



**KTO KARATAY ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ
ELEKTRİK ve ELEKTRONİK MÜHENDİSLİĞİ ANABİLİM DALI
ELEKTRİK ve BİLGİSAYAR MÜHENDİSLİĞİ TEZLİ YÜKSEK LİSANS
PROGRAMI**

**KÜLTÜREL MİRAS GÖRÜNTÜLERİNİN DERİN ÖĞRENME YÖNTEMLERİ
İLE SINIFLANDIRILMASI**

İrem UYANIK

Yüksek Lisans Tezi

**KONYA
Eylül 2021**

KÜLTÜREL MİRAS GÖRÜNTÜLERİNİN DERİN ÖĞRENME YÖNTEMLERİ İLE
SINIFLANDIRILMASI

İrem UYANIK

KTO Karatay Üniversitesi
Lisansüstü Eğitim Enstitüsü
Elektrik ve Elektronik Mühendisliği Anabilim Dalı
Elektrik ve Bilgisayar Mühendisliği Tezli Yüksek Lisans Programı

Yüksek Lisans Tezi

Tez Danışmanı: Dr. Öğr. Üyesi Ali ÖZTÜRK

Konya
Eylül 2021

BİLDİRİM

Enstitü tarafından onaylanan Yüksek Lisans tezimin tamamını veya herhangi bir kısmını basılı veya dijital biçimde arşivleme ve aşağıda belirtilen koşullar dahilinde erişime açma iznini KTO Karatay Üniversitesine verdiğimi bildiririm. Bu izinle, Üniversiteye verilen kullanım hakları dışındaki tüm fikri mülkiyet haklarım bende kalacak ve gelecekteki çalışmalar (makale, kitap, lisans, patent vb.) için tezimin tamamının veya bir bölümünün kullanım hakları yalnızca bana ait olacaktır.

Tezimin bütünüyle kendi çalışmam olduğunu, başkalarının haklarını ihlal etmediğimi ve tezimin tek yetkili sahibi olduğumu beyan ve taahhüt ederim. Telif hakkı bulunan ve sahiplerinden yazılı izinle kullanılması zorunlu olan kaynakları, yazılı izin alarak kullandığımı ve istenildiğinde izinlerin suretlerini Üniversiteye teslim etmeyi taahhüt ederim.

Yükseköğretim Kurulu tarafından yayımlanan “Lisansüstü Tezlerin Elektronik Ortamda Toplanması, Düzenlenmesi ve Erişime Açılmasına İlişkin Yönerge” kapsamında, tezim, aşağıda belirtilen koşullar haricince, YÖK Ulusal Tez Merkezi ve KTO Karatay Üniversitesi Açık Erişim Sisteminde erişime açılır.

Enstitü / Fakülte Yönetim Kurulu kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren 2 yıl ertelenmiştir.¹

Enstitü / Fakülte Yönetim Kurulunun gerekçeli kararı ile tezimin erişime açılması mezuniyet tarihimden itibaren ... ay ertelenmiştir.²

Tezimle ilgili gizlilik kararı verilmiştir.³⁴

09 Eylül 2021

İrem UYANIK

¹ MADDE 6(1) Lisansüstü teze ilgili patent başvurusu yapılması veya patent alma sürecinin devam etmesi durumunda, tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulu iki yıl süre ile tezin erişime açılmasının ertelenmesine karar verebilir.

² MADDE 6(2) Yeni teknik, materyal ve metotların kullanıldığı, henüz makaleye dönüşmemiş veya patent gibi yöntemlerle korunmamış ve internetten paylaşılması durumunda 3. şahıslara veya kurumlara haksız kazanç imkanı oluşturabilecek bilgi ve bulguları içeren tezler hakkında tez danışmanının önerisi ve enstitü anabilim dalının uygun görüşü üzerine enstitü veya fakülte yönetim kurulunun gerekçeli kararı ile altı ayı aşmamak üzere tezin erişime açılması engellenebilir.

³ MADDE 7(1) Ulusal çıkarları veya güvenliği ilgilendiren, emniyet, istihbarat, savunma ve güvenlik, sağlık vb. konulara ilişkin lisansüstü tezlerle ilgili gizlilik kararı, tezin yapıldığı kurum tarafından verilir. Kurum ve kuruluşlarla yapılan işbirliği protokolü çerçevesinde hazırlanan lisansüstü tezlere ilişkin gizlilik kararı ise, ilgili kurum ve kuruluşun önerisi ile enstitü veya fakültenin uygun görüşü üzerine üniversite yönetim kurulu tarafından verilir. Gizlilik kararı verilen tezler Yükseköğretim Kuruluna bildirilir.

⁴ MADDE 7(2) Gizlilik kararı verilen tezler gizlilik süresince enstitü veya fakülte tarafından gizlilik kuralları çerçevesinde muhafaza edilir, gizlilik kararının kaldırılması halinde Tez Otomasyon Sistemine yüklenir.

ETİK BEYAN

KTO Karatay Üniversitesi Lisansüstü Eğitim Enstitüsü Tez Hazırlama ve Yazım Kurallarına uygun olarak Dr. Öğr. Üyesi Ali ÖZTÜRK danışmanlığında tarafımdan üretilen bu tez çalışmasında; sunduğum tüm veri, enformasyon, bilgi ve belgeleri bilimsel etik kuralları çerçevesinde elde ettiğimi, tüm değerlendirme, analiz, bulgu ve sonuçları bilimsel usullere uygun olarak sunduğumu, tez çalışmasında yararlandığım kaynakların tümüne bilimsel normlara uygun biçimde atıfta bulunarak kaynak gösterdiğimi, tezimin kaynak gösterilen durumlar dışında özgün olduğunu bildirir, aksi bir durumda aleyhime doğabilecek tüm hak kayıplarını kabullendiğimi beyan ederim.

09 Eylül 2021

İrem UYANIK

TEŐEKKÖR

Bu alıőmanın gerekleőtirilmesinde, deęerli bilgilerini benimle paylaőan ve tezin birok aőamasındaki yardımlarından dolayı danıőman hocam Sayın Dr. Öęr. Üyesi Ali Öztürk'e ve desteklerini esirgemeyen sevgili aileme iten teőekkürlerimi sunarım.

Eylöl, 2021

İrem UYANIK

ÖZET

İrem UYANIK

KÜLTÜREL MİRAS GÖRÜNTÜLERİNİN DERİN ÖĞRENME YÖNTEMLERİ İLE SINIFLANDIRILMASI

Yüksek Lisans

Konya, 2021

Kültürel miras; bir toplumun veya ulusun ortak geçmişlerini gelenek ve göreneklerini oluşturan aynı zamanda geleceğin de doğru kurulmasına yardım eden tarih ve kültürle ilgili soyut ve somut değerlerdir. Geçmiş ve gelecek arasında bağlantı kurduğundan değer verilmesi, saklanması ve korunması gerekmektedir.

Kültürel miras görüntülerinin sınıflandırılması dijital belgeleme için gereklidir. Çok sayıda görüntü işleneceği için bunların sınıflandırılması sıkıcıdır, hatalara açıktır ve çok fazla zaman alır. Bu sıralama görevlerini kolaylaştıracak otomatik tekniklerin olması, dijital dokümantasyon sürecinin önemli bir bölümünü iyileştirecektir. Ek olarak, mevcut görüntülerin doğru bir şekilde sınıflandırılması, söz konusu miras varlığını inceleme ve yorumlama, daha iyi yönetme ve belirli terimler aracılığıyla daha verimli aramalara kolaylık sağlar.

Literatüre bakıldığında daha çok uluslararası alanda ilgili çalışmalar yapıldığı görülmektedir. Bu çalışmada kültürel miras görüntülerinin doğru bir şekilde sınıflandırılması amacıyla derin öğrenme yöntemleri (AlexNet, GoogleNet, ResNet ve SqueezeNet) kullanılmıştır. Görüntü veri kümeleri ile önce eğitimleri tamamlanmış daha sonra performans testleri gerçekleştirilmiştir. Sonuçlara bakıldığında hem eğitim hem de test süreçlerinde sistemin büyük oranda başarılı olduğu görülmüştür. İleride literatüre yeni ve faydalı çalışmalar eklenmesine olanak sağlamaktadır.

Anahtar Kelimeler

AlexNet, GoogleNet, ResNet, SqueezeNet, Derin öğrenme yöntemleri, Kültürel miras

ABSTRACT

İrem UYANIK

Classification of Cultural Heritage Images Using Deep Learning Methods

Master's

Konya, 2021

Cultural heritage are intangible and tangible values related to history and culture that constitute the common pasts, traditions and customs of a society or nation, and also help to establish the future correctly. Because it connects the past and the future, it needs to be valued, preserved and protected.

Classification of cultural heritage images is essential for digital documentation. Classifying them is tedious, error-prone, and takes too much time, as a large number of images will be processed. Having automated techniques to facilitate these sequencing tasks will improve a significant part of the digital documentation process. In addition, the correct classification of available images facilitates the examination and interpretation of the heritage asset in question, facilitating better management and more efficient searches through specific terms.

Looking at the literature, it is seen that related studies are conducted in the international arena. In this study, deep learning methods (AlexNet, GoogleNet, ResNet and SqueezeNet) were used in order to correctly classify cultural heritage images. With the image datasets, first the training was completed and then the performance tests were carried out. Looking at the results, it was seen that the system was largely successful in both training and testing processes. It allows the addition of new and useful studies to the literature in the future.

Keywords

AlexNet, GoogleNet, ResNet, SqueezeNet, Deep learning methods, Cultural heritage

İÇİNDEKİLER

BİLDİRİM	ii
ETİK BEYAN.....	iii
TEŞEKKÜR.....	iv
ÖZET.....	v
ABSTRACT	vi
İÇİNDEKİLER	vii
TABLolar DİZİNİ	viii
ŞEKİLLER DİZİNİ.....	ix
KISALTMALAR	xi
1.GİRİŞ	1
2.LİTERATÜRDE YAPILMIŞ BENZER ÇALIŞMALAR.....	6
3.KULLANILAN YÖNTEMLER	8
3.1.Yapay Sinir Ağları.....	8
3.2.Derin Öğrenme	10
3.2.1.Derin Öğrenme Katmanları	11
3.2.2.Konvolüsyonel Sinir Ağları.....	13
3.2.3.Çalışmada Kullanılan Derin Öğrenme Modelleri.....	14
3.2.4.Nesne Tanıma	20
3.2.5.Görüntü İşleme	21
3.3.Aktarım Öğrenmesi (Transfer Learning)	22
4.DENEYSEL SONUÇLAR	26
4.1. Veri Kümesinin Oluşturulması.....	26
4.2.Matlab Ortamında CNN Modelleri ile Eğitim Yapılması.....	29
4.2.1.AlexNet.....	29
4.2.2.GoogleNet.....	30
4.2.3.ResNet.....	31
4.2.4.SqueezeNet	33
4.3.Deeptest Sonuçları ve Karşılaştırması.....	35
5. SONUÇ	42
KAYNAKLAR	45
ÖZGEÇMİŞ	52

TABLolar DİZİNİ

Tablo 1. Elimizde mevcut olan ilk verilerin sayısı	27
Tablo 2. 227*227 boyutunda artırılmış verilerin sayısı	27
Tablo 3. 224*224 boyutundaki artırılmış verilerin sayısı	28
Tablo 4. Test için kullanılacak yeni görüntülerin sayısı	28
Tablo 5. Kullanılan yöntemlerin karşılaştırılması.....	34
Tablo 6. Deeptest sonuçları için accuracy değerleri	37
Tablo 7. AlexNet için elde edilen sonuçlar	38
Tablo 8. GoogleNet için elde edilen sonuçlar	39
Tablo 9. ResNet için elde edilen sonuçlar.....	40
Tablo 10. SqueezeNet için elde edilen sonuçlar	41
Tablo 11. Derin öğrenme modellerinin boyutlarının karşılaştırılması.....	42

ŞEKİLLER DİZİNİ

Şekil 1. Unesco'ya göre kültürel mirasın sınıflandırılması	5
Şekil 2. AlexNet mimarisi.....	15
Şekil 3. AlexNet mimarisi.....	16
Şekil 4. GoogleNet inception boyut azaltma modülü (inception modülü).....	16
Şekil 5. GoogleNet ağ mimarisi	17
Şekil 6. Residual blok	18
Şekil 7. Microsoft ResNet ilk 34 katmanın ağ mimarisi	19
Şekil 8. SqueezeNet ağ mimarisi.....	20
Şekil 9. AlexNet'in split_train_test 0,5 için eğitimi sonunda test verilerinden bazıları için tahmin sonuçları	29
Şekil 10. AlexNet'in split_train_test 0,5 için eğitim aşamasına ait doğruluk ve kayıp grafikleri.....	30
Şekil 11. GoogleNet'in split_train_test 0,5 için eğitimi sonunda test verilerinden bazıları için tahmin sonuçları	31
Şekil 12. GoogleNet'in split_train_test 0,5 için eğitim aşamasına ait doğruluk ve kayıp grafikleri.....	31
Şekil 13. ResNet'in split_train_test 0,5 için verilerin eğitimi sonunda test verilerinden bazıları için tahmin sonuçları	32
Şekil 14. ResNet'in split_train_test 0,5 için eğitim aşamasına ait doğruluk ve kayıp grafikleri.....	32
Şekil 15. SqueezeNet'in split_train_test 0,5 için verilerin eğitimi sonunda test verilerinden bazıları için tahmin sonuçları.....	33
Şekil 16. SqueezeNet split_train_test 0,5 için eğitim aşamasına ait doğruluk ve kayıp grafikleri.....	33
Şekil 17. Deeptest sonrası AlexNet'in split_train_test 0,5 için tahmin sonuçları	35
Şekil 18. Deeptest sonrası GoogleNet'in split_train_test 0,5 için tahmin sonuçları	35
Şekil 19. Deeptest sonrası ResNet'in split_train_test 0,5 için tahmin sonuçları	36
Şekil 20. Deeptest sonrası SqueezeNet'in split_train_test 0,5 için tahmin sonuçları.....	36
Şekil 21. 'adam' optimizasyon yöntemi ile AlexNet eğitim aşamasına ait doğruluk ve kayıp grafikleri	38
Şekil 22. 'adam' optimizasyon yöntemi ile GoogleNet eğitim aşamasına ait doğruluk ve kayıp grafikleri	39
Şekil 23. 'adam' optimizasyon yöntemi ile ResNet eğitim aşamasına ait doğruluk ve kayıp grafikleri	40

Şekil 24. 'adam' optimizasyon yöntemi ile SequeezeNet eğitim aşamasına ait doğruluk ve kayıp grafikleri	41
---	----

KISALTMALAR

Kısaltma	Açıklama
3B	3 Boyutlu
ADALINE	Adaptive Linear Neuron
AE	Oto-kodlayıcılar - Autoencoders
CAD	Computer Aided Design
CNN	Convolutional Neural Networks
DBN	Derin İnanç Ağları
DCNN	Deep Convolution Neural Network
DVM	Destek Vektör Makinesi
ESA	Evrişimsel Sinir Ağları
Fast R-CNN	Fast Region Convolutional Neural Network
GPU	Grafic Process Unit
LiDAR	Light Detection and Ranging; veya Laser Imaging Detection and Ranging
MNIST	Modified National Institute of Standards and Technology database
MRI	Manyetik Rezonans Görüntüleme
RBM	Restricted Boltzmann Machines
R-CNN	Region Convolutional Neural Network
ReLU	Rectified Linear Unit
RGB	Red Green Blue
RNN	Recurrent Neural Network
ROI	Region of Interest
SRN	Simple Recurrent Network
SVHN	The Street View House Numbers
TCS	Traction Control System – Çekiş Kontrol Sistemi
YSA	Yapay Sinir Ağları

1.GİRİŞ

Kültürel miras ve yerler bir ulus, hatta bütün dünya için çok önemlidir. Kültürel mirasa sahip mekanları gezip görmek, hakkında bilgi sahibi olmak bir ulusu tanımak demektir. Teknolojinin gelişmesiyle, akıllı telefonlar ve tablet kullanımının yaygınlaşmasıyla insanlar her türlü bilgiye kısa sürede ulaşabilmektedir. Bu tezin amaçlarından birisi, kişilerin çektikleri kültürel miras fotoğraflarını, akıllı telefonlarına veya tabletlerine yüklediklerinde hem hızlı hem de doğru bir şekilde hakkındaki bilgilere ulaşmasını sağlayacak model ya da modellerin belirlenmesidir. Bu açıdan bakıldığında, hedeflenen modelin daha az depolama alanı dolayısıyla daha az bellek gereksinimi ve daha az işlemci gereksinimi duyması beklenir.

Kültürel mirasa sahip mekanların görüntülerinin depolanması ve sınıflandırılması günümüzde kültürel mirasın korunması için önemli bir görevdir. Çok fazla sayıda görüntü işlendiği için var olan tekniklerle yapılan sınıflandırmalar çok zaman ve emek istemektedir. Görüntünün doğru sınıflandırılması kültürel mirasın incelenmesi ve yorumlanması için önem teşkil etmektedir. Bu tezin diğer amacı, kültürel mirasın korunması sürecinde tarihsel miras görüntülerinin sınıflandırılmasında derin öğrenme modelleri için en yüksek doğruluğun elde edilmesini araştırmaktır. Kültürel miras çalışmalarında kullanılacak bir sistemin istemci-sunucu mimarisinde olması mantıklıdır. Sunucu bilgisayarda istemcilerden gelecek anlık görüntü sınıflandırma talepleri için kullanılacak eğitilmiş derin öğrenme modelimiz yer alabilir. Ya da, istemcilerden gelen görüntüler biriktirilerek derin öğrenme modeli çevrim-dışı biçimde eğitilebilir. Modelimiz güçlü bir sunucuda çalışacağı için işlemci hızı ve depolama gereksinimi ikincil olarak düşünülecek önemdedir (Junuz, 2019 (Llamas v.d., 2017).

Bu tez çalışmasında, derin öğrenme modelleri sıfırdan eğitilerek ya da önceden eğitilmiş olanların üzerinde ince ayar yapılarak performansları değerlendirilmiştir. Tüm bunlar kültürel değeri olan yapıların görüntülerinin sınıflandırılması için uygulanmıştır. Yüksek doğruluk değerleri elde edildiği takdirde otomatik bilgilendirmeye ve kültürel mirası koruma çalışmalarına katkı sağlayacağı düşünülmektedir.

Bu tez çalışması genel olarak literatürde “sahne tanınma ve sınıflandırma” kavramı kapsamında yapılan çalışmalar başlığı altında değerlendirilebilir. Daha önce yapılan

'sahne tanıma ve sınıflandırma' işlemlerinde farklı yöntem ve yaklaşımlar kullanılmıştır. Bu yöntem ve yaklaşımlar genel olarak iki temel başlıkta ele alınmıştır:

1. Nesne tanımaya bağlı sahne sınıflandırma

2. Düşük seviyeli görüntü özelliklerini kullanarak sahne sınıflandırma

Nesne tanımaya bağlı sahne sınıflandırması için nesne tanımayı kullanmak, sahne sınıflandırmasına yönelik temel bir yaklaşımdır ve düşük seviyeli özellikleri kullanarak yapılması zor ve karmaşık sahnelerin tanınmasına yardımcı olur. Düşük seviyeli görüntü özelliklerinin sahne sınıflandırması için etkili bir işlem olduğu kanıtlanmıştır, yüksek seviyeli işlemler için, bu tür düşük seviyeli görüntü örnekleri yeterli değildir. Ama nesne bankası örneği kullanılarak yüksek seviyeli görsel tanıma görevlerinde başarılı sonuçlar elde edilebilir (Li vd., 2010)(Felzenszwalb vd., 2010).

Düşük seviyeli görüntü işlemleri; birçok çalışmada düşük seviyeli işlemler kullanılmaktadır. Buradaki zorluk, görüntünün anlamsal bağlamını çıkarmak için başarıyla kullanılacak düşük seviyeli özellikleri bulmaktır. Görüntüden sahne tanıma amacıyla çıkarılabilecek birçok özellik arasında doku, yönelim ve renk yaygın olarak kullanılmış ve farklı veri setleri ve sınıflandırıcılar ile uygulanmıştır. Doğal sahne kategorilerini öğrenmek için yapılan çalışmalarda görüntü tanımlama için düşük seviyeli doku özellikleri kullanılmıştır. Başlangıçta yerel bölgeler farklı ara temalar ve ardından kategoriler halinde sınıflanır. Sahneleri temsil etmek için kod kelimelerinin dağılımından en iyi yararlanan modeli elde etmek için öğrenme algoritması, değiştirilmiş bir Latent Dirichlet modelidir. Bir görüntünün bir temaya zor bir şekilde atandığı geleneksel sahne modellerinin aksine, algoritma bir görüntü ile ilişkilendirilebilecek bir tema koleksiyonu üretir. Renk tabanlı özellikler nadiren kullanılsa da bazı durumlarda sahneleri ayırt etmede çok daha kullanışlı ve güçlüdür. Renk tanımlayıcılar, ayırt etme gücünü artırmak için sahne ve nesne tanıma için kullanılabilir. Teorik ve deneysel sonuçlardan, ışık yoğunluğunun ve açık-koyu renk değişikliklerinin sınıflandırmayı etkilediği görülmüştür (Blei vd., 2003)(Van De Sande vd., 2010).

İç ve dış mekan sınıflandırma; yüksek seviyeli sahne özelliklerinin özellikle iç mekan - dış mekan sahne alma problemi için düşük seviyeli özelliklerin sınıflandırılmasından elde edilebilmektedir. Algoritmalar üç tür özellik çıkarır; ilki, ohta renk uzayındaki

histogram, bir diğreri çok çözünlürlüklü eş zamanlı otoregresif model parametreleri, üçüncüsü ise değişmez katsayılar. Performansın, alt bloklardaki bilgilerin sınıflandırılarak ve bunların toplanıp birleştirilerek geliştiği görülmüştür. Sahnelerin kategorisini tahmin etmek için düşük seviyeli özellikler kullanır. Bu, bayes ağı kullanılarak yapılan basit bir iç-dış mekan sınıflandırmasıdır. Düşük boyutlu renk ve dalgacık dokusu özellikleri, destek vektör makinesi (DVM) sahneleri sınıflandırmak için kullanılır (Szumner ve Picard, 1998)(Serrano vd., 2004).

Anlamsal tipiklik; gerçek dünyadaki doğal sahnelerin kategorize edilmesinde ilginç bir anlamsal tipiklik kavramı vardır. Önerilen tipiklik ölçüsü, bir sahne kategorisine göre bir görüntünün benzerliğini derecelendirmek için kullanılır. Tipiklik, görüntünün belirsizliği için bir ölçü olarak tanımlanır. Bu önemli bir kavramdır çünkü birçok doğal sahne belirsizdir ve kategorizasyon doğruluğu bazen açıklamayı gerçekleştiren belirli bir kişinin fikrini yansıtır. Anlamsal tipiklik ölçüsü, doğal gerçek dünya sahnelerinin kıyıları, nehirler-göller, ormanlar, ovalar, dağlar ve gökyüzü-bulutlar dahil altı sahneye göre benzerliğini bulmak için kullanılır. Tipiklik, ortaya çıkma sıklığının bir sonucudur yani en tipik olarak kabul edilen öğeler kategori içindeki en çok ortak olan özelliklerdir (Rosch ve Mervis, 1975) (Rosch ve Mervis, 1976)(Vogel ve Schiele, 2004).

Yapılandırılmalı tanıma; bu sistemde, modellerin elle işlenmesi yerine, otomatik bir işlemin örnek kümedeki resimler arasındaki ilgili ilişkileri tanımlayan bir dizi şablon oluşturabildiği bir iyileştirme yapılmıştır. Bulanık kısım tabanlı bir model içinde tanımlanmış ve bulanık kümeler de bir görüntüdeki uzamsal tanımlayıcılar için yaygın ve etkili bir şekilde kullanılmıştır ve daha yakın zamanda, bulanık uzaysal ilişkiler deforme olabilen modellere entegre edilmiş ve MRI görüntülerine uygulanmıştır (Miyajima ve Ralescu, 1994)(Colliot vd., 2006).

Kovaryans tanımlayıcı; sahne düzeyinde coğrafi görüntü sınıflandırması için kovaryans tanımlayıcısına dayalı denetimli bir işbirlikçi çekirdek kodlama yöntemi önerilmektedir. Kovaryans tanımlayıcı; renk, uzamsal konum ve gradyan gibi farklı özelliklerin bir kovaryans matrisidir; rotasyon ve ölçek değişmezdir, ancak Riemannian uzayında bulunur ve bu nedenle öklid uzayında kullanılan geleneksel hesaplama ve matematiksel modeller kullanılamaz. Yöntem, üç aşamalı bir süreç olarak görülebilir. İlk adım covd özelliklerini coğrafi sahne görüntüsünden çıkarmaktır. İkinci adımda, kodlama temsil

aşamasında işbirlikçi çekirdek kodlaması denetlenir ve doğrusal sınıflandırma aşaması gerçekleştirilir. Son olarak, sınıflandırma aşamasında, sözlük katsayılarına ve öğrenilen lineer sınıflandırıcıya dayalı olarak bir etiket vektörü türetilir. Bu yöntem, sahne düzeyinde coğrafi görüntü sınıflandırması için verimli bir yöntemdir ve yüksek çözünürlüklü hava görüntüsü veri kümesinde tatmin edici performans sağlamaktadır (Yang vd., 2016).

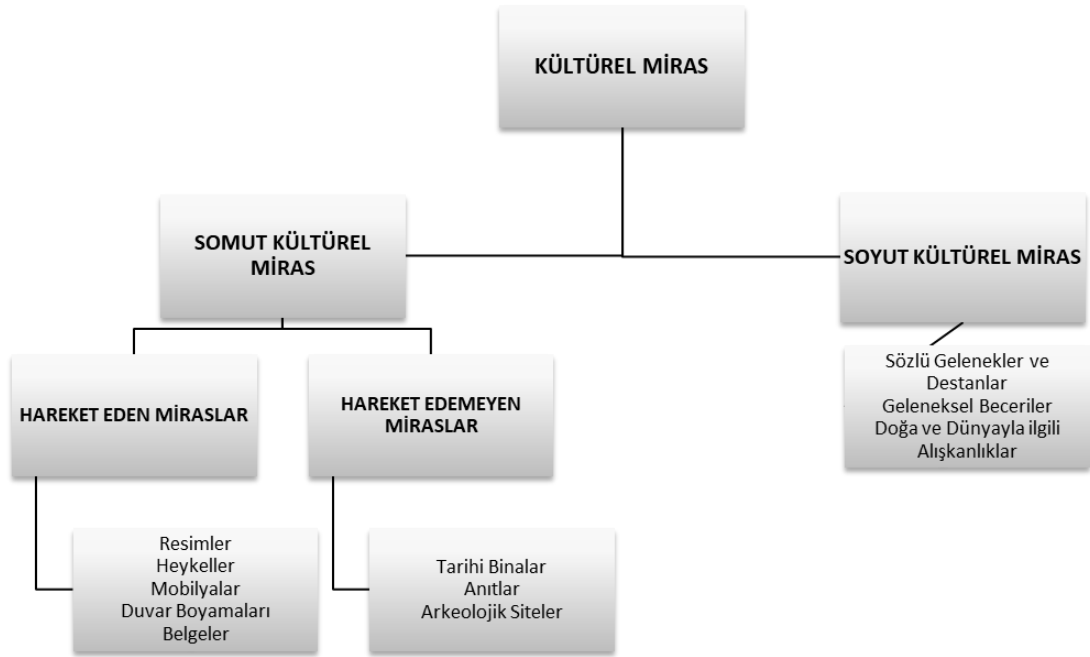
Sahnenin şekli; bu yaklaşımda sahneyi tek bir şekle sahip tek bir nesne olarak ele almayı önermektedir. Sahnenin uzamsal özelliklerini tanımlamaya özel olarak ayrılmış birkaç algısal boyutu kullanarak sahnenin şeklini bulmak için bir hesaplama modeli önerilmektedir. Mekanın şeklini tanımlayan duvarlar, kesit, yükseklik vb. bir küme olarak tanımlanır. Algoritmalar, etiketli eğitim verilerinden temel sahne kategorilerinin spektral imzalarını öğrenir. Daha sonra global özellikler ve spektral özellikler arasındaki ilişkiyi bulmak için bir öğrenme algoritması kullanılır (Oliva ve Torralba 2001).

Sahnenin ötesinde tanıma; amacı, görüntüdeki olayı tanımak-sınıflandırmak ve aynı zamanda görüntü içindeki nesneye ve sahne ortamına bir dizi anlam katmaktır. Olayın, sahneye ve nesneye bağlı olduğu varsayılır, ancak her ikisi de olayı tahmin etme olasılığını etkiler. Sahne tanıma, olay tanımayı büyük ölçüde etkiler. İlk yaklaşım olarak olay tanıma aslında sahne tanımadır. Sahne ve nesne tanımaya ek olarak, olayı doğru bir şekilde tanımlamada görüntünün düzeninin önemini anlaşılmıştır. Görüntünün düzenini tanımlamak için bazı basit ipuçları kullanılır ve tüm sahnenin ne (olay), nerede (sahne) ve kim (nesne) tanınmasını gerçekleştirerek bir görüntüye bütünleştirici ve hiyerarşik etiketler sağlamayı başarırlar. Başka bir çalışmada ise sahne sınıflandırmasında insan performansını ölçmüş ve bunu son teknoloji algoritmalarla karşılaştırmışlardır. Hem insan hem de algoritma sonuçlarında hatalar oluşmuştur, insanlar anlamsal olarak benzer kategoriler arasında gidip gelirken, algoritmalar sahte görsel eşleşmeler nedeniyle anlamsal olarak alakasız sahneler arasında hata yapmışlardır. Gerçek dünya genellikle sahnelerin kombinasyonunu içerdiği için, tüm bir görüntünün tamamını bir sahneyle etiketlemek yerine görüntülerdeki sahne türünü tanıma fikrini önermişlerdir. Bu yeni ve ilginç bir fikirdir ve gelecekteki sahne tanıma algoritmalarının ilerleyebileceği yönlerden biri de olabilir (Murphy vd., 2003)(Fei-Fei ve Perona 2005) (Xiao vd., 2010).

Bu çalışmada ise, kişilerin doğru biçimde ve kısa sürede bilgilendirilmesini sağlamak ve kültürel mirasın korunması çalışmalarına katkı sağlamak amacıyla farklı derin öğrenme modellerinin aktarım öğrenmesi yöntemi ile kullanılarak farklı kültürel miras görüntüleri üzerinde sınıflandırma performansları değerlendirilmiştir.

Kültürel miras; kültürümüzle, tarihimizle ilgili somut ve soyut değerlerin tümüdür (İSMEP, 2014).

UNESCO 'ya göre kültürel miras Şekil 1'deki gibi kategorize edilmektedir.



Şekil 1. Unesco'ya göre kültürel mirasın sınıflandırılması

(Kaynak: Kurniawan v.d., 2011)

Şekil 1'e göre kültürel miras, somut kültürel miras ve soyut kültürel miras olmak üzere ikiye ayrılmaktadır. Somut kültürel ise hareketli miraslar ve hareketsiz miraslar şeklinde iki farklı kategori altında düşünülmektedir. Hareket eden miraslar; resimler, heykeller, mobilyalar, duvar boyamaları ve belgelerden oluşur. Hareket edemeyen miraslar ise; tarihi binalar, anıtlar ve arkeolojik sitelerdir. Kültürel mirasın diğer bir kolu olan soyut kültürel miras ise; sözlü gelenek ve destanlar, geleneksel beceriler, doğa ve dünyayla ilgili alışkanlıklar olarak ayrılır.

2.LİTERATÜRDE YAPILMIŞ BENZER ÇALIŞMALAR

Literatür incelendiğinde, kültürel miras görüntülerinin yapay zeka yöntemleri ile sınıflandırılması ile ilgili hiç ulusal yayın olmadığı ve çok az uluslar arası yayın olduğu görülmektedir.

Jose Llamas, Pedro M. Leronés, Eduardo Zalama, Jaime Gómez-García-Bermejo, kültürel mirasın (KH) dijital olarak belgelenmesi, büyük miktarda görüntünün yorumlanması ve sınıflandırılması konusunda bir çalışma yapmışlardır. Fotoğrafların yorumlanmasına ve semantik terimler kullanılarak arama yapılmasına olanak tanıyan otomatik tekniklerin mevcudiyeti, proje geliştirme çalışmalarını kolaylaştırmıştır. Makalelerinde, kültürel mirasın görüntülerini analiz etmek için derin öğrenme tekniklerinin, özellikle evrişimli sinir ağlarının (CNN'ler) kullanımı değerlendirilmektedir. Bu tekniklerin uygulanmasının, kültürel mirasın dijital belgelenmesinde aranan hedeflere önemli katkılarda bulunabileceği düşünülmektedir (Llamas v.d., 2016).

Montoyo Obeso A. , Garcia Vazquez M. S. , Ramirez Acesto A. A. , Benois-Pineau J., çalışmalarında multimedya içeriğinin otomatik açıklaması, esas olarak sınıflandırma görevleri ve büyük veri sıralaması için geliştirilmiştir. Kültürel mirasın korunması, bu yöntemin uygulanması için büyük önem taşıyan bir alan olduğunu düşünmüşler ve Meksika kültürel mirasının dijital fotoğraflarındaki binaların mimari tarzlarının sınıflandırılması konusunda bir çalışma yapmışlardır. Bu çalışmada, görüntülerdeki görsel dikkati tahmin etmek için belirginlik odaklı bir yaklaşım kullanarak, bunu Meksika binalarının mimari stilini tanımlamak için bir Evrişimsel Sinir Ağı eğitmek için kullanmışlardır. Ayrıca, geleneksel kırılmış görüntü ve belirginlik haritaları altında eğitilen modellerin davranışlarının bir analizini sunulmuş olup bu anlamda, doğrulama veri setinde %97'lik bir sınıflandırma oranına ulaşan belirginlik tabanlı CNN'lerin performansının geleneksel eğitimden daha iyi olduğunu kanıtlamışlardır. Bu teknikle Meksika kültürel mirasının otomatik belgelenmesinde geniş katkı sağlayabileceği düşünülmektedir (Obeso 2016).

Jankovic R.'e göre görüntü sınıflandırma, dijital çağın en önemli görevlerinden biridir. Kültürel miras açısından, görüntü sınıflandırması genellikle çok büyük veri kümeleri kullandığından, iyi doğruluk elde eden ancak aynı zamanda daha az hesaplama yoğun

olan sınıflandırma yöntemleri geliştirmek önemlidir. Bu çalışma, farklı sınıflandırma algoritmalarını eğitmeyi ve test etmeyi amaçlamıştır ve bunları Evrişimli Sinir Ağı'ndan (CNN) elde edilen sonuçlarla karşılaştırılmıştır. Görüntülerden üç tür özellik çıkarılmıştır: (i) kenar histogramı, (ii) renk düzeni ve (iii) JPEG katsayıları. Algoritmalar öznitelik seçimi uygulanmadan önce ve uygulandıktan sonra test edilmiş ve sonuçlar her iki durumda da çok katmanlı algılayıcı için en iyi sınıflandırma performansının elde edildiğini göstermiştir (Jankovic, 2019).

Cosovic M. ve Jankovic R.'e göre kültürel miras görüntü sınıflandırması, dijitalleşme sürecindeki en önemli görevlerden birini temsil etmektedir. Görüntü sınıflandırması için evrişimli sinir ağı kullanılmıştır, aynı mimari 10 kategoriden oluşan tam veri kümesi ve 5 farklı görüntü kategorisine sahip veri kümesine uygulanmıştır. Sonuçlar, her iki mimarinin de iyi performans gösterdiğini ve %90'a varan doğruluk elde ettiğini göstermiştir (Cosovic Jankovic, 2020).

Kulkarni UN., Meena S.M., Gurlahosur S.V., Mudengudi U. Çalışmalarında Hindistan'ın, zengin kültürel mirasa sahip, özellikle de 37'si UNESCO tarafından listelenen miras alanı olan ünlü mimari mekanlara sahip bir ülke olduğu ve bu kültürel varlıkların arşivlenmesi, belgelenmesi ve bilgilerinin paylaşılması görevi oldukça zor olduğu belirtmişleridir. Görüntülerin doğru etiketine (miras alanı) doğru bir şekilde tahmin edilmesi, belirli terimler aracılığıyla daha etkili aramalara olanak tanır ve böylece miras varlıklarının incelenmesine ve anlaşılmasına yardımcı olabileceği söylenmiştir çünkü görüntüleri içeren verilerin sınıflandırılması karmaşıktır ve zaman alıcıdır. Mevcut makine öğrenimi teknikleri, görüntülerin sınıflandırılmasını otomatize etmek için kullanılabilir. Yüksek işlemler bilgi işlem kaynaklarının ortaya çıkışı, son teknoloji makine öğrenimi ve derin öğrenme algoritmaları somut çözümler sağlar. Bu yazıda, Hindistan Dijital Miras Alanı (IHDS) anıt verilerini toplamak, görüntü sınıflandırması gerçekleştirmek ve Görüntü etiketlerinin sorguya dayalı olarak alınmasını sağlamak için bir kitle kaynaklı platform önerilmiştir. Ayrıca, ImageNet Dataset üzerinden önceden eğitilmiş MobileNet V2 mimarisinde IHDS veri seti kullanılarak yeniden eğitilen, transfer öğrenme tabanlı bir görüntü sınıflandırıcı tasarlanmıştır. Önerilen transfer öğrenme algoritması %98.75 çıkarım doğruluğunu göstermiştir (Kulkarni v.d., 2019).

3.KULLANILAN YÖNTEMLER

3.1.Yapay Sinir Ağları

Genel bir tabirle yapay sinir ağları insan beyninin biyolojik sinir yapısını taklit eder. Önceden öğrenilmiş ya da sınıflandırılmış bilgileri kullanarak yeni bilgiler meydana getiren ve karar verebilen bilgisayar programlarıdır. Örüntü tanıma, sistem tanılama, robotik, sinyal işleme, nonlineer denetim alanları gibi birçok alanda kapsamlı bir şekilde kullanılmaktadır.

Tekniksel olarak yapay sinir ağının işlevi, girdi seti olarak kendisine verilen bilgilere karşılık bir çıktı elde edebilmektir. Bunun yapılabilmesi için ağ eğitime girer ve ağ genelleme yapabilecek ve karar verebilecek düzeye ulaşır. Sonrasında alınan eğitim ve kazanılan tecrübe sayesinde çıktıları belirler (Efeoğlu ve Kaynak, 1999) (Efeoğlu vd., 1999)(Narendra ve Parthasarathy, 1990).

Günümüze gelinceye kadar yapay sinir ağları farklı alanlarda kullanılmıştır. Bu uygulama alanlarına; üretim planlama, beyin modellemesi, kalite kontrolü, parmak izi tanıma, otomatik araç denetimi, kredi kartı hilelerini açığa çıkartma, akıllı araçlar ve robotlar için yol saptama, ses tanıma, meteoroloji yorumlama, elektrik işareti tanıma, el yazısı saptama, hastalıkların tanımlanması ve tedavisi, radar ve sonar sinyalleri sınıflandırma, spam maillerin filtrelenmesi olarak örneklerlendirilebilir (Keskenler ve Keskenler, 2017).

Yapay sinir ağı insan beyninin öğrenme sürecinden yola çıkılarak ortaya atılmış bir yaklaşımdır. Bu yaklaşım ilk kez 1943 yılında, insan beynindeki hücrelerin yapısının matematiksel modellemesi örnek alınarak oluşturulmuştur. 1980 yılından sonra yapay sinir ağları konusunda büyük gelişmeler yaşanmıştır bu sebeple yapay sinir ağlarını araştırılırken 1980 öncesi ve sonrası olarak ayırmak gereklidir. Yapay sinir ağlarının tarihsel süreci aşağıdaki gibi özetlendirilebilir;

1940 yılından önce mühendislik değeri olmayan çalışmalar yapılmıştır.

Warren McCulloch ve Walter Pitts (1943): ilk yapay sinir ağı modelidir. Araştırmacılar, ilk yapay sinir ağı hücresinin yapısını oluşturarak YSA'nın temelini attılar. McCulloch ve Pitts, elektrik devrelerini kullanarak bu hücreyi geliştirmişlerdir. Yapay sinir hücreleri sayesinde mantıksal ifadelerin formülize edilebildiği anlaşıldı. Daha sonra iki

veya daha fazla hücrenin birlikte çalışması gerektiğini anlayarak paralel çalıştırma ile öğrenme kurallarını ortaya çıkarmaya başladılar (McCulloch ve Pitts, 1943).

Donald Hebb (1949): “Organization of Behavior” isimli kitabında “Hebbian öğrenme” ile ilişkili başlıca teoriyi ele aldı. Hebb prensibi sayesinde sinir ağının bağlantı sayısı ile öğrenebilmenin bağlantılı olduğu anlaşıldı. 1950’den sonra bazı araştırmacılar Hebb prensibinden faydalanarak yapay sinir ağının etkisini artırıcı bazı çalışmalarda bulunmuşlardır. Aynı dönemlerde IBM laboratuvarlarında yapay sinir ağı ile ilgili simülasyon çalışmaları başarısızlığa uğramıştır (Hebb, 1943).

Frank Rosentblatt (1957): 1957 senesinde “Perceptron” bulunduktan sonra yapay sinir ağı alanındaki buluşlar hızla çoğalmıştır. Bu çalışmadaki perceptron kavramı, en basit haliyle tek katmanlı ve tek çıkışı olan yapay sinir ağıdır (Rosentblatt, 1957).

Widrow ve Hoff (1960): Bernard Widrow ve Marcian Hoff 1959 senesinde Stanford Üniversitesinde ADALINE ve MADALINE olarak adlandırdıkları yapay sinir ağı modellerini geliştirdiler. Bu modeller, yapay sinir ağlarının mühendislik alanına yönelik ilk çalışmalarıdır. ADALINE daha sonraki yapay sinir ağı çalışmalarının temelini oluşturmuştur. MADALINE telefon hatlarında oluşan ses yankılarını ortadan kaldırmada kullanılmıştır ve ilk sinir ağı olup halen kullanılmaktadır. 1963 yılında Widrow’un geliştirdiği ters sarkaç denetleyicisi MADALINE’in en iyi örneklerindedir (Widrow ve Hoff, 1960).

Anderson (1972) ve Kohonen (1972): bu iki çalışmacı birbirlerinden habersiz bir şekilde hemen hemen aynı sonuçlara ulaştıkları “çağrışımli bellek” konulu çalışmalarını yayımladılar. Bu araştırmacılar, 1980 yılına gelindiğinde 10 yıla yakın durağan bir dönemin sonunda emeklerinin karşılığını almışlar ve yapay sinir ağları çalışmalarındaki duraklama süreci sona ermiştir (Anderson, 1972) (Kohonen, 1972).

Fukushima (1979): Fukushima, 1970 yılının sonlarında görüntü işlemeğe yönelik NEOCOGNITRON modelini geliştirdi. Bu model, biyolojik doğruluk yerine sonuçların kullanılabilirliği konusunda pragmatist yaklaşım içeriyordu. Bu modelin ardından daha fazla mühendislik uygulamaları görülmüştür (Fukushima vd., 1979).

Rumelhart, Hinton, & Williams (1986): 1982’de Hopfield tarafından ağların önemli sınıflarının matematik temelleri geliştirilmiştir. Gezgin satıcı probleminin çözülmesi bu yapıya en iyi örnektir. Yapay sinir ağları için bir icat olan çok katmanlı algılayıcıların

yaratılmasında kayda değer bir basamak olan Rummelhart'ın paralel programlama çalışmalarını kapsayan, 2 cilt olarak 1986 ve 1988 yıllarında basılan eserler bu çalışmaların temelini oluşturmuştur (Rumelhart vd., 1986) (Rumelhart ve McClelland, 1986) (Rumelhart ve McClelland, 1988).

Broomhead & Lowe (1988): Broomhead ve Lowe "Radyal tabanlı fonksiyonlar" modelini 1988 yılında oluşturmuşlardır. Bu ağ modeli çok katmanlı algılayıcılara alternatif olarak geliştirilmiştir ve çok daha büyük başarılar elde edilmiştir. Bunlara ek olarak Specht, bu ağların gelişmiş modeli olan "Genel Regrasyon Ağları" ve "Probalistik Ağları" adlı modelleri geliştirmiştir (Broomhead ve Lowe, 1988).

Pan, Z., Rust, A. G., & Bolouri, H. (2000): Çok katmanlı sinir ağı modelinin meydana çıkmasıyla birlikte katmanların sayısını artırarak daha başarılı sonuçlar elde edilmesi için Convolutional Neural Network (CNN) geliştirilmiştir. Buradaki sinir ağı yapısı gizli katmanlar içermekte ve elde edilen sonuçlar epeyce başarılı olmaktadır (Pan vd., 2000).

Hinton & Salakhutdinov (2006): 2006 senesinde Geoffrey Hinton ve Ruslan Salakhutdinov'un birlikte yayınladığı makale ile derin öğrenme terimi ortaya atılmış ve derin öğrenme çalışmaları başlamıştır (Hinton ve Salakhutdinov, 2006).

3.2.Derin Öğrenme

Hinton yaptığı çalışmalar sonucu elde ettiği sonuçlarla birlikte yayınladığı makalede yapay sinir ağlarına taze bir yaklaşım getirmiştir. Bu yaklaşım derin öğrenme (Deep Convolution Neural Network) olarak isimlendirilmiştir. Konvolüsyonel sinir ağları çok katmanlı sinir ağları olarak tanımlanmaktadır. Bu sinir ağı yapısı ile kayda değer çalışmalar yapılmış ve büyük oranda başarılı sonuçlar kazanılmıştır (Hinton ve Salakhutdinov, 2006)(Krizhevsky vd., 2012)(Szegedy vd., 2015).

Derin Öğrenme ilk olarak 2012 senesinde bilim alanında büyük ses getirmiştir. Nesne tanımlama kategorisinde en büyük yarışma olan Büyük Ölçekli Görsel Tanıma Yarışması (ImageNet)(Competition 2012) o sene derin öğrenmede temel mimari varsayılan Evrimsel Sinir Ağı(ESA) ile kazanıldı. Bu yarışmada Derin Öğrenmenin çok büyük bir yükselişe geçmesinin nedeni yarışmada %26,1 olan Top-5 hata oranı %15,3'e düşürülmesidir. Derin Öğrenmedeki ilerlemeler neticesinde Top-5 hata oranı %3,6'ya kadar düşürülmüştür (Krizhevsky vd., 2012).

Derin öğrenmenin tarihi çok geçmişe dayanmasına karşılık son yıllarda başarılı sonuçlar elde edilmiştir. Bunun en önemli nedenlerinden biri elde edilen verilerin fazla olmasıdır.

Son yıllarda, iş dünyasında analitik, büyük veri ve yapay zeka gibi kelimeler büyük ilgi görüyor. Bunun en önemli nedenlerinden ilki, üretilen veri miktarı son zamanlarda büyük ölçüde artış göstermesi, ikincisi ise etkili pazarlama programlarının bir “analitik” yarışına girmesidir.

Her gün çok fazla sayıda veri elde edilmektedir ve elde edilen bu veriler, derin öğrenmenin işlemlerini sağlayan tek kaynaktır. Derin öğrenme algoritmaları eğitim işlevini yapabilmek için çok büyük miktarda veriye ihtiyaç duyar, veri üretimindeki bu artış, son zamanlarda derin öğrenme yöntem ve işlevlerinin artmasına neden olmuştur. Daha fazla veri üretmeye ilave olarak, derin öğrenme algoritmaları, yapay zekanın gelişmesinin yanında, günümüzde kullanılan daha güçlü bilgi işlem gücünden faydalanır.

Derin öğrenme, makinelerin çok farklı ve karmaşık sorunların bile çözülmesini sağlar. Derin öğrenme algoritmalarının iyi performans sağlaması ne kadar çok eğitime girdiğine bağlıdır.

Kişiler ve kurumlar arasında da günden güne artan veri sayısı ve çeşitliliği, yeniliklere yol açma ve geleceğe yön verme konusunda kişiler ve kurumlar arasında da fark edilmiş ve bu teknolojilere ağırlık vermelerine neden olmuştur (Demirden, 2019).

3.2.1. Derin Öğrenme Katmanları

Giriş (Input) Katmanı: Verilerin giriş katmanı olarak kabul edilmektedir. Sinir ağlarında elde edilecek olan veri kümesi ağın mimarisine göre belirlenmelidir. Sinir ağına girecek olan veriler giriş katmanına eğitim için girerler ve ‘girdi verisi’ olarak adlandırılırlar. Veriler arka arkaya gelerek bir veri seti oluşturur (Tamura ve Tateishi, 1997).

Konvolüsyon(Convolution) katmanı: Bu katman konvolüsyonel sinir ağlarının temelini oluşturur. Bu katmanda girdi verisi üzerinde daha önceden belirlenen bir filtrenin girdi verisi üstünde gezdirilerek girdinin belirgin ve farklılık yaratan özelliklerinin açığa çıkmasını hedeflemektedir. Filtreleme işlemi sonunda giriş verisinden daha küçük bir matris ortaya çıkartır (Ciresan, 2011).

Aktivasyon (Relu) katmanı: Konvolüsyon katmanından sonra aktivasyon katmanı gelir. Çok fazla aktivasyon fonksiyonu bulunmasına rağmen derin öğrenme mimarilerinde Relu (Rectified Linear Unit) $f(x) = \max(0, x)$ fonksiyonu kullanılmaktadır. Aktivasyonsonucunda elde edilen değer negatif ise 0 pozitif ise 1 değerini alır (Krizhevsky vd., 2012)(Jarrett vd., 2009).

Havuzlama (Pooling) Katmanı: Havuzlama katmanı genellikle relu katmanından sonra yer alır. Havuzlama katmanı verilerin giriş boyutunu azaltır ve bir sonraki katman için hazırlar. Bu işlem sonrasında veride kayıplar oluşabilmektedir ama verilerde oluşabilecek bu kayıplar ağı hızlanmasına sebep olmaktadır. Böylelikle hesaplama miktarı ve bellek alanı da azalır (Hinton vd., 2012).

Tam Bağlı (Full-Connected) Katman: Bu katmandaki bütün nöronlar tek sıra şeklinde görünür. Katmandaki nöronlar, önceki katmandaki aktivasyonların hepsine bağlı durumdadır, özellikler önceki katmana bağlıdır. Nesneyi belirleyecek olan özelliklerin hangi sınıfla bağlantılı olduğu saptanır. Bir tam bağlı katman bir sınıfla ilişki derecesi yüksek olan belirleyici özelliklere bakar. Bu özellikler göz önüne alınarak hangi sınıfa ait olduğu belirlenir (Lin vd., 2013)(LeCun vd., 1995).

Dropout Katmanı: Çok katmanlı yapay sinir ağlarında sinir ağı eğitilirken aşırı öğrenme denilen ağı ezberlenmesi durumu oluşur. Bu durumun oluşması istenmez. Ağı ezberlenmesini engellemek için ağda ezber yapan bazı düğümlerin yok edilmesi gerekmektedir ve ağı ezberlenmesi ortadan kaldırılmaya uğraşılır (Srivastava vd., 2014).

Dropout katmanı: Tam bağlı katmanlar için bir düzenleme katmanıdır. Dropout işlemi, sinir ağının düzgün hale gelmesi işlemini yerine getirdiği yapılan test performanslarıyla kanıtlanmıştır (Tang, 2013).

Sınıflandırma (Classification) katmanı: Sınıflandırma katmanında işlem yapılacak küme sayısı kadar elde edilir. Sonuçlardan her biri bir sınıfı temsil eder. Son katman olan sınıflandırıcı katman için genellikle softmax sınıflandırıcı kullanılır (Ciresan, 2011).

Yumuşatma (Softmax) Katmanı: Softmax katmanı sınıflandırma işlemi yapmak için tam bağlı katmandan girdi verisini alır. Hangi sınıfa ait olduğuna dair bilgi üretir. Derin öğrenme ağı içerisinde her bir sınıf için olasılık değerini belirler. Bu işlemler için çapraz entropi kullanılır (Tang, 2013).

Normalizasyon (Normalization) Katmanı: Derin konvolüsyonel sinir ağlarının eğitilmesi ve hesaplanması çok fazla zaman almaktadır. Nöronların aktivasyonu normalize edilirse eğitim süresi azalabilir. Çoğunlukla Relu katmanı sonrasında normalleştirme işlemi gerçekleştirilir (Li vd., 2015).

Derin öğrenme mimarileri aşağıdaki gibi sınıflandırılabilir:

- Konvolüsyonel Sinir Ağları (Convolution Neural Network-CNN)
- Tekrarlayan Sinir Ağları (Simple Recurrent Network-SRN)
- Uzun Kısa Vadeli Hafıza Ağları
- Sınırlı Boltzmann makineleri (Restricted Boltzmann Machines-RBM)
- Derin inanç ağları (DBN)
- Oto-kodlayıcılar (Autoencoders-AE)

Bu çalışmada temel olarak sahne sınıflandırma yapılacağı Konvolüsyonel Sinir Ağları kullanılacaktır (Li vd., 2015).

3.2.2.Konvolüsyonel Sinir Ağları

Konvolüsyonel Sinir Ağları (Convolution Neural Network-CNN) çok katmanlı algılayıcıların bir çeşididir. Gelişmiş bir sinir ağı sistemi olan CNN algoritması, hayvanların görme merkezi örnek alınarak ortaya çıkmıştır. Buradaki konvolüsyon, bir nöronun uyarılara verdiği cevap olarak ele alınabilir. CNN, bir veya daha fazla konvolüsyonel katman, alt örnekleme katmanı gibi bir ve ya daha fazla katmandan oluşur(Fukushima vd., 1980) (Hubel ve Wiesel, 1968) (Lecun vd., 1968)

CNN algoritmaları görüntü ve ses işleme, doğal dil işleme, biyomedikal gibi farklı alanlarda kullanılmaktadır. Olabilecek en başarılı sonuçlar görüntü işleme alanında alınmıştır (Cireşan vd., 2012).

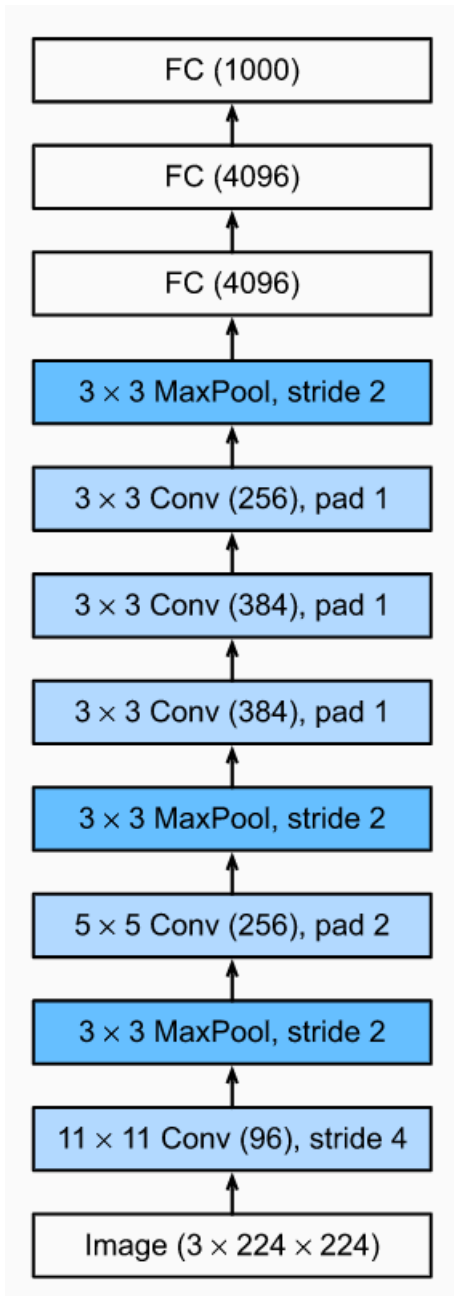
Konvolüsyonel sinir ağı ilk kez LeCun ve arkadaşları tarafından ortaya atılmıştır. Elde edilen bu yapay sinir ağına ise LeNet adı verilmiştir. LeNet sinir ağı çok katmanlıdır ve içerisinde birden fazla konvolüsyon katmanı, tam bağlı katman, aktivasyon katmanı, sınıflandırıcı katman, havuzlama katmanı ve bunlara ek bazı katmanlardan oluşmaktadır. Her katman kendi görevini yerine getirir ve sınıflandırıcı katmanda sonuç elde edilir. En çok sınıflandırma ve tespit işlemleri için kullanılan konvolüsyonel sinir

ağları derin öğrenme yöntemleri içerisinde en çok tercih edilendir. Konvolüsyonel sinir ağları, Imagenet tarafından 2012 ve 2014 yıllarında Büyük Ölçekli Görsel Tanıma yarışmasına katılmış ve en başarılı sonuçlar elde edilmiştir (Lecun vd., 1968) (Girshick vd., 2014).

3.2.3.Çalışmada Kullanılan Derin Öğrenme Modelleri

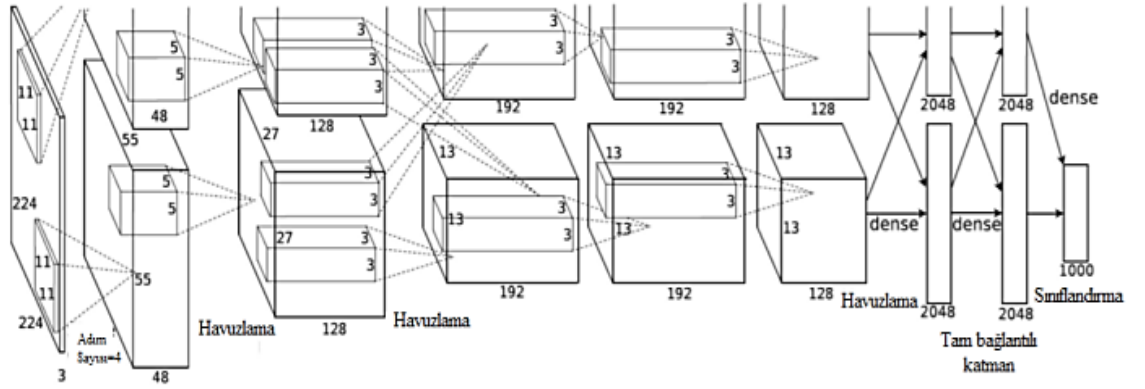
Derin öğrenme algoritmalarından ilki olan LeNet 2006 yılında ortaya çıkmış olsa da, derin öğrenme 2012 yılında yapılan ImageNet yarışmasıyla popüler olmuştur. Yapılan yarışmada nesne tanıma için algoritmalar ortaya konmuştur ve derin öğrenme mimarileriyle yapılan çalışmalarda çok başarılı sonuçlar elde edilmiştir. Bu sebeple daha sonraki yıllarda yarışmacılar derin öğrenme mimarileri üzerinde daha çok çalışma yapmışlar ve farklı modeller ortaya koymuşlardır. Bu modeller derin öğrenmenin yapıtaşdırlar. Çok geniş alanlarda derin öğrenme mimarilerinden faydalanılır. Bu mimariler görüntü tanıma, tespit ve sınıflandırmaya ek olarak gelecek tahmini, ilaç üretimi, sözlük, sinyal işleme, tıbbi alanlar, finans sektörü, savunma sanayi gibi bir çok alanda kullanılmaktadır. Derin öğrenme mimarilerini bilinen yöntemlerden daha iyi sonuç verdiği kanıtlanmıştır (İnik ve Ülker,2017)(Doğan ve Türkoğlu, 2018).

AlexNet her ne kadar derin öğrenmenin 1998 yılında ortaya çıktığı söylene de asıl derin öğrenme, 2012 yılında yapılan ImageNet yarışması ile bilinmiştir. Bu yarışmada AlexNet derin öğrenme mimarisi birinci gelmiştir. Bu mimari ile örüntü tanımda hata oranı % 26,2'den %15,4'e düşmüştür. Bu başarı ile AlexNet en popüler derin öğrenme mimarilerinden biri olmuştur. 25 katmandan oluşan AlexNet derin öğrenme ağı 5 temel konvolüsyon katmanından oluşmaktadır. Çoğunlukla konvolüsyon katmanlarından sonra relu katmanı kullanılmaktadır. Bunun yanı sıra giriş katmanı, normalizasyon katmanı, havuzlama(pooling) katmanı, dropout kamanı, tam bağlı katman, yumuşak bağlı (SoftMax) katman ve çıkış katmanı da bu 25 katmanı oluşturmaktadır (Krizhevsky vd., 2012) (You vd., 2017)(Iandola vd., 2016).



Şekil 2. AlexNet mimarisi

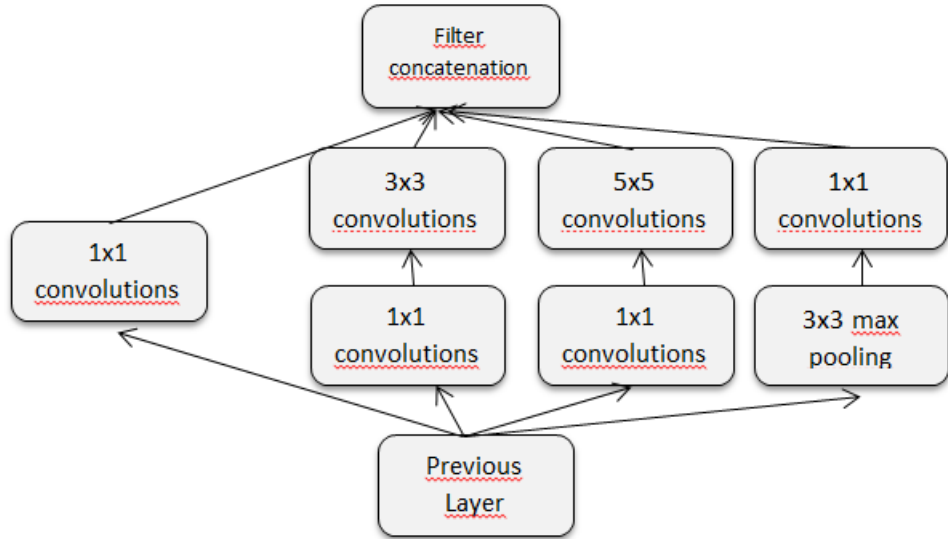
(Kaynak: Zhang v.d., 2019)



Şekil 3. AlexNet mimarisi

(Kaynak: Krizhevsky vd., 2012)

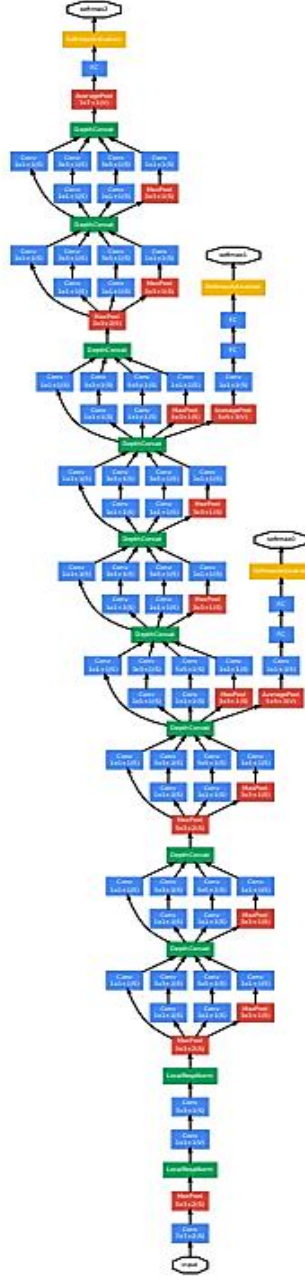
GoogleNet; GoogleNet 2014 yılında yapılan ImageNet yarışmasının kazananı olmuş, %5,7 hata oranı ile çok yüksek bir başarı elde etmiştir. Yapısındaki inception modülleri nedeniyle karmaşık bir yapıya sahiptir. GoogleNet 22 katman derinliğine sahiptir ve toplam 144 katmandan oluşmaktadır.



Şekil 4. GoogleNet inception boyut azaltma modülü (inception modülü)

(Kaynak: Doğan ve Türkoğlu, 2018)

Boyut azaltmak için kullanılan filtreler inception modülü sayesinde oluşturulur. Bu filtreler AlexNet'ten 12 kat daha az parametre içerir. Kullanılan katman sayılarında bağımsız yapı taşlarına göre değişiklik görülebilir (Krizhevsky vd., 2012) (Doğan ve Türkoğlu, 2018)



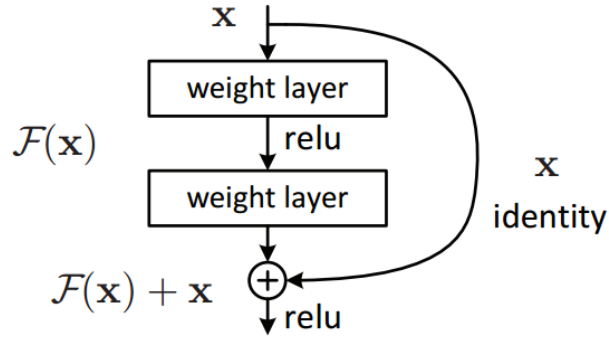
Şekil 5. GoogleNet ağ mimarisi

(Kaynak: Teles de Menezes vd., 2019)

ResNet; Microsoft ResNet 2015 yılında gerçekleştirilen ImageNet yarışmasında birincilik elde etmiştir. İnsanların görüntüyü sınıflandırırken hata oranı %5 ve %10 arasında değişirken ResNet'in hata oranı %3,6'dır. ResNet insandan daha iyi görsel tanıma yapmaktadır. Önceki derin öğrenme mimarilerinden daha derin bir yapıya sahiptir ve katman sayısı en fazla olanıdır. ResNet 152 katmandan oluşan bir mimaridir (Russakovsky vd., 2015).

Bu mimari Residual bloklardan oluşmaktadır. Bu yapı ile öğrenmenin daha hızlı şekilde gerçekleşeceği düşünülmüştür (He vd., 2016).

Residual blok örneği Şekil 6'da, ResNet ağ mimarisi ise Şekil 7'de gösterilmiştir.



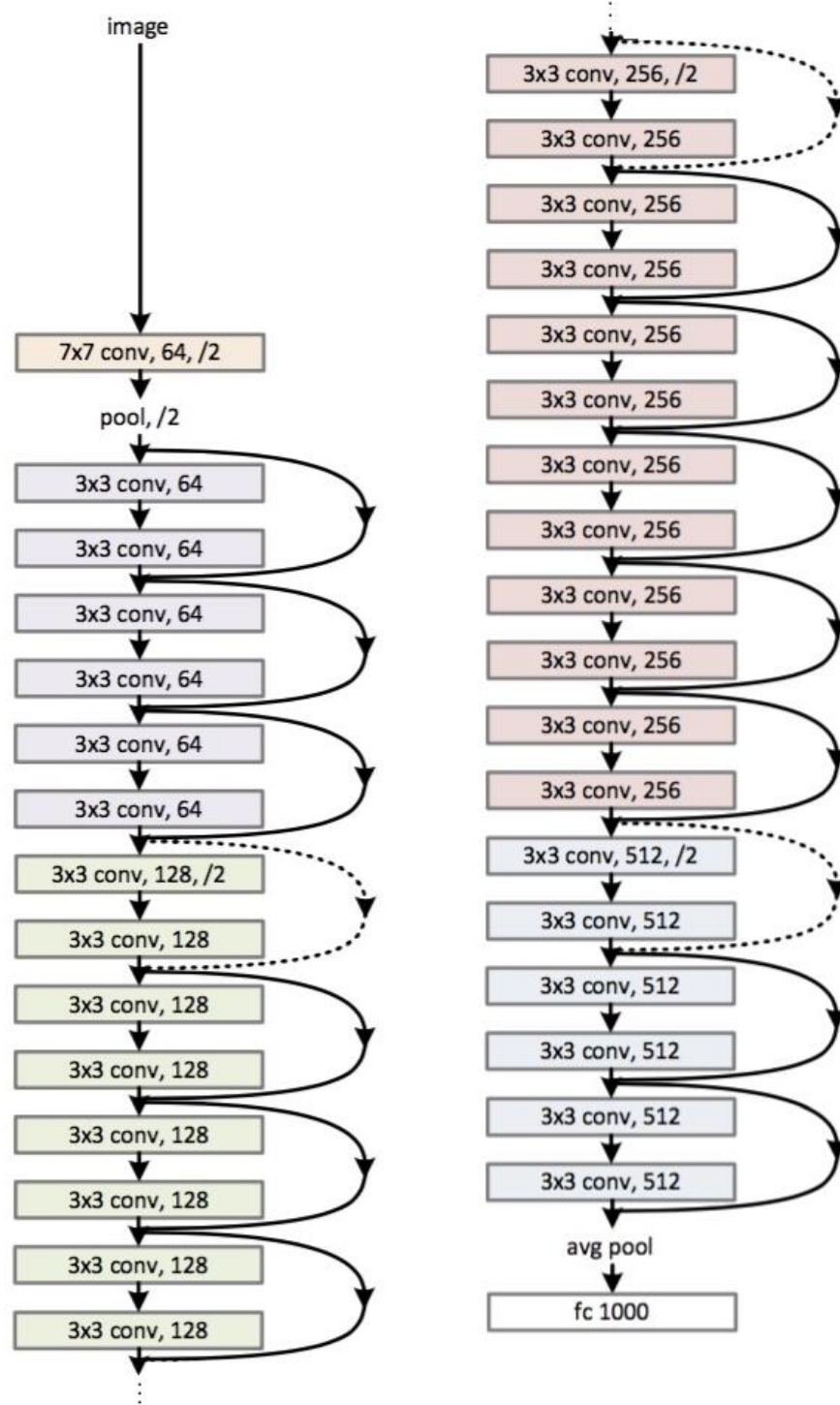
Şekil 6. Residual blok

(Kaynak: He vd., 2016)

SqueezeNet; son teknoloji bir CNN modelidir. Bu mimari 'fire' (yangın modülleri) denilen yapı taşlarından oluşur. Sıkılaştırılmış bir mimari tasarıma sahiptir. Diğer modellere kıyasla daha az parametre içerir. AlexNetten 50 kat daha az parametreye sahiptir. Daha az bant genişliği ve eğitim sırasında daha az iletişim sayesinde bir belleğe yüklenmesi en kolay mimari ağıdır. Geleneksel 3x3 evrişim filtreleri 1x1 olarak yer değiştirmiştir. 1x1'ler 3x3'e göre 9 kat daha az parametre içerir. Daha az sayıdaki parametrelerle en iyi şekilde yararlanır ve doğruluğu en üst düzeye çıkartır (Iandola vd., 2016).

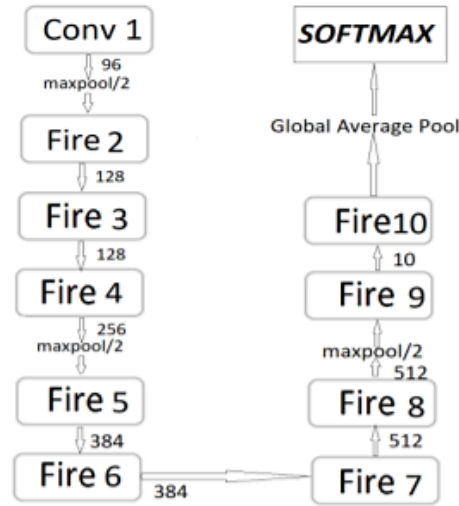
SqueezeNet ağ mimarı Şekil 8'de gösterilmiştir.

34-layer residual



Şekil 7. Microsoft ResNet ilk 34 katmanın ağ mimarisi

(Kaynak: He vd., 2016)



Şekil 8. SqueezeNet ağ mimarisi

(Kaynak: Iandola vd., 2016)

3.2.4. Nesne Tanıma

Görmenin, nesnelerin nasıl bir şey olduğunun belirlenmesine nesne tanıma olarak adlandırılır. Bir nesneyi tanıyabilme, sınıflandırabilme ve değerlendirebilme durumu anlamına gelir. Nesne tanıma, bilgisayarlı görü alanındaki belli başlı problemlerden birisi olup başka alanlarda farklı uygulamalar ile sıklıkla kullanılmakta ve sürekli gelişmektedir. Görüntünün belirlenmesi için görüntüyle birlikte şekil, renk ve benzeri belirleyicileri çıkartılabilecek yüksek düzeyde görüntü işleme gerçekleştirilir. Tanıma sistemleri biyomedikal alanında, gıda alanında, tür olarak ise; optik karakter tanıma, plaka tanıma, yüz tanıma, trafik gözetleme ve parmak izi tanıma gibi çeşitli alanlarda çok geniş bir yer kaplamaktadır (Nabiyev ve Vasıf, 2005) (Andreopoulos ve Tsotsos, 2013) (Kavuncu, 2018).

İnsanların nesnelere tanıyabilmesi, nesnelerin özelliklerini anlamaları ve bu özellikleri daha önceki tanıdık bilgileri ile benzetme ve karşılaştırmalarıyla ortaya çıkar. Nesne tanıma da aynı yöntem kullanılıyor diyebiliriz. Nesne tanıma probleminin çözümü için genellikle sorgu görüntü ile önceden çekilmiş görüntüler eşleştirilir. Öncelikle tanınması gereken nesnelere için data base oluşturulur. Data base sisteme öğretilir ve karşılaştırma işlemleyicisi ile nesne, daha önce öğrendikleriyle karşılaştırma yaptırılıp nesnenin tanınabilmesi sağlanır (Mikolajczyk vd., 2005) (Johan ve Prabuwno, 2011).

Nesne tanımadaki en önemli yer sistem yazılımıdır. Yazılım girdilerinin iyi olması gereklidir ki sistemin nesne tanıması daha iyi bir şekilde sonuç versin. Çevre girdisinde renklendirme ve kamera özellikleri yazılıma ulaştıkça görüntü kalitesini etkileyen en önemli faktördür (Johan ve Prabuwno, 2011) (Erdoğan ve Turgay, 2012).

3.2.5. Görüntü İşleme

Görüntü işleme, terimsel olarak resimsel bilgilerin harekete geçmesi ve çözümlemesidir (Castelman, R. K., 1996). Görüntü sayısallaştırılarak bir mekanizma ortamına yüklendikten sonra resimlerden istenen bilgilere ulaşmak için bazı önemli yöntemlerden geçmesi gerekmektedir (Yaman ve Aktürk, 2001). Resimler karışık olmaya meyillidirler. İnsanlar tarafından anlaşılması ne kadar kolayda olsa bilgisayar programına tanıtması oldukça zordur. Fakat insanların görme sistemini oluşturan ana düzeneğin anlaşılması görüntü işleme için ehemmiyetlidir (Kavuncu, 2018).

Görüntü işleme, verilerin bulunup değerlendirme ve ölçme yönteminden sonra, başka bir cihazda okunabilir bir şekilde dönüştürülmesine yönelik bir çalışma olan “sinyal işlemeden” farklı bir yöntemdir (Gonzalez ve Woods, 2007).

Resimlerin elektronik bir ortamda okunabilmeleri için resimlerin bilgisayarın işlem yapabileceği bir duruma dönüştürülebilmeleri gerekir. Bu işlemin adına sayısallaştırma denir (Dipova, 2018).

Görüntü işleme yöntemleri kullanılarak yapılan çalışmalarda, öncelikle kameradan görüntüler alınır. Alınan görüntüler üstünde, görüntü ön işleme basamakları kullanılmakta ve ilgilenilen görüntü ait özellik çıkartımı gerçekleştirilir. Ortamda bulunan görüntülerin doğru şekilde tespiti, özellik çıkarımı basamağı için oldukça önem teşkil eder. Nesnelere saptanması ya da tanınması hedefiyle yapılan çalışmalarda çeşitli yöntemler önerilir. Nesnelere ait belirleyici özellikler sayesinde etkili ve hızlı nesne tanımayla yönelik çalışmalar, karışık arka plan çıkarımı ile renk, kenar ve köşe tanıma, şekil, örüntü tanıma, şablon eşleme gibi farklı yöntemler kullanılır (Viola ve Jones, 2001) (Hussin vd., 2012) (Sonka vd., 2014).

Görüntü işleme alanında en yaygın işlemlerden biri görüntü sınıflandırma işlemidir. Bu işlemde; eğitim için 60000 el yazısı ve test için 10000 el yazısı rakamlardan oluşan MNIST kullanılır. Farklı türde çok sayıda yöntem bu data kümesi üzerinde test edilir. Şu ana kadarki en başarılı sonuçlardan biri %0.23 hata oranı ile Ciresan’ın çalışmasıdır.

CNN modeli kullanılan çalışmada genişlik normalizasyonu uygulanmıştır (Cireşan vd., 2012)

Nesne tanıma problemleri için de derin öğrenme algoritmaları büyük oranda başarılıdır. Girshick ve ekibi, görüntüdeki nesnelere tanımlamak için bölgeleri ve CNN modelini birlikte kullandıkları, R-CNN denilen algoritmayla çok çeşitli ve zengin özellikler elde etmişlerdir (Girshick vd., 2014).

Ultrason, elektro mikroskop ve bilgisayar içerikli görüntüler, belli bir hedefe yönelmiş görüntü kaynağından alınan farklı içeriklerden oluşur.

Görüntü işlemede görüntünün üstündeki bulanıklık ve kötü görüntü gibi görüntünün kalitesini bozan etkenleri yok etmek için ön hazırlık işlemi uygulanır. Bunun sebebi görüntülere düşük, orta ve yüksek seviye içeren işlemler uygulanır. Düşük seviyedeki işlemlerde görüntüye filtreleme uygulanır. Orta düzey seviyedeki işlemlerde bölme ve tanıma işlemleri uygulanır böylece görüntülerde tanıma ve sınıflandırma gerçekleştirilmiş olur. Yüksek seviye işlemler görüntülerdeki nesnelere tanımda görüntülerin analiz edilmesini içerir. Görüntülerin bilgisayar ortamında işlemlerden geçmesiyle de, görüntülerdeki nesnelere görüntü içeriği detaylandırılır. Bu detaylandırma aşaması ile birlikte görüntü işleme gerçekleştirilmiş olur (Jähne, 2005).

Derin öğrenme yöntemleri, literatürde yer alan nesne tanıma ve görüntü sınıflandırma uygulamalarında ön adım olarak yapılması gereken özellik çıkarımı aşamasını model içindeki evrimsel katmanlarda yaptığı için önemli bir avantaj sağlamaktadır.

3.3.Aktarım Öğrenmesi (Transfer Learning)

Günlük yaşantımızda bir olay karşısında öğrendiğimiz bir bilgiyi sıkça başka bir durumda kullanırız ve bilgiyi aktarırız. Eğer karşılaşılan durumlar benzerse öğrenilen bilgi kolaylıkla aktarılır (Nova Research Lab., 2020).

Derin öğrenme yöntemleri görüntü işleme, doğal dil işleme, sinyal işleme gibi farklı alanlarda kayda değer sonuçlar elde etmektedir, ama derin öğrenme algoritmalarındaki başarıyı örnek sayısı büyük oranda değiştirmektedir. Bu yüzden fazla miktarda veriye sahip olmak veya üretmek bir problem olarak kabul edilmektedir. Bu sorunla mücadele edebilmek için aktarım öğrenmesi yöntemi geliştirilmiştir. Aktarım öğrenmesinde, her bir ağı sıfırdan eğitmek yerine, önceden eğitilmiş modeller kullanılır. Öğrenme aktarımı

ile önceki bilgiler kullanılarak daha az eğitim verisi ile daha yüksek başarı elde eden ve daha hızlı öğrenen modeller ortaya çıkartılır (Şeker, 2018).

Öğrenme aktarımı yapılırken uygulanan ilk aşama neyin aktarılacağıdır. Kaynak veriden hedefe hangi bilgi aktarılınca hedef görevin başarısının artacağına cevap aranan bölümdür. Bilginin ne zaman aktarılacağı aşamasında, verinin hangi durumlarda aktarılacağı saptanır. Bazen kaynak veri ve hedef veri birbirinden çok farklı olabilir, bu gibi durumlarda aktarım başarılı olmayabilir ya da eğitimin başarısını düşürebilir. Bu duruma ‘olumsuz aktarım/negatif transfer’ denir. Nasıl aktarılacağı aşamasında ise hangi yöntemin kullanılacağı esastır. Bir model olduğu gibi kullanılabilir ya da üzerinde değişiklikler yapılabilir (Sarkar, 2018).

Öğrenme aktarımında kaynak görevi kendimiz geliştirebiliriz veya hazır bir model kullanabiliriz. Uygulamada daha çok tercih edilen yöntem hazır bir model kullanılmasıdır. Örneğin ImageNet yarışmasında kullanılan modellerin eğitilmesi günler hatta haftalar alabilir ama eğitilmiş iyi bir modelin öğrenme aktarımı ile bir sonraki sene de kullanılması oldukça yaygın ve mantıklıdır.

CNN gibi iyi çalışan bir derin öğrenme yönteminde aktarım öğrenmesi kullanma sebeplerimizden biri eğitim süresinin daha hızlı olmasıdır. Önceden eğitilmiş model üzerinde ince ayar yapılarak yeni model çok hızlı bir şekilde eğitebilir. İkincisi daha az veri kümesine ihtiyaç duyulmasıdır, az veri ile yüksek performans sağlanabilir. Son olarak eski modele katman eklenmesi gibi basit ve kolay bir işlemle sonra bile çok daha yüksek başarı sağlanabilir.

Öğrenim aktarımı yöntemleri aşağıdaki gibidir:

Tümevarım öğrenim aktarımı; bu senaryoda, kaynak ve hedef etki alanları aynıdır, ancak kaynak ve hedef görevleri birbirinden farklıdır. Algoritmalar, hedef görevi iyileştirmeye yardımcı olur. Kaynak etki alanının etiketli veri içerip içermediğine bağlı olarak ikiye ayrılır. Birincisi kaynak görevde fazla miktarda etiketli verinin bulunmasıdır. Bu durumda, öğrenme aktarımı çok yönlü öğrenme (multi-task learning) gibi çalışır. Ancak çok yönlü öğrenmenin eğitildiği tüm görevlerde başarı göstermesi beklenirken öğrenme aktarımında amaç hedef görev başarısının yüksek olmasıdır. İkincisi ise kaynak verinin etiketli veri içermemesi durumudur. Bu gibi durumlarda tümevarımsal öğrenme aktarımı bu defa kendi kendine öğrenme (self taught learning)

gibi çalışır. Kendi kendine öğrenmede kaynak ve hedef verilerin etiket uzayı farklıdır, bu da kaynak verisinin doğrudan kullanılamayacağı anlamına gelir.

Dönüştürücü öğrenme aktarımı; bu yöntemde, kaynak ve hedef görevler arasında benzerlikler vardır, ancak ilgili etki alanları farklıdır. Bu ayarda, kaynak etki alanında çok sayıda etiketlenmiş veri varken, hedef etki alanında hiç veri yoktur. Hedef ve kaynak veri kümelerine göre dönüştürücü öğrenme aktarımı iki farklı şekilde ifade edilebilir. Bunlar özellik uzayının kaynak ve hedef veri kümeleri için farklı olması ve aynı olması durumlarıdır.

Gözetimsiz öğrenme aktarımı; bu ayar, hedef etki alanındaki denetimsiz görevlere odaklanarak, tümevarım aktarımına benzer. Kaynak ve hedef etki alanları benzerdir, ancak görevler farklıdır. Bu senaryoda, etiketli veriler etki alanlarının hiçbirinde kullanılamaz.

Bu üç transfer kategorisi, transfer öğrenmenin uygulanabileceği ve detaylı olarak çalışılabileceği farklı ayarların ana hatlarını verir. Bu kategoriler arasında neyin aktarılacağı sorusuna cevap vermek için aşağıdaki yaklaşımlardan bazıları uygulanabilir.

Örnek aktarımı; kaynak ve hedef verisinin ortak birçok özelliğe sahip olması durumunda kullanılır. Kaynak verideki bazı örneklerin hedef verisinde kullanılmasıdır. Kaynak verisindeki belirli örneklerle uygun ağırlık değerleri verilmesi ve hedef veri üzerinde kullanılmasıyla başarı artırılabilir.

Özellik temsili aktarımı; Bu yaklaşım, kaynaktan hedef alanlara kullanılacak iyi özellik temsillerini tanımlayarak alan ayrışmasını en aza indirmeyi ve hata oranlarını azaltmayı amaçlamaktadır. Kaynak ve hedef verilerin ortak özelliklerinin olması beklenir.

Parametre aktarımı; Bu yaklaşım, ilgili görevler için modellerin bazı parametreleri paylaştığı veya önceden hiperparametrelerin dağıtımını paylaştığı varsayımı üzerinde çalışır. Hem kaynak hem de hedef görevlerin aynı anda öğrenildiği çoklu görev öğreniminin aksine, aktarım öğrenimi için, genel performansı iyileştirmek için hedef alanın kaybına ek ağırlık uygulayabiliriz (Nova Research Lab, 2020).

İlişki kurma tecrübesinin aktarımı; Kaynak verideki bazı ilişkilerin hedefteki ilişkilere benzetilmesi durumudur. Kaynak veri ve hedef veri arasında bazı ilişkilerin olduğu varsayılır, veri kaynakları arasında ilişkisel bilginin eşleştirilmesi yapılır. Başka bir deyişle, her veri noktasının diğer veri noktaları ile bir ilişkisi olduğu veriler; örneğin, sosyal ağ verileri ilişkisel bilgi aktarım tekniklerini kullanır (Sarkar, 2018).

Derin öğrenme, farklı CNN mimarileri kullanılarak, nesne tanıma ve tanımlama gibi çeşitli bilgisayarla görme görevleri için oldukça başarılı bir şekilde kullanılmıştır.

Bu kısım kadar anlatılan modelleme, yöntem ve yaklaşımlar kullanılarak yapılan çalışma anlatılacak ve derin aktarımla öğrenim modelleri oluşturulacaktır.

4.DENEYSEL SONUÇLAR

4.1. Veri Kümesinin Oluşturulması

Çalışma için gerekli olan veriler fotoğraf, video paylaşım ve barındırma hizmeti sunan bir web sitesi olan 'flickr'dan sağlandı. Bu tezde elimizde 14 farklı tarihi yerin her birinden 19-29 arası fotoğraf vardı fakat bu görüntüler yeterli sayıda olmadığı için veri çoğaltma yolu seçildi.

Veri çoğaltma, küçük veri setlerinde başarı oranını yükseltmek için verilerin çeşitli bozulma etkilerine maruz kalmasıyla sayısının çoğaltılmasıdır. Veri çoğaltma, ön hazırlık sürecinde mevcut eğitim verisi örneklerinin görüntüler için döndürme, öteleme, kırpma, soldurma, ölçeklendirme gibi farklı teknikler ile meydana gelen kopyalarını oluşturuyor ve bu kopyaları mevcut eğitim verisine ilave ederek çok daha geniş bir eğitim verisi setini kullanıma hazır hale getiriyor. Görüntüleri tanımaya çalışan başarılı bir yapay zeka modeli, görüntüler ne kadar düzensiz (sağa veya sola yatmış, ters, ötelenmiş) olsa bile tanıyabilmelidir. Veri artırımı uygulandıktan sonra yeni eğitim veri setinde benzer örneklerin sayısı artmış olur ve sonuç olarak veri artırımı, modelin ezberleme (overfit) sorununun da üstesinden gelmiş olur.

Bu çalışmada da eldeki hazır verileri işleyerek yeni veriler üretip daha başarılı sonuçlar elde etmek adına veri çoğaltma işlemi gerçekleştirildi. Bu işlem keras kütüphanesi ile gerçekleştirildi. Keras, Python programlama dili üzerine yazılmış yüksek seviyeli bir derin öğrenme kütüphanesidir. TensorFlow kütüphanesi üzerinde de çalışabiliyor olması en önemli özelliğidir (Python, 2020).

Her bir kültürel miras için elimizdeki orijinal veri sayıları Tablo 1'de verilmiştir. Veri çoğaltma işlemi ve boyutlandırma sonrasında elde edilen yeni veri kümeleri Tablo 2 ve Tablo 3'te verilmiştir. Bu veriler derin öğrenme yöntemlerinin eğitilmesi için kullanılmıştır. Test aşamasında kullanılacak veri kümesindeki görüntüler eğitim veri kümesinden tamamen farklı görüntülerdir. Test aşamasında kullanılan görüntü sayıları Tablo 4'te verilmiştir.

Tez çalışmasında kullanılan derin öğrenme yöntemlerinin giriş katmanının beklediği resimlerin boyutları 227x227 ve 224x224 olduğu için, elde ettiğimiz resimler bu boyutlara getirildi. Bu işlemle birlikte veri çoğaltma yöntemi de kullanılarak resimler çoğaltıldı. Bunun için Python dilinde geliştirdiğimiz bir betik kullanıldı. Sonuçta elde edilen resim sayıları Tablo 2 ve Tablo 3'te verilmiştir.

Tablo 1. Elimizde mevcut olan ilk verilerin sayısı

Kültürel Miras	Görüntü Sayısı
Dolmabahçe	25
Efes	28
Eşferoğlu	22
Hagia eleni	25
İnce Minare	25
Karatay Medresesi	29
Mevlana Türbesi	22
Nemrut Dağı	24
Pamukkale	19
Peri Bacaları	23
Sahibata Türbesi	24
Sırçalı Medrese	24
Süleymaniye Camii	24
Sumela Manastırı	27

Tablo 2. 227*227 boyutunda artırılmış verilerin sayısı

Kültürel Miras	Görüntü Sayısı
Dolmabahçe	270
Efes	303
Eşferoğlu	239
Hagia eleni	271
İnce Minare	240
Karatay Medresesi	312
Mevlana Türbesi	240
Nemrut Dağı	261
Pamukkale	206
Peri Bacaları	250
Sahibata Türbesi	260
Sırçalı Medrese	252
Süleymaniye Camii	261
Sumela Manastırı	291

Tablo 3. 224*224 boyutundaki artırılmış verilerin sayısı

Kültürel Miras	Görüntü Sayısı
Dolmabahçe	534
Efes	600
Eşferoğlu	472
Hagia eleni	539
İnce Minare	475
Karatay Medresesi	622
Mevlana Türbesi	476
Nemrut Dağı	516
Pamukkale	412
Peri Bacaları	493
Sahibata Türbesi	509
Sırçalı Medrese	736
Süleymaniye Camii	771
Sumela Manastırı	1119

Tablo 4. Test için kullanılan yeni görüntülerin sayısı

Kültürel Miras	Görüntü Sayısı
Dolmabahçe	5
Efes	4
Eşferoğlu	4
Hagia eleni	4
İnce Minare	3
Karatay Medresesi	3
Mevlana Türbesi	3
Nemrut Dağı	3
Pamukkale	3
Peri Bacaları	3
Sahibata Türbesi	3
Sırçalı Medrese	4
Süleymaniye Camii	4
Sumela Manastırı	3

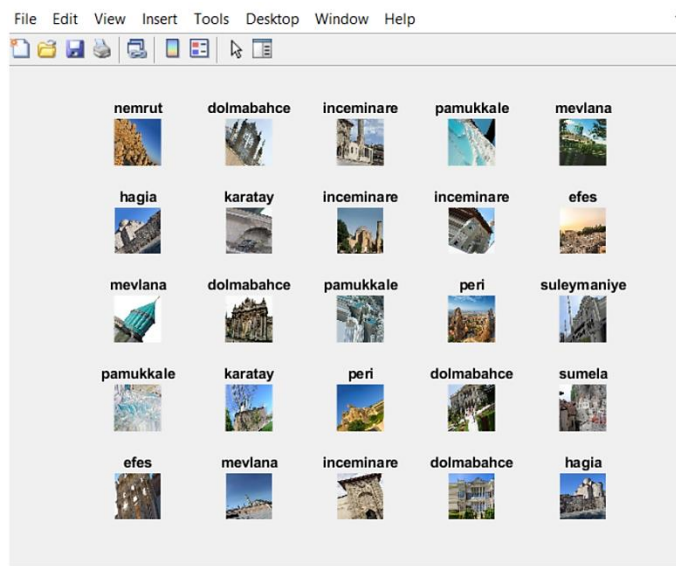
Eđitim ve test iin kullanılacak olan resimler sabit disk zerinde bir klasrde her bir kltrel mirasa ait grntlerin ayrı bir alt klasrde saklandıđı dizin yapısı iinde bulunmaktadır.

Verilerin artırılması ařamasından sonra eđitilmesi ařamasına geilmiřtir. Bunun iin Matlab ortamında derin đrenme modellerinden AlexNet, GoogleNet, ResNet ve SqueezeNet kullanılmıřtır (MATLAB, 2018).

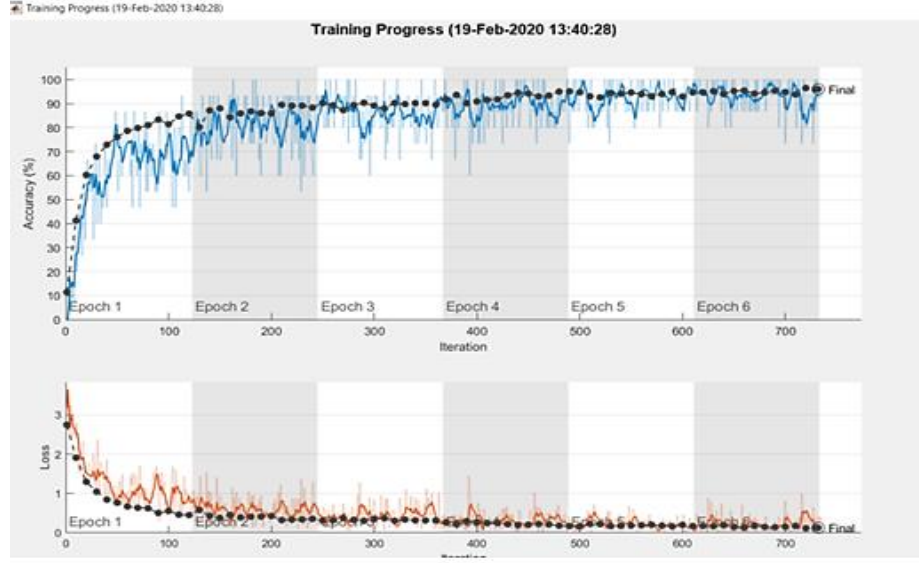
4.2. Matlab Ortamında CNN Modelleri ile Eđitim Yapılması

4.2.1. AlexNet

Alexnet giriř katmanı 227x227 boyutunda resimler beklediđi iin 227x227 boyutunda grntler kullanılmıřtır. Yazılan Matlab kod blođunda split_train_test kısımdaki deđerler 0,5-0,6-0,7-0,8-0,9 olarak deđiřtirilip her biri iin ayrı ayrı alıřtırılmıřtır. rnek olarak; split_train_test iin 0,5 kullanılması durumunda, Őekil 9 ve Őekil 10'da sırasıyla eđitim ařaması sonrası bazı rnek resimler iin elde edilen tahmin sonuları ve eđitim ařamasına ait dođruluk ve kayıp grafikleri verilmiřtir.



Őekil 9. AlexNet'in split_train_test 0,5 iin eđitimi sonunda test verilerinden bazıları iin tahmin sonuları



Şekil 10. AlexNet'in split_train_test 0,5 için eğitim aşamasına ait doğruluk ve kayıp grafikleri

Daha sonra AlexNet 0,6-0,7-0,8-0,9 değerleri için program tekrar çalıştırılıp değerler kaydedilmiştir. Bu çalıştırmalara ait sonuçlar Tablo 5'te verilmiştir.

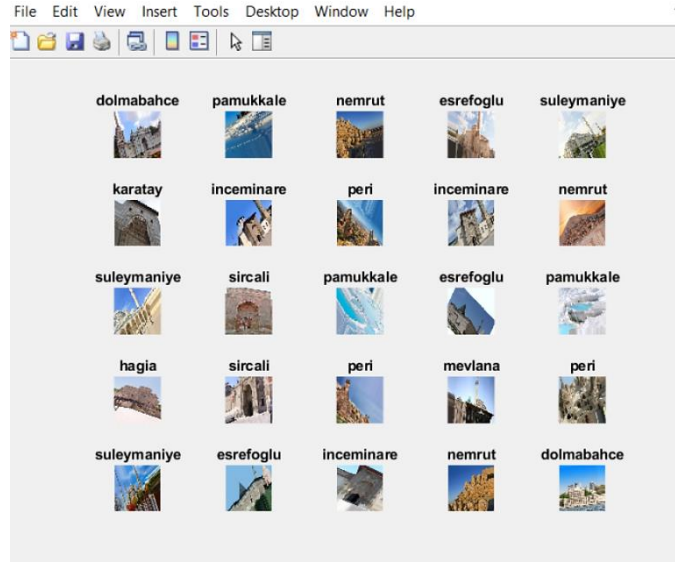
Tablo 5'e tezde kullanılan bütün CNN modelleri belirtilen her bir split_train_test değeri için doğruluk değerleri, eğitim süreleri, epoch sayısı, iterasyon sayısı ve her epochtaki iterasyon sayısı eklenmiş, ayrıntılı olarak karşılaştırmaları yapılmıştır.

4.2.2. GoogleNet

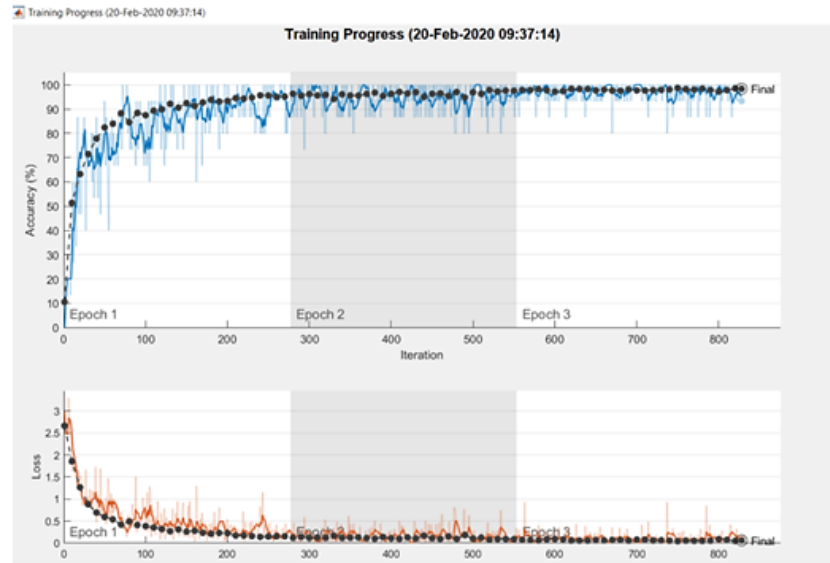
GoogleNet giriş katmanı 224x224 boyutunda resimler beklediği için 224x224 boyutuna getirilen görüntüler kullanılmıştır. Yazılan Matlab kodunda split_train_test kısmındaki değerler 0,5-0,6-0,7-0,8-0,9 olarak değiştirilip her biri için ayrı ayrı çalıştırılmıştır. Örnek olarak split_train_test için 0,5 değeri kullanılması durumunda, Şekil 11 ve Şekil 12'de sırasıyla eğitim aşaması sonrası bazı örnek resimler için elde edilen tahmin sonuçları ve eğitim aşamasına ait doğruluk ve kayıp grafikleri verilmiştir.

Daha sonra GoogleNet 0,6-0,7-0,8-0,9 değerleri için program tekrar çalıştırılmış, her bir değer için bazı örnek resimler için elde edilen tahmin sonuçları ve eğitim aşamasına ait doğruluk ve kayıp grafikleri bulunup kaydedilmiştir.

Bu çalıştırmalara ait her bir değer için elde edilen sonuçlar Tablo 5'te verilmiştir.



Şekil 11. GoogleNet'in split_train_test 0,5 için eğitimi sonunda test verilerinden bazıları için tahmin sonuçları

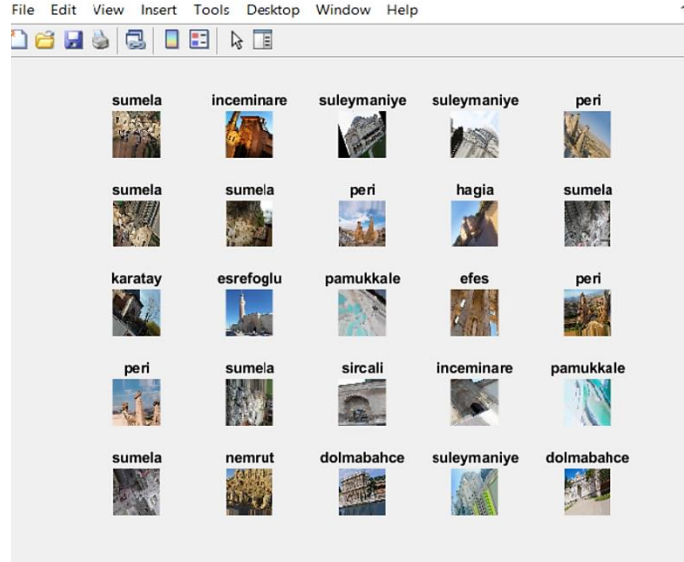


Şekil 12. GoogleNet'in split_train_test 0,5 için eğitim aşamasına ait doğruluk ve kayıp grafikleri

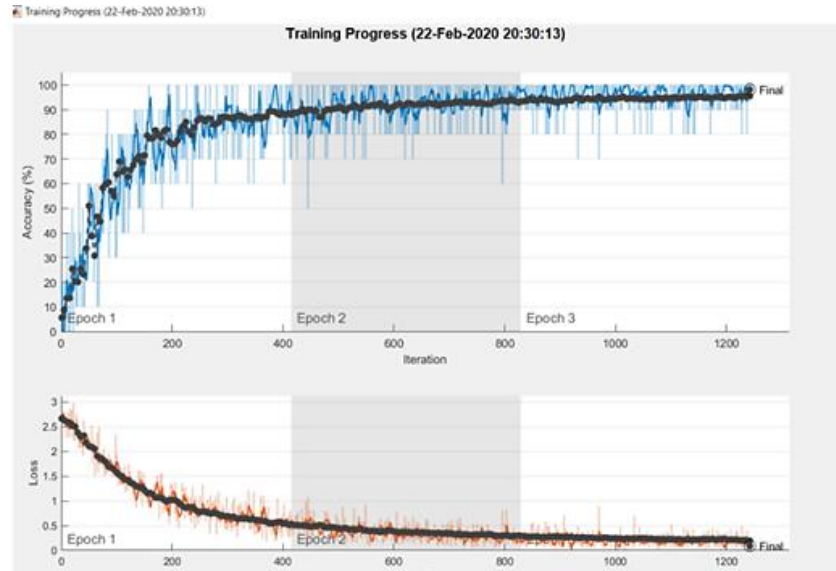
4.2.3.ResNet

ResNet için 'resize_224' dosyası içindeki görüntüler kullanılmıştır. Yazılan kodda split_train_test kısmındaki değerler 0,5-0,6-0,7-0,8-0,9 olarak değiştirilip her biri için ayrı ayrı çalıştırılmıştır. Örnek olarak; split_train_test için 0,5 kullanılması durumunda, Şekil 16 ve Şekil 17'de sırasıyla eğitim aşaması sonrası bazı örnek resimler için elde

edilen tahmin sonuçları ve eğitim aşamasına ait doğruluk ve kayıp grafikleri verilmiştir. Eğitim sonrası alınan sonuçlar aşağıda gösterilen Şekil 13 ve Şekil 14'deki gibidir. Program ResNet 0,6-0,7-0,8-0,9 değerleri için tekrar çalıştırılıp değerler karşılaştırılmıştır. Bu çalıştırmalara ait sonuçlar Tablo 5'de verilmiştir.



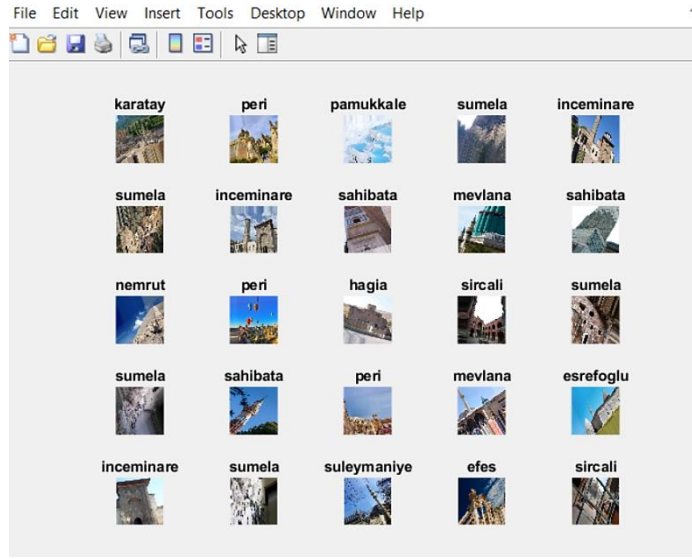
Şekil 13. ResNet'in split_train_test 0,5 için verilerin eğitimi sonunda test verilerinden bazıları için tahmin sonuçları



Şekil 14. ResNet'in split_train_test 0,5 için eğitim aşamasına ait doğruluk ve kayıp grafikleri

4.2.4.SqueezeNet

SqueezeNet için 'resize_227' dosyası içindeki görüntüler kullanılmıştır. Yazılan kodda split_train_test kısmındaki değerler 0,5-0,6-0,7-0,8-0,9 olarak değiştirilip her biri için ayrı ayrı çalıştırılmıştır. Örnek olarak; split_train_test için 0.5 kullanılması durumunda, Şekil 18 ve Şekil 19'da sırasıyla eğitim aşaması sonrası bazı örnek resimler için elde edilen tahmin sonuçları ve eğitim aşamasına ait doğruluk ve kayıp grafikleri verilmiştir. Eğitim sonrası alınan sonuçlar Şekil 15 ve 16'daki gibidir:



Şekil 15. SqueezeNet'in split_train_test 0,5 için verilerin eğitimi sonunda test verilerinden bazıları için tahmin sonuçları



Şekil 16. SqueezeNet split_train_test 0,5 için eğitim aşamasına ait doğruluk ve kayıp grafikleri

Değerler 0,5-0,6-0,7-0,8-0,9 olarak değiştirilip program tekrar çalıştırılıp değişen değerler gözlemlenmiştir. Bu çalıştırmalara ait sonuçlar Tablo 5’te verilmiştir.

Uygulama kısmındaki verilerin yukarıda gösterilen yöntemlerle eğitilmesi işlemi yapılmıştır. Özetle yöntemlerin karşılaştırılması Tablo 5’te gösterilmiştir.

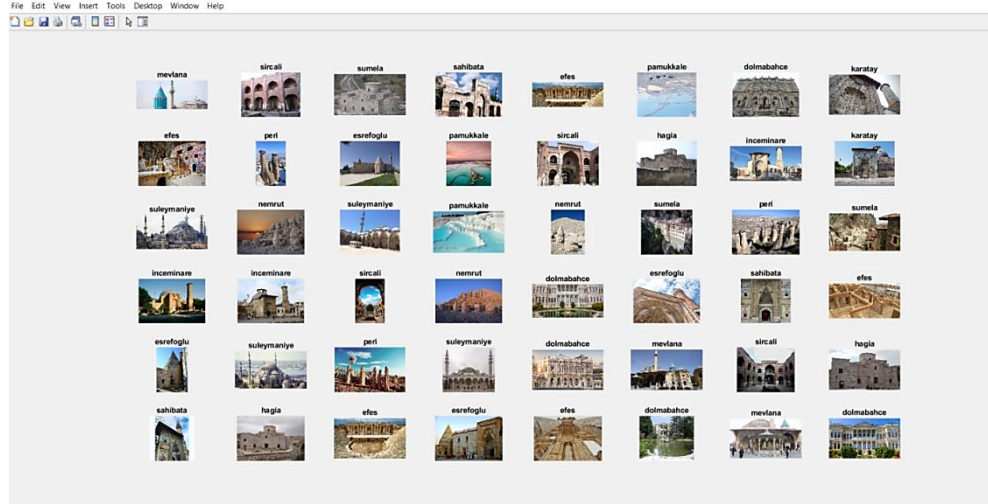
Tablo 5. Kullanılan yöntemlerin karşılaştırılması

	Validation Accuracy Sonuçları	Eğitim Süresi	Epoch Sayısı	İterasyon Sayısı	Her Epochtaki İterasyon
AlexNet0,5	% 96.05	105dk 57sn	6	732	122
AlexNet0,6	% 93.98	109dk 30sn	6	876	146
AlexNet0,7	% 96.26	102dk	6	1020	170
AlexNet0,8	% 95.89	94dk 28sn	6	1170	195
AlexNet0,9	% 93.15	79dk 24sn	6	1314	219
GoogleNet0,5	% 98.38	782dk 6sn	3	828	276
GoogleNet0,6	% 98.49	734dk	3	990	330
GoogleNet0,7	% 98.79	595dk 46sn	3	1158	386
GoogleNet0,8	% 98.73	540dk 13sn	3	1323	441
GoogleNet0,9	% 99.40	384dk 29sn	3	1488	496
ResNet0,5	% 98.11	1801dk 31sn	3	1242	414
ResNet0,6	% 97.01	512dk 13sn	3	744	248
ResNet0,7	% 98.83	1427dk 5sn	3	1737	579
ResNet0,8	% 97.64	352dk	3	993	331
ResNet0,9	% 98.67	259dk 21sn	3	1116	372
SqueezeNet0,5	% 97.63	341dk 56sn	3	828	276
SqueezeNet0,6	% 96.83	327dk 50sn	3	990	330
SqueezeNet0,7	% 97.46	299dk 32sn	3	1158	356
SqueezeNet0,8	% 97.58	247dk 57sn	3	1323	441
SqueezeNet0,9	% 98.31	174dk 45sn	3	1488	496

Tablo 5’te görüldüğü üzere en yüksek başarı sağlayan yöntem %99.40’lık bir başarı ile GoogleNet0,9 olup, başarısı en düşük yöntem ise AlexNet0,9 olmuştur başarı oranı ise %93.15’tir.

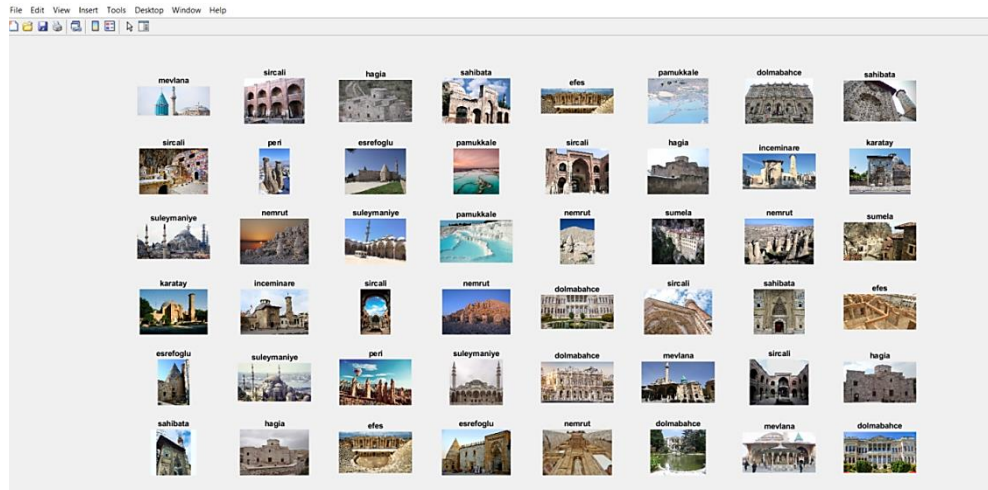
4.3. Deepest Sonuçları ve Karşılaştırması

Sonuç bölümünde, yukarıdaki kullanılan yöntemler çalıştırılarak ‘myalexnet.mat’, ‘mygooglenet.mat’, ‘myresnet.mat’ ve ‘mysqueezenet.mat’ adında yeni modeller oluşturulmuştur. Bu modeller ‘deep_test.m’ matlab dosyası içinde çalıştırılmıştır. Tablo 4’te de görülen eğitime sokulmamış yeni görüntüler kullanılmıştır. Her yöntem ve değer için görüntülerin ‘deeptest’ tahmin sonuçları aşağıdaki gibidir:



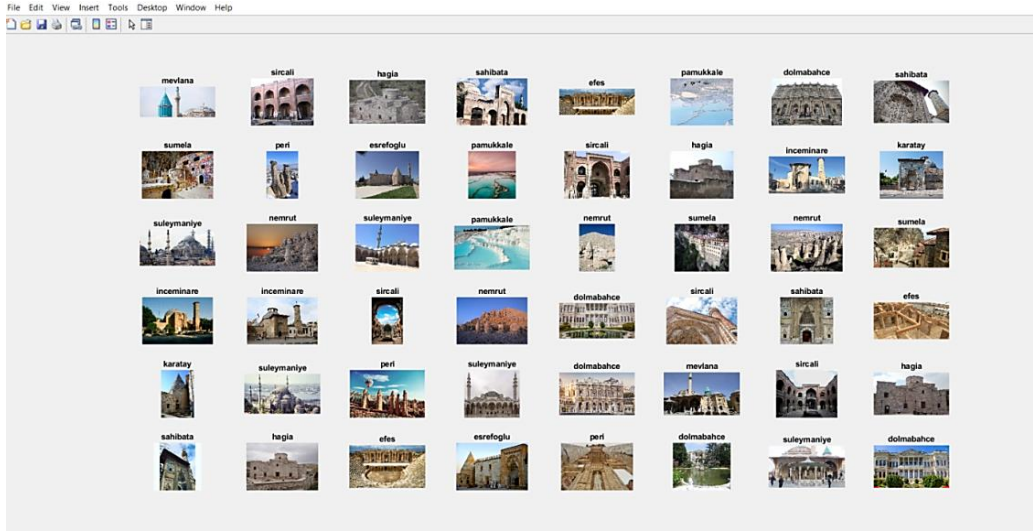
Şekil 17. Deeptest sonrası AlexNet'in split_train_test 0,5 için tahmin sonuçları

Alexnet ‘split_train_test’ değeri 0,5 iken doğruluk değeri 0.8750’dir. 0,6-0,7-0,8 ve 0,9 değerleri için doğruluk değerleri Tablo 6’da gösterilmiştir.



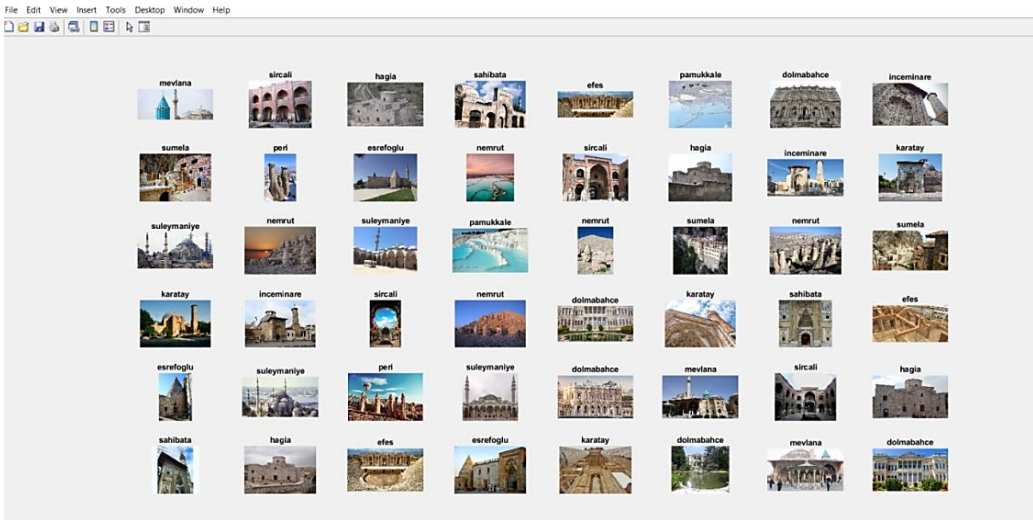
Şekil 18. Deeptest sonrası GoogleNet'in split_train_test 0,5 için tahmin sonuçları

GoogLeNet ‘split_train_test’ değeri 0,5 iken doğruluk değeri 0.8958’dir. 0,6-0,7-0,8 ve 0,9 değerleri için doğruluk değerleri Tablo 6’da gösterilmiştir.



Şekil 19. Deeptest sonrası ResNet’in split_train_test 0,5 için tahmin sonuçları

ResNet ‘split_train_test’ değeri 0,5 iken doğruluk değeri 0.8125’dir. 0,6-0,7-0,8 ve 0,9 değerleri için doğruluk değerleri Tablo 6’dadır.



Şekil 20. Deeptest sonrası SqueezeNet'in split_train_test 0,5 için tahmin sonuçları

Squeezenet ‘split_train_test’ değeri 0,5 iken doğruluk değeri 0.8542’dir. 0,6-0,7-0,8 ve 0,9 değerleri için doğruluk değerleri Tablo 6’da gösterilmiştir.

Tablo 6. Deeptest sonuçları için accuracy değerleri

Kullanılan CNN Mimarisi	'split_train_test' değeri	Accuracy (doğruluk) değeri
Alexnet	0.5	0.8750
Alexnet	0.6	0.8958
Alexnet	0.7	0.8125
Alexnet	0.8	0.8542
Alexnet	0.9	0.8542
Googlenet	0.5	0.8542
Googlenet	0.6	0.7917
Googlenet	0.7	0.8958
Googlenet	0.8	0.8125
Googlenet	0.9	0.8750
Resnet	0.5	0.8125
Resnet	0.6	0.8125
Resnet	0.7	0.7917
Resnet	0.8	0.8125
Resnet	0.9	0.8542
Squeezenet	0.5	0.8542
Squeezenet	0.6	0.7917
Squeezenet	0.7	0.8125
Squeezenet	0.8	0.8333
Squeezenet	0.9	0.8125

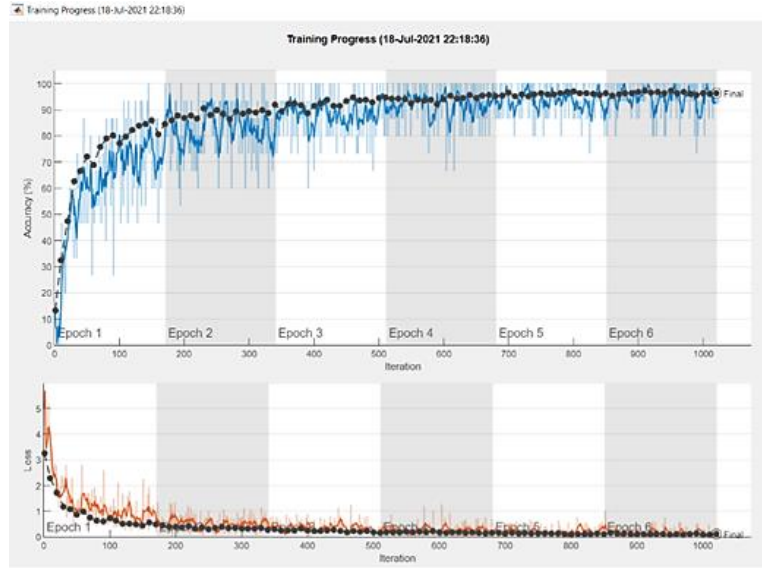
Tablo 6’da verilen sonuçlar Intel Core İ7 2.70 GHz işlemci ve 12 GB belleğe sahip bir dizüstü bilgisayarda elde edilmiştir. Modeller ayrıca, Intel Core i7 2.60GHz işlemcili, 16 GB bellek ve 1920 Cuda çekirdekli 8 GB paylaşımlı sistem belleğine ve 6 GB ayrılmış video belleğe sahip NVIDIA GeForce RTX 2060 grafik kartlı bir dizüstü bilgisayarda da çalıştırılmıştır. GPU içeren bu ikinci bilgisayarda, Tablo 6’da her derin öğrenme yöntemi için en iyi eğitim doğruluğunu veren train_test_split değeri kullanılarak, farklı başlangıç öğrenme oranları kullanılarak eğitim gerçekleştirilmiştir. Her bir yöntem için elde edilen sonuçlar Tablo 7, Tablo 8, Tablo 9 ve Tablo 10’da verilmiştir.

AlexNet en iyi doğruluk değeri sonucunu 0.6 train_test_split için vermiştir. Tablo 7’de, 0.0001 ila 0.0009 arasındaki öğrenme oranları için “sgdm” optimizasyon yöntemi kullanılarak elde edilen eğitim doğruluğu, test doğruluğu ve süreler verilmiştir. Eğitim için diğer parametreler aynen kullanılmıştır. Öğrenme oranları eğitim boyunca sabit tutulmuştur.

Tablo 7. AlexNet için elde edilen sonuçlar

Başlangıç Öğrenme Oranı	Eğitim doğruluğu	Test doğruluğu	Süre
0,0001	0.9624	0.8542	4 dk 43 sn
0.0002	0.9521	0.7917	4 dk 44 sn
0.0003	0.9665	0.7917	4 dk 54 sn
0.0004	0.9227	0.7083	4 dk 57 sn
0.0005	0.9138	0.7083	4 dk 45 sn
0.0006	0.9049	0.7500	4 dk 44 sn
0.0007	0.8891	0.6667	4 dk 49 sn
0.0008	0.8042	0.6458	5 dk 28 sn
0.0009	0.4196	0.3542	4 dk 47 sn

Tablo 7'ye göre en yüksek test doğruluğunu veren öğrenme oranı kullanılarak “adam” optimizasyon yöntemi ile AlexNet yeniden eğitilmiştir. Bu eğitim 4 dk 56 sn sürmüştür ve 0.7858 eğitim doğruluğu, 0.7500 test doğruluğu elde edilmiştir.



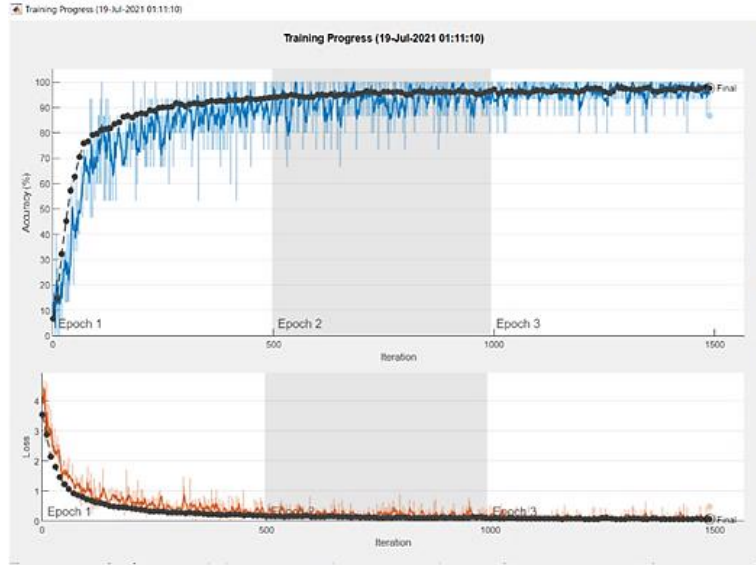
Şekil 21. 'adam' optimizasyon yöntemi ile AlexNet eğitim aşamasına ait doğruluk ve kayıp grafikleri

Tablo 6'ya bakıldığında GoogleNet en başarılı doğruluk değeri sonucunu 0.7 train_test_split için vermiştir. Tablo 8'de, 0.0001 ila 0.0009 arasındaki öğrenme oranları için “sgdm” optimizasyon yöntemi kullanılarak elde edilen eğitim doğruluğu, test doğruluğu ve süreler verilmiştir. Eğitim için diğer parametreler aynen kullanılmıştır. Öğrenme oranları eğitim boyunca sabit tutulmuştur.

Tablo 8. GoogleNet için elde edilen sonuçlar

Başlangıç Öğrenme Oranı	Eğitim doğruluğu	Test doğruluğu	Süre
0,0001	0.9758	0.8542	8 dk 28 sn
0.0002	0.9903	0.8542	9 dk 9 sn
0.0003	0.9927	0.8750	9 dk 13 sn
0.0004	0.9952	0.8542	9 dk 11 sn
0.0005	0.9903	0.8542	9 dk 8 sn
0.0006	0.9940	0.8750	9 dk 5 sn
0.0007	0.9940	0.8542	9 dk 4 sn
0.0008	0.9746	0.8542	9 dk 7 sn
0.0009	0.9976	0.7917	9 dk 6 sn

Tablo 8'e göre en yüksek test doğruluğunu veren öğrenme oranı kullanılarak "adam" optimizasyon yöntemi ile GoogleNet yeniden eğitilmiştir. Bu eğitim 10 dk 39 sn sürmüş ve 0.9021 eğitim doğruluğu, 0.7292 test doğruluğu elde edilmiştir.



Şekil 22. 'adam' optimizasyon yöntemi ile GoogleNet eğitim aşamasına ait doğruluk ve kayıp grafikleri

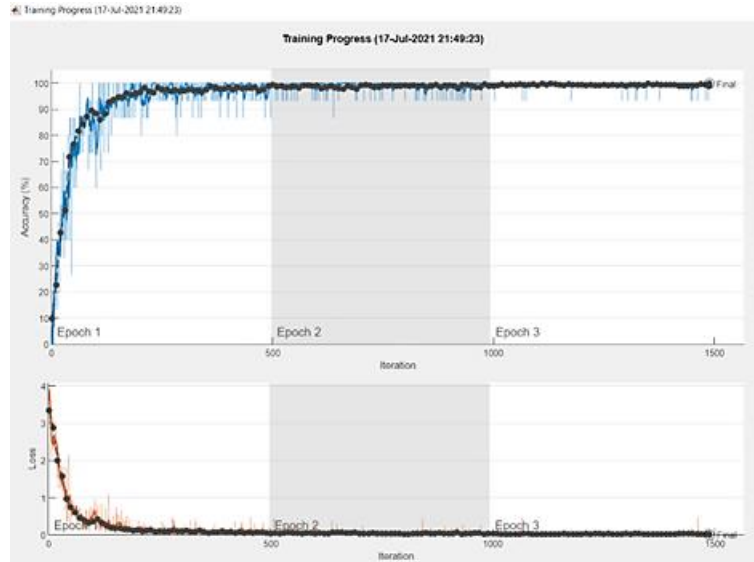
Tablo 6'ya bakıldığında ResNet en yüksek doğruluk değeri sonucunu 0.9 train_test_split için vermiştir. Tablo 9'da, 0.0001 ila 0.0009 arasındaki öğrenme oranları için "sgdm" optimizasyon yöntemi kullanılarak elde edilen eğitim doğruluğu,

test doğruluğu ve süreler verilmiştir. Eğitim için diğer parametreler aynen kullanılmıştır. Öğrenme oranları eğitim boyunca sabit tutulmuştur.

Tablo 9. ResNet için elde edilen sonuçlar

Başlangıç Öğrenme Oranı	Eğitim doğruluğu	Test doğruluğu	Süre
0,0001	0.9927	0.8333	6 dk 36 sn
0.0002	0.9915	0.8333	6 dk 58 sn
0.0003	0.9927	0.8333	6 dk 45 sn
0.0004	0.9940	0.8542	6 dk 26 sn
0.0005	0.9903	0.8542	6 dk 25 sn
0.0006	0.9903	0.8542	6 dk 25 sn
0.0007	0.9976	0.8542	6 dk 24 sn
0.0008	0.9976	0.8542	7 dk 16 sn
0.0009	0.9940	0.8333	7 dk 11 sn

Tablo 9'a göre en yüksek test doğruluğunu veren öğrenme oranı kullanılarak "adam" optimizasyon yöntemi ile ResNet yeniden eğitilmiştir. Bu eğitim 7 dk 49 sn sürmüş ve 0.9819 eğitim doğruluğu, 0.8542 test doğruluğu elde edilmiştir.



Şekil 23. 'adam' optimizasyon yöntemi ile ResNet eğitim aşamasına ait doğruluk ve kayıp grafikleri

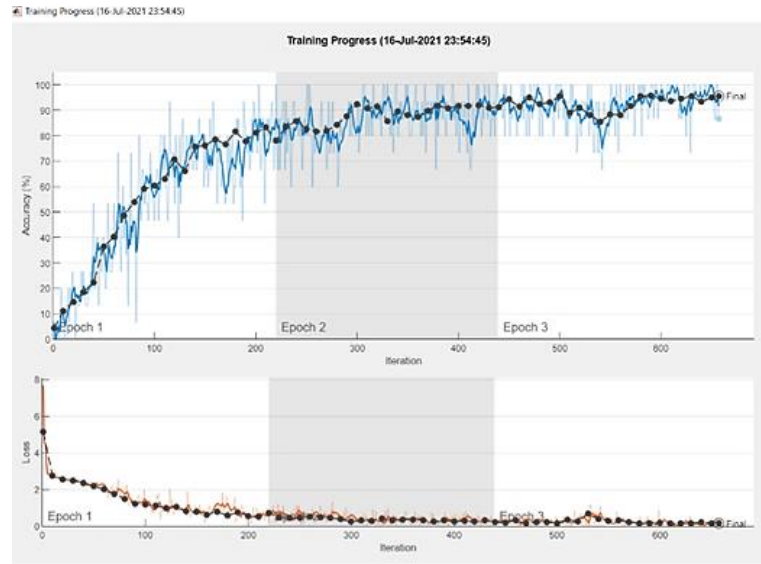
Tablo 6'ya bakıldığında SqueezeNet en yüksek doğruluk değeri sonucunu 0.5 train_test_split için vermiştir. Tablo 10'da, 0.0001 ila 0.0009 arasındaki öğrenme oranları için "sgdm" optimizasyon yöntemi kullanılarak elde edilen eğitim doğruluğu,

test doğruluğu ve süreler verilmiştir. Eğitim için diğer parametreler aynen kullanılmıştır. Öğrenme oranları eğitim boyunca sabit tutulmuştur.

Tablo 10. SqueezeNet için elde edilen sonuçlar

Başlangıç Öğrenme Oranı	Eğitim doğruluğu	Test doğruluğu	Süre
0,0001	0.9178	0.8333	1 dk 34 sn
0.0002	0.9534	0.8125	1 dk 35 sn
0.0003	0.9671	0.8958	2 dk 43 sn
0.0004	0.9534	0.7500	1 dk 35 sn
0.0005	0.9726	0.7917	1 dk 34 sn
0.0006	0.9534	0.8750	1 dk 35 sn
0.0007	0.9068	0.7500	1 dk 34 sn
0.0008	0.9479	0.8542	1 dk 37 sn
0.0009	0.8959	0.7500	1 dk 37 sn

Tablo 10'a göre en yüksek test doğruluğunu veren öğrenme oranı kullanılarak "adam" optimizasyon yöntemi ile SqueezeNet yeniden eğitilmiştir. Bu eğitim 2 dakika 8 saniye sürmüştür ve 0.7890 eğitim doğruluğu, 0.7083 test doğruluğu elde edilmiştir.



Şekil 24. 'adam' optimizasyon yöntemi ile SqueezeNet eğitim aşamasına ait doğruluk ve kayıp grafikleri

5. SONUÇ

Kültürel miras ve yerler bir ulus, hatta bütün dünya için çok önemlidir, bu sebeple de kültürel mirasa sahip mekanların görüntülerinin depolanması ve sınıflandırılması günümüzde kültürel mirasın korunması için oldukça önemlidir. Bu çalışmada görüntünün doğru sınıflandırılması ve kültürel mirasın korunması sürecinde tarihsel miras görüntülerinin sınıflandırılmasında derin öğrenme modelleri için en yüksek doğruluğun elde edilmesi amacıyla çalışmalar yapılmıştır.

Kullanılan yöntemler, uygulama aşamasında gidilen yollar, sınıflandırma işlemleri ve doğruluğu, yapay sinir ağları mimarilerinin karşılaştırılması, alınan sonuçlarda eğitim ve test doğruluk değerlerinin karşılaştırılması ve bu çalışmanın önemi ve literatüre katkısı bu bölümde belirtilmiştir.

Bu makalede, kültürel miras görüntülerinin sınıflandırılmasında aktarım öğrenmesi ile derin öğrenme yöntemlerinin performansları değerlendirilmiştir. Bu amaçla, literatüre göre yarışmalarda başarılı olduğu kanıtlanmış olan derin öğrenme yöntemleri (GoogleNet, AlexNet, ResNet, SqueezeNet) için sonuçlar elde edilmiştir. Toplamda 14 kategoriye ait çoğaltma tekniği ile 11.930 görüntünün elde edildiği bir veri kümesi oluşturulmuştur. Veri kümesi, hem görüntü sayısının artırılması hem de yeni kategorilerin dahil edilmesi ile genişleyebilir ve yeni ve daha kapsamlı çalışmaların önünü açabilir. Bu oluşturulan 11.930 görüntü, tezde bahsi geçen bütün yöntemler ile matlab ortamında ayrı ayrı eğitimleri yapılmıştır. Bu yöntemler ve yöntemlerin üzerlerinde yapılan küçük değişikliklerle alınan sonuçlar kaydedilmiş ve karşılaştırılması yapılmıştır. Derin öğrenme yöntemi modellerinin eğitim sonunda disk üzerindeki boyutu, parametreler için model boyutu ve toplam parametre sayısı Tablo 11'de karşılaştırılmıştır.

Tablo 11. Derin öğrenme modellerinin boyutlarının karşılaştırılması

Derin Öğrenme Yöntemi	Eğitim Sonunda Disk Üzerindeki Boyutu	Parametreler için Model Boyutu	Toplam Parametre Sayısı
AlexNet	1.69 GB	227 MB	61 Milyon
GoogleNet	161.9 MB	44 MB	11.7 Milyon
ResNet	1.35 GB	27 MB	7 Milyon
SqueezeNet	17.24 MB	5.2 MB	1.24 Milyon

Tablo 11'e göre SqueezeNet parametre sayısı ve boyutu diğer derin öğrenme modellerinden daha az olduğu için bellek ve işlemci kısıtlaması olan donanımlar için daha uygundur.

Görüntülerin doğru yorumlanması, bu tür uygulamalara büyük bir katma değer sağlar, çünkü bu tür uygulamalardaki olağan sorun, çok fazla veriye sahip olmak değil, onlardan maksimum miktarda bilgi çıkarmak ve kolayca erişilebilir kılmaktır. Elde edilen doğruluk sonuçları çok tatmin edicidir. Tez çalışmasında bulunan sonuçlar göz önüne alındığında derin öğrenmenin kullanımının kültürel miras görüntülerini sınıflandırma görevlerinde büyük bir ilerleme olacağı düşünülebilir.

En iyi yaklaşımın bulunması her bir ağ ve ağın özelliklerine bağlı farklılıklar gösterir ve ince ayarlamalar sayesinde bulunur. Genel olarak, hesaplama kaynakları sınırlıysa veya mevcut veri kümesi çok büyük değilse, uygulaması genellikle daha basit olan ince ayar tekniklerinin kullanılması tavsiye edilebilir ama zaman ve kaynakların olması durumunda tam eğitim tercih edilebilir.

Belirtildiği gibi, elde edilen doğruluk değerleri yüksektir, ancak mantıksal olarak belirli faktörler dikkate alınarak iyileştirilebilir. En basit seçenek eğitim sürelerini artırmak olabilir, ancak eğitim sağlandıktan sonra ağ eğitimine devam etmenin doğruluğu artırmayacağı ve aşırı uyum nedeniyle azaltabileceği açıktır. Doğruluğu artırmak için bazı olası çözümler, daha doğru etiketlenmiş eğitim görüntülerini toplamak, çok sınıflı bir sınıflandırıcı kullanmak, bağımsız bir test seti kullanmak veya kullanılan mimariyi iyileştirmek olabilir. Bu nedenle, tarihi mekan görüntülerin sınıflandırılması için, daha iyi öğrenme modellerinin kullanımı ile daha fazla eğitim verisinin kullanımı arasında bir denge bulmak önemlidir. Daha büyük, iyi etiketlenmiş veri kümeleri oluşturmak, en az yeni algoritmaların geliştirilmesi kadar önemlidir. Çok sayıda veri kümesinin sürekli gelişimi sayesinde sahne tanıma zorluklarının üstesinden gelmiş ve güzel ilerlemeler kaydedilmiştir. Yakın gelecekte de bu çalışma ve benzer bir çok çalışmanın üstüne eklenerek daha başarılı sonuçlar alınacağı açıktır.

Bu tezde tarihi mekan görüntülerinin sınıflandırılmasında farklı derin öğrenme tekniklerinin uygulanması için kullanılan metodoloji sunulmuş ve elde edilen sonuçlar gösterilmiştir. Bir çok yöntem ve yaklaşım kullanılmış, yararlanılan mimariler üzerinde ince ayarlamalar yapılmıştır. Sonucunda dikkate değer başarılı sonuçlar elde edilmiş ve

karşılaştırmaları yapılmıştır. Bu değerleri elde etmek için derin öğrenmenin en temel mimarilerinden yararlanılmıştır.

Son olarak bu çalışma baz alınarak, ağın sonuçlarını iyileştirme ya da görülen hataları tespit ederek düzeltme sağlanabilir. İleride literatüre yeni ve faydalı çalışmalar eklenmesine olanak sağlamaktadır.

KAYNAKLAR

- Anderson, J. A. (1972). A simple neural network generating on interactive memory. *Mathematical Biosciences*, 14, 197-220.
- Andreopoulos, A., & Tsotsos, J. K. (2013). 50 years of object recognition: Directions forward. *Computer vision and image understanding*, 117(8), 827-891.
- Anwar, S., Hwang, K., & Sung, W. (2015, April). Fixed point optimization of deep convolutional neural networks for object recognition. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 1131-1135). IEEE.
- Blahuta, J., Soukup, T., Čermák, P., Novák, D., & Večerek, M. (2013). Semi-automatic ultrasound medical image recognition for diseases classification in neurology. In *Advances in Intelligent Analysis of Medical Data and Decision Support Systems* (pp. 125-133). Springer, Heidelberg.
- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. *the Journal of machine Learning research*, 3, 993-1022.
- Broomhead, D. S., & Lowe, D. (1988). *Radial basis functions, multi-variable functional interpolation and adaptive networks*. Royal Signals and Radar Establishment Malvern (United Kingdom).
- Castelman, R. K., 1996. Digital image processing. Prentice hall, Englewood Cliffs, New Jersey, USA. Neuman, M. R., H. D. Sapirstein, E. Shwedyk and W. Bushuk. 1989. Wheat grain colour analysis by digital image processing. II. Wheat class discrimination. *Journal of Cereal Science* 10: 183-188.
- Cengil, Emine (2017). 'Derin Öğrenme Tabanlı Yüz Ayırt Etme ve Tanıma', Bilgisayar Mühendisliği Anabilim Dalı, Fırat Üniversitesi, Elazığ.
- Ciregan, D., Meier, U., & Schmidhuber, J. (2012, June). Multi-column deep neural networks for image classification. In *2012 IEEE conference on computer vision and pattern recognition* (pp. 3642-3649). IEEE.
- Ciresan, D. C., Meier, U., Masci, J., Gambardella, L. M., & Schmidhuber, J. (2011, June). Flexible, high performance convolutional neural networks for image classification. In *Twenty-second international joint conference on artificial intelligence* (Vol. 22, No. 1, p. 1237)..
- Colliot, O.; Camara, O.; and Bloch, I. 2006. Integration of fuzzy spatial relations in deformable models application to brain mri segmentation. *Pattern recognition* 39(8):1401– 1414.
- Ćosović, M., & Janković, R. (2020, March). CNN Classification of the Cultural Heritage Images. In *2020 19th International Symposium INFOTEH-JAHORINA (INFOTEH)* (pp. 1-6). IEEE.

- Çağlayan, Ali (2018). ‘Derin öğrenme tekniklerini kullanarak rgb-d nesne tanıma’, Bilgisayar Mühendisliği Anabilim Dalı, Hacettepe Üniversitesi, Ankara.
- De Menezes, R. S. T., Magalhaes, R. M., & Maia, H. (2019). Object recognition using convolutional neural networks. In *Recent Trends in Artificial Neural Networks- from Training to Prediction*. IntechOpen.
- Demirden, A. (2019). Deep Learning Nedir? Erişim Adresi: <https://medium.com/@altugdemirden/deep-learning-836ae89a2be3#:~:text=Deep%20Learning%2C%20veya%20T%C3%BCrk%C3%A7e,i%C3%A7in%20kullan%C4%B1lan%20bir%20algoritmalar%20toplulu%C4%9Fudur>. Erişim Tarihi: 10 Ocak 2021
- DİPOVA, N. (2018). Görüntü Analizi Yöntemlerinin Geoteknik Mühendisliğinde Kullanımı. *Mehmet Akif Ersoy Üniversitesi Fen Bilimleri Enstitüsü Dergisi*, 9(1), 33-44.
- Doğan, F., & Türkoğlu, İ. (2019). Derin öğrenme modelleri ve uygulama alanlarına ilişkin bir derleme. *Dicle Üniversitesi Mühendislik Fakültesi Mühendislik Dergisi*, 10(2), 409-445.
- Efe, M. O., & Kaynak, O. (1999). A comparative study of neural network structures in identification of nonlinear systems. *Mechatronics*, 9(3), 287-300.
- Efe, M. O., Abadoğlu, E., & Kaynak, O. (1999). A novel analysis and design of a neural network assisted nonlinear controller for a bioreactor. *International Journal of Robust and Nonlinear Control: IFAC-Affiliated Journal*, 9(11), 799-815.
- Erdoğan, Turgay (2012), ‘Hareketli Konveyör Üzerinde Kamera Görüntüsü İle Nesne Tanıma ve Nesnelere Yerine Koyma Uygulaması’, Fen Bilimleri Enstitüsü, İzmir.
- Fei-Fei, L., & Perona, P. (2005, June). A bayesian hierarchical model for learning natural scene categories. In *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (Vol. 2, pp. 524-531). IEEE.
- Felzenszwalb, P. F., Girshick, R. B., McAllester, D., & Ramanan, D. (2009). Object detection with discriminatively trained part-based models. *IEEE transactions on pattern analysis and machine intelligence*, 32(9), 1627-1645.
- Fukushima, K., & Miyake, S. (1982). Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition. In *Competition and cooperation in neural nets* (pp. 267-285). Springer, Berlin, Heidelberg.
- Fukushima, K., Miyake, S., & Ito, T. (1983). Neocognitron: a neural network model for a mechanism of visual pattern recognition. *IEEE Transactions on Systems, Man and Cybernetics*, SMC-13.
- Gatys, L. A., Ecker, A. S., & Bethge, M. (2015). A neural algorithm of artistic style. *arXiv preprint arXiv:1508.06576*.

- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580-587).
- Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 580-587).
- Gonzalez R. C., Woods R. E. (2007). *Digital Image Processing*, 3th Ed., (pp. 1-7), A.B.D., New Jersey : Prentice Hall.
- He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).
- Hebb, D. O. (1949). The organization of behaviour. *The first stage of perception: growth of the assembly*, 4, 60-78
- Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, 313(5786), 504-507.
- Hinton, G. E., & Salakhutdinov, R. R. (2006). Reducing the dimensionality of data with neural networks. *science*, 313(5786), 504-507.
- Hinton, G. E., Srivastava, N., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. R. (2012). Improving neural networks by preventing coadaptation of feature detectors. arXiv preprint arXiv:1207.0580.
- Hubel, D. H., & Wiesel, T. N. (1968). Receptive fields and functional architecture of monkey striate cortex. *The Journal of physiology*, 195(1), 215-243.
- Hussin, R., Juhari, M. R., Kang, N. W., Ismail, R. C., & Kamarudin, A. (2012). Digital image processing techniques for object detection from complex background image. *Procedia Engineering*, 41, 340-344.
- Iandola, F. N., Han, S., Moskewicz, M. W., Ashraf, K., Dally, W. J., & Keutzer, K. (2016). SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and < 0.5 MB model size. arXiv preprint arXiv:1602.07360.
- İnik Ö. & Ülker E. (2017) Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri, 419
- Jähne, B. (2005). *Digitale Bildverarbeitung*. Springer-Verlag.
- Janković, R. (2020). Machine learning models for cultural heritage image classification: Comparison based on attribute selection. *Information*, 11(1), 12.
- Jarrett, K., Kavukcuoglu, K., Ranzato, M. A., & LeCun, Y. (2009, September). What is the best multi-stage architecture for object recognition?. In *2009 IEEE 12th international conference on computer vision* (pp. 2146-2153). IEEE.

- Johan, T. M. ve Prabuwono, A. S. (2011). Recognition of bolt and nut using artificial neural network. Pattern analysis and intelligent robotics (ICPAIR), 2011 international conference, 1, 165-170.
- Junuz, E. Classification Methods in Cultural Heritage.
- Kavuncu, S. K. (2018). Makine Öğrenmesi ve Derin Öğrenme: Nesne Tanıma Uygulaması. Kırıkkale Üniversitesi. *Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalı, Yüksek Lisans Tezi*.
- Keskenler, M. F., & Keskenler, E. F. (2017). Geçmişten günümüze yapay sinir ağları ve tarihçesi. *Takvim-i Vekayi*, 5(2), 8-18.
- Kılınç, Hatice (2018). 'Derin öğrenmeli konvolüsyonel sinir ağları (deep learning convolutional neural network) kullanarak fotoğraftan trizomi 21(down sendromu) tespiti', *Biyomedikal Mühendisliği Anabilim Dalı, Çukurova Üniversitesi, Adana*.
- Kohonen, T. (1972). Correlation matrix memories. *IEEE transactions on computers*, 100(4), 353-359.
- Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 25, 1097-1105.
- Kulkarni, U., Meena, S. M., Gurlahosur, S. V., & Mudengudi, U. (2019, September). Classification of cultural heritage sites using transfer learning. In *2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM)* (pp. 391-397). IEEE.
- Kurniawan, H., Salim, A., Suhartanto, H., & Hasibuan, Z. A. (2011, May). E-cultural heritage and natural history framework: an integrated approach to digital preservation. In *International Conference on Telecommunication Technology and Applications (IACSIT)* (pp. 177-182).
- LeCun, Y., & Bengio, Y. (1995). Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10), 1995.
- LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, 86(11), 2278-2324.
- Li, H., Lin, Z., Shen, X., Brandt, J., & Hua, G. (2015). A convolutional neural network cascade for face detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 5325-5334).
- Lin, M., Chen, Q., & Yan, S. (2013). Network in network. arXiv preprint arXiv:1312.4400.
- Llamas, J., Lerones, P. M., Zalama, E., & Gómez-García-Bermejo, J. (2016, October). Applying deep learning techniques to cultural heritage images within the inception project. In *Euro-Mediterranean Conference* (pp. 25-32). Springer, Cham.

- Llamas, J., M Leronés, P., Medina, R., Zalama, E., & Gómez-García-Bermejo, J. (2017). Classification of architectural heritage images using deep learning techniques. *Applied Sciences*, 7(10), 992.
- MATLAB (R2018b)* [Bilgisayar Yazılımı]. The Mathworks, Natick MA.
- Maturana, D., & Scherer, S. (2015, September). Voxnet: A 3d convolutional neural network for real-time object recognition. In *2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* (pp. 922-928). IEEE.
- McCulloch, W. S., & Pitts, W. A. (1943). A logical calculus of the ideas immanent in nervous activity. *Buttetin of Mathematics and Biophysics*, 5, 115-133.
- Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., ... & Van Gool, L. (2005). A comparison of affine region detectors. *International journal of computer vision*, 65(1), 43-72.
- Miyajima, K., and Ralescu, A. (1994). Spatial organization in 2d segmented images: representation and recognition of primitive spatial relations. *Fuzzy Sets and Systems* 65(2- 3):225–236.
- Murphy, K.; Torralba, A.; Freeman, W.; et al. (2003). Using the forest to see the trees: a graphical model relating features, objects and scenes. *Advances in neural information processing systems* 16:1499–1506.
- Nabiyev, Vasif V., Yapay Zeka. Sözkese Matbaacılık. Ankara, 2005.
- Narendra, K. S., & Parthasarathy, K. (1990). Identification and Control of Dynamical Systems Using Neural Networks. *IEEE Transactions on Neural Networks*, 1(1), 4-27.
- Nova Research Lab. (2020). *Öğrenme Aktarımı/Transfer Learning*. Erişim adresi: <https://medium.com/novaresearchlab/%C3%B6%C4%9Frenme-aktar%C4%B1m%C4%B1-transfer-learning-c0b8126965c4>.
- Obeso, A. M., Vázquez, M. S. G., Acosta, A. A. R., & Benois-Pineau, J. (2017, June). Connoisseur: classification of styles of Mexican architectural heritage with deep learning and visual attention prediction. In *Proceedings of the 15th international workshop on content-based multimedia indexing* (pp. 1-7).
- Oliva, A., and Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. *International journal of computer vision* 42(3):145–175.
- Pan, Z., Rust, A. G., & Bolouri, H. (2000). Image redundancy reduction for neural network classification using discrete cosine transforms. In *Neural Networks, 2000. IJCNN 2000, Proceedings of the IEEE-INNS-ENNS International Joint Conference on* (Vol. 3, pp. 149-154).
- Python (Versiyon 3.9.4)* [Bilgisayar Yazılımı]. Python Yazılım Vakfı.

- Richardson Santiago Teles de Menezes, Rafael Marrocos Magalhaes and Helton Maia
Object Recognition Using Convolutional Neural Networks, Submitted: May 6th
2019Reviewed: September 14th 2019 Published: November 7th 2019.
- Rosch, E., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive psychology*, 7(4), 573-605.
- Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychoanalytic Review*, 65, 386-408.
- Rumelhart, D. E., & McClelland, J. L. (1986). Parallel distributed processing, explorations in the microstructure of cognition. Foundations, MIT Press Cambridge, MA, 1.
- Rumelhart, D. E., & McClelland, J. L. (1988). Parallel distributed processing, explorations in the microstructure of cognition, A handbook of models, programs and exercises. MIT Press, Cambridge, MA.
- Rumelhart, D. E., Hinton, D. E., & Williams, R. J. (1986). Learning representation by backpropagating errors. *Nature*, 323(9), 533-536.
- Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., ... & Berg, A. C. (2015). Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3), 211-252.
- Sarkar, D. (2018). A comprehensive hands-on guide to transfer learning with real-world applications in deep learning. *Towards Data Science. Saatavissa: <https://towardsdatascience.com/a-comprehensive-handson-guide-to-transfer-learning-with-real-world-applications-in-deep-learning-212bf3b2f27a>. Hakupäivä, 20, 2020.*
- Schwarz, M., Schulz, H., & Behnke, S. (2015, May). RGB-D object recognition and pose estimation based on pre-trained convolutional neural network features. In *2015 IEEE international conference on robotics and automation (ICRA)* (pp. 1329-1335). IEEE.
- Serrano, N., Savakis, A. E., & Luo, J. (2004). Improved scene classification using efficient low-level features and semantic cues. *Pattern Recognition*, 37(9), 1773-1784.
- Sonka, M., Hlavac, V., & Boyle, R. (2014). *Image processing, analysis, and machine vision*. Cengage Learning.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., & Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *The Journal of Machine Learning Research*, 15(1), 1929- 1958.
- Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. *Cvpr*.

- Szumner, M., and Picard, R. W. (1998). Indoor-outdoor image classification. In Content-Based Access of Image and Video Database, 1998. Proceedings., 1998 IEEE International Workshop on, 42–51. IEEE.
- Şeker, A. (2018, September). Evaluation of Fabric Defect Detection Based on Transfer Learning with Pre-trained AlexNet. In *2018 International Conference on Artificial Intelligence and Data Processing (IDAP)* (pp. 1-4). IEEE.
- Tamura, S. I., & Tateishi, M. (1997). Capabilities of a four-layered feedforward neural network: four layers versus three. *IEEE Transactions on Neural Networks*, 8(2), 251-255.
- Tang, Y. (2013). Deep learning using linear support vector machines. arXiv preprint arXiv:1306.0239.
- Ünal, Fatıma Zehra (2017). Derin Öğrenme ile Yüz Tanıma, Fen Bilimleri Enstitüsü, Ankara Üniversitesi, Ankara.
- Viola, P., & Jones, M. (2001, December). Rapid object detection using a boosted cascade of simple features. In *Proceedings of the 2001 IEEE computer society conference on computer vision and pattern recognition. CVPR 2001* (Vol. 1, pp. I-I). Ieee.
- Vogel, J., and Schiele, B. (2004). A semantic typicality measure for natural scene categorization. In *Joint Pattern Recognition Symposium*, 195–203. Springer.
- Widrow, B., & Hoff, M. E. (1960). Adaptive switching circuits. *WESTCON Convention, Record Part IV*, 96-104.
- Xiao, J.; Hays, J.; Ehinger, K. A.; Oliva, A.; and Torralba, A. (2010). Sun database: Large-scale scene recognition from abbey to zoo. In *Computer vision and pattern recognition (CVPR), 2010 IEEE conference on*, 3485–3492. IEEE.
- Yaman, K., & Aktürk, N. (2001). Görüntü İşleme ile Kişi Yoğunluklarının Belirlenmesi. *UMTS, Selçuk Üniversitesi, Konya*, 12-14.
- Yang, C., Liu, H., Wang, S., & Liao, S. (2016). Scene-level geographic image classification based on a covariance descriptor using supervised collaborative kernel coding. *Sensors*, 16(3), 392.
- Yang, M. H., Kriegman, D. J., & Ahuja, N. (2002). Detecting faces in images: A survey. *IEEE Transactions on pattern analysis and machine intelligence*, 24(1), 34-58.
- Zhang, Y. D., Govindaraj, V. V., Tang, C., Zhu, W., & Sun, J. (2019). High performance multiple sclerosis classification by data augmentation and AlexNet transfer learning model. *Journal of Medical Imaging and Health Informatics*, 9(9), 2012-2021.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı : İrem UYANIK

EĞİTİM DURUMU

Lisans Öğrenimi : 2018, KTO Karatay Üniversitesi, Mühendislik Fakültesi,
Elektrik- Elektronik Bölümü

Yüksek Lisans Öğrenimi : 2021, KTO Karatay Üniversitesi, Mühendislik Fakültesi,
Elektrik- Elektronik Bölümü

Bildiği Yabancı Diller : İngilizce

Tarih: 09 Eylül 2021