algoritmanın performansını olumlu yönde etkilemektedir [12].

Temel öğrenme yaklaşımları verileri değerlendirmede yetersiz kalabilmektedir. Literatürde bu tür problemlerin üstesinden gelmek ve daha iyi bir veri temsili elde etmek için metrik öğrenme yaklaşımını esas alan çalışmalar bulunmaktadır [5, 12-14]. Günümüzde popüler olan ve aynı zamanda alt dallara ayrılmış bir çalışma alanı olarak metrik öğrenme konusu güncelliğini koruyan ve gelişime açık bir konudur.

Bir sınıflandırma probleminde verinin doğru bir şekilde anlaşılır olması için metrik öğrenme yaklaşımıyla veriyi en iyi şekilde ifade eden yeni bir uzaklık metriği odak noktasıdır. Literatürde sıklıkla kullanılan Öklid uzaklık metriği verilerin birbirlerine uzaklığı ve yakınlığıyla doğrudan ilgilidir. Ancak bu metriğin kabiliyeti sınırlıdır. Metrik öğrenme örnekler arasındaki uzaklığı çok daha iyi bir şekilde ayırt edebilmek için öğrenme kabiliyeti yüksek bir uzaklık metriği öğrenme eğilimindedir. Burada esas gaye ise; benzer veri örneklerini birbirine yakınlaştırmak olup, farklı veri örneklerini ise birbirinden uzaklaştırma eğilimi bulunmaktadır.

Metrik öğrenme bir uzaklık metriğine dayalı olmakla birlikte benzer nesneleri (aynı sınıf etiketine sahip) birbirine yakınlaştırmayı hedeflemekle birlikte; farklı nesneleri (farklı sınıf etiketine sahip) ise birbirinden uzaklaştırmayı amaçlamaktadır. Şekil 2.1'de görüldüğü üzere 3 farklı nesne tipi bulunmaktadır. Metrik öğrenmede farklı yaklaşımlar olmakla birlikte temel amaç Şekil 2.1'de görüldüğü üzere bu nesnelerden aynı tipte olanlar

arasındaki uzaklık azaltılmakta, farklı tipteki nesneler arasındaki uzaklık ise arttırılmaktadır.



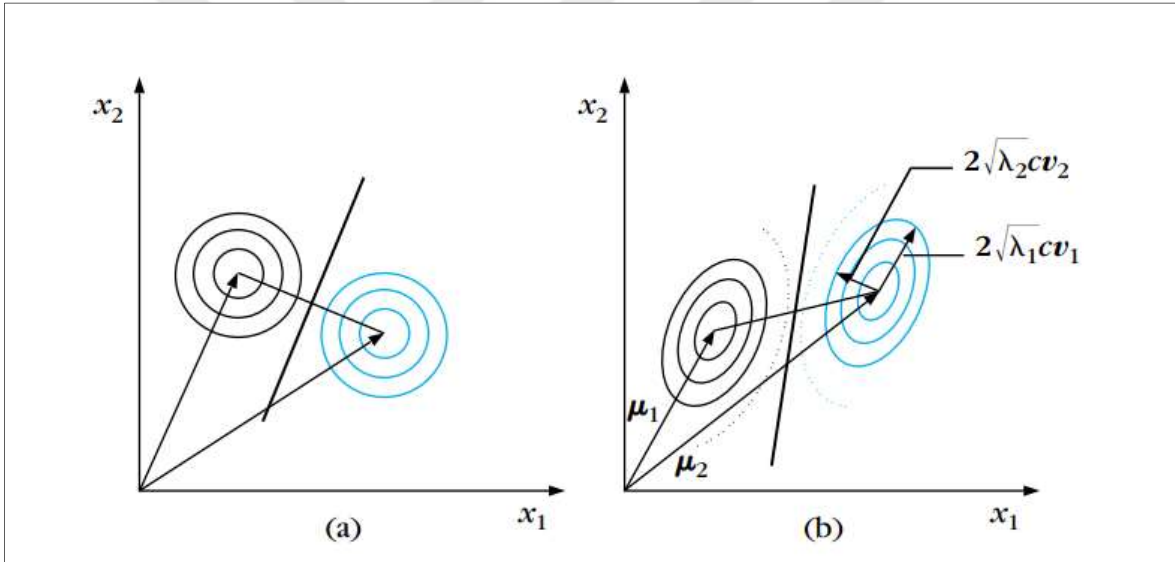
Şekil 2.1. Metrik öğrenme

Literatürde makine öğrenmesinde sıklıkla kullanılan en temel ve basit sınıflandırıcılardan biri olan k en yakın komşu yöntemi de uzaklık metriğine (Öklid uzaklığı) dayalı bir yaklaşımdır. Bu yaklaşımda metrik öğrenmeden ziyade var olan metriği kullanarak aynı ya da farklı tipteki nesnelerin birbirine olan uzaklığına bağlı olarak bir karar verilmektedir. Burada veri üzerinde bir dönüşüm ya da verinin farklı bir uzaya taşınması söz konusu olmayıp aynı düzlem üzerinde karar verme söz konusudur. Ancak elbette bu durumdan farklı olarak verinin farklı bir uzaya taşınıp, yeni bir temsili uzayda verinin temsili ile daha farklı bir görünüm kazandırmak mümkündür [15].

Literatürdeki mevcut çalışmalar, hangi uzaklık metriğinin daha sağlıklı sonuç vereceği ya da veri uzayını hangi ölçütün daha iyi temsil edeceği noktasına odaklanmaktadır [16]. Uzaklık metrikleri aynı zamanda veri örnekleri arasındaki benzerliği tespit etmek amacıyla da kullanılmaktadır. Literatürde de sıklıkla kullanılan bazı uzaklık veya benzerlik metriklerine Öklid, Manhattan, Minkowski, Mahalanobis, Bhattacharya, Chi-Square, Cosine benzerlik, Pearson korelasyon katsayısı örnekler olarak verilebilir [17].

2.1. Uzaklık Metriğine Dayalı Öğrenme Yaklaşımı

Literatürdeki çalışmaların birçoğu incelendiğinde metrik uzaklığı öğrenme teorisinin alt yapısı kurgulanırken Mahalanobis uzaklığı [18] üzerinde durulmaktadır [3-5, 15]. Metrik öğrenme alt yapısında veri örnekleri yeni bir temsili uzaya taşınma gerçekleştirerek bilgisayar yazılımları açısından daha anlamlı bir örüntüye sahip olabilmektedir. Bu yeni bir temsili uzaya taşınma sürecinde bir W dönüşüm matrisi ile uzaklık metriği arasında doğrudan bir ilişki bulunmaktadır. Metrik öğrenmenin Mahalanobis uzaklık metriği üzerine olan teorik alt yapısı da yine bu dönüşüm matrisine dayanmaktadır. Öklid uzaklık metriği ile Mahalanobis uzaklık metriğinin iki boyutlu bir düzlemde iki nokta arasında uzaklığa bakış açısı Şekil 2.2’de görülmektedir.



Şekil 2.2. Uzaklık metrikleri (a) öklid uzaklığı (b) mahalanobis uzaklığı [17]

Matematiksel olarak Mahalanobis uzaklık metriğinin metrik öğrenme arasındaki ilişkiyi açık bir şekilde ifade etmek gerekirse; $X = [x_1, x_2, \dots, x_N] \in \mathbb{R}^{d \times N}$ ile N boyutlu bir eğitim örnekleri kümesi kabul edelim. Burada $x_i \in \mathbb{R}^d$ olup eğitim örneklerini temsil etmektedir. Eğitim örneklerinden x_i ile x_j arasındaki uzaklığı, Öklid uzaklık metriğinden yararlanıp ℓ^2 normu ile ifade edersek aşağıdaki denklem ortaya çıkacaktır.

$$d_{\text{öklid}}(x_i, x_j) = \|x_i - x_j\|_2 \quad (2.1)$$

Aynı uzaklık metriğini denklem 2.2’de ki gibi de ifade edebiliriz.

$$d_{\text{öklid}}(x_i, x_j) = \sqrt{(x_i - x_j)^T (x_i - x_j)} \quad (2.2)$$

Öklid uzaklık metriği ölçeklendirme ve öznitelik boyutlarına karşı duyarlı olup farklı bir ek bilgiyi kullanmamaktadır. Bu durum verinin benzerliğini tam olarak doğru bir şekilde yansıtmasının önünde bir engel olarak karşımıza çıkmaktadır [19]. Bu tür problemlerin üstesinden gelmek için Mahalanobis uzaklık metriğinden yararlanılarak metrik öğrenme bakış açısıyla öğrenme kavramına odaklanılmaktadır.

Temel amacımız olan Mahalanobis uzaklık metriğini nasıl Öklid uzaklık metriği cinsinden ifade edeceğimiz noktasına gelince; öncelikle iki boyutlu düzlemde noktalar arasındaki uzaklığı Mahalanobis uzaklık metriği cinsinden ifade ederek gösterelim.

$$d_{\text{mahalanobis}}(x_i, x_j) = \sqrt{(x_i - x_j)^T M (x_i - x_j)} \quad (2.3)$$

Burada $d_{\text{mahalanobis}}(x_i, x_j)$ uzaklık metriği bazı temel özelliklere sahiptir.

$$d_{\text{mahalanobis}}(x_i, x_j) \geq 0 \quad \text{Uzaklığın negatif olmaması} \quad (2.4)$$

$$d_{\text{mahalanobis}}(x_i, x_i) = 0 \quad \text{Özdeşlerin ayırt edilemezliği} \quad (2.5)$$

$$d_{\text{mahalanobis}}(x_i, x_j) = d_{\text{mahalanobis}}(x_j, x_i) \quad \text{Simetrik olması} \quad (2.6)$$

$$d_{\text{mahalanobis}}(x_i, x_k) \leq d_{\text{mahalanobis}}(x_i, x_j) + d_{\text{mahalanobis}}(x_j, x_k) \quad \text{Üçgen eşitsizliği} \quad (2.7)$$

Bu özellikler benzer yaklaşımlara sahip yöntemlerin olması gereken temel özellikleridir. Denklem 2.3’te ki M parametresi simetrik ve pozitif yarı tanımlı olmak zorundadır. M ’nin pozitif yarı tanımlı olması için belirleyicileri ya da tüm öz değerleri pozitif ya da sıfır olmalıdır [8]. Burada M kovaryans matrisi olup, onu ayrıştırıp Mahalanobis uzaklık metriğini bir ağırlık cinsinden ifade edip yeniden denklemi oluşturursak;

$$M = W^T W \quad (2.8)$$

$$d_{mahalanobis}(x_i, x_j) = \sqrt{(x_i - x_j)^T M (x_i - x_j)} \quad (2.9)$$

$$d_{mahalanobis}(x_i, x_j) = \sqrt{(x_i - x_j)^T W^T W (x_i - x_j)} \quad (2.10)$$

$$d_{mahalanobis}(x_i, x_j) = \sqrt{(W(x_i - x_j))^T (W(x_i - x_j))} \quad (2.11)$$

$$d_{mahalanobis}(x_i, x_j) = \|Wx_i - Wx_j\|_2 \quad (2.12)$$

Denklem 2.12’de ki Mahalanobis uzaklığının nihai hali metrik öğrenmenin temel yaklaşımını göstermektedir. Metrik öğrenmede öğrenme yaklaşımı dönüşüm matrisi olarak ifade edilen M ’nin probleme dayalı olarak en uygun ağırlıklarının elde edilmesine dayanmaktadır. Bu dönüşüm bize sadece temel bir yapıyı sunmakla beraber literatürde farklı yaklaşımlarda söz konusudur.

2.2. Uzaklık Metriğine Dayalı Yöntemler

Sınıflandırma ve kümeleme problemlerinde ayırım gücü yüksek bir yeni dönüşüm uzayı elde etmek öğrenme yönteminde daha iyi tahminler yapmamızı sağlar [20]. Metrik öğrenmeyle örnekler arasındaki benzerlik sınıflar arasında minimum seviyeye indirilir ya da sınıflar içinde maksimum seviyeye çekilir. Bu süreç metriğin veriden ne kadar iyi öğrendiğine bağlı olmaktadır. Uzaklık ölçütüne dayalı öğrenme yaklaşımında öğrenme sürecinde verinin yerel ve global olarak değerlendirme ile doğrusal ve doğrusal olmayan şekilde ele alınma biçimine göre farklı çözüm yöntemleri bulunmaktadır [21, 22]. Bu yöntemlerde veriyi dönüşüm uzayında taşıyıp; yeni bir temsili uzayda haritalama yapılmaktadır. Mahalanobis uzaklık metriğinde olduğu gibi yapılan W dönüşüm matrisi yeni bir temsili uzaya taşınma işlemi ile doğrusal bir dönüşüm sağlanmaktadır. Doğrusal dönüşüme sahip yaklaşımların optimize edilmesi daha kolay olup, aynı zamanda aşırı öğrenmeye karşı daha az eğilim göstermektedirler [23]. Ancak gerçek dünya problemleri genellikle doğrusal olmayan problemlere dayalı veriler içermektedir.

Veriden öğrenme süreci ancak verinin doğasına uygun hareket etmekle mümkündür. Doğrusal dönüşüme sahip yöntemler veriden öğrenme sürecini sınırlı yetenekleriyle gerçekleştirirken; verinin doğrusal olmayan özelliklerini yakalamakta problem yaşamaları mümkündür. Metrik öğrenmede bu tür sorunların üstesinden gelmek için çekirdek fonksiyonları aracılığıyla doğrusal olmayan uzayda temsil kabiliyeti sağlanır [20]. Bu yöntemlerin dezavantajı ise aşırı öğrenmeden etkilenme olabilir. Verinin doğrusal ya da doğrusal olmaması ile ifade edilen durum, verinin taşınacağı uzayda nasıl bir geometrik yapı barındırdığı ile ilgilidir.

Yerel ve global çözümünü ele aldığımızda veriyi bir görüntü olarak tasvir edelim. Burada görüntüyü de bir yüz görüntüsü olarak ifade ettiğimizi düşünelim. Yerel olmakla ifade edilen durum burada göz için; gözün bulunduğu piksellere odaklanma ve her bir farklı kısım ya da küme için komşuluklarına dayalı bir çözüm olarak ele alınmaktadır. Global yaklaşımda ise veri bir bütün olarak değerlendirilmekte olup dönüşüm uzayında genel bir çözüm sunulmaktadır.

2.3. Uzaklık Metriğine Dayalı Öğrenme Algoritmaları

Bu bölümde literatürde genel kabul görmüş olan bazı uzaklık ölçütüne dayalı algoritmalar üzerinde durulacaktır. Öncelikle literatürde genel bir analizi yapılmış akademik makalelerde farklı kabuller [4, 19, 24]. olmakla birlikte gözetimli, gözetimsiz ve yarı gözetimli olarak uzaklık metriği öğrenme algoritmaları üç farklı kategoride ele alınacaktır. Yöntemler ele alınırken veri N boyutlu bir uzayda $X = [x_1, x_2, \dots, x_N] \in \mathbb{R}^{d \times N}$ şeklinde temsil edilmektedir. Uzaklık metriğine dayalı öğrenme ele alınırken genel yaklaşım ise;

$$J(D) = \lambda_1 L(D) + \lambda_2 U(D) \quad (2.13)$$

şeklinde dir.

$L(D)$, verinin etiket bilgisini içeren bilgiyi ifade eden gözetimli öğrenmeyi ifade ederken; $U(D)$ ise veri dağılımıyla ilgilenip gözetimsiz öğrenmeye karşılık gelmektedir. λ parametrelerinin aldığı değere bağlı olarak algoritmalar bu üç kategoriden biriyle sınıflandırılmaktadır. λ parametresi sıfır ya da pozitif bir reel sayı olup; her iki değer in sıfırdan büyük olması durumunda yarı gözetimli bir öğrenme olarak kategorize edilmekte,

herhangi biri sıfır iken ise gözetimli ya da gözetimsiz bir öğrenme olarak değerlendirilmektedir.

2.3.1. Gözetimli uzaklık metriği öğrenme

Gözetimli uzaklık ölçütü öğrenme verinin benzerlik ya da benzersizlik ölçütü üzerinden bir değerlendirme yaparken etiket bilgisinden faydalanmaktadır [25]. Gözetimli öğrenmede $\lambda_2 U(D)$ değeri sıfır olarak alınmakta olup öğrenme sürecinde bir etkisi olmamaktadır. Literatürde birçok gözetimli öğrenme yöntemi olmakla beraber burada belli başlı bazı algoritmalara odaklanılacaktır. Gözetimli uzaklık metriği öğrenme algoritmalarının bazısının doğrusal veya doğrusal olmama durumu ile yerel veya global olma durumlarına göre sınıflandırılmış hali Çizelge 2.1’de verilmektedir.

Çizelge 2.1. Gözetimli uzaklık metriği öğrenme algoritmaları

	Yerel	Global
Doğrusal	NCA [26] ANMM [27] LMNN [5]	LDA [28] RCA [29] ITML [13]
Doğrusal olmayan	KANMM [27]	KLDA [30] KRCA [31]

Goldberger ve arkadaşları [26] tarafından geliştirilen NCA (Neighborhood Component Analysis) yönteminde yerelliğe dayalı bir çözüm sunulmuştur. Yöntem yerelliği sağlama amacıyla komşuluk bilgisinden faydalanmaktadır. Yaklaşım olarak Mahalanobis uzaklık metriğini temel almakta ve dönüşüm uzayını elde ederken Öklid uzaklık ölçütünden yararlanmaktadır. Gradyan temelli bir optimizasyon çözümü sunmaktadır. Dönüşüm yaparken W haritalamasından faydalanmakta ve özdeğer ayrıştırması ile dönüşüm elde edilmektedir. Wang ve Zhang [27], ANMM (Average Neighborhood Margin Maximization) yönteminde sınıfların yerel olarak birbirlerinden en iyi şekilde ayrılabilirdiği yeni bir projeksiyon uzayını elde etmeye odaklanmışlardır. Bu sebeple yöntemde homojen ve heterojen olmak üzere iki farklı komşuluktan yararlanılmıştır. Ortalama komşuluk için her iki yöntemin ortalaması alınıp daha sonra heterojen ortalamadan homojen ortalama farkı alınarak bir marjın değeri elde etmektedir. Dönüşümde kullanılmak üzere $Y = W^T x$

eşitliğinden yararlanılmıştır. Öz değer ayrıştırması temelli olup Öklid uzaklığından yararlanmaktadır. Yazarlar çekirdek yöntem ile aynı zamanda yöntemin doğrusal olmayan projeksiyonuna dayalı bir öğrenme yöntemini de önermişlerdir. Weinberger ve Saul [5], LMNN yöntemi (Large Margin Nearest Neighbor) ile aynı etiket değerine sahip verileri birbirlerine yakınlştırıp, farklı etiket değerlerine sahip verileri birbirinden uzaklaştıracak şekilde yeni bir marjın formülünü önermişlerdir. Yöntemde veriler arası uzaklığı sağlamak amacıyla W matrisinden yararlanılmıştır. Uzaklık hesabında Öklid metriğinden yararlanılmıştır.

LDA (Linear Discriminant Analysis) [28], farklı sınıfları birbirinden daha iyi ayırt etme amacıyla bir projeksiyon çözümü sunmaktadır. Sınıf içinde dağılım matrisi ve sınıflar arası dağılım matrisinden yararlanılmaktadır. Hesaplama W matrisinden yararlanılmakta ve öz değer ayrıştırmasını da içermektedir. Diğer yöntemlerde olduğu gibi Öklid uzaklık metriği tercih edilmektedir. Yöntem Mika ve arkadaşları [30] tarafından çekirdek yaklaşımı kullanılarak doğrusal olmayan bir yapıya dönüştürülmüştür. Shental ve arkadaşları [29], ilgililiği arttırıp, ilgisizliği ise azaltmaya dönük olarak bir dönüşüm uzayı elde etmeyi amaçlamışlardır. Geliştirilen RCA (Relevant Component Analysis) yöntemi yine bir W dönüşüm matrisi üzerinde kurulu olup; çekirdek kullanılarak geliştirilen KRCA (Kernel Relevant Component Analysis) [31], doğrusal olmayan yaklaşıma sahip algoritma da bulunmaktadır. ITML (Information-Theoretic Metric Learning) [13]. Davis ve arkadaşları tarafından geliştirilmiş olup bilgi teorisine dayalı Mahalanobis uzaklığını esas almaktadır. Pozitif tanımlı matris konisi üzerinde bilgi teorisi olarak Bregman matrisinden yararlanılmaktadır.

2.3.2. Gözetimsiz uzaklık metriği öğrenme

Gözetimsiz uzaklık ölçütü öğrenme herhangi bir etiket bilgisine ihtiyaç duymayan ve verinin dağılımı ile ilgilenen bir öğrenmedir. Bu öğrenme tekniğinde veriler bir geometrik uzayda temsil edilebileceği gibi; aynı zamanda en uygun bir ayırmacı ile de veriler ayrıştırılabilirler [4]. Gözetimsiz uzaklık ölçütü öğrenmede $\lambda_1 L(D)$ değeri sıfır olup öğrenme sürecine katkısı yoktur. Literatürde kabul gören bazı gözetimsiz uzaklık ölçütü öğrenme algoritmaları özelliklerine göre Çizelge 2.2’de verilmektedir.

Çizelge 2.2. Gözetimsiz uzaklık metriği öğrenme algoritmaları

	Yerel	Global
Doğrusal	LPP [32]	PCA [33]
Doğrusal olmayan	LE [34] LLE [35] ISOMAP [36]	KPCA [37]

PCA (Principal Component Analysis) [33], algoritması basit ve uygulanması kolay bir yöntem olup veri vektörlerinin temel bileşenleri üzerine kurulu bir yöntemdir. Yöntemde öncelikli olarak verinin ortalamaları bulunmakta, daha sonra veriden farkları alınarak yeni bir X veri vektörü elde edilmektedir. Elde edilen X vektörünün verileri arasındaki ilişkisini inceleme amacıyla kovaryans matrisi elde edilip bu kovaryans matrisi üzerinden öz vektörler elde edilmektedir. Elde edilen öz vektörün verinin orijinal hali ile çarpılması sonucunda dönüşüm matrisi elde edilmektedir ($Y = W^T x$). PCA algoritmasının çekirdek ile haritalanması sonucu doğrusal olmayan dönüşüme sahip algoritması olan KPCA (Kernel Principal Component Analysis) Schölkopf ve Smola tarafından 2002 yılında geliştirilmiştir [37].

LPP (Locality Preserving Projections) [32], verinin yeni bir uzaya dönüşümü yapıldıktan sonra orijinal uzayındaki görüntüsünü korumayı amaçlamaktadır. Yöntem aynı zamanda verinin orijinal uzaydaki yerelliğini de korumaktadır. Komşuluk esasına dayalı olan yerel komşulukları korumakla yükümlüdür. Komşular arası uzaklık hesabında Öklid uzaklığı kullanılmakta ve öz değer ayrıştırması işlemi gerçekleştirilmektedir. Dönüşüm uzayında da W ağırlıklarından yararlanılmaktadır.

LE (Laplacian Embedding) [34] yöntemi, yerelliği koruma için yine LPP'deki yerellik denkleminde faydalanmaktadır. LPP'deki yöntemle benzer şekilde Laplasyan matrisinden yararlanılmaktadır. LLE (Locally Linear Embedding) [35], yöntemi de yine yerelliğe dayalı bir yaklaşım olup; verinin komşulukları arasındaki değere göre yeni bir X verisi elde etmektedir. Algoritma komşulukları kullanırken doğrusal bir denklem kullandığı için denklemin isminde "Linear" kavramı olmakla beraber esasında doğrusal olmayan yapıya sahip bir yöntemdir. "Linear" kavramı yerel komşuluklardan elde edilen verideki değerlerin ağırlık matrisi ile çarpılıp "1" değerine eşitlenmesini ifade etmektedir.

Algoritma diğer yöntemlerde olduğu gibi dönüşüm uzayını elde ederken W matrisinden yararlanmaktadır. Son aşamada elde edilen dönüşüm uzayı öz vektörlerden elde edilen dönüşüm uzayıdır. Isomap (Isometric Feature Mapping) [36], geometrik mesafelere dayalı bir çözüm olup ağırlıklı bir graf üzerinde bu uzaklıklar kullanılmaktadır. Diğer geometrik algoritmalarından farklı olarak benzerlik uzaklığı yerine benzersizlik üzerine kuruludur. Yine bu yöntemde de komşuluklar üzerine kurulu bir yapı olup, bu yüzden yerel bir yaklaşıma sahip olup aynı zamanda doğrusal olmayan bir yapıya sahiptir. Algoritmada ağırlıklar belirlenirken komşuluk oluyorsa bir ağırlık bulunurken; eğer veriler arasında komşuluk bulunmuyorsa ağırlık 0 olarak kabul edilir.

2.3.3. Yarı gözetimli uzaklık metriği öğrenme

Yarı gözetimli uzaklık ölçütü öğrenme hem etiket bilgisini kullanan hem de veri dağılımını birlikte inceleyen bir öğrenme yaklaşımıdır. Öğrenme tekniğinde λ_1 ve λ_2 değerleri sıfır olmayıp, öğrenme yöntemine katkıları her iki yönteminde bazı sınırlar doğrultusunda dengelenmesi ile elde edilmektedir. Literatürde kabul gören bazı yarı gözetimli uzaklık metriği öğrenme algoritmaları özelliklerine göre Çizelge 2.3'te verilmektedir.

Çizelge 2.3. Yarı gözetimli uzaklık metriği öğrenme algoritmaları

	Yerel	Global
Doğrusal	LRML [38]	LRML [38]
Doğrusal olmayan	KLRML [38] SSDR [39]	KLRML [38] SSDR [39]

LRML (Laplacian Regularized Metric Learning) [38] yönteminde, hem etiket bilgisinden hem de veri dağılımından yararlanılmaktadır. Öğrenme sürecinde etiket bilgisi kullanımında hem ilgililik hem de ilgisizlik özellikleri dikkate alınmış ve hesaba katılırken ilgililik değeri, ilgisizlik değerinden çıkarılma şeklinde dikkate alınmıştır. Gözetimsiz kısım ise veriyi pürüzsüzleştirmeye odaklanmıştır. Yöntem gözetimsiz öğrenme kısmında yerel olup, gözetimli kısmında ise global bir yapıya sahiptir. Yöntem çekirdek yapısı ile aynı zamanda doğrusal olmayan bir yapıya büründürülmüştür. Yang ve arkadaşları [39], tarafından geliştirilen SSDR (Semi-Supervised Dimensionality Reduction) uzaklık öğrenme algoritması verinin dönüşümü amacıyla kullandığı matriste PCA'daki kovaryans

matristen yararlanabileceği gibi LE’de ki Laplacian matrisinden de yararlanabilmektedir. Bu matris hesabında PCA’daki mantığı kullanırsa yöntem global bir yöntem olurken, Laplacian matris kullanıldığında ise yerel bir çözüm olmaktadır. Uzaklık metriği olarak Öklid uzaklığından yararlanılmaktadır.

2.4. Uzaklık Metriğinde Ağırlık Hesaplama

Uzaklık metriğine dayalı öğrenme algoritmalarını analiz ettiğimizde bir W matrisine odaklanıldığını ve dönüşüm uzayına geçişte temel parametrenin bu W matrisini elde etmek yönünde olduğu açıktır. W matrisi ya da w ağırlık değerleri farklı algoritmalar ve yaklaşım alanlarında veriyi en iyi temsil eden ayırıcı olarak önem değeri oldukça yüksektir. Verileri en iyi şekilde analiz etmede en uygun W matrisini ya da w ağırlığını bulmak çözümlerdeki temel yaklaşımdır. Bu amaçla Çizelge 2.4’te farklı yaklaşımlara ait W değerlerinin nasıl elde edildiğini gösteren hesaplamalar verilmektedir.

Çizelge 2.4. W ağırlık hesaplama

Yöntem	W kullanımı
Perceptron	<u>w güncelleme:</u> $w(t+1) = w(t) + \rho x(x), \quad x(t) \in \omega_1 \text{ ve } w^T(t)x_t \leq 0$ $w(t+1) = w(t) - \rho x(x), \quad x(t) \in \omega_2 \text{ ve } w^T(t)x_t \geq 0$ $w(t+1) = w(t)$ <u>Sınıf Hesabı:</u> $w^{*T}x > 0, \quad \forall x \in \omega_1$ $w^{*T}x < 0, \quad \forall x \in \omega_2$
Pocket Algoritması	<u>w güncelleme:</u> $h > h_s \text{ ise } w_s = w(t+1) \text{ ve } h_s = h$ En iyiyi tutma <u>Sınıf Hesabı:</u> $w^{*T}x > 0, \quad \forall x \in \omega_1$ $w^{*T}x < 0, \quad \forall x \in \omega_2$
Geriye Yayılım Algoritması	<u>w güncelleme:</u> $w_j^r = [w_{j0}^r, w_{j1}^r, \dots, w_{jk_{r-1}}^r]$ $w_j^r(\text{yeni}) = w_j^r(\text{eski}) + \Delta w_j^r$ $\Delta w_j^r = -\mu \sum_{i=1}^N \delta_j^r(i) y^{r-1}(i)$
SVM	<u>w hesaplama:</u> $w = \sum_{i=1}^N \lambda_i y_i x_i$ $\sum_{i=1}^N \lambda_i y_i = 0$ <u>Sınıf Hesabı:</u> $w^T x + w_0 \geq 1, \quad \forall x \in \omega_1$ $w^T x + w_0 \leq -1, \quad \forall x \in \omega_2$

Çizelge 2.4. (devam) W ağırlık hesaplama

LDA	<u>w hesaplama:</u> $S_W = \sum_{i=1}^M P_i \Sigma_i$ $\Sigma_i = E[(x - \mu_i)(x - \mu_i)^T]$ $P_i = \frac{n_i}{N}$ $w = S_W^{-1}(\mu_1 - \mu_2)$ <u>Dönüşüm Uzayı:</u> $y = w^T x$
Laplacian EigenMaps	<u>W hesaplama ve dönüşüm uzayı:</u> Öklid metriği ile her bir niteliğin birbirine uzaklığı bulunur. $W(i, j) = \begin{cases} \exp\left(-\frac{\ x_i - x_j\ ^2}{\sigma^2}\right) & \text{Komşuluk ilişkisi varsa} \\ 0 & \text{Komşuluk ilişkisi yoksa} \end{cases}$ $D_{ii} = \sum_j W(i, j)$ $L = D - W$ $Lv = \lambda Dv$ $y_i^T = [v_1(i), v_2(i), \dots, v_m(i)] \quad \text{Öz vektörler dönüşüm uzayı}$ $E_L = \sum_{i=1}^n \sum_{j=1}^n (y_i - y_j)^2 W(i, j) \quad \text{Yakınlık kriteri}$
LLE	<u>w hesaplama:</u> Öklid uzaklık metriği ile her bir metriğin birbirine olan uzaklığı bulunur $\argmin E_W = \sum_{i=1}^N \left\ x_i - \sum_{j=1}^n W(i, j) x_{i_j} \right\ ^2$ Komşuluk yoksa ağırlıkları sıfırla, tüm ağırlıkların toplamı sıfır olacak $\sum_{j=1}^n W(i, j) = 1$ Y dönüşüm uzayını ağırlıkları kullanarak öz vektörden hesapla

Çizelge 2.4. (devam) W ağırlık hesaplama

Ayrık Fourier Dönüşümü – Bir Boyutlu	<u>W hesaplama:</u> $W_N \equiv \exp\left(-j\frac{2\pi}{N}\right)$ $W_N = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & W_N & W_N^2 & \dots & W_N^{N-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \dots & W_N^{(N-1)(N-1)} \end{bmatrix}$ <u>Dönüşüm Uzayı:</u> $Y = W^H x$
Ayrık Fourier Dönüşümü – İki Boyutlu	<u>W hesaplama:</u> $W_N \equiv \exp\left(-j\frac{2\pi}{N}\right)$ $W_N = \frac{1}{\sqrt{N}} \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & W_N & W_N^2 & \dots & W_N^{N-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & W_N^{N-1} & W_N^{2(N-1)} & \dots & W_N^{(N-1)(N-1)} \end{bmatrix}$ <u>Dönüşüm Uzayı:</u> $Y = W^H x W^H$

3. DERİN ÖĞRENME

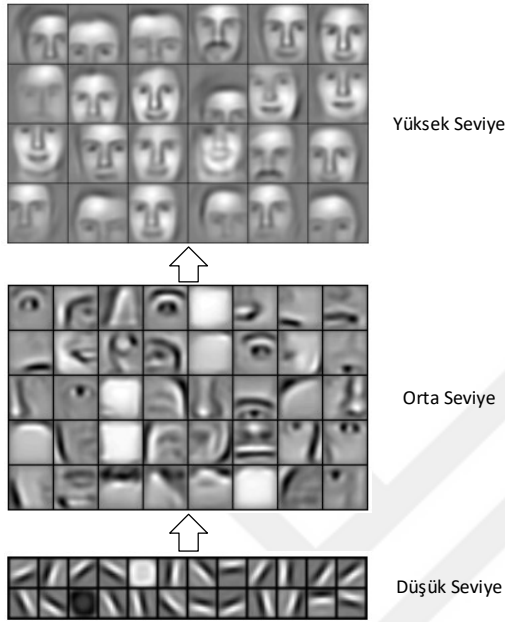
3.1. Derin Öğrenme

Gelişen teknolojiyle beraber verilerin dijital ortamlara aktarılması sürecinde hacimsel ve çeşitlilik anlamında çok büyük artışlar olmuştur. Bilgisayar bilimciler için artan verinin önemi oldukça fazladır. Çünkü verinin artışı ve çeşitliliği veri işleme sürecinde algoritmanın başarısını da pozitif yönden etkilemektedir. Böylece veri daha anlaşılır olmakla birlikte; makine öğrenmesi ve yapay zeka yaklaşımlarıyla daha başarılı tahminler yapmak mümkün olabilmektedir. Ancak veri artışı olumlu bir gelişme olmakla beraber, artan veriyi kullanan algoritmalar açısından negatif yönü de bulunmaktadır. Hem verinin niteliği (görüntü boyutları gibi) hem de örnek sayısının fazlalığından dolayı veriyi işlemek oldukça fazla zaman almaktadır.

Gelişen teknoloji veri artışını sağladığı gibi veri işleme açısından da pozitif gelişmelerin önünü açmıştır. Artan veriyi paralel bir şekilde işleme kolaylığı hız performansı açısından oldukça değerlidir. Günümüz bilgisayar ekran kartlarının çekirdek sayısının binlerle ifade edilmesi ve performans artışıyla beraber paralel mimari teknolojisini etkin bir şekilde kullanmanın yolları açılmıştır. Artan verinin hızlı, etkin ve başarılı bir şekilde kullanımıyla beraber derin öğrenme kavramı ortaya çıkmıştır. Derin öğrenme, makine öğrenmesinin bir alt dalı olup, beyin fonksiyonları ve yapısından esinlenilerek ortaya çıkmış bir öğrenme yaklaşımıdır [40]. Derin öğrenmedeki “derin” kavramı oluşturulan modelde kullanılan katmanların sayısı ile doğrudan ilgilidir. Modeldeki katman sayısı, yapılar karmaşıktıkça verinin işleme kapasitesi artmakla birlikte, çözüm iyileştirilmesi mümkün olmaktadır.

Derin öğrenme doğrudan ham veriyi kullanıp, veriyi daha soyut bir seviyeye taşıyan temsili öğrenme yöntemlerindendir [7]. Burada amaç veriyi daha iyi bir şekilde ifade eden daha yüksek soyutlama seviyelerine ulaştırmaktır. Bir görüntü verisinde derin katmanlı bir mimari üzerinden soyutlama seviyesini anlatmak gerekirse (Şekil 3.1), resimdeki nesnelerin kenarları derin öğrenmenin düşük seviyeli katmanında belirlenir. Orta katmanda, görüntüdeki motifler, kenarlardaki küçük değişikliklerden bağımsız olarak önceki katmandan yararlanılarak belirlenir. Yüksek seviyeli katmanda ise, tespit edilen motiflerin daha büyük kombinasyonları bir araya getirilir ve öğrenme sürecinde veriyi en

uygun temsil eden niteliklerin elde edilmesi hedeflenir. Veriyi tanımlayacak daha iyi bir temsili seviyeye taşıyan derin öğrenme algoritmaları, artan veri hacmi ve hesaplamaya etkin çözümler sunmaktadır.



Şekil 3.1. Derin öğrenme seviyeleri [41]

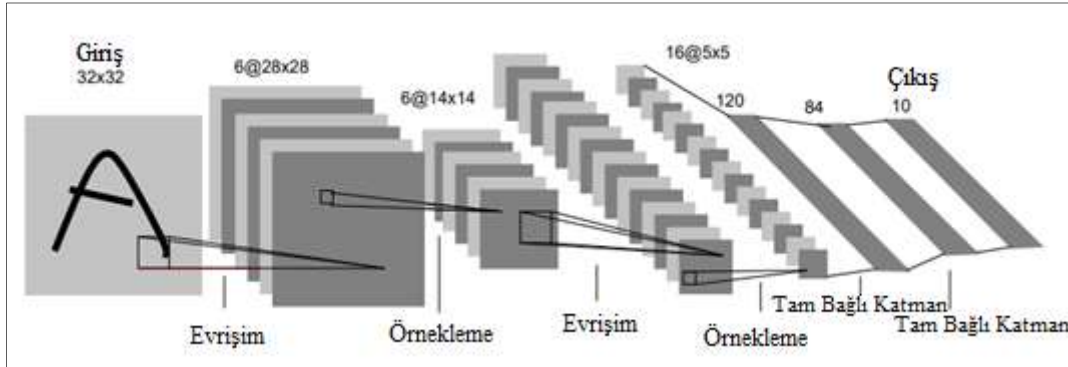
Geleneksel makine öğrenmesi yöntemlerinin ham veriyi doğrudan kullanma başarıları sınırlıdır. Bu sebeple öncesinde bazı ön işlemler gerekebilmektedir. Mevcut makine öğrenmesi yöntemleri; önişleme, öznitelik çıkartma gibi yaklaşımlarla öznitelikler üzerinde uzmanlık gerektiren bazı işlemler içerebilmektedir. Bu işlemler doğrudan bir sınıflandırıcı yapısında olmamaktadır. Derin öğrenme ise bu noktada geleneksel yöntemlerden ayrılarak doğrudan sınıflandırıcı yapısında verinin kabiliyetlerini ortaya çıkarabilmektedir. Geleneksel makine öğrenmesi yöntemleri bu bakış açısıyla derin öğrenmeden ayrılmaktadır. Derin öğrenmede yüksek bir başarıyı elde etmenin yolu yeterli sayıda verinin ağı sunulmasıyla da doğrudan ilgilidir. Her ne kadar gelişen teknoloji ve paralel mimari bakış açısıyla hızlı sonuçların alınması hedeflense de derin öğrenme ağlarında parametrelerin fazlalığı ve ağı derinliği öğrenme sürecini yavaşlatabilmektedir. Günümüzde derin öğrenme alanında NVIDIA'nın Cuda desteğiyle birçok derin öğrenme kütüphanesi geliştirilmiştir.

Bir sinir ağı, nöron adı verilen bir veya daha fazla düğüm içeren katmanlardan oluşur. Düğüm ve katman sayısı, probleme bağlı olarak değişir. Bir ağı en iyi şekilde eğitmek için

en önemli faktör, katmanlarla katman sayısı arasında bağlantı sağlayan iyi yapılandırılmış ağırlıklardır. Temsili öğrenme yöntemi olarak derin öğrenme, son yıllarda makine öğrenimi ve bilgisayarla görme alanındaki araştırmacıların dikkatini çekmiştir. Ham veriler üzerinden verinin yeni bir temsilini sağlayan derin öğrenme, nesne tanıma [42], yüz doğrulama [43], insan hareketi tanıma [44], görüntü sınıflandırma [45] gibi birçok konuda görüntüler üzerinde oldukça yüksek başarılar elde etmiştir.

3.2. Evrişimli Sinir Ağları

Evrişimli sinir ağları, özellikle görüntü verilerinde karmaşıklığı ve soyutlamayı arttıran daha güçlü öznitelikler elde etmek için vazgeçilmez bir sinir ağı modelidir [46]. Çok büyük sayıda veri parametresinin ölçeklendirilmesi hız ve başarımındaki temel faktörlerden birisidir. Tek bir ağ yapısı modeli üzerinden örnekleri sınıflandırmaktadır. Evrişimli sinir ağlarının kökeni çok eskilere dayanmasına rağmen günümüz teknolojisinin gelişimiyle popülerliği oldukça artmıştır. LeCun tarafından önerilen evrişimli sinir ağı modellerinden biri Şekil 3.2’de görülmektedir. Evrişimli sinir ağı modelleri temelde 3 yapıdan oluşur. Bu adımlar evrişim katmanı, örnekleme katmanı ve tam bağlı katmandır.



Şekil 3.2. Evrişimli sinir ağı modeli [47]

3.2.1. Evrişim katmanı

Giriş katmanından alınan verilerin işlenerek örnekleme (pooling) katmanına iletildiği katmandır. Ağ sinir modelinin bu katmanında filtreleme, veri paylaşımı, ağırlıklandırma, yanlılık, ölçeklendirme, yeni kanallar oluşturma temel adımlardır. Katmanda kullanılacak olan veriyi bir görüntü olarak ele alırsak verinin filtrelenmesi bu adımda

gerçekleşmektedir. Görüntü üzerinde kullanılacak olan filtre paylaşımlı bir yapıya sahip olup; uygulanacak olan işlemlerde hacimsel azaltma sağlanmaktadır. Bu işlem de önemli bir zaman kazanımı sağlayacaktır. Filtre boyutunun çok küçük olması işlem karmaşıklığının artmasına neden olacaktır. Filtre boyutunun büyük olması durumunda ise, işlem sayısı daha az olmasına karşın; istenen performans elde edilemeyebilir. Bu sebeple probleme özgü bir filtreleme yaklaşımı sunulmalıdır. Filtreleme komşuluk esasına dayalı olup kullanılacak olan ağırlık değerleri ve yanlılık parça parça uygulanıp katmandaki her bir derinlik için paylaşımlı olup; derinlik sayısınca farklı filtreler kullanılmaktadır.

Evrişim katmanının en temel işlevlerinden biri de ölçeklendirme ve yeni kanallar (derinlikler) oluşturulmasıdır. Bu işlemi de yine görüntü üzerinde dolaşarak kullanılan filtre parametrelerine bağlı olarak sağlamaktadır. Ölçeklendirme yaklaşımında kullanılacak olan modele bağlı olarak görüntü boyutu azaltılabileceği gibi; filtre boyutuna bağlı olarak giriş görüntüsünün 0 piksel değerleri ile genişletilerek aynı kalması da sağlanabilir. Giriş görüntü boyutu W , Filtre boyutu F , Dolgu P , Adım uzaklığı S olmak üzere ölçeklendirme sonrasında oluşacak olan yeni görüntü boyutu aşağıdaki denklem aracılığıyla hesaplanacaktır. Ayrıca görüntüde kullanılacak olan farklı filtre sayısı kadar yeni kanallar (derinlikler) oluşturulacaktır.

$$\text{Yeni Görüntü Boyutu} = \frac{W-F+2P}{S} + 1 \quad (3.1)$$

3.2.2. Örnekleme katmanı

Örnekleme katmanında amaç veri boyutunu azaltmak ve evrişim katmanına göre daha iyi öznitelikler elde etmektir. Bazı çalışmalarda örnekleme katmanı kullanılmadan da başarılı sonuçlar elde edilebildiği görülmüştür. Örnekleme yapmak yerine evrişim katmanındaki adımı büyük tutmakla varyasyonel oto kodlayıcı ve üretken çekişmeli ağlarda daha önemli başarılar elde ettiklerini iddia etmektedirler. Bu katmanda örnekleme (pooling) için literatürde birçok bakış açısı olmakla birlikte en fazla kabul gören yaklaşımlar ortalama örnekleme ve maksimum örneklemedir. Yine bir görüntü üzerinde dolaşmakla birlikte filtre boyutu ve gezilen penceredeki değerler üzerinden işlemler yapılarak veri hacmi azaltılmaktadır. Görüntünün boyutunda (yükseklik ve genişlik) bir küçülme yaşanırken; derinliğinde ise bir değişiklik olmamaktadır. Görüntü boyutu W , Filtre boyutu F , Adım

uzaklığı S olmak üzere aşağıdaki denklem aracılığıyla görüntünün yeni boyutu hesaplanacaktır.

$$\text{Yeni Görüntü Boyutu} = \frac{W-F}{S} + 1 \quad (3.2)$$

3.2.3. Tam bağlı katman

Sinir ağlarında bir katmandaki bütün nöronların bir sonraki katmanın bütün nöronlarına bağlı olduğu bir katman bulunmaktadır. Bütün giriş ve çıkışların birbirine bağlı olduğu bu katmanda her bir nöronunda bir ağırlığı bulunmaktadır. Bu katmana tam bağlı katman denilmektedir. Sınıflandırma ya da örüntüyü tanıma işlemi öncesi son katman burasıdır. Katmanın sinir ağlarındaki karşılığı olarak sınıflandırma hatası hesaplanarak ağıın tekrar eğitim sürecine de bu katmandan başlanmaktadır. Tam bağlı katman sonrasında genellikle sinir ağlarında çok sınıflı problemleri çözmek için kullanılan en temel ve basit sınıflandırıcı olarak softmax fonksiyonundan yararlanılmaktadır. Sınıf sayısı kadar çıkış değeri üretilir ve elde edilen değer ise örüntü tahminidir. Aşağıda denklemleri verilmiş olan softmax fonksiyonundan yararlanarak her bir sınıf için tahmin değerleri üretilir. Bu tahmin değerleri de ağıımızı tekrar eğitmek için kullanılacak olan kayıp fonksiyonunda çapraz bilgi kullanılır. Kullanılacak olan kayıp fonksiyonu aracılığıyla da parametreler güncellenerek daha başarılı bir sınıflandırma yapılması amaçlanır.

$$\text{softmax}(x)_i = \frac{e^{x_i}}{\sum_j e^{x_j}} \quad (3.3)$$

3.2.4. Aktivasyon fonksiyonları

Gerçek dünya problemlerinin doğrusal olmamasından dolayı; bir problemi tanımlamak ya da çözmek için de doğrusal olmayan bir yaklaşımla probleme yaklaşmak gerekmektedir. Bu yaklaşım tarzı ile probleme dayalı bir çözüm örneği sunulmaktadır. Böylece daha sağlıklı sonuçların alınması hedeflenip; sınıflandırma problemlerinde tahmin kabiliyetinin artırılması hedeflenmektedir. Derin öğrenme aktivasyon fonksiyonları sayesinde doğal olmayan verilere daha iyi yaklaşımlar sunmayı amaçlamaktadır [48, 49]. Bir sinir ağıının eğitiminde, niteliklerin ağırlıkla çarpılmasından ve yanlılıkla toplanmasından sonra bir aktivasyon işlemi yapılır. Böylece aktivasyon fonksiyonu sonucunda elde edilen değer, bir

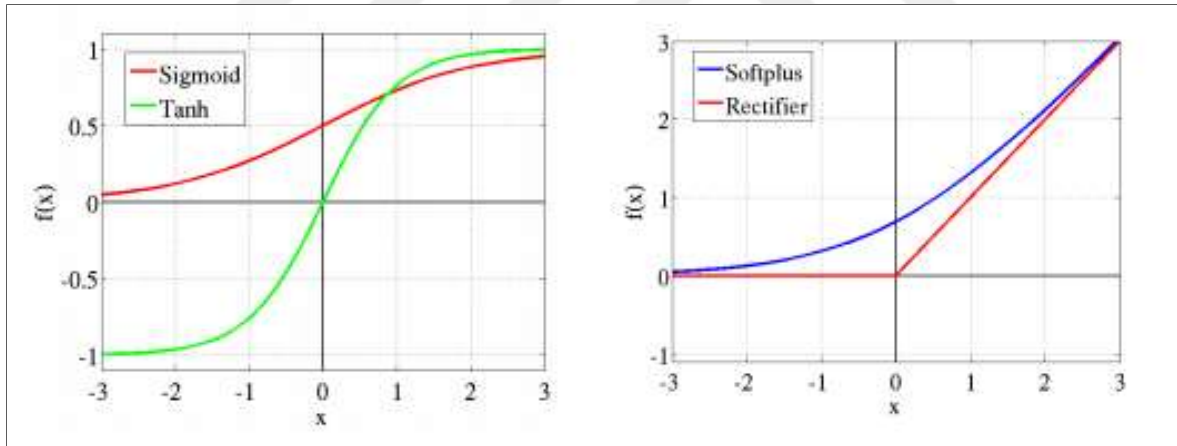
sonraki katmana çıkış değeri olarak aktarılır. Literatürde sıklıkla tercih edilen doğrusal olmayan aktivasyon fonksiyonlarından bazıları Sigmoid, Tanh, SoftPlus, ReLU'dur [50]. Bu aktivasyon fonksiyonlarına ait denklemler ile veriyi ele alış biçimlerinin grafiksel gösterimleri aşağıda verilmektedir.

$$f_{\text{Sigmoid}}(x) = \frac{1}{1 + e^{-x}} \quad (3.4)$$

$$f_{\text{Tanh}}(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (3.5)$$

$$f_{\text{SoftPlus}}(x) = \log(1 + e^x) \quad (3.6)$$

$$f_{\text{ReLU}}(x) = \max(0, x) \quad (3.7)$$

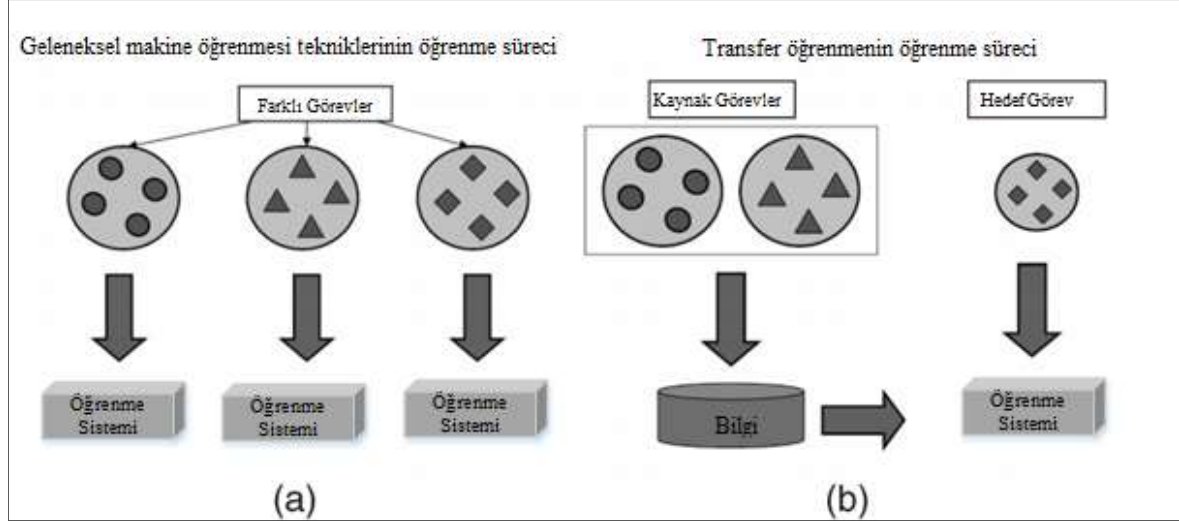


Şekil 3.3. Aktivasyon fonksiyonları [50]

3.3. Transfer Öğrenme

Derin öğrenmede yüksek öznitelikleri yakalamak için çok sayıda derin katmandan oluşan doğrusal olmayan bir yapı tercih edilir. Özellikle derin öğrenmenin popülerliğini kazandığı bilgisayarlı görü yarışmalarıyla beraber bu alana olan ilgiler artmıştır. Bu yarışmalarda kullanılan veri kümeleri çok sayıda kategori ve milyonlarca görüntüden oluşmaktadır [51]. Yarışmada nesneleri sınıflandırma ve tespit etme gibi görevler bulunmaktadır. 2010 yılından itibaren gerçekleştirilen en iyi sonucu tahmin etmeye odaklı yarışmada hem

kategori sayısının fazla olması hem de örnek sayısının fazla olmasından dolayı yarışmacıların rekabeti için çok cezbedicidir. Bu yarışmalarda kullanılan bazı sinir ağı modelleri üstün ayırt edici özellikleriyle günümüzde birçok çalışmada transfer öğrenme amacıyla kullanılmaktadır.



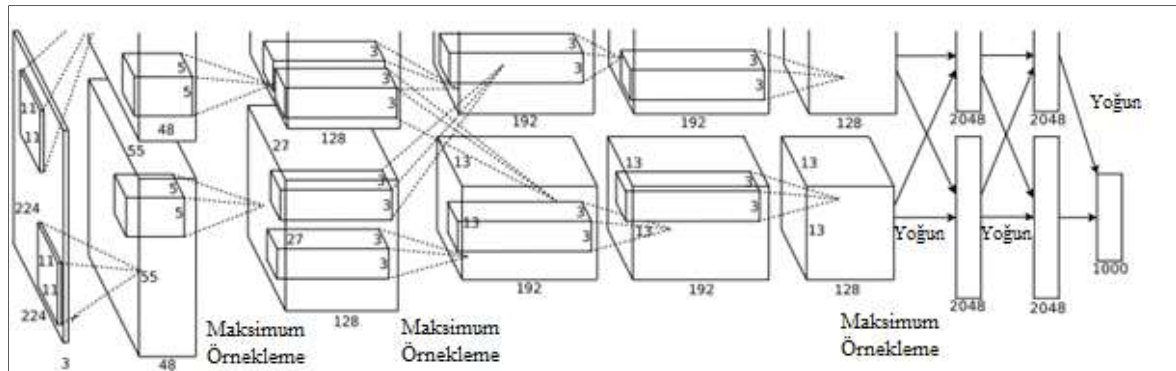
Şekil 3.4. Transfer öğrenme ile makine öğrenmesi farkı [52]

Transfer öğrenme, Şekil 3.4'te görüldüğü gibi önceden elde edilmiş bilginin başka bir probleme aktarılıp uygulanmasıdır. İnsanların doğal yaşamlarından örnek verecek olursak; çok iyi bir müzik aleti çalan birinin bilgilerini başka birine aktarması ve sonra o kişinin de müzik aletini iyi bir şekilde çalma becerilerini kazanmasıdır. Transfer öğrenmenin motivasyonu, gündelik hayatta olduğu gibi daha önceden öğrenilmiş, tecrübe edilmiş bilginin aktarılması ve aktarılan bu bilgiyle hem daha iyi hem de daha hızlı çözümler elde edilmesidir [52].

Transfer öğrenme, derin sinir ağlarında ayırt ediciliği yüksek öznelilikleri elde etmede oldukça başarılıdır [53]. Transfer öğrenme, bir alanda başarı ile uygulanan bir ağ modelinin benzer başka bir alandaki probleme bu işlemlerin uygulanıp transfer edilmesi sürecinden oluşmaktadır. Bu alanda yapılan çalışmalarda modelin uygulanabilirliği modelin daha önce uygulandığı ağlardaki öğrenme kabiliyetinden gelmektedir. Çünkü transfer öğrenmede kullanılan ağ modelleri, çok sayıda kategoriye sahip ve milyonlarca görüntünün bulunduğu, hem nitelik hem de nicelik anlamında ağ eğitimi için yeterli bilgi taşıyan veri kümelerine uygulanmaktadır. Transfer öğrenme süreciyle bu başarıyla

uygulanmış modellerden alınan parametreler yeni bir veri üzerinde daha anlamlı sonuçlar elde etmemizi sağlar. Özellikle transfer öğrenmenin uygulanacağı hedef verinin sınırlı olduğu durumlarda transfer öğrenme vazgeçilmez bir araç olacaktır.

AlexNet [45], ILSVRC 2012 yarışmasında başarısını kanıtlamış bir ağ modelidir. Ağ modelinin eğitim sürecinde, 1000 farklı sınıf içerisinde 1,2 milyon görüntü kullanılmıştır. Önerilen ağ modelinde her bir katmanın önemli olduğu belirtilmektedir. Ağda 1 evrişim katmanı kaldırıldığında ağın performansını negatif etkilediğini belirtmişlerdir. Bu sebeple derinliğin önemi özellikle vurgulanmıştır.



Şekil 3.5. AlexNet [45]

VGG16 [54] ve VGG19 [54] ağ modelleri çok başarılı ve rekabetçiliği yüksek derin sinir ağı modelleridir. Bu modeller görüntü verilerinde ağı eğitimi sürecinde sonuca daha hızlı yakınsamasını ve sonuçları beklenenden daha fazla arttırmayı amaçlamaktadır. Aynı zamanda çok karmaşık bir yapısı olup, milyonlarca parametre ile ağ modeli eğitimi sürecini gerçekleştirmektedir. Ağ modelinde transfer öğrenmenin uygulanması için giriş verisinin orijinalinde kullanılmakta olan 224×224 boyutlu görüntülere sahip olması gerekmektedir.

ResNet [55] ağ modeli de literatürde sıklıkla tercih edilen ve başarıyla uygulanan öğrenme ağ modellerinden biridir. ResNet geliştiricileri ağ modellerinin VGG'ye göre daha derin ancak daha az karmaşıklığa sahip olduğunu belirtmektedirler. ILSVRC 2015 yarışmalarında ağ modeli düşük hata oranıyla ilk sırayı almıştır. ResNet ağ modeli VGG-19 ağ modeline göre daha derin bir ağ yapısına sahiptir.

DenseNet [56], ağ modeli yine literatürde transfer öğrenme amacıyla sıklıkla kullanılan ağ

modellerinden biridir. DenseNet pek çok açıdan ilgi uyandırıcı avantajlara sahiptir. Bu avantajlar ise, eğimin yok olma problemini azaltır, özellik yayılımını güçlendirir, özelliğin yeniden kullanımını teşvik eder ve parametre sayısını önemli ölçüde azaltır [56]. Ağ modeli ImageNet veri kümesinde ResNet'e oranla daha iyi sonuçlar elde etmiştir.

MobileNet [57], hafif derin sinir ağları oluşturmak için derinlemesine ayrılabilir kıvrımlar kullanan aerodinamik bir mimariye dayanan bir ağ modelidir. Az sayıda parametre kullanım avantajıyla hem daha hızlı hem de daha küçük bir ağ modelidir. MobileNet nesne algılama, yüz nitelikleri ve büyük ölçekli coğrafi konumlandırma gibi uygulama alanlarında başarısını kanıtlamış bir ağ modelidir.





4. DERİN METRİK ÖĞRENME

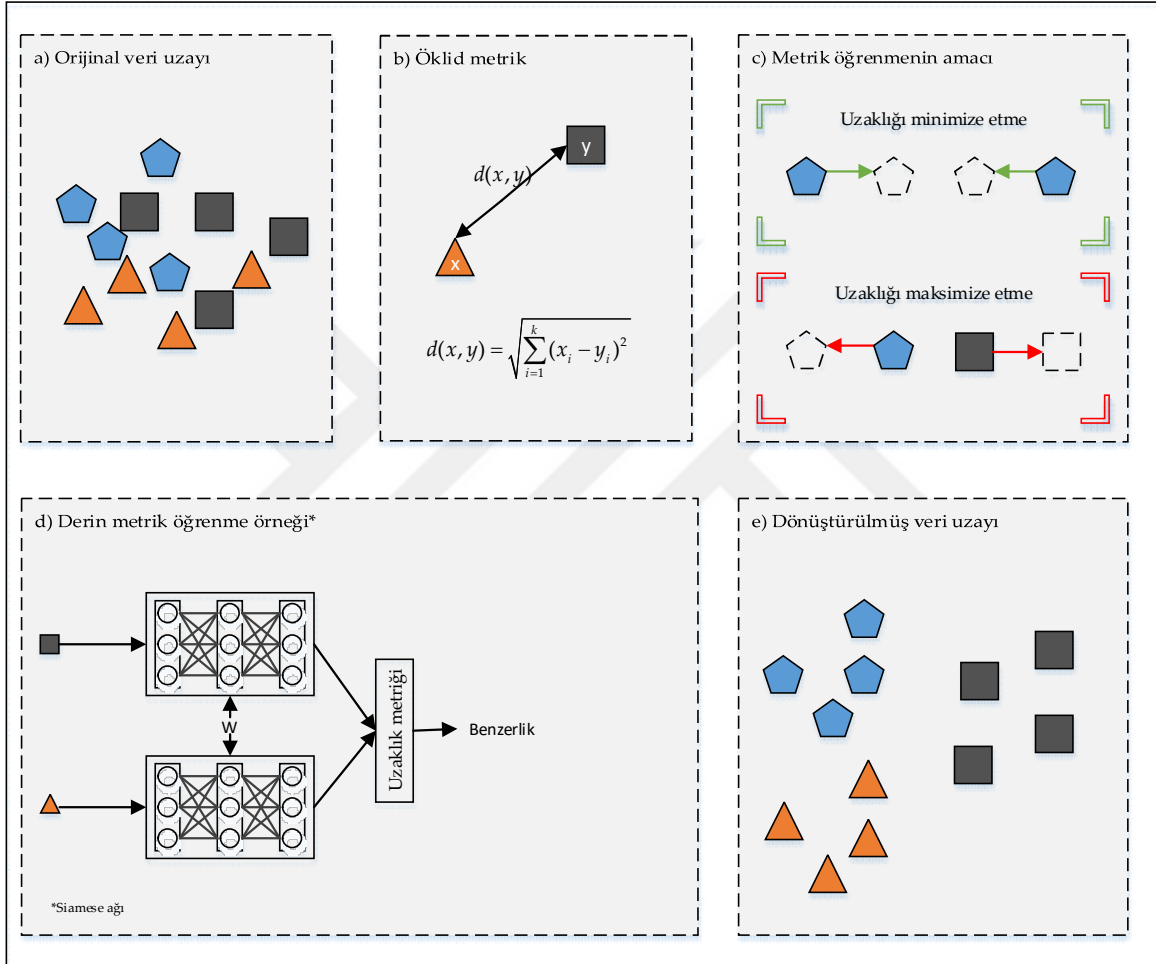
Literatürde sınıflandırma için kullanılan temel benzerlik metrikleri Öklid, Mahalanobis [18], Matusita [58] Bhattacharyya [59] ve Kullback-Leibler [60] metrikleridir. Her ne kadar bu metrikler literatürde sıklıkla kullanılsa da veriler üzerinde doğru tahminler yapabilme yetenekleri sınırlıdır. Literatürdeki temel metrik öğrenme yaklaşımları doğrudan bir metrik kullanmak yerine Mahalanobis metriğine dayalı olarak verinin yeni bir temsilini yakalamaya odaklanmaktadır. Bu yeni temsilde dönüştürülmüş veri uzayında veriyi daha iyi temsil eden yüksek ayırım gücüne sahip özneliklerin yakalanması hedeflenmektedir.

Mevcut metrik öğrenme yaklaşımları çekirdek fonksiyonu kullanmaksızın genel itibariyle verinin doğrusal dönüşümü üzerine kuruludur. Bu sebeple bu yaklaşımların verinin doğrusal olmayan bilgisini ortaya çıkarmada pek başarılı oldukları söylenemez [61]. Bundan dolayı bu yaklaşımlarda verinin daha iyi temsili için beklenen çıktılar elde edilemeyebilir. Mevcut yöntemler çekirdek tabanlı yaklaşımlarla bu sorunların üstesinden gelmeye çalışsa da ölçekleme gibi bazı problemlerden dolayı problemin çözümü tam olarak beklendiği gibi gerçekleşmeyebilir [62]. Geleneksel metrik öğrenmeye karşılık derin öğrenme yöntemleri bu tür eksikliklerin üstesinden gelmekte, aktivasyon fonksiyonları ve ağı yapısı nedeniyle oldukça başarılıdır.

Mevcut derin öğrenme yaklaşımlarının çoğu verinin yeni bir temsili uzayda uzaklık metriğine odaklanmaktan ziyade derin mimari alt yapısına dayanmaktadır. Ancak son birkaç yılda uzaklık tabanlı yaklaşımların derin öğrenme içerisindeki popülerliği artmıştır [63-67]. Örnekler arasındaki benzerlik yaklaşımını, uzaklık metriğine dayanarak çözmeye çalışan derin metrik öğrenme kavramı ortaya konulmuştur [9]. Genel derin metrik öğrenme yaklaşımında aynı anda birden fazla giriş verisi ağı sunulmaktadır. Derin metrik öğrenmede ağı sunulan bu verilerden benzer olanları arasındaki uzaklık azaltılmaya çalışılırken; farklı görüntüler arasındaki uzaklık ise artırılmaya çalışılmaktadır [68, 69].

Şekil 'de derin metrik öğrenme süreci tek bir görselle ifade edilmeye çalışılmıştır. Bu görselden de görüleceği üzere veriler sınıflandırmaya tabi tutulmadan önce mevcut uzaylarında karmaşık bir yapıdadırlar. Doğrudan basit bir Öklid metriği ile sınıflandırma yapmak pek başarılı sonuçlar doğurmayabilir. Ancak burada uzaklık metriğini istediğimiz

amaç doğrultusunda kullanıp metrik öğrenmeye odaklanılacaktır. Metrik aynı sınıftan örnekleri yakınlaştırma ve farklı sınıftan örnekleri ise birbirinden uzaklaştırmaya dayanmaktadır. Siamese ağ örneği bu amaçla aynı anda iki görüntüyü temel alıp veriyi en iyi temsil eden ağırlıkları bulmaya odaklanmaktadır. Böylece verinin yeni dönüşüm uzayında daha iyi bir temsili elde edilmektedir.

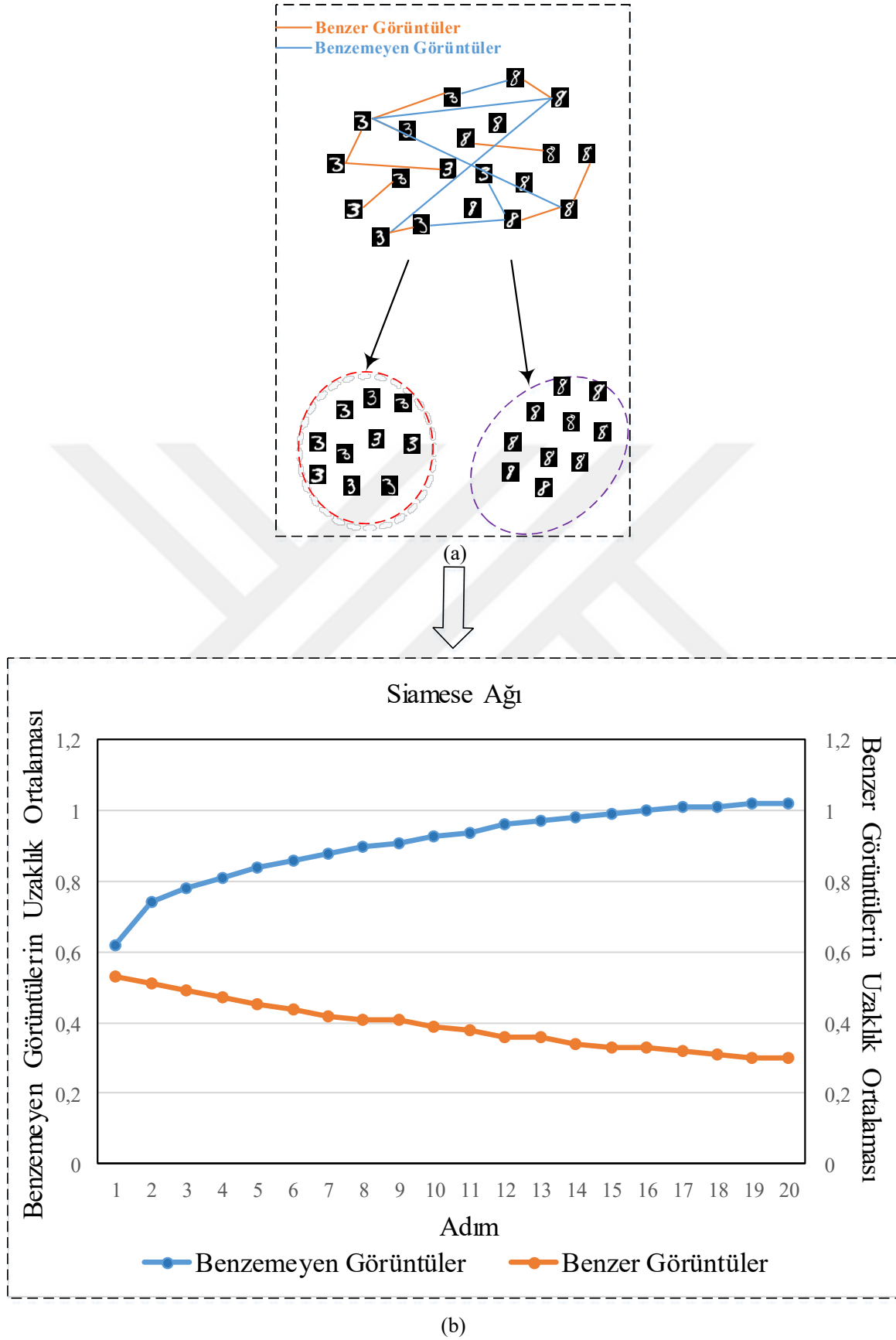


Şekil 4.1. Derin metrik öğrenme

Şekil 4.1.'deki senaryoyu gerçekleştirmek üzere MNIST veri kümesi [70] kullanılarak bu senaryonun doğruluğu test edilmeye çalışılmıştır. MNIST veri kümesi el ile yazılmış 0-9 arasındaki rakamlara ait görüntülerden oluşmaktadır. Şekil 4.2'de bu veri kümesindeki 3 ve 8 rakamlarını kullanarak verinin Siamese ağı eğitimi boyunca kontrastif kayıp [71] fonksiyonu ile nasıl hareket ettiği izlenmiştir. Şekil 4.2'de görüldüğü üzere ağ eğitimi süresince benzer görüntüler arasındaki uzaklık azalırken; farklı görüntüler arasındaki uzaklığın ise arttığı görülmüştür. Uzaklık değerleri görüntüler arasındaki uzaklıkların

ortalamasını ifade etmektedir. Şekil 4.2 (a)'da temsili olarak 3 ve 8 rakamlarının dağılımları ve derin metrik öğrenme yaklaşımıyla hedeflenen durum gösterilmektedir. Burada beklenti ağı eğitimi süresince 3 ve 8 rakamlarının kendi sınıfları içindeki uzaklıklarının azalması olurken; aynı zamanda 3 ve 8 rakamlarının birbirinden uzaklaşarak yeni bir temsilin elde edilmesidir. Şekil 4.2 (b)'de hedeflenen durumun adım adım hedefe ne kadar doğru bir şekilde gittiği açıkça görülmektedir. Burada benzer sınıflara ait örneklerin (3 örneklerinin kendi içinde ve 8 örneklerinin kendi içinde) kendi sınıfları içerisinde ki uzaklığının azaldığı görülürken; farklı sınıflara ait olan 3 ve 8 örnekleri arasındaki uzaklığın ise adım adım arttığı açık bir şekilde görülmektedir. Bu durum bize derin metrik öğrenmenin gücünü uygulamalı olarak göstermektedir.

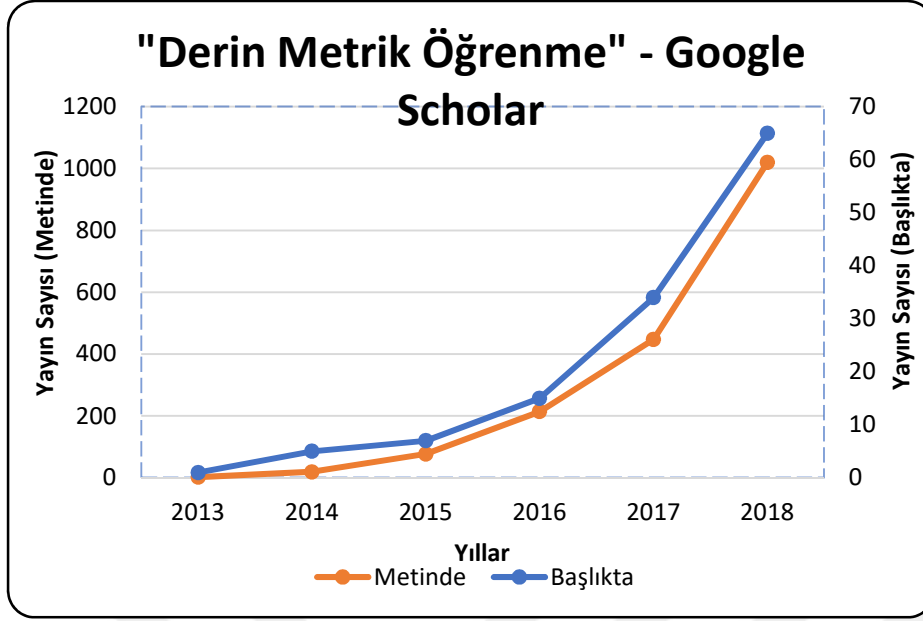




4.1. Derin Metrik Öğrenme Uygulama Alanları

Derin metrik öğrenme, derin öğrenmede var olan doğrusal olmayan derin mimari kurgusu ile metrik öğrenmede bulunan benzerliğe dayalı yaklaşımı bir araya getirip; ham veriden bir uzaklık ilişkisi kurmayı amaç edinmektedir. Derin metrik öğrenme (DML) çok geniş bir alanda problemlere çözümler sunmaktadır. Detayları aşağıda verilecek olup, DML kişiyi yeniden tanımlama [72-74] video anlama [75, 76] 3 boyutlu modelleme (3B) [64, 77], medikal veriler [78] yüz doğrulama ve tanıma [49, 68, 79] imza doğrulama [80], metin anlama [81, 82], ses sinyalleri [83, 84] vb. birçok alanda başarılı bir şekilde uygulanmaktadır. Derin metrik öğrenme üzere yapılan çalışma alanlarından da görüleceği üzere son yıllarda oldukça ilgi çeken bir çalışma alanı olarak karşımıza çıkmaktadır. Şekil 4.3' te 2013-2018 yılları arasında bu alana olan ilgi görülmektedir. Bu şekilde, akademik yayınları tarayan bir arama motoru olan Google Scholar'da anahtar kelime olarak “derin metrik öğrenme” kavramının yayın başlıkları ve yayın metinleri içerisinde görünme sıklığı görülmektedir. Buradan da görüleceği üzere derin metrik öğrenmenin üssel şekilde artan bir çalışma konusu olduğu görülmektedir.

Video görüntülerini anlamlandırmada son kullanıcıyı ilgilendiren video etiketleme, video tavsiyesi, video arama gibi çeşitli makine öğrenmesi için çeşitli uygulama alanları bulunmaktadır. Bu tür problemlerin metrik öğrenme yaklaşımlarıyla çözümleri mümkündür. Lee ve arkadaşları [75], derin metrik öğrenme yaklaşımlarından Triplet öğrenme yaklaşımından yararlanarak videolardan hem ses ve hem de görüntü özniteliklerini kullanarak, son kullanıcıya video tavsiyesi ve video bilgisi sunmuşlardır. Huang ve arkadaşları [76], video gözetiminde insan lokalizasyonunu harekete geçirmek için derinlemesine çok kanallı artık (residual) ağ tabanlı bir metrik öğrenmeye odaklanmışlardır. Önerilen metrik öğrenme tabanlı yaklaşımın mevcut yöntemlerle karşılaştırıldığında çok daha iyi sonuçlar elde edildiği gösterilmiştir. Hu ve arkadaşları [85], ön tanımlı metriklerin görsel nesnelerde önemli değişimlerin etkisiyle yetersiz olduğunu ifade etmektedirler. Bu sebeple görsel izleme için ön tanımlı bir metrik yerine benzer sınıftaki pozitif örnekler arasındaki uzaklığı minimize eden ve farklı sınıflardaki negatif örnekler arasındaki uzaklığı maksimize eden bir derin metrik öğrenmeyi amaçlamışlardır. Yapılan çalışma sonuçlarından da görüleceği üzere derin metrik öğrenme bu alanda üstün başarılar sunmuştur.



Şekil 4.3. Literatürde derin metrik öğrenme ile ilgili yayın sayıları

Makine öğrenmesinde görsel görüntülerde bir başka problem kişiyi yeniden tanımlama problemidir. Bu çalışma alanında kişi videolarda kişi görüntüsünü tekrar tanıma ya da kişiyi izleme gibi görevler bulunmaktadır. Evrişimli sinir ağları (CNN) tabanlı yaklaşımların son yıllarda görüntü verilerinde elde ettiği başarılarla kişi yeniden tanımlama problemlerinde de sıklıkla tercih edilmektedir [86]. Kişi yeniden tanımlama görevlerinde aynı kişiden farklı şekillerde alınan görüntülerden oluşmaktadır. Problemin zorluğu olarak kişinin hareketleri ve farklı duruşları bir problem olarak ortaya çıkmaktadır. Bu tür problemlere özgü uzaklık metriği öğrenmesi mümkündür [72, 87]. Bu alanda literatürde sıklıkla karşılaştırma amacıyla kullanılan bazı bilinen veri kümeleri bulunmaktadır (CUHK01, CUHK03, Market-1501, MARS ve VIPER). DML bu görevler için uçtan uca öğrenme yaklaşımı kolaylığı sağlamaktadır [74]. Ahmed ve arkadaşlarının yaptıkları çalışmada [88], öncelikle her bir giriş görüntüsün orta seviyeli özelliklerini kullanarak aynı anda ağa sunulan iki giriş görüntüsü için yerel ilişkiyi tespit eden çapraz girişli komşuluk farklarını hesaplayan bir derin mimari katmanı bulunmaktadır. Bu katmanda elde edilen yüksek seviyeli özellikler sonraki katmanda tekrar ele alınıp bir en iyi (optimal) uzaklık metriği öğrenme hedeflenmektedir. Ding ve arkadaşları [89] ise, Triplet kayıp fonksiyonunu kullanarak iki farklı görüntü örneği arasındaki uzaklığı maksimuma çıkarmayı amaçlamışlardır. Triplet giriş verilerini oluşturmada olası tüm üçlülere kullanmak yerine etkili bir üçlü oluşturma şeması ve optimize edilmiş GDA (Gradient descent algorithm) kullanmışlardır.

Medikal veriler üzerinde yapılan çalışmalar derin metrik öğrenmede ilgi duyulan konulardan bir diğeridir. Medikal görüntü verilerinde veriyi sınıflandırma, veriyi tespit etme, veri bölütleme ve veri kaydetme gibi bazı temel konular bulunmaktadır [90]. DML'in benzerlik yaklaşımıyla bu tür konular güncel uygulama alanları olarak karşımıza çıkmaktadır. Eğitim süreci sonunda öğrenilen derin bir metrik aracılığıyla verinin daha yüksek seviyeli bir temsili mümkün olmaktadır. Annarumma ve arkadaşları [78], geleneksel DML yaklaşımlarıyla benzer bir mimariyi kullanan ML2 kayıp fonksiyonundan yararlanarak gömülü uzaydaki verinin global bilgisini yakalamaya çalışmışlardır. Radyolojik benzerlik üzerine kurulu olan yaklaşım kümeleme ve görüntü elde edinim görevleri için çok etiketli röntgen filmlerinde çok başarılı sonuçlar ortaya koymuştur.

DML'in 3B şekillerde elde edinim için etkin bir ayırıcı güç olarak kullanılması yönünde yapılan birçok çalışma bulunmaktadır [64, 77, 91-94]. 3B şekillerde elde edinimde hem eskiz hem de 3B şekil yapısı metrik kayıp fonksiyonunda doğru bir metrik elde etmek için kullanılmaktadır [91, 93]. Bir ağ modelinde birden fazla kayıp fonksiyonu ile daha iyi (optimal) çözümler hedeflenmesi mümkündür. Bu amaçla yapılan çalışmalarda Dai ve arkadaşları [64], hem korelasyon kaybı hem de ayırım kaybını bir arada kullanarak birden fazla metrikle daha iyi bir temsile yakınsama yapılmasını amaçlamışlardır. [64]'deki çalışmadan farklı olarak [92]'de metrik kayıp fonksiyonu eğitim boyunca ayrıca gizli katmanda da kullanılmıştır. He ve arkadaşları [94], daha dayanıklı ve ayırıcı öznelilikler elde etmek için Triplet kayıp fonksiyonu ve merkez kayıp fonksiyonunu 3B görüntü elde edinim görevleri için kullanmışlardır. Lim ve arkadaşları [77], 3B görüntü sınıflarını tespit etmek amacıyla benzer ve farklı görüntülerin uzaklıklarını Triplet kayıp fonksiyonuyla karşılaştırma yapmak amacıyla kullanmışlardır.

Bilgisayarlı görünümün en önemli ilgi alanlarından biri olan yüz tanıma ve doğrulama problemleri için DML son yıllarda çok rekabetçi sonuçlar ortaya koymaktadır [49, 68, 79, 83, 95]. Hu ve arkadaşları [49], yüz doğrulama için sinir ağlarının rekabetçi gücünü kullanarak yüz çiftlerinin hiyerarşik doğrusal olmayan bilgisini kullanan yeni bir ayırım gücü yüksek DML modeli önermişlerdir. Yine ilgi çekici bir konu olan yüz görüntülerinden akrabalık ilişkisi çıkarma konusunda iki yüz görüntüsünü ele alıp inceleyerek DML tabanlı modelleme çözümleri sunulmaktadır [95]. Schroff ve arkadaşları [79] tarafından geliştirilen FaceNet modelinde verilere yüz benzerliği ölçümüne karşılık gelen Öklid uzayında yeni bir temsil yeteneği kazandırılmıştır. Çalışmada güncellenen bir

Triplet öğrenme yaklaşımı tercih edilmiştir. Önerilen model yüz tanıma, yüz doğrulama ve yüz kümeleme gibi problemlere oldukça başarılı çözümler sunmuştur. Bununla birlikte DML ile yüz ifadesini tanıma [96] ve yüz görüntüsünden yaş doğrulama [97] gibi çalışmalar bulunmaktadır.

Her ne kadar DML alanında ya da derin öğrenme alanında görüntü verileri popüler olmasına karşın; literatürde yoğun bir şekilde çalışılan bir diğer konu metin anlama ve bilgi elde edinimi konularıdır. Mueller ve arkadaşları [81], Siamese ağ modelinde LSTM mimarisini kullanarak cümleler arasındaki anlamsal benzerliği tespit etmeye çalışmışlardır. Benajiba ve arkadaşları [82], yine anlamsal örüntülerin benzerliklerini öğrenmek amacıyla benzer bir ağ modelinden yararlanmışlardır. Mueller'in yönteminden farklı olarak ağ modelini eğitmek için SQL yapısal uzaklığından ortalama kare hatasını (MSE) kullanarak bir regresyon fonksiyonu önermişlerdir. Zhu ve arkadaşları [98], cümle temsili öğrenmede bir fark oluşturmak için; cümle temsiliinde temel ve destekleyici bileşenleri kullanarak bağımlılık tabanlı Siamese-LSTM ağ modelini önermişlerdir. Ein-Dor ve arkadaşları [99], cümleler arasında tematik bir benzerlik ilişkisi kurmaya odaklanmışlardır. Bunun için öncelikle Wikipedi makalelerinden zayıf denetimli üçlü cümleler üretmişlerdir. Daha sonra, Wikipedia cümlelerini yüksek kalitede cümle yerleştirmeleriyle kümelemek için Triplet ağ modelinden yararlanmışlardır.

Görüntü ve metin alanında olduğu gibi DML ses sinyallerinden anlamlı örüntüler elde etmede rekabetçi sonuçlar elde edilmektedir [83]. Narayanaswamy ve arkadaşları [84], konuşmacının kimliğine göre ses sinyallerini ayırmak için Triplet ve Quadruplet ağlarından yararlanmışlardır. Yazarlar yaptıkları çalışmada farklı örnekleme stratejileri kullanmışlardır. Bilgisayarlı görü problemlerinde genellikle çok iyi sonuçlar veren yarı negatif madencilik örneklemenin kendi problemlerinde ancak sabit marjın ve Triplet ağ modelinde faydalı olduklarını öne sürmüşlerdir. Wang ve arkadaşları [100], prototipik ağ kaybını [101] kullanarak; konuşmacı doğrulama ve konuşmacı tanımlama görevleri için Triplet kayıp fonksiyonu ile kıyaslamalı bazı deneyler yapmışlardır. Yazarlar konuşmacı tanıma için iki farklı veri kümesinde çok başarılı sonuçlar elde etmişlerdir. Yazarlara göre, prototipik ağ kaybının sonuçları, Triplet kaybın sonuçlarından daha iyi olup aynı zamanda daha hızlı bir eğitim süresine sahiptir.

Yukarıda farklı alanlarda derin metrik öğrenme üzerine kategorize edilmiş çalışmalar

olmasına rağmen, çok daha farklı alanlarda yapılmış çalışmalar bulmak mümkündür. Derin metrik öğrenmenin uygulandığı diğer ilgi çekici konular ise; müzik benzerliği [102], kalabalık regresyonu [103], benzer bölge araştırması [104], hacimsel görüntü tanıma [105], örnek bölütleme [106], kenar tespiti [107], renklere duyarlı keskinleştirmedir [108], Bu çalışmaları incelediğimizde derin metrik öğrenmenin, farklı alanlardaki yüksek performansı nedeniyle literatüre çok önemli katkılar sağladığını görmek mümkündür.

Çizelge 4.1’de ki karşılaştırmalı sonuçlarda yer alan metriklerden R (recall) bilgi elde edinimi görevlerinde başarılı şekilde tespit edilmiş örneklerin oranını ifade etmektedir. NMI (Normalized Mutual Information) sınıf etiketlerinin entropisini ve küme etiketlerinin entropisini kullanarak kümelemenin kalitesini belirler. F1 hem yanlış pozitif hem de yanlış negatifleri dikkate alan bir metriktir. 3 boyutlu şekil elde etme performanslarını ölçmek için kullanılan metrikler olarak ilgili eşleştirmelerin oranı FT (First Tier) ve ikinci bağlayıcı ST (Second Tier), en yakın komşu (NN) , F1 ile aynı ölçüm bilgisi olan E (Emeasure), indirimli kümülatif kazanç (DCG) ve ortalama hassasiyettir (mAP). Anlamsal benzerlik için kullanılan metrikler Pearson korelasyon katsayısı (r), iki değişkenin monotonik ilişkisini değerlendiren ρ ve ortalama kare hatası (MSE)’dir. Konuşma doğrulama için kullanılan metrikler ise, eşit hata oranı (EER) ve minimum karar maliyet fonksiyonu (MDCF)’dir.

Çizelge 4.1. Derin metrik öğrenme problemlerinin karşılaştırmalı sonuçları

Veri Kümesi	Referans	Görev ve Karşılaştırmalı Sonuçlar					Değerlendirme Protokolü	CNN LSTM	Yıl
		Görüntü Kümeleme (%)		Görüntü Elde Ediminin Recall@R (%)					
		NMI	F1	R=1	R=2	R=4	R=8		
CUB-200-2011 [109]	[73]	56.2	22.7	46.5	58.1	69.8	80.2	11788 görüntüyle 200 sınıf	CNN 2016
	[110]	60.3	27.2	50.9	63.3	74.2	83.2	İlk 100 sınıf eğitim için	CNN 2016
	[111]	61.1	29.4	54.7	66.3	76.0	83.9	(5864 görüntü)	CNN 2017
	[112]	59.2	-	48.1	61.4	71.8	81.9	Kalan 100 sınıf test için	CNN 2017
	[113]	-	-	57.1	68.8	78.7	86.5	(5,924 görüntü)	CNN 2018
CAR-196 [114]		NMI	F1	R=1	R=2	R=4	R=8		
	[73]	55.1	21.5	48.3	61.1	71.8	81.1	16,185 görüntüyle 198 sınıf	CNN 2016
	[110]	63.9	33.5	71.1	79.7	86.4	91.6	İlk 98 sınıf eğitim için	CNN 2016
	[111]	63.2	32.2	71.4	81.4	87.5	92.1	(8,054 görüntü)	CNN 2017
	[112]	59.0	-	58.1	70.6	80.2	87.8	Kalan 98 sınıf test için	CNN 2017
Online Products [73]	[113]	-	-	81.4	88.0	92.7	95.7	(8,131 görüntü)	CNN 2018
		NMI	F1	R=1	R=10	R=100	R=1000		
	[73]	87.4	24.7	63.0	80.5	91.7	97.5	120053 görüntüyle 22634 ürün ilk	CNN 2016
	[110]	88.1	28.1	67.7	83.7	92.9	97.8	11318 ürün kategorisi eğitim için	CNN 2016
	[111]	88.6	29.9	70.9	85.0	93.5	98.0	(59,551 görüntü)	CNN 2017
	[112]	89.4	-	67.0	83.6	93.2	-	Kalan 11316 ürün kategorisi test için	CNN 2017
	[113]	-	-	74.8	88.3	94.8	98.4	(60,502 görüntü)	CNN 2018

Çizelge 4.1. (devam) Derin metrik öğrenme problemlerinin karşılaştırmalı sonuçları

Veri Kümesi	Referans	Görev ve Karşılaştırmalı Sonuçlar		Değerlendirme Protokolü	CNN LSTM	Yıl
		<i>Kişi Yeniden Tanımlama</i>				
		<u>R=1</u>	<u>R=5</u>			
Market-1501 [115]	[116]	59.47	80.73	1,501 kişi görüntüsü 750 kişi görüntüsü eğitim için 751 kişi görüntüsü test için	CNN	2016
	[72]	83.55	92.37	-	CNN	2018
	[87]	84.26	93.59	1,501 kişi görüntüsü 750 kişi görüntüsü eğitim için 751 kişi görüntüsü test için	CNN	2019
	[117]	88.20	-	1,501 kişi görüntüsü 750 kişi görüntüsü eğitim için 751 kişi görüntüsü test için	CNN	2019
		<u>R=1</u>	<u>R=5</u>			
CUHK03 [118]	[116]	65.77	92.85	1360 kişi görüntüsü 1160 kişi görüntüsü eğitim için 100 kişi görüntüsü test için	CNN	2016
	[72]	68.63	92.28	-	CNN	2018
	[87]	39.64	-	1367 kişi görüntüsü 767 kişi görüntüsü eğitim için 700 kişi görüntüsü test için	CNN	2019
	[117]	82.75	96.59	1360 kişi görüntüsü 1160 kişi görüntüsü eğitim için 100 kişi görüntüsü test için	CNN	2019

Çizelge 4.1. (devam) Derin metrik öğrenme problemlerinin karşılaştırmalı sonuçları

Veri Kümesi	Referans	Görev ve Karşılaştırmalı Sonuçlar							Değerlendirme Protokolü	CNN LSTM	Yıl
SHREC'13 [119]		<i>3B şekil elde edilini</i>									
		<u>NN</u>	<u>FT</u>	<u>ST</u>	<u>E</u>	<u>DCG</u>	<u>Map</u>				
	[64]	65.0	63.4	71.9	34.8	76.6	67.4	1258 şekil ve 7200 eskiz, 90 grup	CNN	2017	
	[92]	73.0	71.5	77.3	36.8	81.6	74.4	Her sınıftaki eskiz sayısı 80	CNN	2018	
	[94]	76.3	78.7	84.9	39.2	85.4	80.7	Her bir gruptaki 50 eskiz eğitim için 30 eskiz test için	CNN	2018	
SHREC'14 [119]		<u>NN</u>	<u>FT</u>	<u>ST</u>	<u>E</u>	<u>DCG</u>	<u>Map</u>				
	[64]	27.2	27.5	34.5	17.1	49.8	28.6	171 sınıf içinde gruplanmış 13680 eskiz ve 8987 3B model, Her sınıftaki eskiz sayısı 80	CNN	2017	
	[92]	40.3	32.9	39.4	20.1	54.4	33.6	Her bir gruptaki 50 eskiz eğitim için 30 eskiz test için	CNN	2018	
	[94]	58.5	45.5	53.9	27.5	66.6	47.7		CNN	2018	
		<i>Yüz Doğrulama</i>									
LFW [120]		Doğruluk									
	[49]	90.68							<i>10 kat:</i> Her bir katta 300 eşleştirilmiş çiftler ve 300 eşleştirilmemiş çiftler <i>Görüntü Kısıtlı</i>	-	2014
	[95]	94.50								-	2017
	[121]	93.27								-	2018
YTF [122]		Doğruluk									
	[49]	82.34							<i>10 kat:</i> Her bir katta 250 eşleştirilmiş çiftler ve 250 eşleştirilmemiş çiftler <i>Görüntü Kısıtlı</i>	-	2014
	[95]	82.50								-	2017

Çizelge 4.1. (devam) Derin metrik öğrenme problemlerinin karşılaştırmalı sonuçları

Veri Kümesi	Referans	Görev ve Karşılaştırmalı Sonuçlar				Değerlendirme Protokolü	CNN LSTM	Yıl
		<i>Anlamsal Metinsel Benzerlik</i>						
		<i>L</i>	<i>P</i>	<i>MSE</i>				
SICK [123]	[81]	0.88	0.83	0.22		9927 cümle çifti 5000 eğitim için 4927 test için	LSTM	2016
	[98]	0.83		0.34			LSTM	2018
	[99]	0.81	0.72	0.33			LSTM	2018
		<i>Konuşmacı Doğrulama</i>						
		<i>EER (%)</i>		<i>MDCF</i>				
NIST i-vector [124]	[11]	2.85		0.30		Her biri 5-i vektör ile kaydedilen 1306 konuşmacı Toplamda 9634 test i-vektör ve 12582004 deneme eğitim ve test kümesi rasgele bölünmüş Tüm I-vektörleri 600 boyutlu	-	2015
	[83]	2.69		0.27			-	2019
		<i>EER (%)</i>		<i>MDCF</i>				
VCTK [125]	[11]	12.26		-		İlk 90 konuşmacı eğitim, doğrulama ve test kümesine bölünmüştür. 18 konuşmacı görülmeyen küme olarak kullanılmıştır.	LSTM	2015
	[100]	10.77		-			LSTM	2019
		<i>EER (%)</i>		<i>MDCF</i>				
VoxCeleb2 [126]	[11]	15.92		-		101 konuşmacı içeren alt küme seçimi 71 konuşmacı eğitim ve doğrulama için, Kalan 30 konuşmacı görülmeyen küme olarak kullanılmıştır	LSTM	2015
	[100]	13.68		-			LSTM	2019

Son yıllarda dergilerde / konferanslarda yayınlanan çalışmaları göstermek için Çizelge 4.1'de oluşturulmuş ve farklı konu başlıkları için kıyaslama veri kümelerini kullanarak benzer bir değerlendirme protokolü kullanan çalışmalar özetlenmiştir. Çizelge 4.1' de sunulan karşılaştırmalı sonuçlar, derin metrik öğrenmenin birçok farklı disiplinde çok başarılı sonuçlar elde ettiğini göstermektedir.

4.2. Derin Metrik Öğrenmede Örnek Seçimi

Derin metrik öğrenme ağ modeli 3 temel yapıdan oluşmaktadır. DML ağa sunulacak olan giriş verisi örnekleri, derin ağ modelinin yapısı ve kayıp fonksiyonundan oluşmaktadır. Derin ağ modelinde kayıp fonksiyonunun önemi çok fazla olmasına rağmen, ağa sunulacak verilerin de bir o kadar önemi vardır. Çünkü bu ağ modellerinde ağa birden fazla örnek aynı anda sunulmaktadır. Sunulan örnek çeşitliliği ve fazlalığı ağın performansını pozitif yönden etkilerken, ağ eğitimi zamanı açısından negatif yönde de etkileyebilir. Ağa sunulacak her ikili ya da üçlü vb. örnek tiplerinin her birinin anlamlı olma durumu da yoktur. Bu sebeple ağın başarısını olumlu yönde etkileyecek örnek ikili ya da üçlülerinin sunulmasında fayda bulunmaktadır.

Siamese veya Triplet gibi ağlarda bir ağa sunulacak eğitim örneklerini rasgele bir şekilde seçmek en kolay yoldur. Siamese ağının gömülü uzayında yeni temsilini elde etmek için yapılan ilk çalışmalarda bilgilendirici örnek çiftlerine pek önem verilmemiştir [71, 127]. Ancak bu durumun ağın öğrenme sürecini yavaşlatacağı ve başarı oranlarını da negatif yönde etkileyebileceği bir gerçektir [128]. Seçilecek olan bilgi değeri yüksek örnekler ile hem ağın eğitim başarısını arttıracak hem de gereksiz yere fazla örnek kullanıp ağın eğitim zamanını arttırmanın önüne geçmek mümkündür. Bu tür problemleri çözmek ve ayırım gücü daha yüksek modeller önermek için sert negatif madencilik gibi yaklaşımlar kullanılmaktadır [128, 129].

Triplet ağında ağa 3 örnek aynı anda sunulur. Bu örnekler çapa örneği, çapa örneği ile aynı sınıftan olan pozitif örnek ve çapa örneğinden farklı sınıfta olan negatif örnektir. Kolay üçlü diye tabir ettiğimiz rasgele örneklerin bir kısmının ağ eğitiminde zayıf ayırım gücünden dolayı neredeyse hiç pozitif etkisi yoktur [130]. Bu etkisiz üçlü örnekler aynı zamanda hem zaman kaybı hem de kaynakların gereksiz kullanımına neden olur. Bu sebeple bu tür problemleri aşmak için; rasgele üçlü örnekler belirlemek yerine bilgi değeri

yüksek olan örnek üçlülere kullanmak ağın kararlılığını arttırmak için çok önemlidir [130, 131].

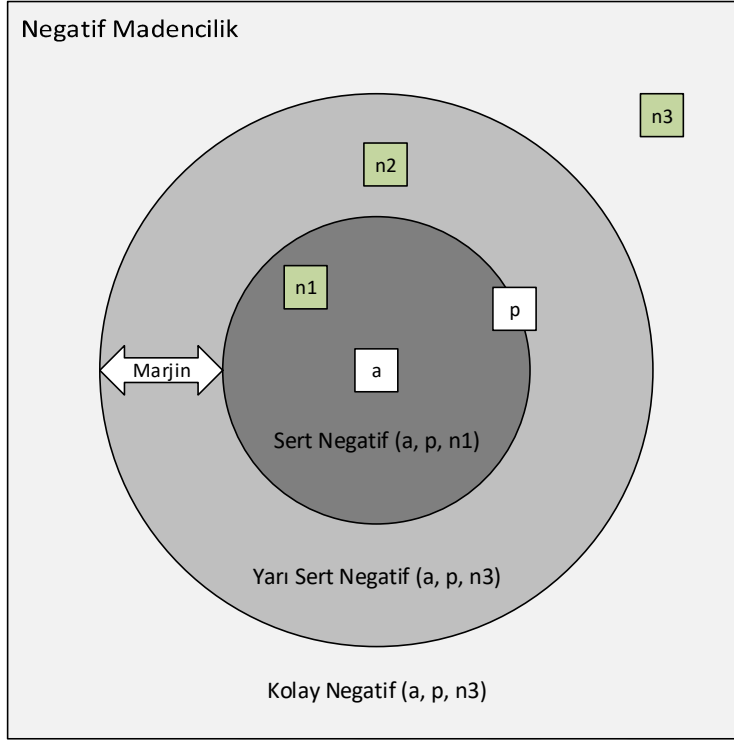
Sert negatif örnekler eğitim verisinde belirlenmiş olan FP (False Positive) örneklerine karşılık gelmektedir. Sert negatif madencilik yaklaşımında çapa (a) ile negatif örnek (n) arasındaki uzaklık, çapa ile pozitif örnek (p) arasındaki uzaklıktan daha azdır. Buna ait denklem şöyledir:

$$d(a, n) < d(a, p) \quad \text{Sert Negatif} \quad (4.1)$$

Yarı negatif madencilik yönteminde bir marjin aracılığıyla negatif örnekler tespit edilir [79]. Çapanın negatif örneğe olan uzaklığı, çapanın pozitif örneğe olan uzaklığından daha fazladır. Çapanın negatife olan uzaklığı, çapanın pozitif olana bir marjin değeri eklediğimiz değerinden ise daha az olmaktadır. Bu Triplet [11] ağ modeli yaklaşımında örnekler arasında yumuşak bir geçiş bulunmaktadır. Kolay negatif ise çapanın negatife olan uzaklığının çok yüksek olduğu durumu ifade etmektedir. Aşağıda bu örnek seçimlerine ait uzaklık denklemleri olmakla birlikte görsel anlamda Şekil 4.4'te bu durum daha açıklayıcı bir şekilde ifade edilmektedir.

$$d(a, p) < d(a, n) < d(a, p) + \text{marjin} \quad \text{Yarı Sert Negatif} \quad (4.2)$$

$$d(a, p) + \text{margin} < d(a, n) \quad \text{Kolay Negatif} \quad (4.3)$$



Şekil 4.4. Negatif madencilik

Negatif örneklerin çapaya çok yakın olması durumunda gradyan, yüksek varyansa ve düşük sinyal görüntü oranına sahip olur [132]. Bu durumun engellenmesi için DWS (Distance Weighted Sampling) yaklaşımıyla gürültülü örneklerden kaçınmak önerilmiştir [132]. Önerilen bu yaklaşım ile yarı sert negatif madencilikten daha geniş bir aralıkta örnek gruplarının kullanımı mümkün olmuştur. Literatürde negatif örnek madenciliği yerine, negatif sınıf madenciliğini [110] öneren yaklaşımlar da bulunmaktadır. Bu yaklaşımda Triplet ağındaki bir negatif örnek için, her bir sınıf örneği kullanılmıştır. Bunu gerçekleştirmek için GSS (Greedy Search Strategy) ile çoklu negatif örneklerin seçilmesine imkan tanınmıştır.

Literatürde negatif madencilik dışında pozitif madenciliği hedefleyen yaklaşımlar da bulunmaktadır. Shi ve arkadaşları [133], kişiyi yeniden tanımlama görevlerinde ağırlık sınırlandırmalı bir derin metrik öğrenmeye çalışmışlardır. Bu metriği öğrenmeye çalışırken sadece sınıflar arası ilişki değil aynı zamanda sınıflar içi ilişkinin de önemli olduğunu bu nedenle pozitif madenciliği önermişlerdir. Önerilen yaklaşımla da eğitim sürecinde aşırı öğrenmenin önüne geçeceklerini belirtmektedirler.

Özetlemek gerekirse; DML’de her ne kadar iyi matematiksel modeller ve mimariler

kullanılırsa kullanılsın, ağın eğitim yeteneği ağa sunulan örneklerin ağı temsil kabiliyetiyle sınırlı olacaktır. Ağa sunulacak olan bilgili örnekler hem öğrenmeyi iyileştirecek hem de ağın daha iyi bir temsil yeteneği kazanmasını sağlayacaktır. Bu nedenle derin metrik öğrenme için örnekler arasındaki ilişkinin etkisi dikkatlice incelenmelidir. Örnek seçimi, derin metrik öğrenme modelini uygulamadan önce ağ modelinin başarısını artırmak için bir ön işleme adımı olarak çok faydalı olacaktır.

Literatürdeki sonuçlar, derin metrik öğrenmede negatif madencilik çalışmalarının yüksek etkiye sahip olduğuna işaret etmektedir. Bilgili örnekler seçmenin diğer bir yararı göz önüne alındığında; aşırı öğrenmeden kaçınmak mümkün olabilir. Ağın eğitim sürecinde benzer örüntüler benzer etkileşimlere sahiptir. Bu nedenle bu durum daha yavaş öğrenmeye ya da yerel en iyi (optimal) çözümle sınırlanmaya neden olabilir.

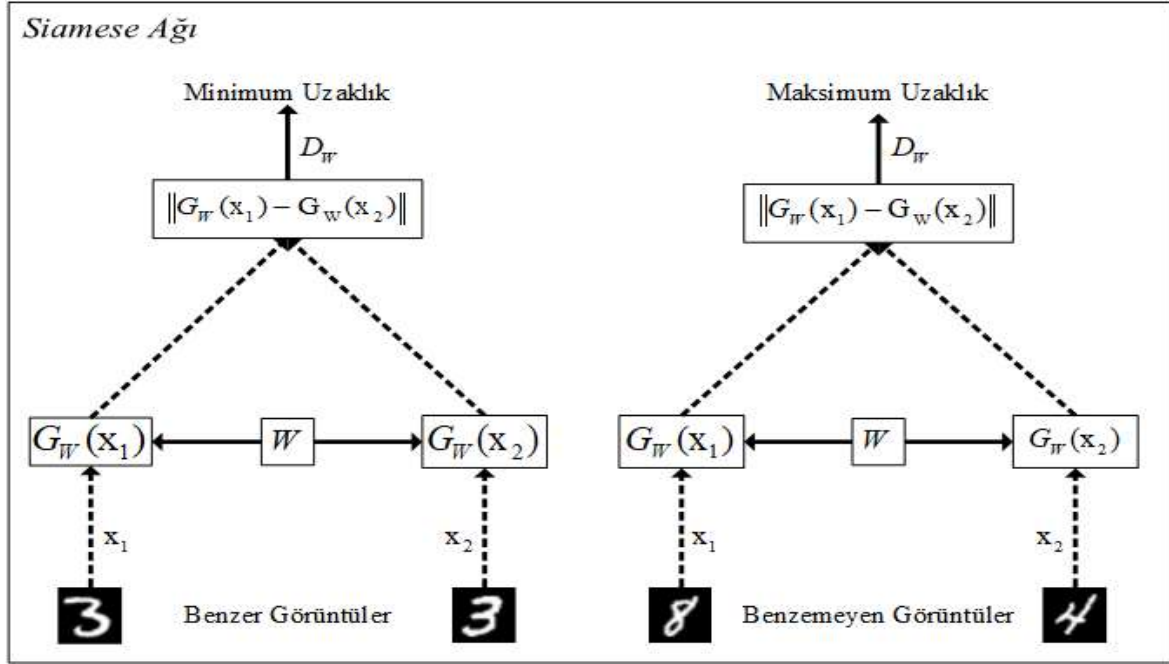
Siamese ağındaki olası tüm ikili örnekler ya da Triplet ağında, olası tüm üçlü örneklerin sayısı, zaman kaybı ve kaynakların gereksiz kullanımıyla sonuçlanmaktadır. Siamese ağında $O(n^2)$, Triplet ağında ise $O(n^3)$ zaman karmaşıklığı durumları söz konusu olabilmektedir. Bu sorunun üstesinden gelmek için sadece değerli ikili ya da üçlü örneklerle ilgilenmek yeterlidir. Böylece, bilgilendirici örnek seçildikten sonra performansta önemli gelişmeler sağlanabilir.

4.3. Derin Metrik Öğrenmede Kayıp Fonksiyonları

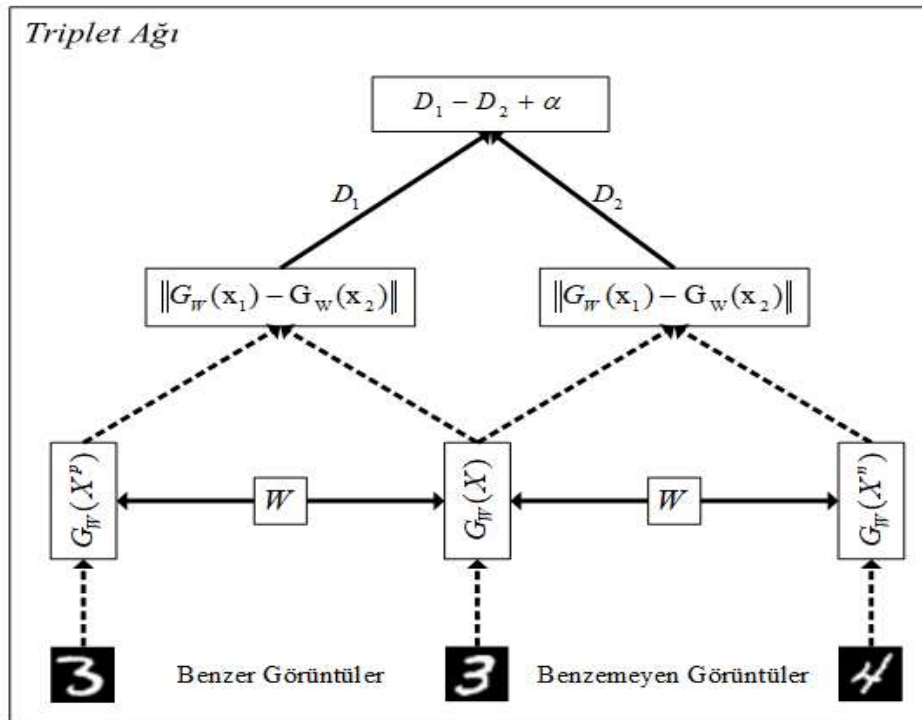
Bu bölümde, literatürde derin metrik öğrenmeyi uygulamak için kullanılan bazı kayıp fonksiyonları vurgulanacaktır. Bu fonksiyonların kullanılma şekli de tanıtılacak ve farklılıkları ile ilgili bazı detaylar verilecektir. Bu işlevler, benzerliklerine bakarak nesneler arasındaki mesafeyi artırmamızı veya azaltmamızı sağlar. Amaç, farklı nesneler arasında en yüksek temsili özellik gösterimini elde etmektir.

Başlangıçta, Siamese ağı, imza doğrulaması için sinir ağları yapısı ile birlikte kullanılmıştır [80]. [80]'den farklı olarak, Siamese ağı enerji bazlı modeller için ayırt edici bir öğrenme çerçevesinden öğrenmeye dayanmaktadır [10]. Bu yaklaşımda, Siamese ağına iki eş görüntü verilir ve bu görüntülerden öğrenmenin bir sonucu olarak ikili bir değer elde edilir. Görüntüler “0” olduğunda aynı sınıf olarak kabul edilir; eğer “1” ise, farklı bir sınıf olarak kabul edilirler. Bir derin metrik öğrenme yaklaşımı olan Siamese ağı, bir ağ modelini

eğitmek için pozitif ve negatif örnekler de dahil olmak üzere ikili görüntüler alır (Şekil 4.5) [134]. Bu çift görüntülerin arasındaki mesafe bir kayıp fonksiyonu ile hesaplanır. Siamese ağında kayıp fonksiyonunu hesaplamak için Kontrastif kayıp fonksiyonundan yararlanılır [71, 135]. Bu yaklaşım, derin metrik öğrenme alanında çalışan araştırmacılara esin kaynağı olan bir yaklaşımdır.



(a)

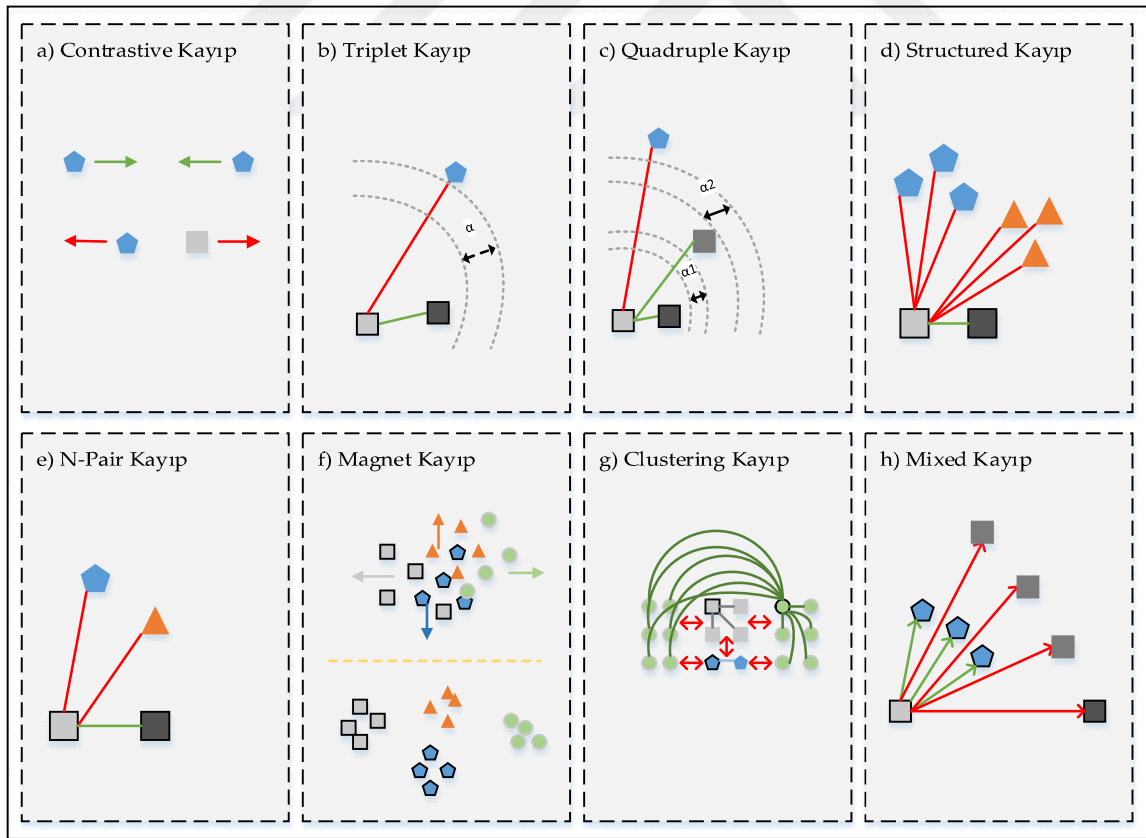


(b)

Şekil 4.5. Siamese ve triplet ağı

Şekil 4.6 (a)'da görülebileceği gibi, Siamese ağı, sınıflandırma performansını iyileştirmek için nesneler arasındaki mesafeyi en üst düzeye çıkarmak veya en aza indirmek için çok başarılı bir modeldir. Bir sinir ağının performansını olumlu yönde etkileyen paylaşılan ağırlıklar, Şekil 4.5'te gösterildiği gibi derin metrik öğrenmede görüntüler arasında anlamlı bir desen elde etmek için kullanılır. Bu ağırlıkların zaman ve hafıza açısından da önemli avantajları vardır. Siamese ağı kendi yapısı içerisinde evrişimli sinir ağı yapısında barındırabilmektedir. Böylece evrişimli sinir ağı aynı anda renk, doku bilgilerini ve doğrudan görüntü piksellerinden öğrenme avantajına sahip esnek bir yapıdadır [136].

Benzerlik yaklaşımıyla Wen ve arkadaşları [137], softmax kaybını ve merkez kaybını yüz tanıma için birleştirmişlerdir. Merkez kaybı, kontrastif kaybı gibi her sınıfın derin özellikleri için derin özellikler ile sınıf merkezleri arasındaki mesafeleri en aza indirmek için bir merkez bulmayı hedeflemektedir. Softmax kaybı ise farklı sınıflar arasındaki mesafeleri arttıran derin özellikler bulmayı amaçlamaktadır.



Şekil 4.6. Kayıp fonksiyonları

Siamese ağından esinlenilen Triplet ağı, pozitif, negatif ve çapa örneklerinden oluşan üç örnek içerir [11]. Triplet ağı örüntü tanıma işlemindeki nesneleri karşılaştırmak için Öklid uzayını kullanır ve bu yaklaşım doğrudan metrik öğrenmeyle ilgilidir. Denklem 4.6'da görülebileceği gibi, Triplet kaybı paylaşımlı ağırlıkları kullanarak aynı ve farklı sınıfların çift örnekleri arasındaki benzerlik üzerine odaklanır. Sınıflandırma, çift örneklerin benzerliğini karşılaştırarak gerçekleştirilir (Şekil 4.6. (b)). Üçlü ağlar, hem sınıf içi hem de sınıf arası ilişkileri kullanarak daha yüksek bir ayrımcılık gücü sağlar.

Ge ve arkadaşları [113], sınıf düzeyinde bir hiyerarşik ağaç kullanan üçlü kayıplara dayanan yeni bir kayıp fonksiyonu önermişlerdir. Hiyerarşik üçlü kaybı, daha bilgilendirici örnekleme yaklaşımı sayesinde çevrimiçi uyarlanabilir ağaç güncellemesiyle ağı performansı artırır. Benzer şekilde, Wang ve arkadaşları [111], derin metrik öğrenmeyi geliştirmek için yeni bir açısız kayıp fonksiyonu önermişlerdir. Siamese ve Triplet ağlardan farklı olarak, açısız kayıp üçlü noktaların negatif noktasındaki açısız kısıtlamaya odaklanır. Açısız kayıp, negatif noktayı pozitif kümenin merkezinden uzağa iter ve daha sonra dönme ve ölçekten bağımsız metriğe sahip bir açı kullanırken pozitif noktaları birbirine yaklaştırır (Denklem 4.10).

Ni ve arkadaşları [138], her ağ eğitim sürecinde giriş olarak dörtlü örnekler kullanırken nesneler arasında daha iyi bir yakınlık derecesi elde etmişlerdir. Denklem 4.9'da gösterildiği gibi, Triplet kayıp için alternatif bir yaklaşım olarak dörtlü kayıp fonksiyonuna X örneğine benzer yeni bir giriş örneği eklenmiştir. Şekil 4.6.(c)'de Quadruple kaybın iki farklı marjin değerinden faydalandığını göstermektedir. Bu marjin değerleri, benzer örnekler arasında ve farklı örnekler arasında daha anlamlı ilişkiler kurmayı amaçlamaktadır. Histogram kayıp fonksiyonu [116], eğitim için Quadruple ağ modelinde olduğu gibi dörtlü ağ örnekleri kullanır. Pozitif ve negatif çiftlerin benzerlik dağılımlarını hesaplamak için histogramlardan yararlanır ve diğer kayıpların aksine herhangi bir ayar parametresine ihtiyaç duymaz. Diğer kayıp fonksiyonlarına kıyasla, CUHK03 ve Market-1501 gibi kişi yeniden tanımlama veri kümeleri için deneysel çalışmalarda daha iyi sonuçlar elde etmiştir.

Yao ve arkadaşları [117], SVM'de olduğu gibi, öğrenme riskini minimize etme kısıtlamasından yararlanmak üzere kişinin yeniden tanımlanması görevleri için kısmi kayıp fonksiyonunu önermişlerdir. Kısmi kayıp fonksiyonu, yalnızca belirli bir noktaya

odaklanmak yerine bütünün farklı parçalarını ayrı ayrı kullanmayı amaçlamaktadır. Yerel kalıpların her birini ayrı bir parça olarak değerlendirmek için beş parçaya bölünmüş bir resmin her kısmı için kısmi kayıp değeri hesaplanmaktadır.

Kayıp fonksiyonları, derin metrik öğrenme modellerinin en önemli kısımlarından biridir. Dönüşüm uzayında yeni bir temsil elde etmek için iki eş görüntüyü dikkate alarak bazı kayıp fonksiyonlarını denklemler yolu ile açıklamaya çalışalım. X_1 ve X_2 şeklinde iki tane ağa sunulacak örneğimiz olduğunu kabul edelim. Bu iki örnek arasındaki uzaklık aşağıdaki denklemdeki gibi olacaktır.

$$D_W(X_1, X_2) = \|G_W(X_1) - G_W(X_2)\|_2 \quad (4.4)$$

Burada $G_W(X_1)$ ve $G_W(X_2)$, X_1 ve X_2 'nin yeni temsili uzaydaki karşılığıdır. D_W ise temsili uzaydaki bu iki örnek arasındaki uzaklığı hesaplayan bir fonksiyondur. Şekil 4.5.'te görüleceği üzere D_W benzer görüntüler için minimize edilecek olup, farklı görüntüler için ise maksimize edilecektir. Bu uzaklık fonksiyonundan yararlanarak Siamese ağ modelinde kayıp fonksiyonunu hesaplamak istersek;

$$L_{Contrastive} = (1 - Y) \frac{1}{2} (D_W)^2 + (Y) \frac{1}{2} \{\max(0, m - D_W)\} \quad (4.5)$$

Denklemdeki Y etiket verilerini temsil etmektedir. Eğer veri kümesindeki örnekler aynı sınıftan ise Y değeri 1 olup, farklı sınıflardan ise Y değeri 0 olmaktadır. $L_{Contrastive}$ 'deki m değeri ise ön tanımlı bir marjın değeridir. Siamese ağ modelinden farklı olarak Triplet ağ modelleri üç girişe sahiptir. Çapa örneği X , çapayla aynı sınıftan olan X^p ve çapadan farklı sınıfta olan X^n örneğidir. Triplet kayıp fonksiyonuna ait denklem ise şu şekildedir:

$$L_{Triplet} = \max(0, \|G_W(X) - G_W(X^p)\|_2 - \|G_W(X) - G_W(X^n)\|_2 + \alpha) \quad (4.6)$$

Denklemdeki α margin değeri olup, ön tanımlı bir değerdir. Triplet ağından esinlenilerek geliştirilen Quadruple ağ modelinde ise, Triplet ağ modelinden farklı olarak yeni bir X^s giriş örneği bulunmaktadır.

$$D_1 = \max(0, \|G_W(X) - G_W(X^p)\|_2 - \|G_W(X) - G_W(X^s)\|_2 + \alpha_1) \quad (4.7)$$

$$D_2 = \max(0, \|G_W(X) - G_W(X^s)\|_2 - \|G_W(X) - G_W(X^n)\|_2 + \alpha_2) \quad (4.8)$$

$$L_{Quadruple} = D_1 + D_2 \quad (4.9)$$

Açısal (angular) kayıp fonksiyonu ise giriş örnekleri arasındaki ilişkiyi incelerken örnekler arasındaki açısal değerden yararlanmaktadır. $L_{Angular}$ denklemi ise şöyledir:

$$L_{Angular} = \max(0, \|G_W(X) - G_W(X^p)\|_2 - 4 \tan^2 \alpha \|G_W(X^n) - G_W(X^c)\|_2) \quad (4.10)$$

X^c , X ve X^p giriş örneklerinin ortasını ifade etmektedir.

$$X^c = (X + X^p)/2 \quad (4.11)$$

Siamese ve Triplet ağ modelleri gibi geleneksel derin metrik öğrenme modelleri, her eğitim adımında eğitim örneklerinin yapısal düzenlemesini ihmal edebilmesi söz konusudur [139]. Ancak, sınırlı eğitim örneklerine sahip küçük boyutlu veri kümeleri için bu durum bir dezavantajdır. Derin yapısal metrik öğrenme yaklaşımı [73], küme içindeki ikili uzaklıklar vektörünün, derin ağdaki özel bir yapısal kayıp ile ikili mesafeler matrisinin kaldırılmasını önermektedir. Bu nedenle, içeriksel bilginin tüm avantajını toplu eğitim örnekleri içinde kullanmak mümkündür. Şekil 4.6. (d)'de görülebileceği gibi, yapısal kayıp aynı anda birçok örnek arasındaki benzerlik ilişkisini ele almaktadır. Önerilen algoritmanın gücü, yeterli sayıda örneğin bir arada değerlendirilmesinden gelir.

Sohn [110], Siamese ve Triplet Ağların yavaş yakınsamasını ve zayıf yerel optimasyonu problemlerini ele almak için çok sınıflı bir N-çifti kaybı önermiştir. Triplet kaybı, her eğitim aşamasındaki güncellemede yalnızca bir negatif örnekle ilgilidir ve diğer negatif örneklerle etkileşime girmez. Üçlü kayıbdan farklı olarak, N çifti kaybı, çapa örneğini karşılaştırmak için N-1 negatif sınıf örneklerinden yararlanır (Şekil 4.6. (e)) Sınıf sayısı sınırlıysa, N-çifti kaybını kullanmak gereksiz olabilir, çünkü sınıf sayısı azalırsa çapa örneği ile yapılan karşılaştırma sayısı da azalacaktır. Bu durumda N çifti kaybının avantajının kaybolmasına neden olacaktır.

Literatürdeki diğer tüm kayıp fonksiyonlarının aksine, Wang ve arkadaşları [140], örnekler arasındaki benzerliği yakalamak için genel çift ağırlıklandırma çerçevesinde öz-benzerlik, negatif bağıl benzerlik ve pozitif bağıl benzerlikten faydalanmışlardır. Bu kayıp fonksiyonu, kendi kendine benzerliği ve göreceli benzerlikleri hesaba katar, bu durumda modelin bilgilendirici çift örneklerini daha verimli bir şekilde toplamasını ve onların daha iyi bir şekilde ağırlıklandırmasını mümkün kılar.

Siamese, Triplet ve n-ikili ağlar için eğitim aşaması sürecinden önce eğitim verilerinin hazırlanması gerekmektedir. Bu işlem diskte daha fazla alana ihtiyaç duymakta ve zaman almaktadır. Song ve arkadaşları [112], bu ağların aynı zamanda verilerin yerel görünümü ile de ilgilendiğini iddia etmektedir. Yazarlar bu sorunların üstesinden gelmek için kümeleme kaybını kullanırken yeni bir derin metrik öğrenme yöntemi önermişlerdir. Şekil 4.6. (g)'de görülebileceği gibi, kümelenme kaybı, küme merkezlerini kullanarak örnekleri bir küme içerisinde bir araya getirmeyi amaçlamaktadır. Ayrıca, farklı kümelerin birbirine yaklaşmasını da önlemektedir. Yapısal tahmine dayanarak önerilen yöntem, normalize edilmiş ortak metriktен yararlanır.

Rippel ve arkadaşları [141], Triplet kaybının bir veri kümesini eğitmek için aynı anda üç giriş örneğini bir arada değerlendirdiğini vurgulamaktadır. Bu durum ağın öğrenme süresini azaltmasına karşın, bu yaklaşımın düşük performansa ve yetersiz eğitime neden olduğunu belirtmektedir. Bu nedenle, kümelenmenin üst üste binmesini cezalandıran ve birden fazla kümeyi ayırmak için kümedeki en yakın komşuları değerlendiren magnet kaybını önermişlerdir. Önerilen yöntem yerel ayrımcılığa sahiptir ve genel olarak optimizasyon prosedürü ile tutarlıdır. Şekil 4.6. (f)'te görüldüğü üzere yerel komşulukları kullanırken benzer örneklerin en yakın kümede nasıl birbirine yaklaştığını göstermektedir.

Triplet kayıptan esinlenen karışık kayıp fonksiyonu [142], çapa ve negatif örneğe ek olarak, örnekler arasındaki benzerlik ilişkisini kurmak için üç pozitif örnek ve üç negatif örneği bir arada kullanır. Şekil 4.6. (h)'ta görüleceği üzere, negatif örnekleri çapadan uzağa itmeyi amaçlarken, kayıp fonksiyonunun pozitif örneklerini çapaya yaklaştırmayı amaç edinmektedir. Literatürde popüler olup derin metrik öğrenmede kullanılan bazı kayıp fonksiyonları, Çizelge 4.2'de detaylı olarak özetlenmiştir.

Çizelge 4.2. Kayıp fonksiyonları

Metrik	Örnek Seçimi	Konu	Veri Kümesi	Amaç	Yıl
Kontrastif Kayıp [71]	Sert negatif	Görüntü tanıma Nesne tanıma	MNIST [70] NORB [143]	Benzenmeyen nesnelerin çiftleri için daha yüksek bir değer elde etmeyi amaçlayan ve benzer nesnelerin çiftleri için daha düşük bir değer elde etmeyi amaçlayan bir kayıp fonksiyonu hesaplar.	2006
Triplet Kayıp [11]	Kolay örnekleme	Görüntü tanıma Nesne tanıma	MNIST [70] CIFAR10 [144] SVHN [145] STL10 [146]	Çapa pozitif örnekler ile çapa negatif örnekler arasındaki uzaklık farkını hesaplar ve benzer nesneleri yakınlaştırmayı amaçlar	2014
Histogram Kayıp [116]	Kolay örnekleme	Görüntü tanıma Görüntü elde edinimi Kişi yeniden tanıma	CUB-200-2011 [109] Online Products [73] Market-1501 [115] CUHK03 [118]	Daha az örtüşen pozitif ve negatif çiftlerin benzerlik dağılımlarını amaçlar.	2016
Structured Kayıp [73]	Sert negatif	Görüntü elde edinimi	CUB-200-2011 [109] Online Products [73] CAR-196 [114]	Eğitim boyunca her bir eğitim grubu içindeki yükseltilmiş yoğun çiftli matrisini kullanarak yeni bir metrik öğrenme algoritmasını hedefler.	2016
N-Pair Kayıp [110]	Çoklu negatif sınıf	Görüntü elde edinimi Görüntü kümeleme Yüz doğrulama Yüz tespit etme Nesne tanıma Nesne doğrulama	CUB-200-2011 [109] Online Products [73] Flower-610 [110] CAR-196 [114] LFW [120] Car-333 [147]	Her eğitim aşamasında pozitif bir örneği çoklu negatif örneklerden uzaklaştırmaya odaklanan üçlü kaybı geliştirmeyi amaçlamaktadır	2016
Magnet Kayıp [141]	Sert negatif	Görüntü tanıma Görüntü etiketleme	Stanford Dogs [148] Oxford-IIIT Pet [149] Oxford 102 Flowers [150] Object Attributes [151]	en yakın kümelerin bulunduğu yerel komşuluktan ele alıp, örtüşmelerini cezalandırmayı amaçlar	2016

Çizelge 4.2. (devam) Kayıp fonksiyonları

Metrik	Örnek Seçimi	Konu	Veri Kümesi	Amaç	Yıl
Angular Kayıp [111]	Çoklu negatif	Görüntü elde edimini Görüntü kümeleme	CUB-200-2011 [109] Online Products [73] CAR-196 [114]	Üçlü üçgenlerin negatif örneğindeki açığı sınırlamaya odaklanılır.	2017
Quadruple Kayıp [139]	Yarı-sert negatif	Hasta benzerliği	The Ischemic Heart [138] The Cerebrovascular [138]	Hastalar arasındaki benzerlik derecesini etkili bir şekilde yakalamayı amaçlar	2017
Clustering Kayıp [112]	Kolay örnekleme	Görüntü elde edimini Görüntü kümeleme	CUB-200-2011 [68] Online Products [69] CAR-196 [73]	Bir kümeleme kalitesi metniği ile dönüşüm uzayında global yapıyı hesaba katan yapısal tahmine dayalı yeni bir metrik öğrenme yaklaşımı hedeflemektedir.	2017
Hierarchical Triplet Kayıp [113]	Çapa komşu örnekleme	Görüntü elde edimini Yüz tanıma	CUB-200-2011 [109] Online Products [73] CAR-196 [114] LFW [120] In-Shop Clothes Retrieval [152]	Çevrimiçi sınıf düzeyinde bir ağaç güncellemesi ile bilgilendirici örnekler toplamayı ve global veri bağlamını yakalamayı amaçlar	2018
Mixed Kayıp [142]	Sert farkında çevrimiçi örnek madencilik	Görüntü elde edimini	Fashion Collocation Dataset [142]	Sinir ağında zamanla çoklu pozitif ve negatif örnekler beslemeyi amaçlar	2018
Part Kayıp [117]	Kolay örnekleme	Kişi yeniden tanıma	Market-1501 [115] CUHK03 [118] VIPeR [153]	Görüntüleri K bölüme ayırarak test için temsili öğrenme risklerini ve eğitim için deneysel sınıflandırma risklerini azaltmayı amaçlar	2019
Multi-Similarity Kayıp [140]	Genel eş ağırlıklandırma	Görüntü elde edimini	CUB-200-2011 [109] Online Products [73] CAR-196 [114] In-Shop Clothes Retrieval [152]	Bilgilendirici eşleştirilmiş örnekler toplamayı hedefler ve bu çiftleri hem kendi hem de göreceli benzerliklerine göre ağırlıklandırır	2019



5. METRİK ÖĞRENME YAKLAŞIMIYLA SINIFLANDIRMA

Nörolojik problemler insan yaşamını doğrudan etkileyen problemlerdendir. Parkinson hastalığı da insan yaşamını doğrudan etkilemekle beraber sağlıklı bir yaşam sürdürmeyi engellemektedir. Ayrıca Parkinson hastalığı en yaygın görülen nörolojik düzensizliklerden biri olmaktadır [154]. Parkinson konuşma verilerinde bulunan düzensizliklerin tespiti literatürdeki çalışmalarda ilgi çeken konulardan biridir [155, 156]. Tez çalışmasının bu bölümünde Parkinson konuşma verilerinin metrik öğrenme yaklaşımıyla nasıl daha başarılı bir şekilde sınıflandırılabilceği konusu ele alınmaktadır.

Günümüz problemlerinde makine öğrenmesi yaklaşımlarında veriyi sınıflandırma amacıyla veri uzayında verinin en iyi temsilini bulmak en temel problemidir. Gerçek dünya problemlerinde verilerin makine öğrenmesi teknikleri ile verinin en uygun temsili veri uzayını bulmak oldukça önemlidir. Bir yüz tanıma problemini ele aldığımızda yüz görüntüsünün çeşitli problemleri bulunmaktadır. Bu problemler görüntünün arka planı, ifade şekli, verilen poz, ışığın yansıması, ölçeklendirme, saç stili, gözlük vb.'dir. Makine öğrenmesi teknikleri bu gibi etmenlerden etkilendirmekte ve bu durumda veriyi doğru sınıflandırmada açısından bir problem ortaya çıkarmaktadır [157]. Her bir veri örüntüsünün kendine has bazı problemleri olmakla beraber herhangi bir veri uzayı üzerinde en uygun uzaklık metriğinin nasıl olacağını (benzerlik veya benzersizlik) belirlemek; problemin başarımını doğrudan etkileyen bir unsur olarak karşımıza çıkmaktadır. Veri uzayının daha iyi bir temsilini elde etmek ya da gürültülerden temizleme ve pürüzsüzleştirme amacıyla bir dönüşüm uzayından yararlanılarak verileri daha iyi bir temsili uzaya taşımak mümkündür [158].

k en yakın komşu yöntemi basit ve etkili bir sınıflandırıcıdır. Ancak bu sınıflandırma yönteminin veriler sınıflandırma yetenekleri sınırlıdır. Sınıflandırma başarımını elde etmek amacıyla veri kümesi üzerinde doğrusal ya da doğrusal olmayan bir dönüşüm işlemi gerçekleştirerek verilerin yeni bir uzaya taşıma işlemi başarı oranıyla doğrudan ilgilidir. Bu sebeple verileri en iyi temsil yeteneğine sahip yeni bir uzaya taşınması işlemi için optimizasyon yöntemlerinden yararlanmak mümkündür. Bu amaçla yapılan tezin bu bölümünde genetik algoritmanın hem gücünden hem de hızından yararlanılmıştır. Problemin çözümünde genetik algoritma; sınıflandırma işlemi gerçekleştirilmeden önce

dönüşüm matrisindeki ağırlık matrisinin belirlenmesi amacıyla kullanılmaktadır. Bu amaç doğrultusunda veriyi en iyi yansıtacak ağırlık dönüşüm matrisinin elde edilmesi amaçlanmaktadır. Daha sonra bu ağırlık matrisi kullanılarak yeni bir uzaya taşınmış olan veri üzerinde sınıflandırma işlemi gerçekleştirilmektedir.

5.1. Materyal Ve Metot

Bu bölümde çalışmada kullanılan veri kümesi ve o veri kümesi üzerinde önerilen metrik öğrenme yaklaşımı hakkında bilgi sunulmaktadır.

5.1.1. Veri kümesi

Veri kümesi üzerinde performans ölçümü gauss dağılımına sahip veri kümelerinde daha başarılı sonuçlar elde edilmesi beklenmektedir. Bu bölümde öncelikli olarak iki sınıflı ve iki niteliğe sahip bir veri kümesi için 1000 farklı örnek oluşturulmuştur. Önerilen yöntemin veri kümesinde sınıfların birbirinden ayırt edilebilme durumuna göre 6 farklı senaryo oluşturulmuştur. Rasgele oluşturulmuş veri kümesi üzerinde belirli bir sınıflandırma başarısı elde edildikten sonra gerçek dünya verileri üzerinde de yöntemin geçerliliği test edilmek istenmiştir. Bu amaçla çalışmada ayrıca parkinson çoklu ses kaydetme veri kümesi [159] üzerinde deneyler gerçekleştirilmiştir.

Parkinson çoklu ses kaydetme veri kümesi 20 sağlıklı, 20 hasta (toplam 40) birey içermekle birlikte; her bir hastadan 26 farklı ses örneği alınarak oluşturulmuş bir veri kümesidir. Bu sesler; harf, rakam, kısa cümle ve kelimelerden oluşmaktadır. Çalışmada kullanılan veri kümesi 1040 örnekten oluşmaktadır. Her bir ses örneğinden 26 öznitelik çıkartılmıştır. Böylece 1040×26 boyutundaki bir veri kümesi üzerinde çalışılmıştır.

5.1.2. Önerilen yöntem

Literatürdeki dönüşüm uzaylarında kullanılan matrisler pozitif tanımlı olarak verilmektedir. Matris dönüşümünde gerçekleştirilen bazı işlemlerde yöntem pozitif tanımlı matris içerdiğinden olası çözümlerde pozitif tanımlı matris zorunluluğu doğurmaktadır (Kovaryans matrisi). Ayrıca verinin geri dönüşümü içinde bu matrislerden yararlanılmaktadır. Ancak bu tek başına bir çözüm değildir. Bu bağımlılığı aşarak çözüm

üretmek ve başarı oranında artış sağlamak mümkündür. Dönüşüm işlemi veya yeni bir uzaya taşıma işlevi denklem 5.1’de verilen karesel bir matris ile sağlanmakta ve 5.2’de verilen denklem ile yeni bir uzaya taşınma işlemi gerçekleştirilmektedir.

$$W = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \quad (5.1)$$

$$y = Wx \quad (5.2)$$

a :Dönüşümde kullanılacak sayısal değerler

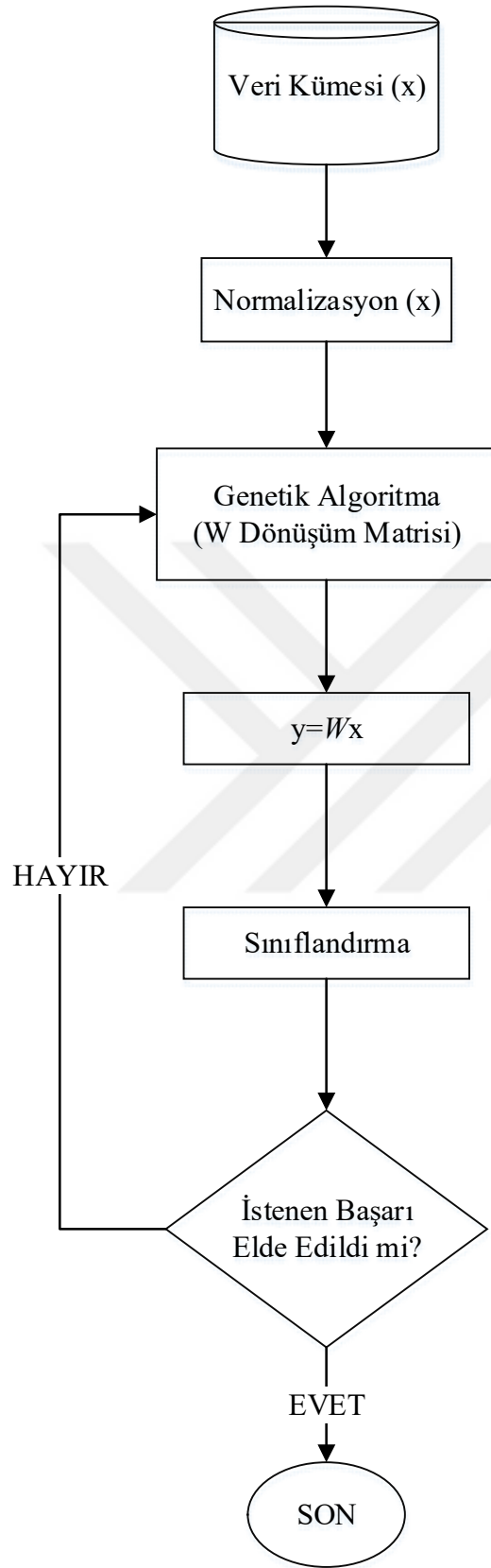
W :Dönüşüm matrisi

x :İşlenmemiş veri kümesi

y :Dönüşüm sonrası elde edilen veri kümesi

Verinin yeni bir uzaya taşınma işlemi için önerilen adımlar Şekil 5.1’de görülmektedir. Bunun için veriler denklem 5.3’te görüldüğü üzere bir ön işleme adımı olarak minimum maksimum normalizasyon işlemine tabi tutulmaktadır. Daha sonrasında -1 ve +1 sayı aralığı içerisinde veriler belirli aralıklara bölünmekte ve her bir veri bir gen olarak kodlanmaktadır. Dönüşüm matrisi boyutu veri kümesinin öznitelik sayısına eşittir ve öznitelik sayısının artışı doğrudan problemin çalışma karmaşıklığını arttırmaktadır. Bu yüzden iteratif bir artış yerine hızlı ve başarılı bir yöntem olan genetik algoritma tercih edilmiştir. Tüm arama uzayında en uygun karesel dönüşüm matrisi bulunana kadar algoritma çalışmasını sürdürecektir ve hedeflenen başarı elde edilince algoritma sonuçlandırılacaktır.

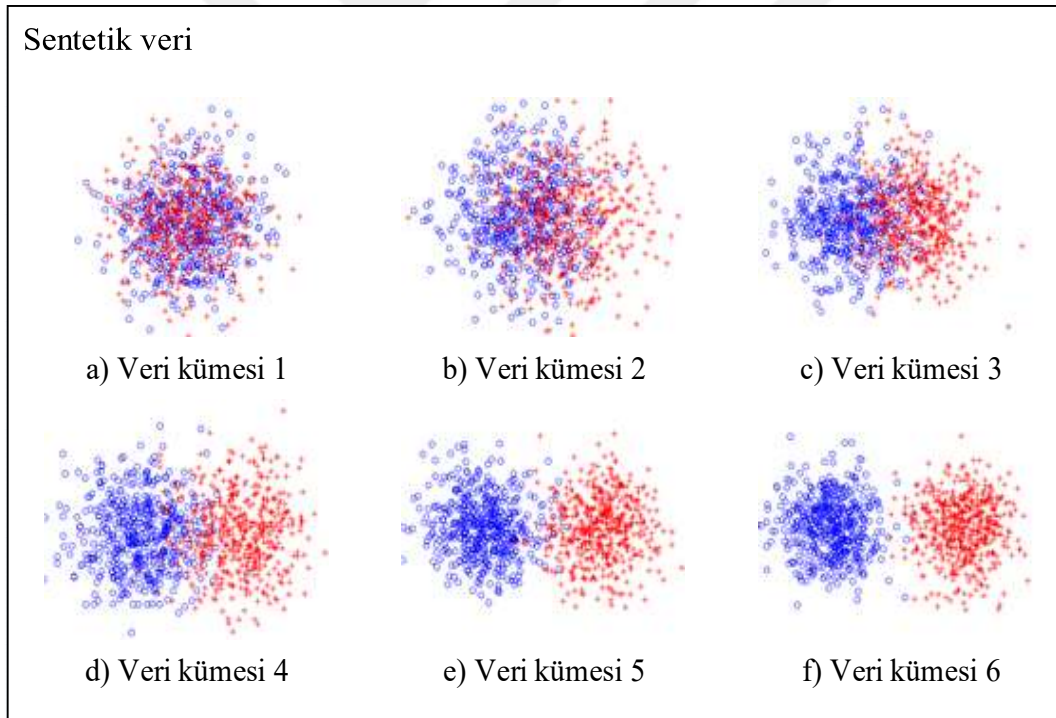
$$x_i = \frac{x_i - x_{min}}{x_{max} - x_{min}} \quad (5.3)$$



Şekil 5.1. Önerilen yöntem [160]

5.2. Sentetik Veri Üzerinde Metrik Öğrenme Yaklaşımı

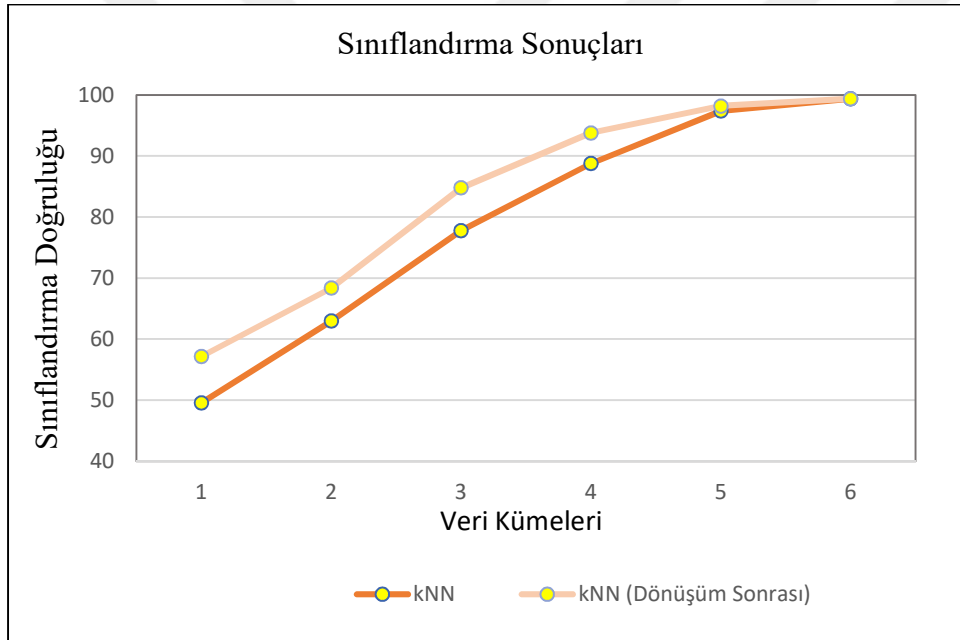
Bir dönüşüm matrisinin gauss dağılımına sahip bir veri kümesi üzerindeki etkilerini görebilmek için altı farklı rasgele veri kümesi oluşturulmuştur. Şekil 5.2’de görüleceği üzere iki farklı sınıfa sahip olan bu veri kümesinde sınıfların birbirinden belirgin olarak ayrılmadığı durumdan ayrışmanın belirginleştirildiği duruma kadar senaryolar oluşturulmuştur. Farklı ayrışma durumları ile etkinin veri kümesine bağlılığını ölçme amaçlanmaktadır. Veri kümeleri görsel olarak iki boyutlu bir düzlemde görülebilmesi amacıyla iki öz nitelik kullanılarak oluşturulmuştur. Aynı zamanda oluşturulacak olan düzenekte yapılacak işlem hacminden dolayı iki öz nitelik ile işlemleri gerçekleştirmek zaman karmaşıklığı açısından büyük önem arz etmektedir. Çünkü öz nitelik sayısının artması doğrudan doğruya bir zaman problemini beraberinde getirmektedir.



Şekil 5.2. Sentetik veri kümesi örnekleri

Oluşturulan veri kümesi üzerinde bir dönüşüm matrisi ile yeni bir uzaya taşınma işlemi hedeflenmiştir. Veri dönüşümünü gerçekleştirmek içinde denklem 5.1’de verilmiş olan karesel bir matris olan W matrisinden yararlanılması amaçlanmıştır. W matrisinin elemanları -1 ile +1 değerleri arasında 0,1 sayı aralıklarıyla değerler almakta ve tüm arama uzayı ardışık bir şekilde incelenerek en uygun W matrisi belirlenmesi amaçlanmaktadır.

Optimal W matrisi belirleme işlemi için ise; veri kümelerinin k en yakın komşu sınıflandırıcı algoritması kullanılarak sınıflandırma performansı değerlendirilmiştir. Gerçekleştirilen deneylerde eğitim ve test kümeleri birbirinden bağımsız olup, komşuluk değeri 3 olarak tercih edilmiştir. Değerlendirme ölçütü sınıflandırma doğruluğu olmakla birlikte; elde edilen en iyi (optimal) W matrisinden yararlanılarak yeni veri kümesi ve ham veri kümesi üzerindeki sınıflandırma sonuçları her bir veri kümesi için Şekil 5.3'te karşılaştırmalı olarak verilmektedir. Sonuçları incelediğimizde başarımdaki artışlar veri kümesinin dağılımına bağlı olarak farklılık göstermekle birlikte; uygulanan tüm veri kümeleri üzerinde en iyi (optimal) W matrisi ile olumlu sonuçlar elde edildiği görülmektedir.



Şekil 5.3. Sentetik veri kümeleri sınıflandırma sonuçları

5.3. Parkinson Çoklu Ses Kaydetme Veri Kümesinde Uygulama

Öznitelik sayısı az olan bir veri kümesinde deneylerin başarılı sonuçlar vermesinden sonra gerçek dünya veri kümesinde de yöntemin analizi gerçekleştirilmiştir. Mevcut gerçek dünya verisi 26 öznitelik, 2 sınıf ve 1040 örnek verisinden oluşmaktadır [159]. Bu nedenle bu veri kümesi üzerinde bir önceki düzenekte olduğu gibi ardışık bir artış oldukça büyük bir işlem hacmini beraberinde getirmektedir. Bu sebeple hem çözüme daha hızlı ulaşım hem de veri kümesini en iyi yansıtacak olan dönüşüm matrisini belirlemek için optimizasyon yöntemlerinde başarısını kanıtlamış olan genetik algoritma tercih edilmiştir.

k en yakın komşu gibi sınıflandırıcı algoritmaları öznelitliklerin aldığı değerlerin çok büyük ya da çok küçük olmasından doğrudan etkilenmektedir. Bu sebeple öznelitliklerin bireysel değerlerinin sınıflandırıcı üzerindeki etkisini kırmak için normalizasyon gerekli bir işlemdir. Çalışmada kullanılan veri kümesinde normalizasyon için literatürde de sıklıkla tercih edilen minimum maksimum normalizasyon tercih edilmiştir.

Veri kümesini dönüşüm uzayında en iyi yansıtacak dönüşüm matrisini tespit etmek amacıyla herhangi bir sınırlama getirilmemiş ve pozitif ya da negatif tanımlı matrisler dönüşüm matrisinde kullanılmıştır. Dönüşüm matrisinin elemanları -1 ile +1 arasındaki sayılardan oluşmakta olup 0,05 sayı aralıklarındaki sayılar genetik algoritmanın gen havuzuna yerleştirilmiştir. Böylece hem daha geniş bir arama uzayı elde edilmesi amaçlanmış hem de başarımın artırılması hedeflenmiştir. Optimal W matrisi belirleme işlemi için k en yakın komşu yönteminden yararlanılmıştır. Deneylerde k değeri 3 olarak belirlenmiştir. Yapılan çalışmanın güvenilirliğini arttırmak için literatürde de sıklıkla kullanılan k-kat çapraz geçerleme yöntemi kullanılmıştır. Ayrıca karşılaştırma amacıyla da başlangıçta rasgele oluşturulan eğitim ve test indeksleri belirlenmiş ve deney boyunca sabit tutulmuştur. Deneylerin test oranını yüksek tutma amaçlanmış ve bu amaçla veriler %20 eğitim ve %80 test şeklinde 5-kat çapraz geçerlemeye uygun indeksler belirlenmiştir. Veri kümesindeki bir bireye ait genetik kodlamanın nasıl yapıldığı Şekil 5.4'te görülmektedir. Burada karesel matris vektöre dönüştürülmüş ve çaprazlama ve mutasyonlar vektör üzerinden gerçekleştirilmiştir. Yapılan işlemler sonrasında tekrar matris biçimine dönüşüm işlemi gerçekleştirilmiştir.

$$W = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \Rightarrow a_{11}a_{12} \cdots a_{1n}a_{21} \cdots a_{nn}$$

Şekil 5.4. Genetik kodlama

Yapılan çalışmada genetik algoritma için probleme uygun olarak parametrelerde çeşitli değişiklikler yapılmıştır. Her bir parametrenin sınıflandırma algoritması üzerinde farklı etkileri olmaktadır. Bu parametre değerlerine bağlı olarak sonuçlar Çizelge 5.1'de görülmektedir. Yapılan deneyler neticesinde önerilen yöntem ile en iyi başarı oranı %67,79

olarak elde edilmiştir.

Çalışmada Parkinson çoklu ses kaydetme veri kümesinde herhangi bir dönüşüm işlemi gerçekleştirilmeden sınıflandırma başarımı %59,71 olarak elde edilmiştir. Aynı veri kümesinde verileri yeni bir dönüşüm uzayına aktarıp daha iyi bir dönüşüm uzayı elde edilmesini amaçlayan PCA yöntemi de analiz edilmiştir. PCA dönüşümü sonrasında yine k en yakın komşu yöntemi kullanılmıştır. Gerçekleştirilen deneylerde belirlenen sabit indekslerle aynı başarımlar elde edilip sınıflandırma işleminde bir artış gerçekleşmemiştir. Karşılaştırmalı olarak Çizelge 5.2’de görülen sonuçlarda ise hem temel kNN hem de PCA dönüşümü uygulandıktan sonra kNN ile sınıflandırmadan daha iyi bir başarımlar elde edildiği görülmekte ve sonuç olarak yöntemin başarısı kanıtlanmaktadır. Sonuç olarak Parkinson konuşma veri kümesinde başarı oranı %59,71’den %67,79’a yükselmiştir. Ayrıca sınıflandırma başarımının maksimum olduğu en iyi W dönüşüm matrisi negatif tanımlı bir matris olarak elde edilmiştir.

Çizelge 5.1. Deneysel sonuçlar

Başarı Oranı	İterasyon Sayısı	Birey Sayısı	Elit Birey Sayısı	Yeni Rasgele Birey Sayısı
66,08	100	1000	10	100
66,92	100	1000	20	100
66,78	100	1000	30	100
67,36	100	1000	50	100
67,76	100	2000	50	200
67,52	100	1000	50	100
67,79	100	500	50	50

Çizelge 5.2. Karşılaştırmalı sonuçlar

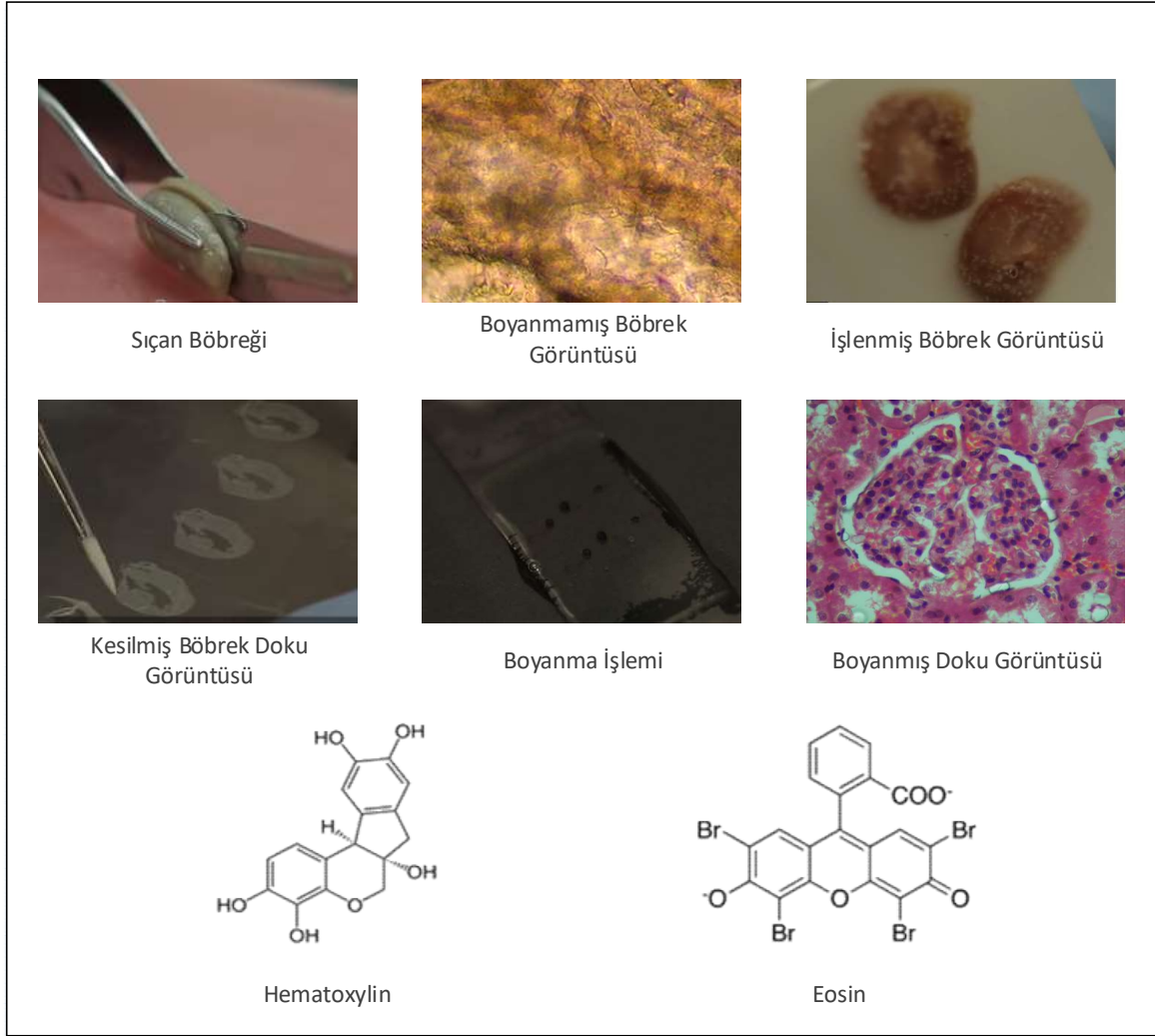
Yöntem	Başarı Oranı
kNN	%59,71
PCA+kNN	%59,71
Önerilen Yöntem	%67,79

6. HİSTOPATOLOJİK GÖRÜNTÜLERİN DERİN METRİK ÖĞRENME İLE SINIFLANDIRILMASI

6.1. H & E Patolojik Görüntü

Patolojik bir görüntüyü boyama yapmadan mikroskopik olarak incelemek donuk gri görüntülerden farksız bir görüntü ortaya çıkaracaktır. Bu durum istenmeyen bir durum olup patolojik görüntüden anlamlı örüntüler çıkarabilmek için öncelikli olarak bir boyama işleminden geçirilmesi gerekmektedir. Bu sebeple çok farklı sayıda boyama tekniği olmakla birlikte Hematoxylin ve Eosin kimyasalları kullanılarak gerçekleştirilen boyama tekniği, histolojide en temel boyama yöntemlerinden biridir. Patologlar H & E boyalama tekniğini kullanarak farklı hücre parçalarını birbirinden ayırt etmeyi amaçlamaktadırlar. H &E kimyasalının ile boyama sonucunda mikroskopik görüntüde üç ana renk oluşmaktadır. Bu renkler; beyaz, mavi ve pembedir. Hematoxylin bazik bir kimyasal olup asidik dokularla etkileşime girer. Örneğin patolojik bir görüntüdeki nükleik asitler asidik özelliğe sahiptir ve bu maddelerle etkileşiminden çekirdeklerin boyamasının mavi olarak gerçekleşmesini sağlayan Hematoxylin kimyasalıdır. Eosin ise asidik bir kimyasal olup negatif yüklüdür. Patolojik görüntüdeki pembe renkli kısımlar ise proteinlerdeki amino grupların pozitif yüklü olup, negatif yüklü olan Eosin ile birleşiminden oluşur [161].

H & E patolojik bir boya dokuma tekniği olup boya işlemi olmadan önce anlamsız olan görüntülerin patologlar tarafından daha iyi bir yorumlama sunmasını amaçlamaktadır. Bu teknikle hücredeki dokuların her biri daha ince detaylarla görüntülenmektedir. Mevcut halinde bir iyileşme yapılmasına karşın patologların uzman görüşlerinin yeterli geldiği yine de söylenemez. Sadece şekil üzerinden yapılan uzman görüşü değerlendirmelerinde de ancak %50 -75 arasında uzman görüşü tespitinin doğru bir şekilde yapılabilmesi mümkün olmaktadır [162, 163]. Patologlara yardımcı olmak için bilgisayar destekli tanı sistemleri ile zararlı ya da hasarlı görüntülerin tespiti mümkün olmaktadır.

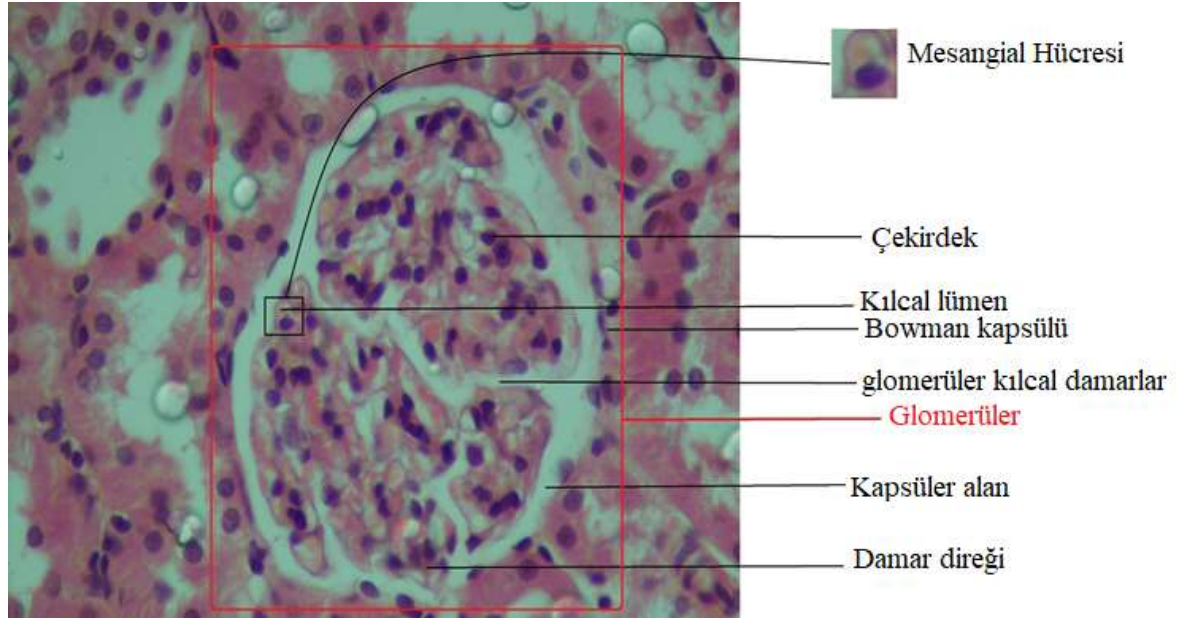


Şekil 6.1. Patolojik görüntü elde etme süreci

6.2. H & E Patolojik Görüntü Veri Kümesi

Bu çalışmada H & E kimyasalları ile boyanmış olan histopatolojik görüntüler kullanılmıştır [164]. Bu Histopatolojik görüntüler tez çalışması kapsamında elde edilen görüntüler olmayıp, özgün veri kümesi daha önce başka bir amaçla kullanılmıştır. Bu tez çalışmasında ise hazır veri kümesi derin metrik öğrenme yaklaşımlarıyla deneysel çalışmalarda kullanılmıştır. Histopatolojik görüntülerde iki grup bulunmaktadır. İlk grup sağlıklı gruptur. İkinci grup, 100 mg / kg / gün intraperitoneal Alüminyum Klorür ($AlCl_3$) enjekte edilen hasarlı gruptur. Elde edilen bir histopatolojik görüntü, mavi, beyaz ve pembe bileşenlerini içermektedir. Mavi bileşenler, hücre çekirdeğini temsil ederken beyaz ve pembe bileşenler, diğer doku yapılarını temsil etmektedir. Veri kümesinde 360 hasarlı ve 779 sağlıklı görüntü bulunmaktadır. Görüntüler 2560×1920 piksel boyutlarında RGB

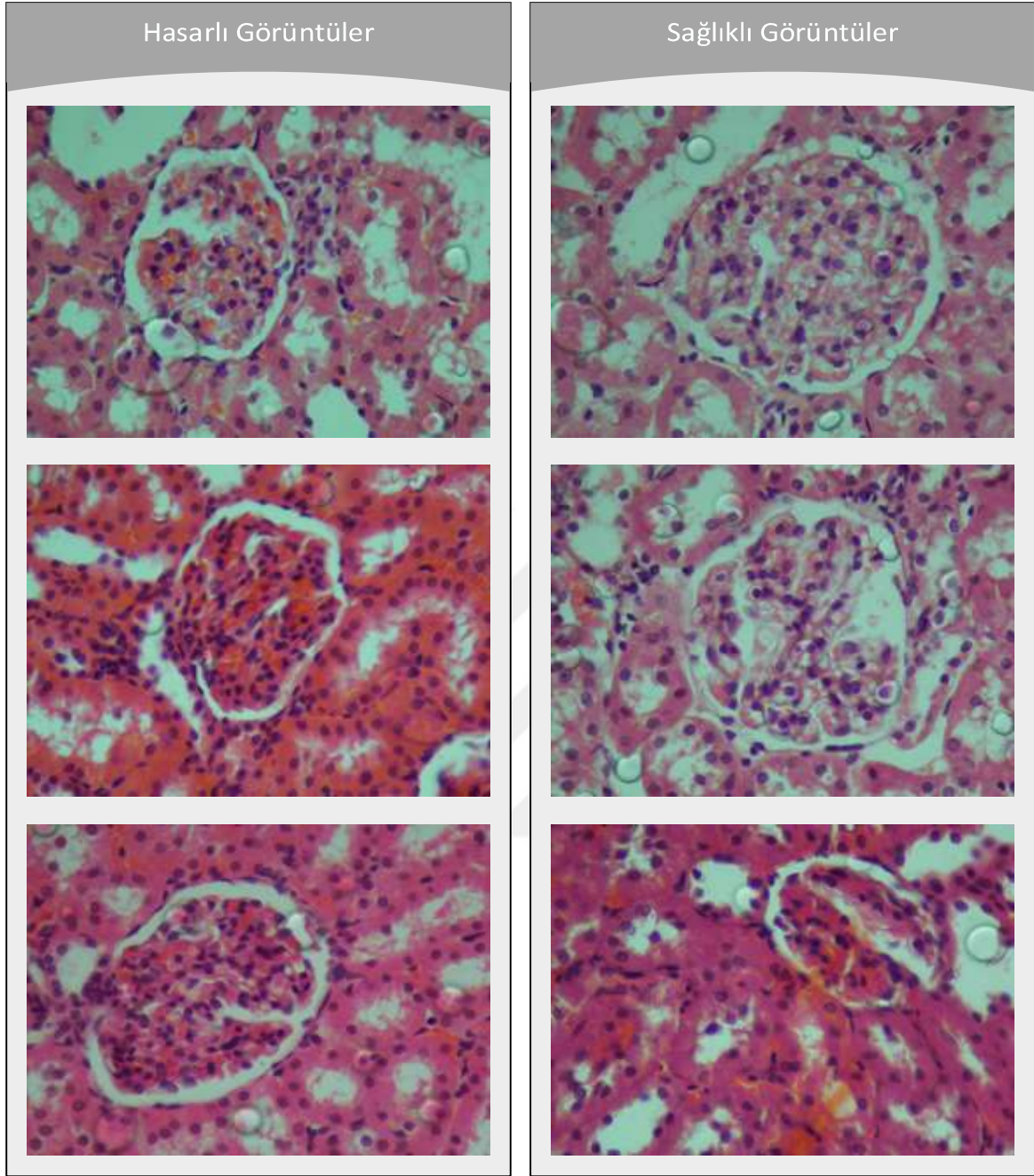
renk uzayına sahiptir. Bir histopatolojik görüntü ve bu histopatolojik görüntüde bulunan bazı bileşenler Şekil 6.2’de görülmektedir.



Şekil 6.2. Bir patolojik görüntü

6.3. Derin Metrik Öğrenme ile Histopatolojik Görüntüleri Sınıflandırma

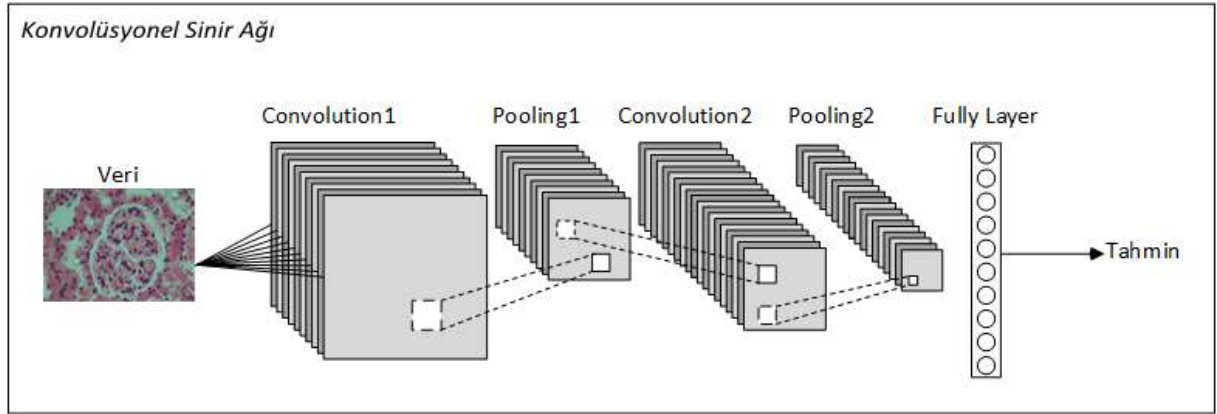
Patolojik veri kümesine ait olan hasarlı ve sağlıklı görüntülere ait örnek görüntüler Şekil 6.3’de görülmektedir. Görüntülerden de görüleceği üzere hasarlı ve sağlıklı veriler birbirine oldukça çok benzemektedir. Uzman görüşleriyle tespitin ne derece zor olduğu da mevcut görüntülerden anlaşılmaktadır.



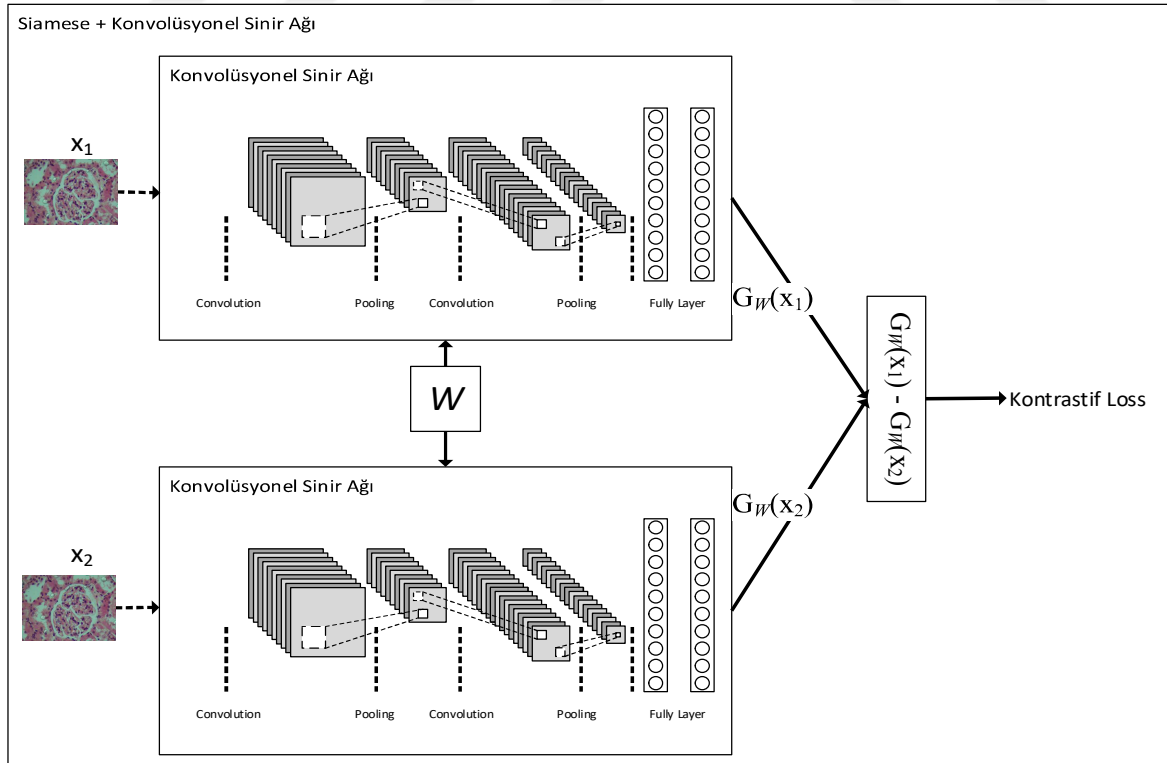
Şekil 6.3. Patolojik veri kümesi örnekleri

Şekil 6.4’te evrişimsel bir sinir ağı modeli örneği verilmektedir. Evrişimsel sinir ağlarından görüntüler her adımda bir görüntü alınacak şekilde ağı sunulmaktadırlar. Her bir görüntü evrişim katmanı, örnekleme katmanı, tam bağlı katman ve SoftMax sınıflandırma süreçlerinden geçerek bir sınıflandırma işlemi yapılmaktadır. Derin metrik öğrenmede ise süreç biraz daha farklıdır. Görüntüler ağı paylaşımlı ağırlıklılıklarla birden fazla görüntüler şeklinde sunulurlar. Yine evrişim katmanı, örnekle katmanı ve tam bağlı katman gibi yaklaşımları içermektedir. Son adımda ise iki görüntünün birbirine benzerliğine

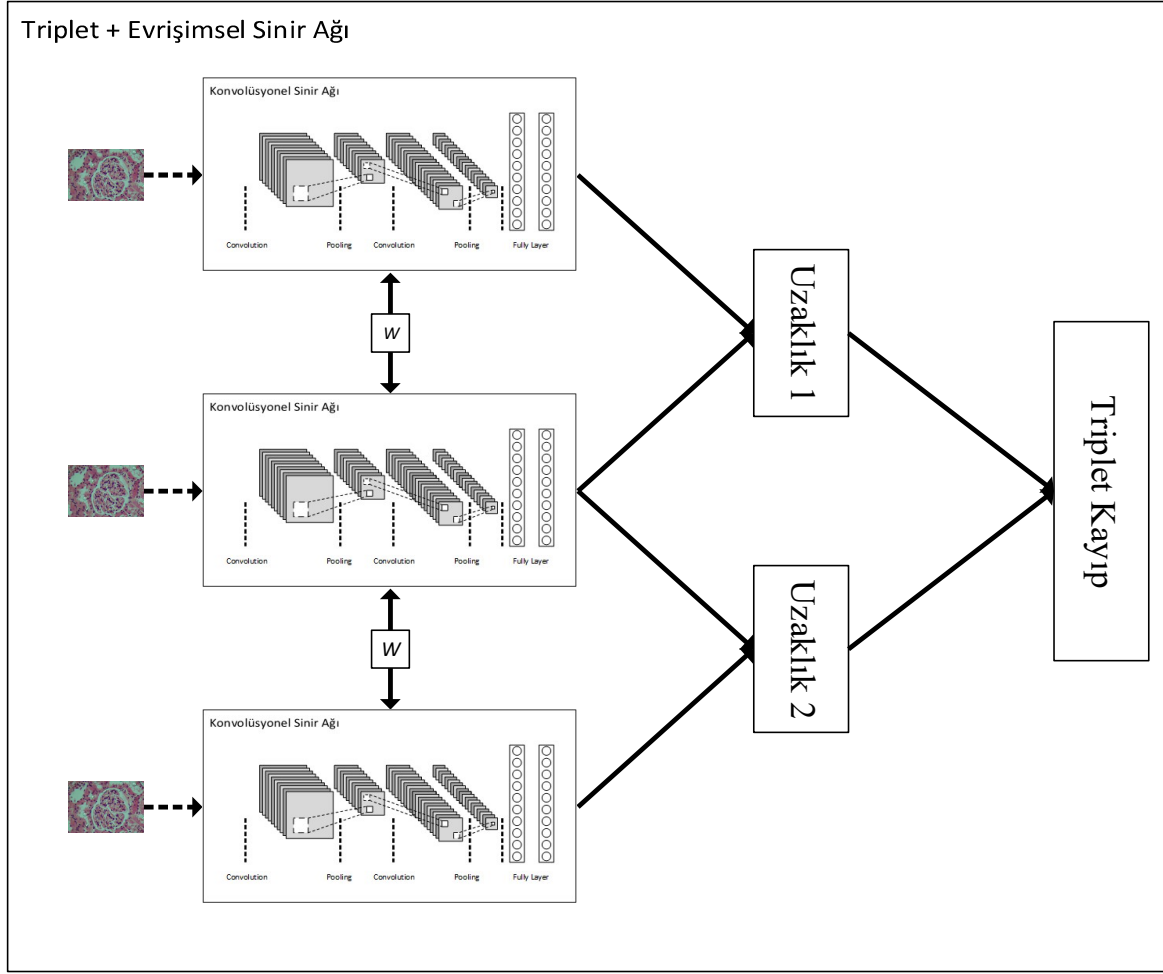
odaklanmaktadır. Benzerlik tespiti için ise kayıp fonksiyonları kullanılmaktadır. Bu şekilde bir yaklaşımı ifade eden Siamese ve evrişimsel sinir ağının bir arada kullanıldığı örnek bir model Şekil 6.5'te görülmektedir. Şekil 6.5'te önerilen ağ modelinin çıkışında tam bağlı katmanda 128 nöron bulunmaktadır. $G_W(x_1)$ ve $G_W(x_2)$ 128 boyutlu birer vektör dizisidir. Bu vektör dizilerinin birbirine olan uzaklığı kontrastif kayıp fonksiyonunda kullanılmaktadır



Şekil 6.4. Evrişimsel sinir ağı modeli



Şekil 6.5. Siamese + Evrişimsel sinir ağı

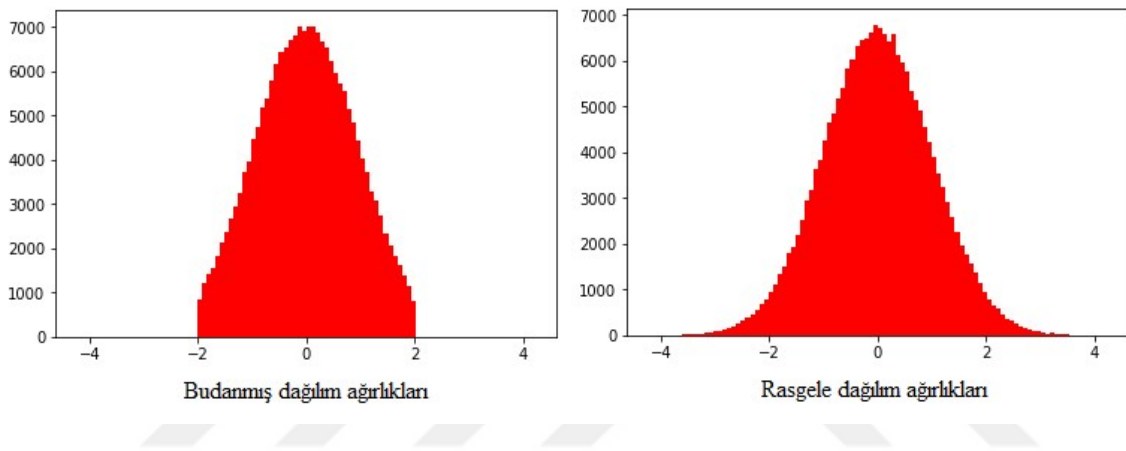


Şekil 6.6. Triplet + Evrişimsel sinir ağı modeli

Derin öğrenmede ağıın başarısında birçok parametre söz konusudur. Bu amaçla bunların her biri ele alınmıştır. Örneğin ağıın başlangıç değerlerinde kullanılan fonksiyon ve başlangıç değerlerin oluşum biçimi (ortalama, standart sapma), değerlerin budanması vs. gibi parametreler dikkate alınarak bazı Siamese ağ modellerinde ağıın başlangıç ağırlıkları değerleri ortalaması 0 ve standart sapması 0.08 olacak şekilde ayarlanmıştır. Elbette yine bu değerler model değişimiyle birlikte yeniden ayarlanması gereken parametreler olacaktır.

Yapılan deneylerde ağıın başlangıç performansını etkileyen ve en iyi (optimal) sonuca yakınsamasını etkileyen bir diğer faktör olan budanmış ağırlık değerlerinin alınmasının ağ performansı açısından daha uygun olduğu sonucuna varılmıştır. Ağıın başlangıç değerlerinin 0'a yakın olması ağıın performansı üzerinde oldukça etkilidir. Bu yapılan deneylerle de görülmüş ve budanmış dağılım ile ağıın performansında daha iyi sonuçlar elde edilmiştir. Aşağıdaki şekilde budanmış ve rasgele dağılım örneğine ait histogramlar

görülmektedir. Budanmış dağılımı sayesinde özellikle ölü nöronların oluşmasının engellenmesi istenmektedir. Çünkü kullanılan aktivasyon fonksiyonunda değerler çok büyük ya da çok düşük olduğunda ağırların öğrenme süreci engellenmekte ve ağırların öğrenme eğilimi durmaktadır ve bu bizim için istenmeyen durumdur. Özellikle özgün veri kümesi üzerinde budanmış ağırlıkların önemi çok daha etkili olmuştur. Yapılan deneylerde başlangıç ağırlıklarının öneminin görülmesinden dolayı transfer öğrenmenin etkinliği göz önüne alınarak ağırların eğitimi için VGG-16, VGG-19 ve ResNet50 gibi ağırların modelleri kullanılmış ve bu ağırların modellerindeki başlangıç ağırlıkları da hesaba katılmıştır.



Şekil 6.7. Budanmış ve rasgele dağılım ağırlıkları

6.3.1. Çeşitli derin metrik öğrenme modelleri ile sınıflandırma

Histopatolojik veri kümesi 911 eğitim ve 228 test örneği olmak üzere başlangıçta rasgele olarak ayrılmıştır. 911 eğitim örneğinin 623'ü sağlıklı, 288'i hasarlı görüntüdür. 228 test örneğinin ise, 156'sı sağlıklı ve 72'si hasarlı görüntülerden oluşmaktadır. Patolojik veri kümesi öncelikle Şekil 6.5'teki ağırların modeli üzerinde denenmiştir. Bu ağırların modelinde kullanılan parametreler Çizelge 6.1'de görülmektedir. Yapılan deneylerde aktivasyon fonksiyonu olarak literatürde de sıklıkla tercih edilen ReLU tercih edilmiştir. 3 farklı optimizasyon yöntemi test edilmiştir. Deneysel sonuçlar neticesinde 3 optimizasyon yönteminde de benzer sonuçlar elde edilmiştir.

Çizelge 6.1. Parametreler

Veri Kümesi	Patolojik Veri
Eğitim Örneği Sayısı (1 adımdaki)	64
Görüntü Boyutu	$80 \times 60 \times 3$
Optimizasyon Yöntemleri	Adam Optimizer Gradient Descent RMSProp Optimizer
Aktivasyon Fonksiyonu	ReLU

Görüntü	Evrisim Maksimum Örnekleme	Evrisim Maksimum Örnekleme	Evrisim Maksimum Örnekleme	Evrisim Maksimum Örnekleme	Evrisim Maksimum Örnekleme	Tam Bağlı Katman	Tam Bağlı Katman
80x60x3	3x3 Filtre 16 Kanal	3x3 Filtre 32 Kanal	3x3 Filtre 64 Kanal	3x3 Filtre 128 Kanal	3x3 Filtre 256 Kanal	1024 Nöron	128 Nöron
80x60x3	3x3 Filtre 32 Kanal	3x3 Filtre 64 Kanal	3x3 Filtre 128 Kanal	3x3 Filtre 256 Kanal	3x3 Filtre 512 Kanal	1024 Nöron	128 Nöron
80x60x3	3x3 Filtre 64 Kanal	3x3 Filtre 128 Kanal	3x3 Filtre 256 Kanal	3x3 Filtre 512 Kanal		1024 Nöron	128 Nöron
80x60x3	3x3 Filtre 16 Kanal	3x3 Filtre 32 Kanal	3x3 Filtre 64 Kanal	3x3 Filtre 128 Kanal		1024 Nöron	128 Nöron
80x60x3	3x3 Filtre 32 Kanal	3x3 Filtre 64 Kanal	3x3 Filtre 128 Kanal	3x3 Filtre 256 Kanal		1024 Nöron	128 Nöron
80x60x3	3x3 Filtre 64 Kanal	3x3 Filtre 128 Kanal	3x3 Filtre 256 Kanal	3x3 Filtre 512 Kanal		512 Nöron	128 Nöron
320x240x3	3x3 Filtre 16 Kanal	3x3 Filtre 32 Kanal	3x3 Filtre 64 Kanal	3x3 Filtre 128 Kanal	3x3 Filtre 256 Kanal	512 Nöron	128 Nöron
320x240x3	3x3 Filtre 32 Kanal	3x3 Filtre 64 Kanal	3x3 Filtre 128 Kanal	3x3 Filtre 256 Kanal	3x3 Filtre 512 Kanal	512 Nöron	128 Nöron
320x240x3	3x3 Filtre 64 Kanal	3x3 Filtre 128 Kanal	3x3 Filtre 256 Kanal	3x3 Filtre 512 Kanal		1024 Nöron	128 Nöron
320x240x3	3x3 Filtre 16 Kanal	3x3 Filtre 32 Kanal	3x3 Filtre 64 Kanal	3x3 Filtre 128 Kanal		1024 Nöron	128 Nöron
320x240x3	3x3 Filtre 32 Kanal	3x3 Filtre 64 Kanal	3x3 Filtre 128 Kanal	3x3 Filtre 256 Kanal		1024 Nöron	128 Nöron
320x240x3	3x3 Filtre 64 Kanal	3x3 Filtre 128 Kanal	3x3 Filtre 256 Kanal	3x3 Filtre 512 Kanal		1024 Nöron	128 Nöron
80x60x3	5x5 Filtre 16 Kanal	5x5 Filtre 32 Kanal	5x5 Filtre 64 Kanal	5x5 Filtre 128 Kanal		1024 Nöron	128 Nöron
80x60x3	5x5 Filtre 32 Kanal	5x5 Filtre 64 Kanal	5x5 Filtre 128 Kanal	5x5 Filtre 256 Kanal		1024 Nöron	128 Nöron
80x60x3	5x5 Filtre 64 Kanal	5x5 Filtre 128 Kanal	5x5 Filtre 256 Kanal	5x5 Filtre 512 Kanal		1024 Nöron	128 Nöron
80x60x3	5x5 Filtre 16 Kanal	5x5 Filtre 32 Kanal	5x5 Filtre 64 Kanal	5x5 Filtre 128 Kanal		1024 Nöron	128 Nöron
80x60x3	5x5 Filtre 32 Kanal	5x5 Filtre 64 Kanal	5x5 Filtre 128 Kanal	5x5 Filtre 256 Kanal		1024 Nöron	128 Nöron
80x60x3	5x5 Filtre 64 Kanal	5x5 Filtre 128 Kanal	5x5 Filtre 256 Kanal	5x5 Filtre 512 Kanal		1024 Nöron	128 Nöron
80x60x3	5x5 Filtre 16 Kanal	5x5 Filtre 32 Kanal	5x5 Filtre 64 Kanal	5x5 Filtre 128 Kanal		1024 Nöron	128 Nöron
80x60x3	5x5 Filtre 32 Kanal	5x5 Filtre 64 Kanal	5x5 Filtre 128 Kanal	5x5 Filtre 256 Kanal		1024 Nöron	128 Nöron
80x60x3	5x5 Filtre 64 Kanal	5x5 Filtre 128 Kanal	5x5 Filtre 256 Kanal	5x5 Filtre 512 Kanal		1024 Nöron	128 Nöron

Şekil 6.8. Denenmiş ağ modelleri

Özgün histopatolojik veri üzerinde Şekil 6.8’de görülen çeşitli ağ modelleri eğitim süreçlerinde farklı ağ modelleri denenmiş ve sonuçlar irdelenmiştir. Bu ağ modellerinin değerlendirmesi için test veri kümesindeki sonuçlara odaklanılmıştır. Şekil 6.5’te görüldüğü üzere Siamese ağ modelinde veriler ikili görüntüler eşliğinde ağı sunulmaktadır. Ağ modelinde görüntülerin eğitimi süreci öncesinde ikili görüntüler oluşturulurken dengeli bir sınıf dağılımı hedeflenmiştir. Bu ikili görüntüler oluşturulurken, minimum ikili maksimum fayda prensibiyle 911 eğitim görüntüsünden 1148 tane ikili görüntü oluşturulmuştur. Oluşturulan görüntülerde hem aynı sınıfların birbiriyle hem de farklı sınıfların birbiriyle eşleşmesi dengeli bir şekilde oluşturulmuştur. Tüm eğitim örnekleri en az 1 defa ağı sunulan ikililerin içerisinde yer almıştır. Yapılan çalışmalarda ağı sunulacak görüntüler rasgele seçilmiş olup ağın performansı bu şekilde değerlendirmeye alınmıştır. Ancak her bir adımda ağ modelinde kullanılacak olan ikili görüntülerin dengeli bir şekilde yer alması hedeflenmiştir. Bu amaçla her bir sınıf için kendi sınıfından ve diğer sınıftan olmak üzere görüntülerin ağıda yer alması hedeflenmiştir.

Çizelge 6.2 Siamese eğitim ağı örnekleri görülmektedir. Burada alınan ikili görüntülerden; aynı sınıftan örnekler bir arada alındığında örneklerin birbirine yaklaştırılması hedeflenirken; farklı sınıftan örnekler bir arada alındığında ise örneklerin birbirinden uzaklaştırılması hedeflenmektedir. Eğitim örneklerinin başarısı hedeflenirken buna odaklanırken; test işleminde ise daha farklı bir bakış açısı söz konusudur. Burada problemi bir sınıflandırma problemi olarak ele almak istediğimizde; her bir test örneği bir eğitim örneği ile karşılaştırılmaktadır (Çizelge 6.3). Test örneğinin X_1 veya X_2 ’ye yerleştirilmesinin bir önemi yoktur. Test aşamasında da çift taraflı aynı ağırlıklar kullanılmaktadır. Yapılan deneylerde 228 test görüntüsünün her biri en az 1 defa kullanılarak eğitim görüntüleriyle eşleştirilmiştir. Bu 228 test görüntüsünden 4560 tane farklı ikili test görüntüsü oluşturulmuştur. Sınıflandırma performansında 4560 ikili görüntünün ortalamasına bakılmıştır. Algoritmanın ağ doğruluğu hesabında da aşağıdaki denklemden yararlanılmaktadır.

$$(\text{etiket} \times (\text{tahmin} < \text{eşik değeri})) + (1 - \text{etiket}) \times (\text{tahmin} > \text{eşik değeri}) \quad (6.1)$$

Çizelge 6.2. Siamese eğitim ağı örnekleri

X_1	X_2
Örnek 5 (Sınıf1)	Örnek8 (Sınıf1)
Örnek 68 (Sınıf1)	Örnek 93 (Sınıf2)
Örnek 36 (Sınıf2)	Örnek 49 (Sınıf2)
Örnek 63 (Sınıf2)	Örnek 27 (Sınıf1)

Çizelge 6.3. Siamese test ağı örnekleri

X_1	X_2
Test Örnek 1 (Sınıf1)	Eğitim Örnek8 (Sınıf1)
Test Örnek 5 (Sınıf1)	Eğitim Örnek 93 (Sınıf2)
Test Örnek 78 (Sınıf2)	Eğitim Örnek 49 (Sınıf2)
Test Örnek 36 (Sınıf2)	Eğitim Örnek 27 (Sınıf1)

Çizelge 6.4 Triplet eğitim ağı örnekleri görülmektedir. Triplet eğitim örnekleri oluşturulurken her bir eğitim örneği en az 1 defa kullanılmak koşuluyla 911 eğitim görüntüsünden 2000 farklı üçlü görüntü eğitim amacıyla kullanılmıştır. Burada alınan üçlü görüntülerden; çapa ile aynı sınıftan olan pozitif örnekler bir arada alındığında örneklerin birbirine yakınlştırılması hedeflenirken; çapa ile farklı sınıftan örnekler bir arada alındığında ise örneklerin birbirinden uzaklaştırılması hedeflenmektedir. Eğitim örneklerinin başarısı hedeflenirken buna odaklanırken; test işleminde ise daha farklı bir bakış açısı söz konusudur. Burada problemi bir sınıflandırma problemi olarak ele almak istediğimizde; her bir test örneği bir çapa örneğine eşitlenmektedir (Çizelge 6.5). Test sürecinde 228 test örneğinin her biri çapaya yerleştirilmiş, diğer pozitif ve test örnekleri ise rasgele olarak eğitim kümesi örneklerinden alınarak karşılaştırılma yapılmıştır. Karşılaştırılma sonucunda çapa örneği hangi örneğe daha yakınsa veri o örneğin etiketiyle

tahmin edilmektedir. Test doğruluğu da test verisinin gerçek etiketiyle kıyaslanması sonucunda elde edilmektedir.

Çizelge 6.4. Triplet eğitim ağı örnekleri

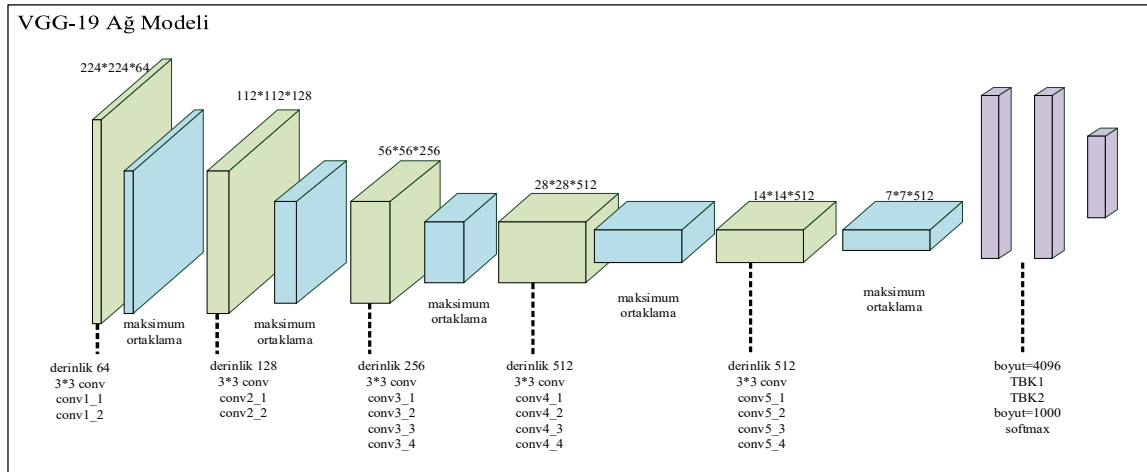
Çapa	Pozitif	Negatif
Örnek 43 (Sınıf1)	Örnek 28 (Sınıf1)	Örnek87 (Sınıf2)
Örnek 56 (Sınıf1)	Örnek 5 (Sınıf1)	Örnek 93 (Sınıf2)
Örnek 4 (Sınıf2)	Örnek 97 (Sınıf2)	Örnek 49 (Sınıf1)
Örnek 26 (Sınıf2)	Örnek 12 (Sınıf2)	Örnek 27 (Sınıf1)

Çizelge 6.5. Triplet test ağı örnekleri

Çapa	Pozitif	Negatif
Test Örnek 0 (Sınıf1)	Eğitim Örnek 81 (Sınıf1)	Eğitim Örnek 9 (Sınıf2)
Test Örnek 26 (Sınıf1)	Eğitim Örnek 53 (Sınıf1)	Eğitim Örnek 38 (Sınıf2)
Test Örnek 41 (Sınıf2)	Eğitim Örnek 34 (Sınıf2)	Eğitim Örnek 186 (Sınıf1)
Test Örnek 78 (Sınıf2)	Eğitim Örnek 11 (Sınıf2)	Eğitim Örnek 13 (Sınıf1)

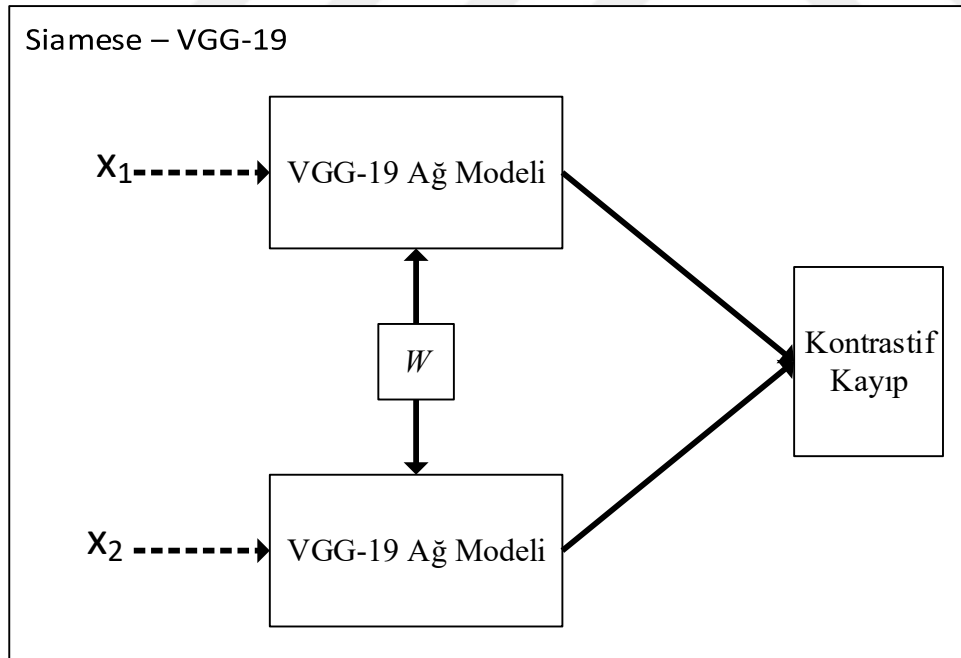
6.3.2. Rekabetçi ağ modellerini kullanarak derin metrik öğrenme modelleri ile sınıflandırma

Yapılan farklı deneylerde çeşitli ağ modelleri denenerek belirli başarılar elde edilmiştir (Şekil 6.8). Daha sonra ağ eğitimini gerçekleştirmek için rastgele bir ağ modelini kullanmak yerine görüntü veri kümelerinde başarısını kanıtlamış bilinen ağ modelleri üzerinde deneyler yapılmıştır. Örnek bir yaklaşım olarak derin metrik öğrenme problemi aşağıdaki Şekil 6.9’da görülen VGG-19 ağ modelinden yararlanılmaktadır.

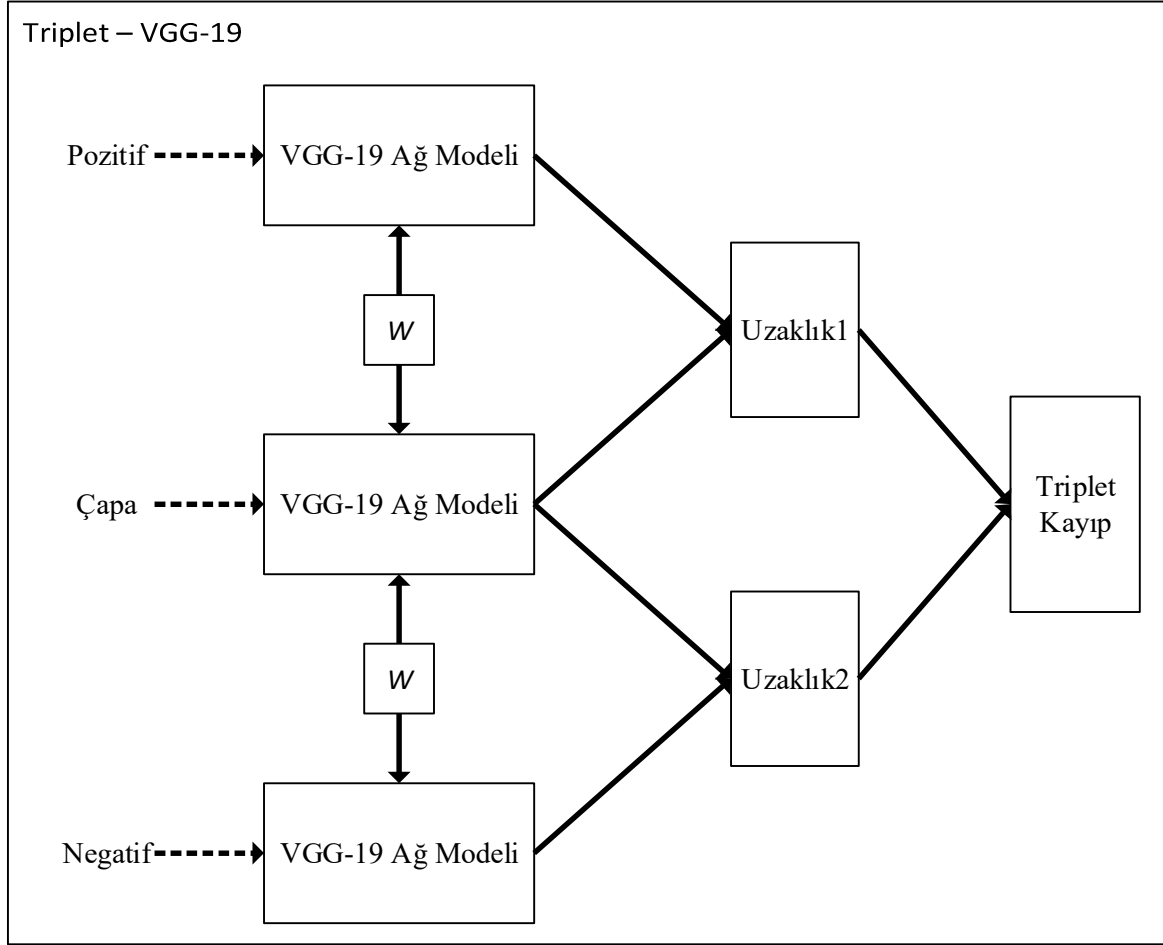


Şekil 6.9. VGG-19 Ağ modeli

Şekil 6.10’da görülen Siamese ağ modelinde ve Şekil 6.11’de görülen Triplet ağ modelinde VGG-19 ağ modeli bu ağ modellerine yerleştirilmiştir. Ağ modeline aynı anda birden fazla görüntüler sunularak, ağ modeli içerisinde aynı anda sunulan görüntüler için paylaşılan (aynı) ağırlıklar kullanılmıştır.



Şekil 6.10. Siamese VGG-19 ağ modeli



Şekil 6.11. Triplet VGG -19 ağ modeli

Eğitim aşamasında VGG-19 ağ modelinin hem gücü hem de daha önceden çok fazla örnekten oluşan bir eğitim kümesinden elde edilen ağırlıkları, başlangıç ağırlıkları olarak veri kümesinin eğitimi için kullanılmıştır. Burada ağ modelinin başarısı ile ağıın daha önceden eğitilmiş başlangıç ağırlıklarından faydalanılmıştır.

Ağın eğitim ve test örnekleri oluşturulurken daha önceki bölümde kullanılan yaklaşım üzerinden gidilmiş ve ağın başarısında test örneklerinin doğruluğu sınıflandırma başarısı olarak değerlendirmeye alınmıştır.

6.3.3. Deneysel sonuçlar

Derin metrik öğrenme ile Histopatolojik verilerin sınıflandırılması sürecinde Siamese ve Triplet ağ modelleri ile evrişimli sinir ağı modelleri birleştirilmiştir. Yapılan çalışmalarda farklı farklı ağ modellerinde Siamese ve Triplet modelleri denenmiştir. Bu deneysel

çalıřmalarda çok farklı parametreler, farklı sayıda evriřimli ve maksimum örnekleme katmanları, tam baėlı katman modelleri denenmiřtir. Bu modeller dıřında rekabetçi aė modelleri (VGG19 gibi) Siamese ve Triplet aėlarıyla bütünleřtirilmiř ve sınıflandırma amacıyla kullanılmıřtır. Bunun dıřında bu aė modellerinin daha önceden farklı bir veri kümesi üzerinde başarılı bir řekilde elde edilen bařlangıç aėlıkları yoluyla da transfer öğrenme ile aėın bařarısının arttırılması hedeflenmiřtir.

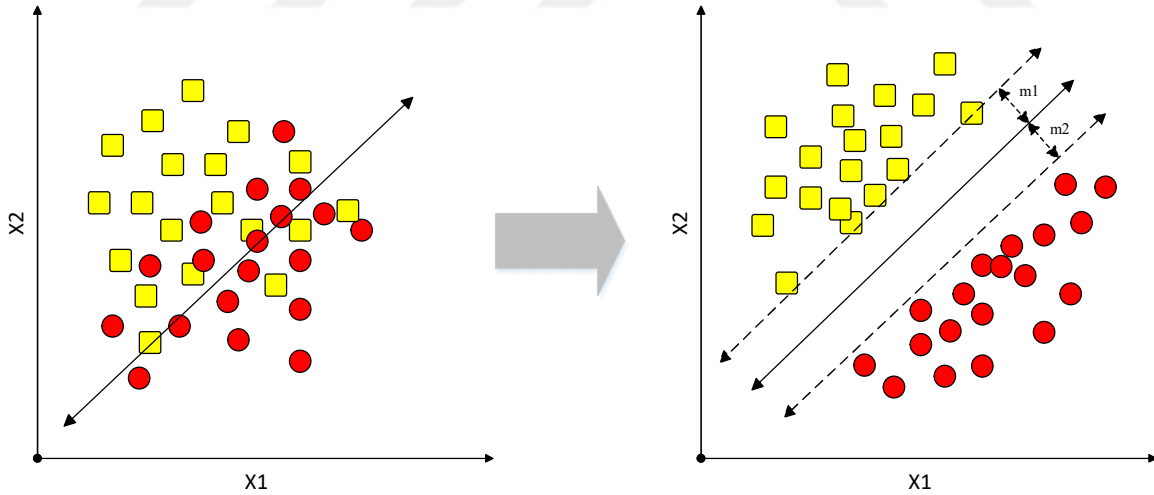
Siamese aė modeli verileri sınıflandırırken kayıp fonksiyonu yapısı gereėince aynı sınıftan örnekleri bir noktada bütünleřtirmeye odaklanmaktadır. Aynı zamanda da farklı sınıftan verileri birbirinden uzaklařtırmayı hedeflemektedir. Ancak Triplet aėından farklı olarak aynı sınıftan örnekleri kayıp fonksiyonunda yalnızca bir arada deėerlendirme imkanına sahip olduėundan dolayı benzer sınıf örneklerini daha dar bir küme içerisine yerleřtirmektedir. řekil 6.13 Siamese eğitim sonuçları ve řekil 6.14 Siamese test sonuçları bunun açık bir örneėidir. Bu řekiller aynı zamanda verinin sayısal sonuçları dıřında aynı zamanda görsel olarakta modelin bařarısını ortaya koymaktadır.

řekillerde elde edilen görsellerdeki her bir nokta Siamese ve VGG-19 aė modelinin sonucunda elde edilen iki boyutlu özniteliklerin gösterimidir. Eğitim görsellerindeki her bir nokta eğitim sürecinde elde edilen aėlıklardan kullanılması sonucunda 911 eğitim görüntüsünün her birinin iki boyutlu yeni temsilini ifade etmektedir. Her bir periyot (epoch) ařamasında verilerin nasıl birbirinden ayrıřtıėı görölmektedir. Test görsellerindeki her bir nokta ise eğitim sürecinde elde edilen aėlıklardan kullanılması sonucunda 228 test görüntüsünün her birinin iki boyutlu yeni temsilini ifade etmektedir. Her bir periyot (epoch) sonrasında eğitim sürecinde elde edilen aėlıklar kullanılarak test verilerinde farklı sınıflar birbirinden başarılı bir řekilde ayrıřmaktadır. Siamese aė modeli farklı sınıf verilerini ele alırken uzaklařtırma yaparken, aynı sınıf verilerini ele alırken ise dar bir kümeye yerleřtirerek benzerliėi arttırmaktadır.

Triplet aė modelinde veriler birbirinden ayrıřtırılırken aynı destek vektör makinelerinde olduėu gibi marjin ile sınıflar arası ayırım yapılmaktadır. řekil 6.12'de bu durumu ifade eden görsel görölmektedir. Bu görselden de anlařılacaėı üzere bir marjin vasıtasıyla farklı sınıflar birbirinden ayrılmaktadır. řekillerde elde edilen görsellerdeki her bir nokta Triplet ve VGG-19 aė modelinin sonucunda elde edilen iki boyutlu özniteliklerin gösterimidir. Derin metrik öğrenmenin temel amacı, farklı sınıfları mümkün olduėu kadar birbirinden

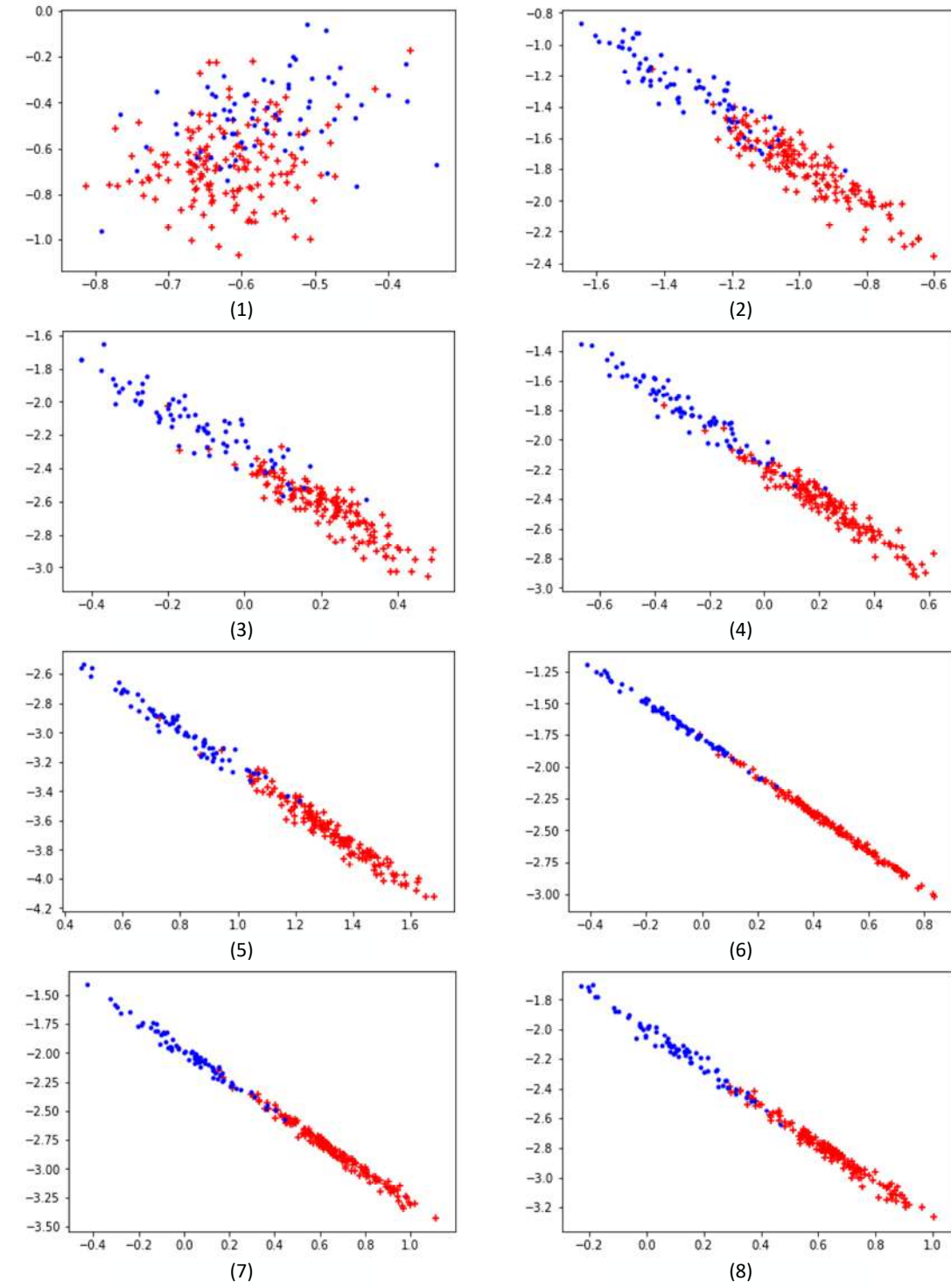
ayırmak ve aynı sınıf örnekleri arasında benzerliği arttırmaktır. Bu yaklaşımın doğru bir şekilde analiz edilmesi için görseller bize daha iyi bir projeksiyon sunmaktadır. Destek vektör makinelerindeki gibi bir ayırım yaklaşımı derin metrik öğrenme yöntemlerinden Triplet ağ modeliyle benzerlik göstermektedir.

Triplet ağ modelinde Siamese ağ modelinden farklı olarak benzer sınıflar dar bir kümeye sıkıştırılmak yerine sınıflar arasında marjin esaslı bir ayırım hedeflenmektedir. Şekil 6.15 Triplet eğitim sonuçları ve Şekil 6.16 Triplet test sonuçlarında bu durum açık bir şekilde görülmektedir. Triplet ağ modelinde ağın her bir eğitim periyodunda (epoch) elde ettiği ağırlıklar kullanılarak ağın eğitim süreci analiz edilmiştir. Eğitim görsellerindeki her bir nokta eğitim sürecinde elde edilen ağırlıkların kullanılması sonucunda 911 eğitim görüntüsünün her birinin iki boyutlu yeni temsilini ifade etmektedir. Test görsellerindeki her bir nokta ise eğitim sürecinde elde edilen ağırlıkların kullanılması sonucunda 228 test görüntüsünün her birinin iki boyutlu yeni temsilini ifade etmektedir. Görsel sonuçlarla da histopatolojik görüntülerin derin metrik öğrenme modelleriyle başarılı bir şekilde sınıflandırılabilirliği görülmektedir.

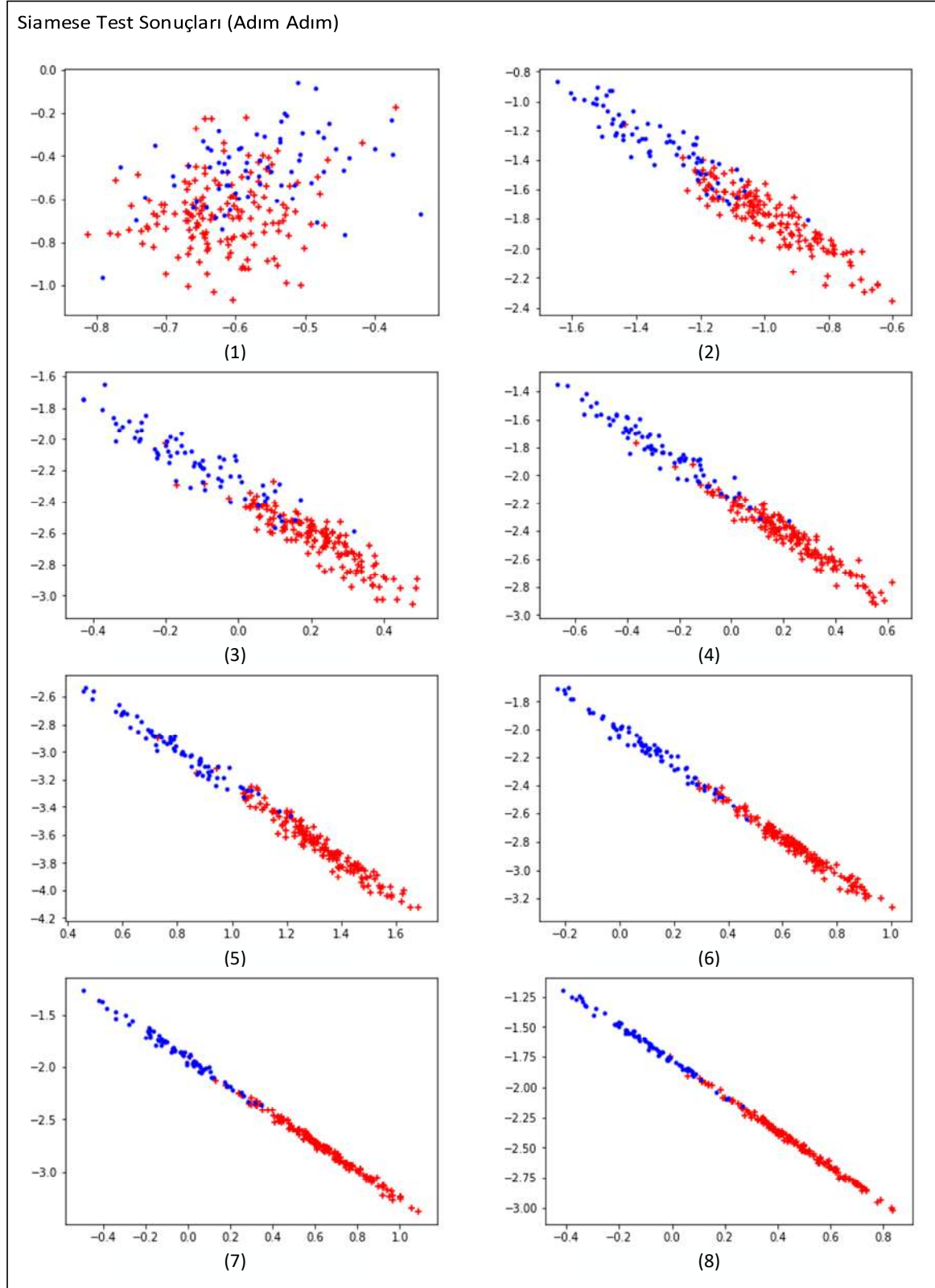


Şekil 6.12. Destek vektör makinesi marjin ayırımı

Siamese Eğitim Sonuçları (Adım Adım)

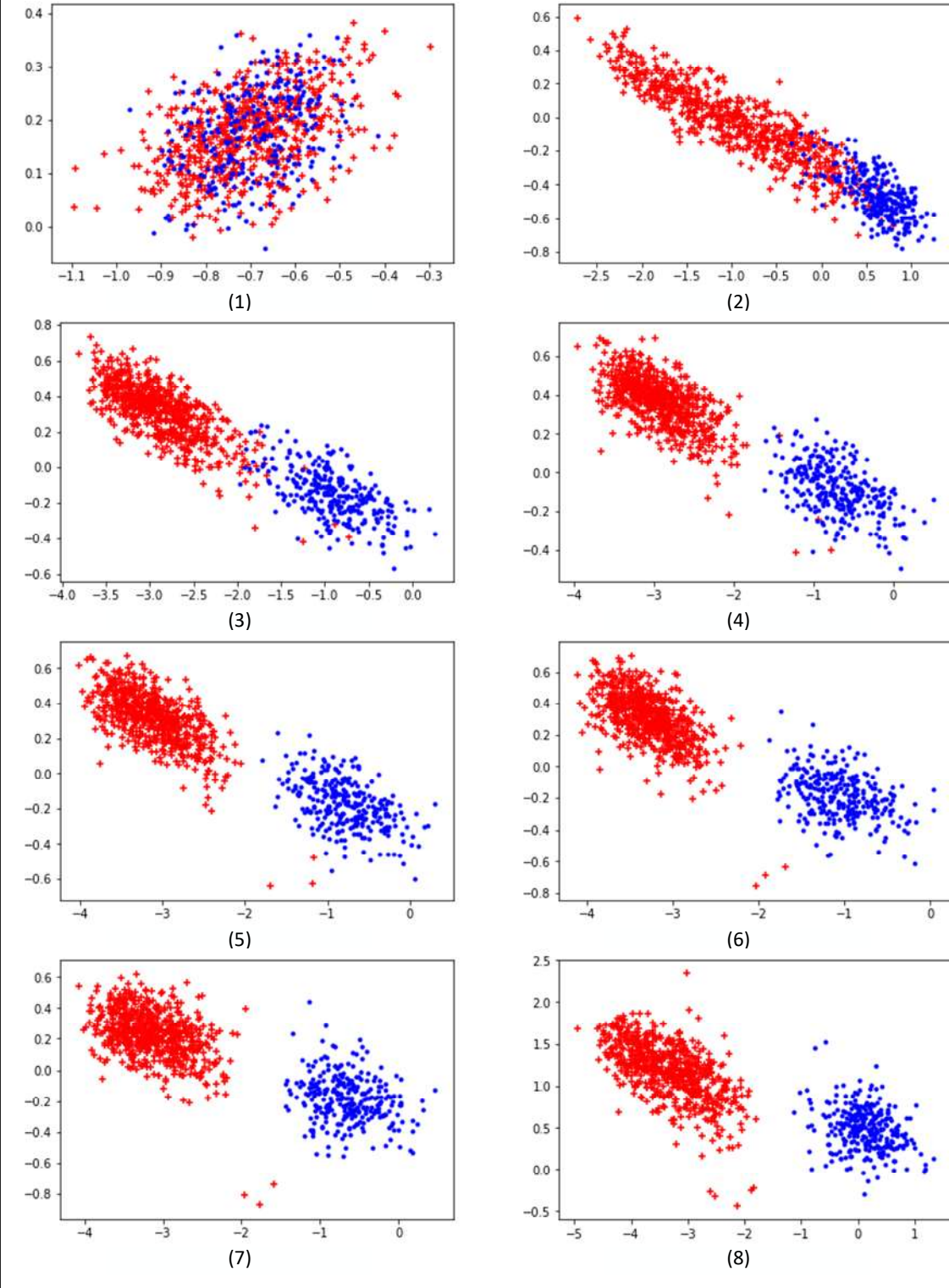


Şekil 6.13. Siamese eğitim sonuçları

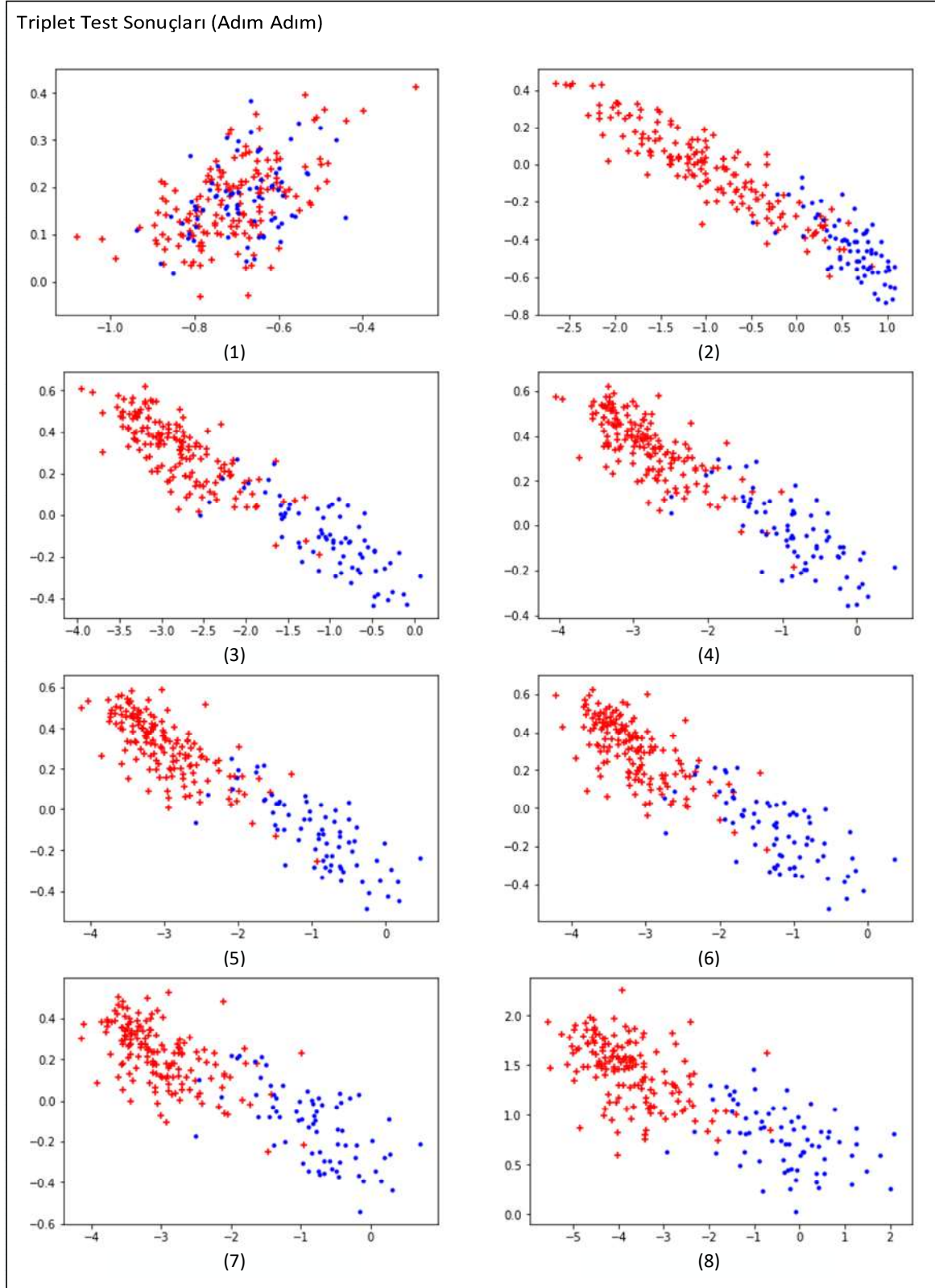


Şekil 6.14. Siamese test sonuçları

Triplet Eğitim Sonuçları (Adım Adım)



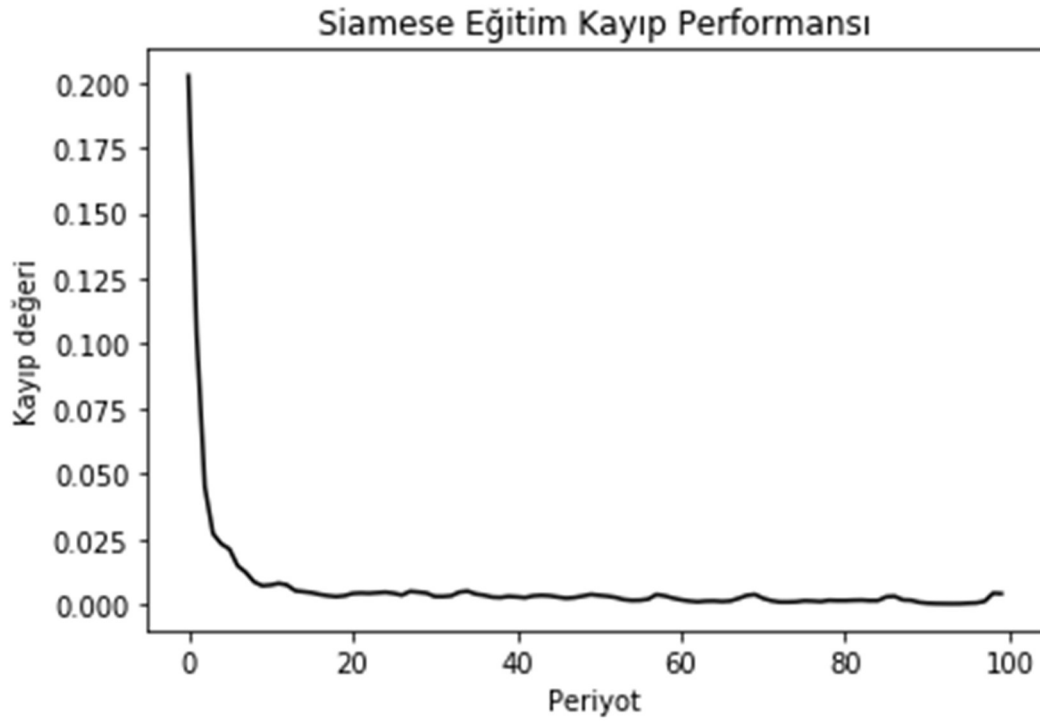
Şekil 6.15. Triplet eğitim sonuçları



Şekil 6.16. Triplet test sonuçları

Şekil 6.17’de Siamese+VGG19 ağ modelinin her bir periyot (epoch) süreci sonucunda oluşan kayıp değerinin ortalama sonuçları görülmektedir. Ağ modelinde önceden eğitilmiş

ağırlıklar kullanıldığından yakınsamanın hızlı bir şekilde gerçekleştiği açıktır. Ağıın eğitim sürecini grafik üzerinden analiz ettiğimizde 20. periyot (epoch) ile beraber ağıın belirli bir öğrenme sağladığı sonucuna varmaktayız. Sonrasında ağıın eğitim sürecinde parametre iyileştirmeleri sonucunda minimum değışiklikler olmaktadır.



Şekil 6.17. Siamese eğitim kayıp performansı

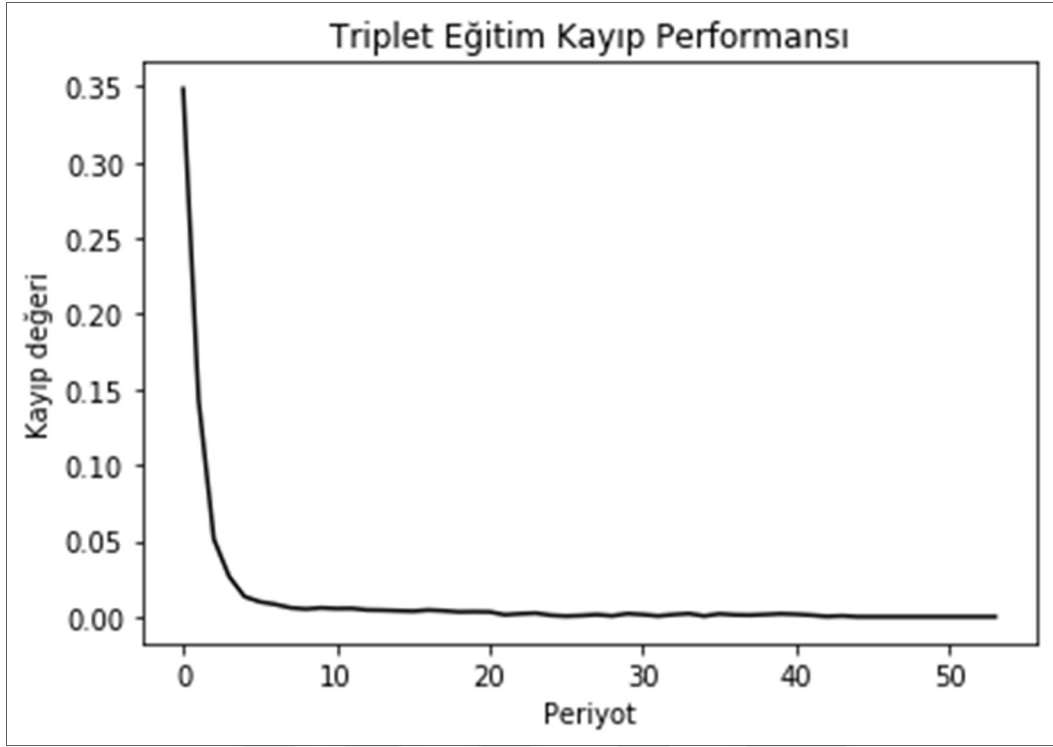
Çizelge 6.6’da Siamese ağ modeli için farklı ağ modelleri sonuçları görölmektedir. VGG-19 ağ modeli derin metrik öğrenme ile histopatolojik verileri sınıflandırmada özgün veri kümesinde oldukça başarılı sonuçlar elde etmiştir. Önerilen ağ modelinde önceden eğitilmiş ağırlıklar söz konusu olmadığından rasgele ağırlık değeriyle ağ eğitim sürecine başlamıştır. Sonuçlara analiz ettiğimizde VGG-19 ağ modeli ile %95,39 gibi yüksek bir başarı oranı elde edilmiştir. Dikkat çekici bir diğer husus olarak hem VGG-19 hem de MobileNet50 ağ modellerinde önceden farklı bir veri kümesi üzerinde eğitilmiş ağırlıklar ile daha başarılı sonuçlar elde edildiği sonucuna varılmıştır. Transfer öğrenmenin gücü burada açık bir şekilde görölmektedir.

Çizelge 6.6. Siamese sonuçları

	Siamese Ağ Modeli Sonuçları			
	<i>Önceden eğitilmiş ağırlıkları kullanmadan eğitime başlama</i>		<i>Önceden eğitilmiş ağırlıklarla eğitime başlama</i>	
	Periyot (Epoch)	Doğruluk	Periyot (Epoch)	Doğruluk
Önerilen Model	130	81,03	-	-
VGG-19	39	87,82	21	95,39
MobileNet50	18	82,78	26	85,41

Şekil 6.18’de Triplet+VGG19 ağ modelinin her bir periyot (epoch) süreci sonucunda oluşan kayıp değerinin ortalama sonuçları görülmektedir. Triplet ağ modelinde, Siamese ağ modeline göre ağ performansının maksimum seviyeye erişiminin daha kısa sürede gerçekleştiği görülmektedir. Ağın eğitim sürecini grafik üzerinden analiz ettiğimizde 6. periyot (epoch) ile beraber ağın belirli bir öğrenme sağladığı sonucuna varmaktayız. Sonrasında ağın eğitim sürecinde parametre iyileştirmeleri sonucunda minimum değişiklikler olmaktadır.

Çizelge 6.7’yi analiz ettiğimizde önerilen modelin Siamese ağ modeline göre %91,22 ile çok daha başarılı bir sonuç elde ettiğini görmekteyiz. İlginç bir sonuç olarak ResNet50 ağ modelinde başlangıçta önceden eğitilmiş olan ağırlıklarla ağ eğitime başladığında daha kötü sonuçlar elde ettiğini görmekteyiz. Başarı oranı %94,29’dan %89,91’e düşmüştür. Bu durum Resnet50 ağ modelinin öğrendiği ağırlıkların özgün ağ modelinde öğrenilen ağırlıklara etkisi olmayıp hata negatif olarak etkilediğini görmekteyiz. VGG-19 ağ modelinde de ağın önceden eğitilmiş ağırlıkla ya da rasgele oluşturulmuş ağırlıklara başlaması arasında sınıflandırma performansı açısından çok az bir fark olduğu sonucuna varılmıştır. Bu durum göstermektedir ki Triplet ağı marjin uzaklığı ayrımıyla çok iyi bir eğitim gerçekleştirmektedir. Sonuç itibarıyla en yüksek başarının yine Siamese ağ modelinde olduğu gibi VGG-19 ağ modeliyle elde edildiği sonucuna varılmıştır. %96,92 gibi yüksek bir başarı oranı elde edilmiştir.



Şekil 6.18. Triplet eğitim kayıp performansı

Çizelge 6.7. Triplet sonuçları

	Triplet Ağ Modeli Sonuçlar			
	<i>Önceden eğitilmiş ağırlıkları kullanmadan eğitime başlama</i>		<i>Önceden eğitilmiş ağırlıklarla eğitime başlama</i>	
	Periyot	Doğruluk	Periyot	Doğruluk
Önerilen Model	121	91,22	-	-
VGG-19	10	96,05	6	96,92
ResNet50	54	94,29	99	89,91

Çizelge 6.7’de Triplet ağ modeli için farklı ağ modelleri sonuçları görülmektedir. VGG-19 ağ modeli derin metrik öğrenme ile histopatolojik verileri sınıflandırmada oldukça başarılı sonuçlar elde etmiştir.



7. HİSTOPATOLOJİK GÖRÜNTÜLERDE ÇEKİRDEK TESPİTİ VE SINIFLANDIRMA

Mevcut histopatolojik veri kümesinde literatürde de karşılaşılan bir diğer problem çekirdeklerin sınıflandırılmasıdır [165,166]. Bu sebeple mevcut veri kümesinde öncelikli olarak çekirdeklerin kümeleme ile tespiti önemlidir. Burada amaç çekirdek tespitinin tek başına yapılmasıyla histopatolojik görüntüler üzerinde derin metrik öğrenme ile daha yüksek başarımlar elde etmeye odaklanmaktadır.

Çekirdekler mavi renge sahip olduğu için bu verilerin kümelenmesinde mavi renge odaklanılmaktadır. Bu sebeple çekirdekleri tespit etmek için bir kümeleme algoritması olan k-ortalamlar algoritmasından yararlanılmıştır. k-ortalamlar algoritması maksimum benzerlik ilkesiyle benzer noktaları Öklid uzaklığını kullanarak aynı küme merkezi içerisine almayı hedeflemektedir. Histopatolojik görüntüde 3 farklı temel bileşen olduğundan dolayı amaç 3 farklı bileşeni birbirinden ayırmaktır. Patolojik görüntülerdeki mavi kısımlar hematoxylin ile boyanan çekirdeklerdir. Pembe kısımlar ise Eosin ile boyanan pembe kısımlardır. Bir diğer bölüm ise beyaz renk kanalıdır.

Girdi: veri kümesi $X = [x_1, x_2, \dots, x_N] \in \mathbb{R}^{d \times N}$

- 1) Görüntüyü oku: $veri = X$
- 2) RGB renk uzayından $L*a*b$ renk uzayına geçiş işlemi yap.
- 3) k-ortalamları kullanarak $a*b*$ renklerini sınıflandır.
yerel minimumdan kaçınmak için birkaç kez k-ortalamları tekrarla.
- 4) k-ortalamlar kümeleme sonuçlarından yararlanarak görüntüdeki her bir pikseli etiketle.
- 5) Oluşturulan etiketlerle görüntüyü renklendir.
- 6) Sadece çekirdeğe odaklanarak çekirdek bölütleme işlemini gerçekleştir.
- 7) İlk adıma dön ve tüm görüntüler için bu adımları tekrarla

Şekil 7.1. Çekirdek bölütleme adımları

Şekil 7.1’de histopatolojik görüntüde ki her bir görüntü için bölütleme işlemi adımları görülmektedir. Bölütleme işlemi her bir görüntü için 1 kez yapılmaktadır ve bu durum tüm görüntüler için ayrı ayrı yapılmaktadır. Tezin bu bölümünde ise sadece mavi renkli çekirdekler tespit edilmektedir. Diğer bölümler tespit edilmesine karşın bu çalışmada kullanılmamaktadır.

k-ortalamlar algoritması ile histopatolojik bir görüntüde, 3 farklı dokuyu birbirinden ayırt etmek mümkündür. Bu bileşenler mavi renkli çekirdek ile pembe renkli protein amino grupları ve beyaz renkli kanal bileşenleridir. Çekirdek bölütleme işlemi Şekil 7.1’de detaylı bir şekilde adım adım ifade edilmektedir. Şekil 7.2’de histopatolojik görüntünün 3 farklı bileşeninin ne kadar başarılı bir şekilde tespitinin yapıldığı ortadadır. Şekil 7.2 (a)’da özgün veri kümesindeki Histopatolojik görüntülerden birisi görülmektedir. Bu görüntüden elde edilen . Şekil 7.2 (b) beyaz kanalları, Şekil 7.2 (c) protein amino gruplarını, Şekil 7.2 (d) ise hücre çekirdeklerini temsil etmektedir. Her bir görüntüden bu şekilde 3 farklı görüntü elde edilerek siyah beyaz ikili görüntülere dönüştürülmüştür.

Şekil 7.1’de işlem adımları verilen bölütleme ile tespit edilen çekirdeklerin daha sonra ikili bir görüntüye dönüştürülme işlemi yapılmaktadır. Bunun için ilk başta mavi renkli çekirdekler bir küme elemanı olarak tespit edilmiştir (Şekil 7.3.(b)). Daha sonra elde edilen mavi renkli görüntü gri seviyeli bir görüntüye dönüştürülmüştür. Son adımda ise gri seviyeli görüntü siyah beyaz ikili görüntüye dönüştürülmüş ve derin metrik öğrenme ağ modeli için hazır hale getirilmiştir.

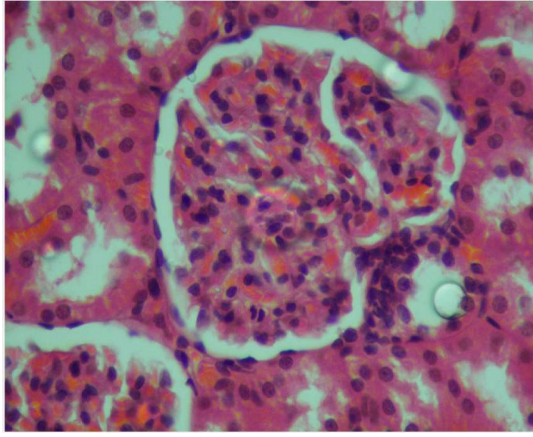
Elde edilen çekirdek görüntüleri daha sonrasında sınıflandırma yapılması amacıyla derin metrik öğrenme ağına gönderilmektedir. 6. Bölümde elde edilen sonuçlar neticesinde VGG-19 ağ modeli ile Siamese ve Triplet ağ modellerinin birleşimiyle çok daha başarılı sonuçlar elde edildiği görülmüştür. Bu sebeple Şekil 7.10’daki Siamese ağ modeli ile Şekil 7.11’deki Triplet ağ modeli ile çekirdek görüntüleri derin metrik öğrenme ile sınıflandırılmıştır.

Çekirdek bölütleme sonrası yapılan sınıflandırma sonuçları Çizelge 7.1’de görülmektedir. Sonuçları değerlendirdiğimizde çekirdek bölütleme sonrasında derin metrik öğrenme sınıflandırmasının tüm görüntünün ağa sunulmasıyla karşılaştırıldığında daha kötü sınıflandırma sonucu elde ettiğini görmekteyiz. Önceden eğitilmiş ağırlık değerleriyle ağın eğitimini başlattığımızda Siamese ağının 181 periyot (epoch) sonrasında %86,623 sınıflandırma başarısı elde ettiği, Triplet ağ modelinin ise 132 periyot (epoch) sonrasında %89,912 sınıflandırma başarısı elde ettiğini görmekteyiz. Sonuç olarak Triplet ağ modelinin, Siamese ağ modeline kıyasla daha iyi bir sınıflandırma işlemi yaptığını görmekteyiz.

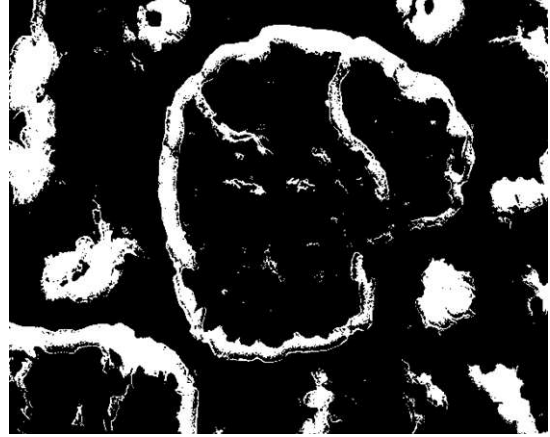
Çizelge 7.1. Çekirdek sınıflandırma sonuçları

	Periyot	Doğruluk (%)
Siamese – VGG19	181	86,623
Triplet - VGG19	132	89,912

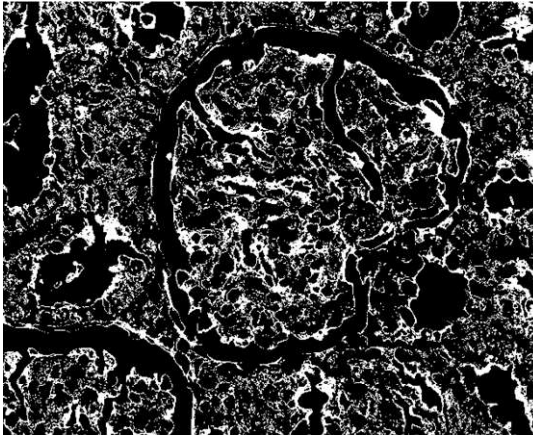
Histopatolojik görüntü bölütleme



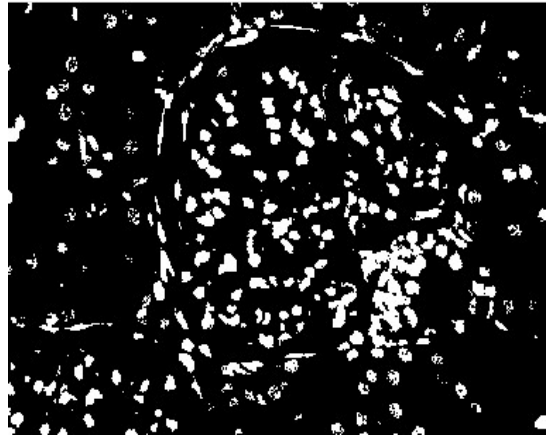
(a)



(b)



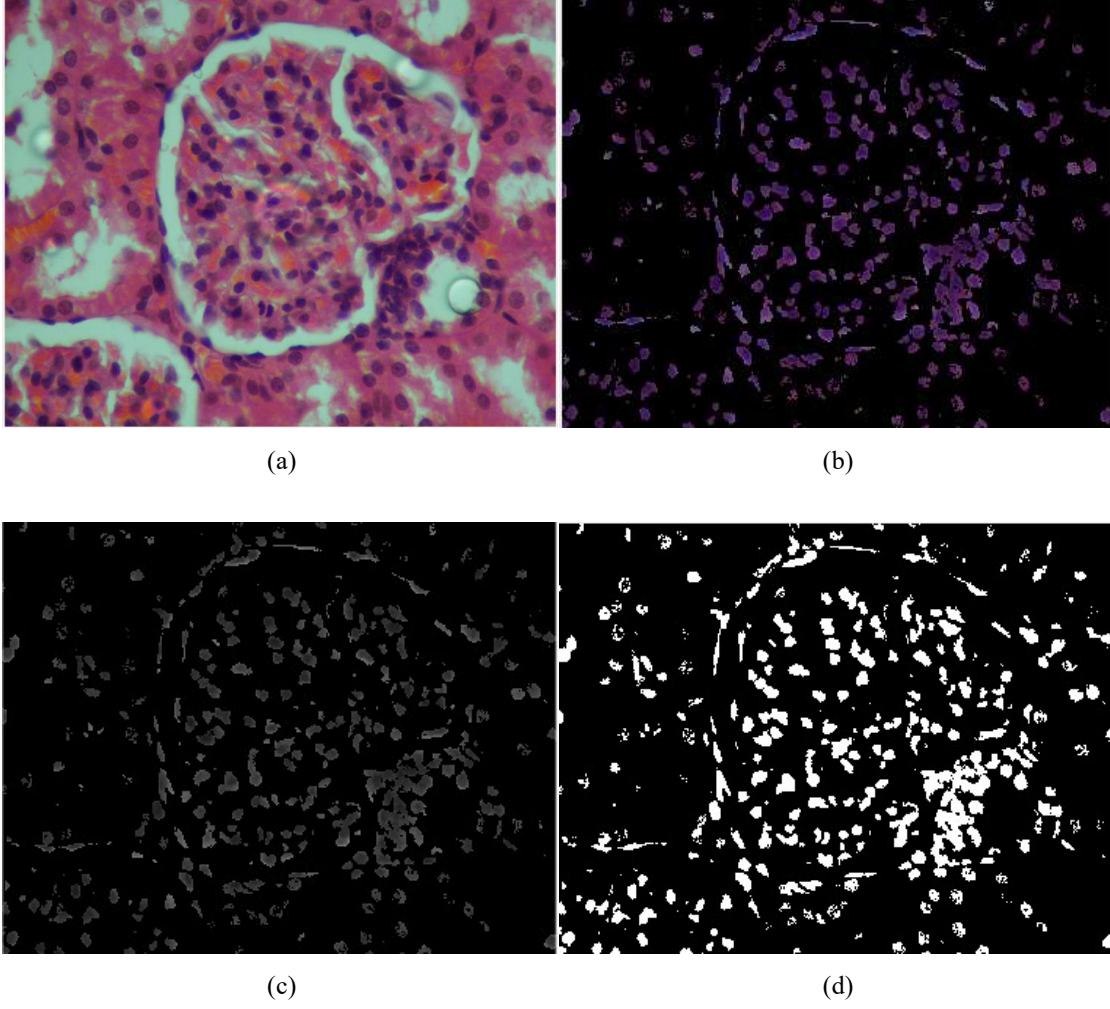
(c)



(d)

Şekil 7.2. Histopatolojik görüntü bölütleme

Çekirdek bölütleme



Şekil 7.3. Çekirdek bölütleme adımları

8. SONUÇ VE DEĞERLENDİRMELER

Son yıllarda, uzaklık bilgisine dayalı olan derin metrik öğrenme, araştırmacılar için önemli bir çalışma alanı olmaktadır. Halen, bu konudaki akademik çalışmalar literatüre çok önemli katkılar sağlamaktadır. Derin metrik öğrenmenin amacı, örnekleri kullanırken iki veya daha fazla nesnenin benzerliğini veya farklılığını hesaplayan bir benzerlik metriği öğrenmektir. Yüz tanıma, yüz doğrulama, kişinin yeniden tanımlanması, 3B şekil alma, anlamsal metinsel benzerlik, konuşmacı doğrulaması gibi görüntü, video, metin ve ses görevlerinde benzerlik üzerine yapılmış olan önemli çalışmalar bulunmaktadır.

Derin metrik öğrenmede literatürdeki son çalışmaların çoğunda Siamese ve Triplet ağları esin kaynağı olmuştur. Bu ağ yapıları, literatürde sıklıkla karşılaştırma amacıyla kullanılan veri kümeleri ve özellikle bilgi edinim problemlerinde yüksek etkinliklerini kanıtlamıştır. Bu çalışmalar, bilgilendirici giriş verisi örnekleri, ağ modelinin yapısı ve kayıp fonksiyonu olmak üzere üç ana bölümden oluşur.

Örnekleme stratejisi DML yöntemlerinin başarısında en temel faktörlerdendir. Verilere bağlı örnekleme stratejisi başlangıçta iyi ayarlanmamışsa, örnekler iyi bir şekilde belirlenmemişse, belli başlı bazı problemler için istenen sonuçlar elde edilemeyebilir. Ek olarak, zayıf örnekleme stratejisi yakınsama yavaşlamasına neden olabilir. Aynı zamanda iyi belirlenmemiş giriş verisi örnekleri ağın eğitim sürecinde yerel en iyiye (optimal) takılma riski taşımaktadır. Hem yavaş yakınsama hem de daha önemlisi en iyi (optimal) çözümü bulamama bir algoritmanın performansında en önemli kriterlerdendir. Bu sebeple örnekleme stratejisinin ağın eğitim sürecindeki etkisinin önemi anlaşılmaktadır.

Örnekleme stratejisindeki sorunları çözmek için, sert veya yarı sert negatif madencilik stratejisi, bilgilendirici girdi örnekleri elde etmek için önemli katkılar sağlamaktadır. Aynı zamanda aynı sınıf örnekleri arasındaki ilişkiyi maksimum seviyeye çıkarmak için benzer örnekler arasında pozitif madencilik ile öneriler sunmak da mümkündür. Bu tür yaklaşımlar ile bir ağa sunulacak daha değerli girdi örneklerini tespit etmek için örnekler arasındaki uzaklığı bir marjin değeri ile tespit etmek basit ve etkili bir çözümdür.

Bu tez çalışmasında örnekleme stratejisine gerek duyulmadan çok başarılı sonuçlar elde

edilmiştir. Başarımın arttırılması için bu strateji zorunlu olmamasına karşın bazı verilerde bunu gerçekleştirmek mümkün olmayabilir. Bu noktada önemli olan metrik öğrenme yaklaşımının en önemli hedefi olan veriye özgü çözümler sunulmasıdır. Bu noktada tez çalışmasında elde edilen sonuçlardan özellikle rekabetçi ağ modellerinin derin metrik öğrenmeye entegre edilmesiyle problemin başarısına önemli katkılar sunduğu açıktır.

Derin metrik öğrenmede kullanılan katmanlar, parametreler, paylaşılan ağırlıklar ve ağ modeli, optimize edilmesi gereken önemli bileşenlerdir. Bu öğrenme modelinde, ortak ağırlıkların benzerlik ilişkisinin elde edilmesinde önemli bir rolü bulunmaktadır. Paylaşılan ağırlıklar yalnızca sınıf içi ve sınıflar arası ilişkileri düzenleme görevini yerine getirmekle kalmaz, aynı zamanda bir ağı eğitirken daha verimli bellek kullanımı ve daha az zaman harcamayı sağlar.

Yapılan deneysel çalışmalarda da görüldüğü üzere transfer öğrenmenin ağ modelinin performansında pozitif etkileri olduğu görülmüştür. Transfer öğrenme ile önerilen derin modellerin kullanılması ağın en iyi (optimal) çözümü elde etmesine katkı sunmaktadır. Ancak daha da önemlisi başka bir veri kümesi üzerinde elde edilmiş olan ağırlıklarla ağın performansına verilen maksimum faydalar görülmüştür. Yapılan deneylerde de ağın başlangıç ağırlıkları dağılımının ağ performansı üzerinde etki düzeyi oldukça yüksektir. Transfer öğrenme kullanılarak yapılan deneylerde daha önceden belirlenmiş olan ağırlıklarla ağın en iyi çözüme olan yakınsamasının çok daha hızlı olduğu sonucuna varılmıştır.

Derin metrik öğrenmede bir diğer temel bölüm, ağın performansını önemli ölçüde belirleyen kayıp fonksiyonudur. İlgili literatürde yapılmış son çalışmalarda, verinin dönüştürülmüş uzaydaki yeni temsilinde daha yüksek ayırt edici özellikler elde etmek için verilere bağlı olarak bir uzaklık metriği elde etmek amaçlanmaktadır.

Yapılan çalışmaları incelediğimizde ağın performansında bu üç hususun bir arada olması gerektiği ve her birinin analizinin önemli olduğu yapılan deneysel çalışmalarda da görülmüştür. Bu sebeple bu alanda yapılacak çalışmalarda öğrenme süreci için olaya bütünsellikle yaklaşmanın önemi ortaya çıkmıştır.

Derin metrik öğrenme başarısını sınıf sayısının çok; ancak her bir sınıftan az sayıda

örneğin olduğu (kişiyi yeniden tanımlama görevleri gibi) veri kümelerinde kanıtlamıştır. Yapılan deneylerde de görüldüğü üzere örnek sayısının az olduğu histopatoloji veri kümesinde de çok yüksek bir başarıım yüzdesi elde edilmiştir. Buradan da görüleceği üzere derin metrik öğrenme uygulama alanı geniş olup, sınıflandırma ve kümeleme problemlerine önemli katkılar sunmaktadır.

Deneysel çalışmalarda kullanılan veri kümesi az sayıda örnek içermektedir. 360 hasarlı ve 779 sağlıklı örnekten oluşan patolojik veri kümesi aynı zamanda dengesiz bir dağılıma sahiptir. Derin metrik öğrenmede ağa sunulacak olan giriş görüntüleri ikili ya da üçlü bir şekilde sunulduğu için doğal yapısı gereği bir veri arttırma yaklaşımına sahiptir. Böylece az örneğe sahip olunmasına karşın, verinin doğal bir artışından söz etmek mümkündür. Verinin sınıflara dengesiz dağılımı da öğrenme süreci için bir diğer problem olup yine ağ yapısının doğallığı ile sınıflar arasında denge kurmak mümkündür. Bu şekilde bir yaklaşımla tez çalışmasında elde edilen kazanımla DML'in bu tür problemlerin üstesinden kolay bir şekilde geleceği kanıtlanmış ve literatüre katkı sunulmuştur.

Sonuç olarak yapılan deneysel çalışmalardan da görüleceği üzere derin metrik öğrenmenin histopatolojik veri kümesinin problemlerine uygun bir çözüm olduğu görülmüştür. Elde edilen yüksek başarı yüzdesi de bu durumun bir kanıtıdır. Literatüre sunulan bu katkı ile bu alanda da DML yaklaşımlarının başarılı bir şekilde kullanılabileceği görülmüştür. Yapılan literatür çalışmalarına bakıldığında da, derin metrik öğrenme alanında oldukça değerli çalışmalar bulunduğu ve bu alanda çalışan araştırmacıların sayısının gün geçtikçe arttığını görmekteyiz.

Gelecek çalışma planlarında Siamese ve Triplet yaklaşımının pozitif noktalarına odaklanılacaktır. Tespitler sonucu sayısal olarak elde edilen ve görsellerle de desteklenen Siamese ağ modelinde bulunan aynı sınıflar arasındaki uzaklığı minimize eden yaklaşım ile Triplet ağ modelinde bulunan farklı sınıflar arasındaki uzaklığı maksimize eden marjın tabanlı yaklaşımın bir arada değerlendirilerek daha başarılı bir benzerlik yaklaşımı sunulması hedeflenmektedir.



KAYNAKLAR

1. Mitchell, T.M. (1999). Machine learning and data mining. *Communications of the ACM*, 42(11), 30-36.
2. Géron, A. (2017). *Hands-on machine learning with Scikit-Learn and TensorFlow*. O'Reilly Media, 205.
3. Globerson, A. and Roweis, S.T. (2006). *Metric learning by collapsing classes*. Advances in neural information processing systems, 451-458, Vancouver, British Columbia, Canada.
4. Wang, F. and Sun, J. (2015). Survey on distance metric learning and dimensionality reduction in data mining. *Data Mining and Knowledge Discovery*, 29(2), 534-564.
5. Weinberger, K.Q. and Saul, L.K. (2009). Distance metric learning for large margin nearest neighbor classification. *Journal of Machine Learning Research*, 10, 207-244.
6. Schmidhuber, J. (2015). Deep learning in neural networks: An overview. *Neural networks*, 61, 85-117.
7. LeCun, Y., Bengio, Y. and Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436-444.
8. Kaya, M. and Bilge, H.Ş. (2019). Deep Metric Learning: A Survey. *Symmetry*, 11(9), 1066.
9. Lu, J., Hu, J. and Zhou, J. (2017). Deep metric learning for visual understanding: An overview of recent advances. *IEEE Signal Processing Magazine*, 34(6), 76-84.
10. Chopra, S., Hadsell, R. and LeCun, Y. (2005). *Learning a similarity metric discriminatively, with application to face verification*. IEEE Conference on Computer Vision and Pattern Recognition, 539-546, San Diego, CA, USA.
11. Hoffer, E. and Ailon, N. (2015). *Deep metric learning using triplet network*. International Workshop on Similarity-Based Pattern Recognition, 84-92, Copenhagen, Denmark.
12. Xing, E.P., Jordan, M.I., Russell, S.J. and Ng, A.Y. (2003). *Distance metric learning with application to clustering with side-information*. Advances in neural information processing systems, 521-528, Vancouver, British Columbia, Canada.
13. Davis, J.V., Kulis, B., Jain, P., Sra, S. and Dhillon, I.S. (2007). *Information-theoretic metric learning*. 24th international conference on Machine learning, 209-216, Corvallis, Oregon, USA.
14. Nguyen, H.V. and Bai, L. (2010). *Cosine similarity metric learning for face verification*. Asian Conference on Computer Vision, 709-720, Queenstown, New Zealand.

15. Suárez, J.L., García, S. and Herrera, F. (2018). A Tutorial on Distance Metric Learning: Mathematical Foundations, Algorithms and Software. *Preprint arXiv:1812.05944*.
16. Bellet, A. and Habrard, A. (2015). Robustness and generalization for metric learning. *Neurocomputing*, 151, 259-267.
17. Theodoridis, S. and Koutroumbas, K. (1998). *Pattern recognition*, Elsevier, 30.
18. Mahalanobis, P.C. (1936). On the generalized distance in statistics, *National Institute of Sciences of India*, 1-8.
19. Moutafis, P., Leng, M. and Kakadiaris, I.A. (2016). An overview and empirical comparison of distance metric learning methods. *IEEE transactions on cybernetics*, 47(3), 612-625.
20. Weinberger, K.Q., Blitzer, J. and Saul, L.K. (2006). *Distance metric learning for large margin nearest neighbor classification*. Advances in neural information processing systems, 1473-1480, Vancouver, British Columbia, Canada.
21. Yang, L., Jin, R., Sukthankar, R. and Liu, Y. (2006). *An efficient algorithm for local distance metric learning*. The Twenty-First National Conference On Artificial Intelligence (AAAI), 543-548, Boston, Massachusetts, USA.
22. Chen, J., Zhao, Z., Ye, J. and Liu, H. (2007). *Nonlinear adaptive distance metric learning for clustering*. 13th ACM SIGKDD international conference on Knowledge discovery and data mining, 123-132, San Jose, California, USA.
23. Bellet, A., Habrard, A. and Sebban, M. (2013). A survey on metric learning for feature vectors and structured data. *Preprint arXiv preprint arXiv:1306.6709*.
24. Zhang, Y. and Yeung, D.-Y. (2010). *Transfer metric learning by learning task relationships*. 16th ACM SIGKDD international conference on Knowledge discovery and data mining, 1199-1208, Washington, DC, USA.
25. Ying, Y. and Li, P. (2012). Distance metric learning with eigenvalue optimization. *Journal of machine Learning research*, 13, 1-26.
26. Goldberger, J., Hinton, G.E., Roweis, S.T. and Salakhutdinov, R.R. (2005). *Neighbourhood components analysis*. Advances in neural information processing systems, 513-520, Vancouver, British Columbia, Canada.
27. Wang, F. and Zhang, C. (2007). *Feature extraction by maximizing the average neighborhood margin*. IEEE Conference on Computer Vision and Pattern Recognition, 1-8, Minneapolis, MN, USA.
28. Martínez, A.M. and Kak, A.C. (2001). Pca versus lda. *IEEE transactions on pattern analysis and machine intelligence*, 23(2), 228-233.

29. Shental, N., Hertz, T., Weinshall, D. and Pavel, M. (2002). *Adjustment learning and relevant component analysis*. European conference on computer vision, 776-790, Copenhagen, Denmark.
30. Mika, S., Ratsch, G., Weston, J., Scholkopf, B. and Mullers, K.-R. (1999). *Fisher discriminant analysis with kernels*. Neural networks for signal processing IX: IEEE signal processing society workshop, 41-48, Madison, WI, USA.
31. Tsang, I.W., Cheung, P.-M. and Kwok, J.T. (2005). *Kernel relevant component analysis for distance metric learning*. IEEE International Joint Conference on Neural Networks, 954-959, Montreal, Quebec, Canada.
32. He, X. and Niyogi, P. (2004). *Locality preserving projections*. Advances in neural information processing systems, 153-160, Vancouver, British Columbia, Canada.
33. Wold, S., Esbensen, K. and Geladi, P. (1987). Principal component analysis. *Chemometrics and intelligent laboratory systems*, 2(1-3), 37-52.
34. Belkin, M. and Niyogi, P. (2002). *Laplacian eigenmaps and spectral techniques for embedding and clustering*. Advances in neural information processing systems, 585-591, Vancouver, British Columbia, Canada.
35. Roweis, S.T. and Saul, L.K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500), 2323-2326.
36. Tenenbaum, J.B., De Silva, V. and Langford, J.C. (2000). A global geometric framework for nonlinear dimensionality reduction. *Science*, 290(5500), 2319-2323.
37. Schölkopf, B., Smola, A. and Müller, K.-R. (1997). *Kernel principal component analysis*. International conference on artificial neural networks, 583-588, Lausanne, Switzerland.
38. Hoi, S.C., Liu, W. and Chang, S.-F. (2010). Semi-supervised distance metric learning for collaborative image retrieval and clustering. *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, 6(3), 18.
39. Yang, X., Fu, H., Zha, H. and Barlow, J. (2006). *Semi-supervised nonlinear dimensionality reduction*. International conference on Machine learning, 1065-1072, Pittsburgh, Pennsylvania, USA.
40. Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*. MIT Press, 13.
41. Lee, H., Grosse, R., Ranganath, R. and Ng, A.Y. (2009). *Convolutional deep belief networks for scalable unsupervised learning of hierarchical representations*. 26th annual international conference on machine learning, Montreal, Quebec, Canada, 609-616.
42. Donahue, J., Jia, Y., Vinyals, O., Hoffman, J., Zhang, N., Tzeng, E. and Darrell, T. (2014). *Decaf: A deep convolutional activation feature for generic visual*

- recognition*. 31st International conference on machine learning, 647-655, Beijing, China.
43. Huang, G.B., Lee, H. and Learned-Miller, E. (2012). *Learning hierarchical representations for face verification with convolutional deep belief networks*. IEEE Conference on Computer Vision and Pattern Recognition, 2518-2525, Providence, Rhode Island, USA.
 44. Le, Q.V., Zou, W.Y., Yeung, S.Y. and Ng, A.Y. (2011). *Learning hierarchical invariant spatio-temporal features for action recognition with independent subspace analysis*. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 3361-3368, Colorado Springs, CO, USA.
 45. Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2012). *Imagenet classification with deep convolutional neural networks*. Advances in neural information processing systems, 1097-1105, Lake Tahoe, Nevada, USA.
 46. LeCun, Y., Boser, B.E., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W.E. and Jackel, L.D. (1990). *Handwritten digit recognition with a back-propagation network*. Advances in neural information processing systems, 396-404, Denver, CO, USA.
 47. LeCun, Y., Haffner, P., Bottou, L. and Bengio, Y. (1999). Object recognition with gradient-based learning. *Shape, contour and grouping in computer vision*. Springer, 319-345.
 48. Yin, X. and Chen, Q. (2016). *Deep metric learning autoencoder for nonlinear temporal alignment of human motion*. IEEE International Conference on Robotics and Automation (ICRA), 2160-2166, Stockholm, Sweden.
 49. Hu, J., Lu, J. and Tan, Y.-P. (2014). *Discriminative deep metric learning for face verification in the wild*. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1875-1882, Columbus, Ohio, USA.
 50. Glorot, X., Bordes, A. and Bengio, Y. (2011). *Deep sparse rectifier neural networks*. Fourteenth International Conference on Artificial Intelligence and Statistics, 315-323, Fort Lauderdale, Florida, USA.
 51. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., . . . Bernstein, M. (2015). Imagenet large scale visual recognition challenge. *International journal of computer vision*, 115(3), 211-252.
 52. Pan, S.J. and Yang, Q. (2010). A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10), 1345-1359.
 53. Lu, J., Behbood, V., Hao, P., Zuo, H., Xue, S. and Zhang, G. (2015). Transfer learning using computational intelligence: a survey. *Knowledge-Based Systems*, 80, 14-23.

54. Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *Preprint arXiv preprint arXiv:1409.1556*.
55. He, K., Zhang, X., Ren, S. and Sun, J. (2016). *Deep residual learning for image recognition*. IEEE Conference on Computer Vision and Pattern Recognition, 770-778, Las Vegas, Nevada, USA.
56. Huang, G., Liu, Z., Van Der Maaten, L. and Weinberger, K.Q. (2017). *Densely connected convolutional networks*. IEEE Conference on Computer Vision and Pattern Recognition, 4700-4708, Honolulu, Hawaii, USA.
57. Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., . . . Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. *Preprint arXiv preprint arXiv:1704.04861*.
58. Matusita, K. (1955). Decision rules, based on the distance, for problems of fit, two samples, and estimation. *The Annals of Mathematical Statistics*, 26(4), 631-640.
59. Aherne, F.J., Thacker, N.A. and Rockett, P.I. (1998). The Bhattacharyya metric as an absolute similarity measure for frequency coded data. *Kybernetika*, 34(4), 363-368.
60. Elgammal, A., Duraiswami, R. and Davis, L.S. (2003). *Probabilistic tracking in joint feature-spatial spaces*. IEEE Conference on Computer Vision and Pattern Recognition, 781-788, Madison, Wisconsin, USA.
61. Yu, J., Yang, X., Gao, F. and Tao, D. (2016). Deep multimodal distance metric learning using click constraints for image ranking. *IEEE transactions on cybernetics*, 47(12), 4014-4024.
62. Cai, X., Wang, C., Xiao, B., Chen, X. and Zhou, J. (2012). *Deep nonlinear metric learning with independent subspace analysis for face verification*. 20th ACM international conference on Multimedia, 749-752, Nara, Japan.
63. Duan, Y., Lu, J., Feng, J. and Zhou, J. (2017). Deep Localized Metric Learning. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(10), 2644-2656.
64. Dai, G., Xie, J., Zhu, F. and Fang, Y. (2017). *Deep Correlated Metric Learning for Sketch-based 3D Shape Retrieval*. The Thirty-First AAAI Conference on Artificial Intelligence, 4002-4008, San Francisco, California, USA.
65. Li, Z. and Tang, J. (2015). Weakly supervised deep metric learning for community-contributed image retrieval. *IEEE Transactions on Multimedia*, 17(11), 1989-1999.
66. Kumar, V.B., Harwood, B., Carneiro, G., Reid, I. and Drummond, T. (2017). Smart Mining for Deep Metric Learning. *Preprint arXiv preprint arXiv:1704.01285*.
67. Gundogdu, E., Solmaz, B., Koç, A., Yücesoy, V. and Alatan, A.A. (2017). *Deep distance metric learning for maritime vessel identification*. 25th Signal Processing and Communications Applications Conference (SIU), 1-4, Antalya, Türkiye.

68. Liu, J., Deng, Y., Bai, T., Wei, Z. and Huang, C. (2015). Targeting ultimate accuracy: Face recognition via deep embedding. *Preprint arXiv:1506.07310*.
69. Hoffer, E. and Ailon, N. (2016). Semi-supervised deep learning by metric embedding. *Preprint arXiv:1611.01449*.
70. LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
71. Hadsell, R., Chopra, S. and LeCun, Y. (2006). *Dimensionality reduction by learning an invariant mapping*. IEEE Conference on Computer Vision and Pattern Recognition, 1735-1742, New York, NY, USA.
72. Chen, M., Ge, Y., Feng, X., Xu, C. and Yang, D. (2018). Person re-identification by pose invariant deep metric learning with improved triplet Loss. *IEEE Access*, 6, 68089-68095.
73. Oh Song, H., Xiang, Y., Jegelka, S. and Savarese, S. (2016). *Deep metric learning via lifted structured feature embedding*. IEEE Conference on Computer Vision and Pattern Recognition, 4004-4012, Las Vegas, Nevada USA.
74. Hermans, A., Beyer, L. and Leibe, B. (2017). In Defense of the Triplet Loss for Person Re-Identification. *Preprint arXiv:1703.07737*.
75. Lee, J., Abu-El-Haija, S., Varadarajan, B. and Natsev, A.P. (2018). *Collaborative Deep Metric Learning for Video Understanding*. 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 481-490, London, United Kingdom.
76. Huang, W., Ding, H. and Chen, G. (2018). A novel deep multi-channel residual networks-based metric learning method for moving human localization in video surveillance. *Signal Processing*, 142, 104-113.
77. Lim, I., Gehre, A. and Kobbelt, L. (2016). *Identifying style of 3D shapes using deep metric learning*. Symposium on Geometry Processing, 207-215, Berlin, Germany.
78. Annarumma, M. and Montana, G. (2018). *Deep metric learning for multi-labelled radiographs*. 33rd Annual ACM Symposium on Applied Computing, 34-37, Pau, France.
79. Schroff, F., Kalenichenko, D. and Philbin, J. (2015). *Facenet: A unified embedding for face recognition and clustering*. IEEE Conference on Computer Vision and Pattern Recognition, 815-823, Boston, Massachusetts, USA.
80. Bromley, J., Guyon, I., LeCun, Y., Säckinger, E. and Shah, R. (1994). *Signature verification using a "siamese" time delay neural network*. Advances in Neural Information Processing Systems, 737-744, Denver, Colorado, USA.

81. Mueller, J. and Thyagarajan, A. (2016). *Siamese recurrent architectures for learning sentence similarity*. Thirtieth AAAI Conference on Artificial Intelligence, 2786-2792, Phoenix, Arizona, USA.
82. Benajiba, Y., Sun, J., Zhang, Y., Jiang, L., Weng, Z. and Biran, O. (2019). *Siamese Networks for Semantic Pattern Similarity*. IEEE 13th International Conference on Semantic Computing (ICSC), 191-194, Newport Beach, California, USA.
83. Chen, X., He, L., Xu, C. and Liu, J. (2019). Distance-Dependent Metric Learning. *IEEE Signal Processing Letters*, 26(2), 357-361.
84. Narayanaswamy, V.S., Thiagarajan, J.J., Song, H. and Spanias, A. (2019). *Designing an Effective Metric Learning Pipeline for Speaker Diarization*. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 5806-5810, Brighton, United Kingdom.
85. Hu, J., Lu, J. and Tan, Y.-P. (2016). Deep metric learning for visual tracking. *IEEE Transactions on Circuits and Systems for Video Technology*, 26(11), 2056-2068.
86. Zheng, L., Yang, Y. and Hauptmann, A.G. (2016). Person re-identification: Past, present and future. *Preprint arXiv:1610.02984*.
87. Yang, X., Zhou, P. and Wang, M. (2018). Person reidentification via structural deep metric learning. *IEEE Transactions on Neural Networks and Learning Systems*, 30(10), 2987-2998.
88. Ahmed, E., Jones, M. and Marks, T.K. (2015). *An improved deep learning architecture for person re-identification*. IEEE Conference on Computer Vision and Pattern Recognition, 3908-3916, Boston, Massachusetts, USA.
89. Ding, S., Lin, L., Wang, G. and Chao, H. (2015). Deep feature learning with relative distance comparison for person re-identification. *Pattern Recognition*, 48(10), 2993-3003.
90. Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., . . . Sánchez, C.I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, 42, 60-88.
91. Wang, F., Kang, L. and Li, Y. (2015). *Sketch-based 3d shape retrieval using convolutional neural networks*. IEEE Conference on Computer Vision and Pattern Recognition, 1875-1883, Boston, Massachusetts, USA.
92. Dai, G., Xie, J. and Fang, Y. (2018). Deep Correlated Holistic Metric Learning for Sketch-Based 3D Shape Retrieval. *IEEE Transactions on Image Processing*, 27(7), 3374-3386.
93. Xie, J., Dai, G., Zhu, F., Shao, L. and Fang, Y. (2018). Deep nonlinear metric learning for 3-D shape retrieval. *IEEE transactions on cybernetics*, 48(1), 412-422.

94. He, X., Zhou, Y., Zhou, Z., Bai, S. and Bai, X. (2018). Triplet-Center Loss for Multi-View 3D Object Retrieval. *Preprint arXiv:1803.06189*.
95. Lu, J., Hu, J. and Tan, Y.-P. (2017). Discriminative deep metric learning for face and kinship verification. *IEEE Transactions on Image Processing*, 26(9), 4269-4282.
96. Liu, X., Kumar, B.V., You, J. and Jia, P. (2017). *Adaptive Deep Metric Learning for Identity-Aware Facial Expression Recognition*. IEEE Conference on Computer Vision and Pattern Recognition Workshops, 522-531, Honolulu, Hawaii, USA.
97. Liu, H., Lu, J., Feng, J. and Zhou, J. (2018). Label-sensitive deep metric learning for facial age estimation. *IEEE Transactions on Information Forensics and Security*, 13(2), 292-305.
98. Zhu, W., Yao, T., Ni, J., Wei, B. and Lu, Z. (2018). Dependency-based Siamese long short-term memory network for learning sentence representations. *PloS one*, 13(3), e0193919.
99. Dor, L.E., Mass, Y., Halfon, A., Venezian, E., Shnayderman, I., Aharonov, R. and Slonim, N. (2018). *Learning Thematic Similarity Metric Using Triplet Networks*. 56th Annual Meeting of the Association for Computational Linguistics, 49-54, Melbourne, Australia.
100. Wang, J., Wang, K.-C., Law, M.T., Rudzicz, F. and Brudno, M. (2019). *Centroid-based deep metric learning for speaker recognition*. IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 3652-3656, Brighton, United Kingdom.
101. Snell, J., Swersky, K. and Zemel, R. (2017). *Prototypical networks for few-shot learning*. Advances in Neural Information Processing Systems, 4077-4087, Long Beach, California, USA.
102. Lu, R., Wu, K., Duan, Z. and Zhang, C. (2017). *Deep ranking: triplet matchnet for music metric learning*. IEEE International Conference on Acoustics, Speech and Signal Processing, 121-125, New Orleans, Los Angeles, USA.
103. Wang, Q., Wan, J. and Yuan, Y. (2017). Deep Metric Learning for Crowdedness Regression. *IEEE Transactions on Circuits and Systems for Video Technology*, 28(10), 2633-2643.
104. Liu, Y., Zhao, K. and Cong, G. (2018). *Efficient Similar Region Search with Deep Metric Learning*. 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 1850-1859, London, United Kingdom.
105. Wang, X. and Liu, M. (2018). *Multi-View Deep Metric Learning for Volumetric Image Recognition*. IEEE International Conference on Multimedia and Expo (ICME), 1-6, San Diego, California, USA.

106. Fathi, A., Wojna, Z., Rathod, V., Wang, P., Song, H.O., Guadarrama, S. and Murphy, K.P. (2017). Semantic instance segmentation via deep metric learning. *Preprint arXiv:1703.10277*.
107. Cai, S., Huang, J., Ding, X. and Zeug, D. (2017). *Semantic edge detection based on deep metric learning*. International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS), 707-712, Xiamen, China.
108. Xing, Y., Wang, M., Yang, S. and Jiao, L. (2018). Pan-sharpening via deep metric learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145(A), 165-183.
109. Wah, C., Branson, S., Welinder, P., Perona, P. and Belongie, S. (2011). *The caltech-ucsd birds-200-2011 dataset*, Computation & Neural Systems Technical Report, CNS-TR-2011-001, Pasadena, CA, USA.
110. Sohn, K. (2016). *Improved deep metric learning with multi-class n-pair loss objective*. Advances in Neural Information Processing Systems, , 1857-1865, Barcelona, Spain.
111. Wang, J., Zhou, F., Wen, S., Liu, X. and Lin, Y. (2017). *Deep Metric Learning With Angular Loss*. IEEE Conference on Computer Vision and Pattern Recognition, 2593-2601, Honolulu, Hawaii, USA.
112. Song, H.O., Jegelka, S., Rathod, V. and Murphy, K. (2017). *Deep metric learning via facility location*. IEEE Conference on Computer Vision and Pattern Recognition, 2206-2214, Honolulu, Hawaii, USA.
113. Ge, W. (2018). *Deep metric learning with hierarchical triplet loss*. European Conference on Computer Vision (ECCV), 272-288, Munich, Germany.
114. Krause, J., Stark, M., Deng, J. and Fei-Fei, L. (2013). *3d object representations for fine-grained categorization*. IEEE International Conference on Computer Vision Workshops, 554-561, Sydney, Australia.
115. Zheng, L., Shen, L., Tian, L., Wang, S., Wang, J. and Tian, Q. (2015). *Scalable person re-identification: A benchmark*. IEEE International Conference on Computer Vision, 1116-1124, Santiago, Chile.
116. Ustinova, E. and Lempitsky, V. (2016). *Learning deep embeddings with histogram loss*. Advances in Neural Information Processing Systems, 4170-4178, Barcelona, Spain.
117. Yao, H., Zhang, S., Hong, R., Zhang, Y., Xu, C. and Tian, Q.J.I.T.o.I.P. (2019). Deep representation learning with part loss for person re-identification, *IEEE Transactions on Image Processing*, 28(6), 2860-2871.
118. Li, W., Zhao, R., Xiao, T. and Wang, X. (2014). *Deepreid: Deep filter pairing neural network for person re-identification*. IEEE Conference on Computer Vision and Pattern Recognition, 152-159, Columbus, Ohio, USA.

119. Li, B., Lu, Y., Godil, A., Schreck, T., Bustos, B., Ferreira, A., . . . Matsuda, T. (2014). A comparison of methods for sketch-based 3D shape retrieval. *Computer Vision and Image Understanding*, 119, 57-80.
120. Howell, A.J. and Buxton, H. (1996). *Towards unconstrained face recognition from image sequences*. Second International Conference on Automatic Face and Gesture Recognition, 224-229, Killington, Vermont, USA.
121. Hu, J., Lu, J. and Tan, Y.-P. (2017). Sharable and individual multi-view metric learning. *IEEE transactions on pattern analysis and machine intelligence*, 40(9), 2281-2288.
122. Wolf, L., Hassner, T. and Maoz, I. *Face recognition in unconstrained videos with matched background similarity*. IEEE Conference on Computer Vision and Pattern Recognition, 529-534, Colorado Springs, Colorado, USA
123. Marelli, M., Bentivogli, L., Baroni, M., Bernardi, R., Menini, S. and Zamparelli, R. (2014). *Semeval-2014 task 1: Evaluation of compositional distributional semantic models on full sentences through semantic relatedness and textual entailment*. 8th international workshop on semantic evaluation (SemEval 2014), 1-8, Dublin, Ireland.
124. Greenberg, C.S., Bansé, D., Doddington, G.R., Garcia-Romero, D., Godfrey, J.J., Kinnunen, T., . . . Reynolds, D.A. (2014). *The NIST 2014 speaker recognition i-vector machine learning challenge*. Odyssey: The Speaker and Language Recognition Workshop, 224-230, Joensuu, Finland.
125. Veaux, C., Yamagishi, J. and MacDonald, K. (2017). CSTR VCTK corpus: English multi-speaker corpus for CSTR voice cloning toolkit, Edinburgh, United, Kingdom.
126. Chung, J.S., Nagrani, A. and Zisserman, A. (2018). Voxceleb2: Deep speaker recognition. *Preprint arXiv:1806.05622*.
127. Bell, S. and Bala, K. (2015). Learning visual similarity for product design with convolutional neural networks. *ACM Transactions on Graphics (TOG)*, 34(4), 98.
128. Simo-Serra, E., Trulls, E., Ferraz, L., Kokkinos, I., Fua, P. and Moreno-Noguer, F. (2015). *Discriminative learning of deep convolutional feature point descriptors*. IEEE International Conference on Computer Vision, 118-126, Santiago, Chile.
129. Bucher, M., Herbin, S. and Jurie, F. (2016). *Hard negative mining for metric learning based zero-shot classification*. European Conference on Computer Vision, 524-531, Amsterdam, The Netherlands.
130. Cui, Y., Zhou, F., Lin, Y. and Belongie, S. (2016). *Fine-grained categorization and dataset bootstrapping using deep metric learning with humans in the loop*. IEEE Conference on Computer Vision and Pattern Recognition, 1153-1162, Venice, Italy.

131. Movshovitz-Attias, Y., Toshev, A., Leung, T.K., Ioffe, S. and Singh, S. (2017). No fuss distance metric learning using proxies. *arXiv preprint arXiv:1703.07464*.
132. Wu, C.-Y., Manmatha, R., Smola, A.J. and Krähenbühl, P. (2017). *Sampling matters in deep embedding learning*. Proc. IEEE International Conference on Computer Vision (ICCV), 2840-2848, Venice, Italy.
133. Shi, H., Zhu, X., Liao, S., Lei, Z., Yang, Y. and Li, S.Z. (2015). Constrained deep metric learning for person re-identification. *Preprint arXiv:1511.07545*.
134. Filković, I., Kalafatić, Z. and Hrkać, T. (2016). *Deep metric learning for person Re-identification and De-identification*. Information and Communication Technology, Electronics and Microelectronics (MIPRO), 1360-1364, Opatija, Croatia.
135. Jeong, Y., Lee, S., Park, D. and Park, K. (2018). Accurate Age Estimation Using Multi-Task Siamese Network-Based Deep Metric Learning for Frontal Face Images. *Symmetry*, 10(9), 385.
136. Yi, D., Lei, Z., Liao, S. and Li, S.Z. (2014). *Deep metric learning for person re-identification*. 22nd International Conference on Pattern Recognition (ICPR), 34-39, Stockholm, Sweden.
137. Wen, Y., Zhang, K., Li, Z. and Qiao, Y. (2016). *A discriminative feature learning approach for deep face recognition*. European Conference on Computer Vision, 499-515, Amsterdam, The Netherlands.
138. Ni, J., Liu, J., Zhang, C., Ye, D. and Ma, Z. (2017). *Fine-grained Patient Similarity Measuring using Deep Metric Learning*. ACM on Conference on Information and Knowledge Management, 1189-1198, Singapore.
139. Gong, Z., Zhong, P., Yu, Y. and Hu, W. (2018). Diversity-Promoting Deep Structural Metric Learning for Remote Sensing Scene Classification. *IEEE Transactions on Geoscience and Remote Sensing*, 56(1), 371-390.
140. Wang, X., Han, X., Huang, W., Dong, D. and Scott, M.R. (2019). *Multi-Similarity Loss with General Pair Weighting for Deep Metric Learning*. IEEE Conference on Computer Vision and Pattern Recognition, 5022-5030, Long Beach, California, USA.
141. Rippel, O., Paluri, M., Dollar, P. and Bourdev, L. (2016). *Metric learning with adaptive density discrimination*. International Conference on Learning Representations, 1-5, San Juan, Puerto Rico, USA.
142. Chen, L. and He, Y. (2018). *Dress Fashionably: Learn Fashion Collocation With Deep Mixed-Category Metric Learning*. Thirty-Second AAAI Conference on Artificial Intelligence, 2103-2110, New Orleans, Los Angeles, USA.

143. LeCun, Y., Huang, F.J. and Bottou, L. (2004). *Learning methods for generic object recognition with invariance to pose and lighting*. IEEE Conference on Computer Vision and Pattern Recognition, 97-104, Washington, DC, USA.
144. Krizhevsky, A. (2009). *Learning Multiple Layers of Features from Tiny Images*. (Master's thesis, University of Toronto, 2009), Toronto, Ontario, Canada.
145. Netzer, Y., Wang, T., Coates, A., Bissacco, A., Wu, B. and Ng, A.Y. (2011). Reading digits in natural images with unsupervised feature learning, 1-9, Granada, Spain.
146. Coates, A., Ng, A. and Lee, H. (2011). *An analysis of single-layer networks in unsupervised feature learning*. Proceedings of the fourteenth international conference on artificial intelligence and statistics, 215-223, Fort Lauderdale, Florida, USA.
147. Xie, S., Yang, T., Wang, X. and Lin, Y. (2015). *Hyper-class augmented and regularized deep learning for fine-grained image classification*. IEEE conference on computer vision and pattern recognition, 2645-2654, Boston, Massachusetts, USA.
148. Khosla, A., Jayadevaprakash, N., Yao, B. and Li, F.-F. (2011). *Novel dataset for fine-grained image categorization: Stanford dogs*. First Workshop on Fine-Grained Visual Categorization (FGVC) in IEEE Conference on Computer Vision and Pattern Recognition, 1-2, Providence, Rhode Island, USA.
149. Parkhi, O.M., Vedaldi, A., Zisserman, A. and Jawahar, C. (2012). *Cats and dogs*. 2012 IEEE conference on computer vision and pattern recognition, 3498-3505, Providence, Rhode Island, USA.
150. Nilsback, M.-E. and Zisserman, A. (2008). *Automated flower classification over a large number of classes*. 2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing, 722-729, Bhubaneswar, India.
151. Russakovsky, O. and Fei-Fei, L. (2010). *Attribute learning in large-scale datasets*. European Conference on Computer Vision, 1-14, Crete, Greece.
152. Liu, Z., Luo, P., Qiu, S., Wang, X. and Tang, X. (2016). *Deepfashion: Powering robust clothes recognition and retrieval with rich annotations*. IEEE Conference on Computer Vision And Pattern Recognition, 1096-1104, Las Vegas, Nevada, USA.
153. Gray, D., Brennan, S. and Tao, H. (2007). *Evaluating appearance models for recognition, reacquisition, and tracking*. IEEE International Workshop on Performance Evaluation for Tracking and Surveillance (PETS), 1-7, Rio de Janeiro, Brazil.
154. Rajput, A. and Birdi, S. (1997). Epidemiology of Parkinson's disease. *Parkinsonism & related disorders*, 3(4), 175-186.

155. Khan, T., Westin, J. and Dougherty, M. (2014). Classification of speech intelligibility in Parkinson's disease. *Biocybernetics and Biomedical Engineering*, 34(1), 35-45.
156. Tsanas, A., Little, M.A., McSharry, P.E., Spielman, J. and Ramig, L.O. (2012). Novel speech signal processing algorithms for high-accuracy classification of Parkinson's disease. *IEEE transactions on biomedical engineering*, 59(5), 1264-1271.
157. Guillaumin, M., Verbeek, J. and Schmid, C. (2009). *Is that you? Metric learning approaches for face identification*. IEEE 12th international conference on computer vision, 498-505, Kyoto, Japan.
158. Cao, Q., Ying, Y. and Li, P. (2013). *Similarity metric learning for face recognition*. IEEE International Conference on Computer Vision, 2408-2415, Portland, Oregon, USA.
159. Sakar, B.E., Isenkul, M.E., Sakar, C.O., Sertbas, A., Gorgen, F., Delil, S., . . . Kursun, O. (2013). Collection and analysis of a Parkinson speech dataset with multiple types of sound recordings. *IEEE Journal of Biomedical and Health Informatics*, 17(4), 828-834.
160. Kaya, M. and Bilge, H.Ş. (2017). *Classification of Parkinson speech data by metric learning*. International Artificial Intelligence and Data Processing Symposium (IDAP), 1-5, Malatya, Turkey.
161. İnternet: H&E staining. URL: https://www.histology.leeds.ac.uk/what-is-histology/H_and_E.php, Son Erişim Tarihi: 22/12/2019.
162. Elmore, J.G., Longton, G.M., Carney, P.A., Geller, B.M., Onega, T., Tosteson, A.N., . . . Schnitt, S.J. (2015). Diagnostic concordance among pathologists interpreting breast biopsy specimens. *Jama*, 313(11), 1122-1132.
163. Armitage, J.O. (1997). A clinical evaluation of the International Lymphoma Study Group classification of non-Hodgkin's lymphoma. *Blood*, 89(11), 3909-3918.
164. Serin, F., Erturkler, M. and Gul, M. (2017). A novel overlapped nuclei splitting algorithm for histopathological images. *Computer methods and programs in biomedicine*, 151, 57-70.
165. Feng, Y., Zhang, L. and Yi, Z. (2018). Breast cancer cell nuclei classification in histopathology images using deep neural networks. *International journal of computer assisted radiology and surgery*, 13(2), 179-191.
166. Alom, M.Z., Yakopcic, C., Taha, T.M. and Asari, V.K. (2018). Microscopic Nuclei Classification, Segmentation and Detection with improved Deep Convolutional Neural Network (DCNN) Approaches. *Preprint arXiv:1811.03447*.



ÖZGEÇMİŞ

Kişisel Bilgiler

Soyadı, adı : KAYA, Mahmut
 Uyuğu : T.C.
 Doğum tarihi ve yeri : 15.08.1986, Palu
 Medeni hali : Evli
 E-mail : mahmutkaya@siirt.edu.tr



Eğitim

Derece	Eğitim Birimi	Mezuniyet Tarihi
Doktora	Gazi Üniversitesi / Bilgisayar Mühendisliği	Devam ediyor
Yüksek lisans	Gazi Üniversitesi / Bilgisayar Mühendisliği	2014
Lisans	Fırat Üniversitesi / Bilgisayar Mühendisliği	2010
Lise	Palu Lisesi	2004

İş Deneyimi

Yıl	Yer	Görev
2019-Halen	Siirt Üniversitesi	Araştırma Görevlisi
2012-2019	Gazi Üniversitesi	Araştırma Görevlisi
2010-2012	Siirt Üniversitesi	Araştırma Görevlisi

Yabancı Dil

İngilizce

Yayınlar

1. Kaya M., Bilge, H. Ş. (2019). Deep Metric Learning: A Survey. *Symmetry*, 11(9), 1066
2. Aral R. A., Keskin Ş. R., Kaya M., Hacıömeroğlu M. (2018) *Classification of TrashNet Dataset Based on Deep Learning Models*, IEEE BigData 2018, Seattle, USA.
3. Kaya M., Bilge H. Ş., *Classification of Parkinson Speech Data by Metric Learning*, IEEE 2nd International Conference on Artificial Intelligence and Data Processing (IDAP17), Malatya, Türkiye.

4. Kaya M., Bilge H. Ş. (2016). *İstatistiksel ve Sarmal Yöntemlere Dayalı Hibrit Öznitelik Seçme Yaklaşımı*, IEEE 24th Signal Processing and Communication Application Conference (SIU), Malatya, Türkiye.
5. Kaya M., Arıöz U. (2015). *Laplacian Skor ile Öznitelik Ağırlıklandırma*, IEEE 23th Signal Processing and Communication Application Conference (SIU), Zonguldak, Türkiye.
6. Kaya M., Bilge H. Ş., (2014). *Classification of pancreas tumor dataset using adaptive weighted k nearest neighbor algorithm*, IEEE International Symposium on INnovations in Intelligent SysTEms and Applications (INISTA), Alberobello, İtalya.
7. Güzel, C., Kaya, M., Yıldız, O., Bilge, H.Ş. (2013). *Breast Cancer Diagnosis Based On Naive Bayes Machine Learning Classifier With Knn Missing Value Imputation*, 3rd World Conference on Innovation and Computer Sciences (INSODE), Antalya, Türkiye.
8. Kaya, M., Bilge, H.Ş., Yıldız, O., (2013). *Gen ifadelerinde Öz Nitelik Seçimi Ve Boyut İndirgeme*, IEEE 21th Signal Processing and Communication Application Conference (SIU), Girne, Kıbrıs.

Hobiler

Yüzme, Futbol



GAZİ GELECEKTİR...