



T.C.

BARTIN ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

AKILLI SİSTEMLER MÜHENDİSLİĞİ

ANABİLİM DALI

YÜKSEK LİSANS TEZİ

YÜZ GÖRÜNTÜLERİNDEN DERİN ÖĞRENMEYE DAYALI

STRES TESPİTİ

HASRET YAŞAR

DANIŞMAN

DR. ÖĞR. ÜYESİ EVRİM GÜLER

BARTIN-2023



T.C.

BARTIN ÜNİVERSİTESİ

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ

AKILLI SİSTEMLER MÜHENDİSLİĞİ ANABİLİM DALI

YÜZ GÖRÜNTÜLERİNDEN DERİN ÖĞRENMEYE DAYALI STRES TESPİTİ

YÜKSEK LİSANS TEZİ

Hasret YAŞAR

BARTIN-2023

BEYANNAME

Bartın Üniversitesi Lisansüstü Eğitim Enstitüsü tez yazım kılavuzuna göre Dr. Öğr. Üyesi Evrim GÜLER danışmanlığında hazırlamış olduğum “YÜZ GÖRÜNTÜLERİNDEN DERİN ÖĞRENMEYE DAYALI STRES TESPİTİ” başlıklı yüksek lisans tezimin bilimsel etik değerlere ve kurallara uygun, özgün bir çalışma olduğunu, aksinin tespit edilmesi halinde her türlü yasal yaptırımını kabul edeceğimi beyan ederim.

28.08.2023

Hasret YAŞAR

ÖNSÖZ

Değerli görüş ve önerileriyle çalışmalarına katkı sağlayan tez danışmanım Dr. Öğr. Üyesi Evrim GÜLER'e ve Dr. Öğr. Üyesi Aslı EYECİOĞLU ÖZMUTLU'ya,

Yüksek lisans çalışmalarım kapsamında beni her zaman destekleyen sevgili eşim; Emrah YAŞAR'a memnuniyetimi iletiyorum.

Hasret YAŞAR

ÖZET

Yüksek Lisans Tezi

YÜZ GÖRÜNTÜLERİNDEN DERİN ÖĞRENMEYE DAYALI STRES TESPİTİ

Hasret YAŞAR

Bartın Üniversitesi

Lisansüstü Eğitim Enstitüsü

Akıllı Sistemler Mühendisliği Anabilim Dalı

Tez Danışmanı: Dr. Öğr. Üyesi Evrim GÜLER

Bartın-2023, sayfa: 42

Endüstri ve güvencesiz çalışma alanlarında, bireyler işçi veya personel olsun çalışma ortamlarında mesleki görevlerinin bir ön koşulu olarak refahlarını kaçınılmaz olarak etkileyen çok sayıda faktöre maruz kalırlar. Bununla birlikte, bu stres faktörlerinin birey üzerindeki etkisini ölçmek, sayısal ölçümlere bağlı olmasından dolayı zor bir görev haline gelmektedir. Ayrıca kişilerin stres değeri buldukları ortamın etkilerine bağlı olarak değişiklikler göstermektedir ve psikolojik olarak bireyler üzerinde farklı etkiler oluşturmaktadır. Aşırı stres, kişinin fiziksel ve zihinsel sağlığı, üretkenlik ve iş performansında, odaklanma ve karar verme ile uygulama yeteneklerinde, işe bağlılığı ve iş memnuniyeti üzerinde önemli ölçüde negatif etkiler oluşturmaktadır. Bunun ışığında, çalışanların taşıdığı hem fiziksel hem de psikolojik gerilimleri, kendi mesleki alanlarında üstlendikleri belirli görevlere göre uyarlanmış ölçülebilir ölçümlerle ifade etmek çok büyük önem taşıyor. Tehlikeli sınıfta yer alan işletmelerin (örneğin, yüksekte çalışma, maden ocak işlerinde çalışma, vb.) yüksek dikkat ve enerji isteyen iş guruplarında çalışanın stres analizi, duyu/durum tespiti çalışanların anlık olarak yüz görüntülerinin otomatik olarak alınması ile stres durum belirlenmesi bu çalışmada yapılacaktır. Çalışanın durumu, elde edilen anlık yüz resimleri üzerinde görüntü işleme teknolojileri ile işe başlamadan önce tespit edilip uygun pozisyonlarda çalışması için öneri vermeyi hedeflenmektedir. Daha önceki çalışmalarda, çalışanın stres seviyesini tespit etmek için kalp atış hızı, ten rengi değişimi, vücut sıcaklığı

gibi bileşenlerle ölçümler yapıldığı gözlemlenmiştir. Fakat, bu araştırmanın amacı, sadece yüz resimlerine bakarak derin öğrenme tekniklerinden başarı oranı yüksek seviyede olan özelleştirilmiş Evrişimli Sinir Ağları (ESA) modeli kullanılarak ne kadar başarılı stresin ölçülebildiğini görmekle birlikte, alınan yüz görüntüleri üzerinden analizler yapılarak çalışanın ne kadar stresli olduğu tespit edilmek istenmektedir. Çalışmada, FER2013 veri seti ile 7 farklı duygu tespiti ve yüz üzerindeki 68 noktadan göz, ağız, kaş aktiviteleri referans alınarak çalışanların duygu durumları ile stres analizi sonuçları başarılı bir şekilde elde edilmektedir.

Anahtar Kelimeler: Derin öğrenme, duygu analizi, kaş hareketi, stres tespiti, yüz tanıma

ABSTRACT

M. Sc. Thesis

DEEP LEARNING-BASED STRESS DETECTION TECHNIQUES FROM FACE IMAGES

Hasret YAŞAR

Bartın University

Graduate School

Department of Intelligent Systems Engineering

Thesis Advisor: Asst. Prof. Dr. Evrim GÜLER

Bartın-2023, pp: 42

The most important element of industry and hazardous work areas, people, workers or staffs are both exposed to various factors in the working environment and also affected by work conditions as a requirement of their jobs. It cannot be calculated clearly that the stress of these factors affecting the human/employee can be expressed with numerical values. In addition, the stress value of people varies depending on the effects of the environment they are in and creates different effects on individuals psychologically. The detrimental impact of excessive stress on an individual's overall well-being, effectiveness at work, cognitive abilities, and job satisfaction cannot be overstated. Hence, it becomes imperative to quantify the physical and psychological strains experienced by employees, tailored to their specific work environments. In this study, stress analysis, emotion/state detection of the employees in the work groups that require high attention and energy will be made by automatically taking instant facial images of the employees in the enterprises in the dangerous class (for example, working at height, working in mining works, etc.). It is aimed to determine the status of the employee before starting the work with image processing technologies on the instant face pictures obtained and to give suggestions for them to work in appropriate positions. In previous studies, it was observed that measurements were made with components such as heart rate, skin color change, and body temperature to determine the

stress level of the employee. The objective of this study is to assess the extent to which stress can be quantified using the specialized Convolutional Neural Networks (CNN) model, a highly effective deep learning method. The focus lies solely on evaluating facial images to ascertain the employee's level of stress. In the study, with the FER2013 data set, 7 different emotion detections and eye, mouth, and eyebrow activities from 68 points on the face are taken as reference, and the results of the employees' emotional states and stress analysis are successfully obtained.

Keywords: Deep learning, emotion analysis, eyebrow movement, face recognition, stress detection

İÇİNDEKİLER

BEYANNAME	ii
ÖNSÖZ	iii
ÖZET	iv
ABSTRACT	vi
İÇİNDEKİLER.....	viii
ŞEKİLLER DİZİNİ.....	ix
TABLolar DİZİNİ.....	x
1. GİRİŞ.....	1
2. MATERYAL VE METOT	6
2.1. Yapay Zekâ	6
2.2. Makine Öğrenmesi	6
2.3. Görüntü İşleme	7
2.4. Derin Öğrenme	7
2.5. Evrişimli Sinir Ağı	8
2.5.1. Aktivasyon Fonksiyonları	9
2.5.2. Havuzlama.....	13
2.5.3. Tam Bağlı Tabaka ve Softmax Tabakası	14
2.5.4. ESA Parametreleri	15
2.6. Optimizasyon Algoritmaları.....	16
2.7. Model Özellikleri	18
2.7.1. VGG16	18
2.7.2. VGG19	20
2.7.3. ResNet.....	21
2.7.4. MobilNetV2	22
2.8. Stres.....	24
2.9. Dlib Kütüphanesi (68 nokta)	25
2.10. Çözüm Akışı.....	26
2.11. Kullanılan Veri Seti	27
2.11.1. Veri Setinin Hazırlanması.....	28
2.11.2. Kullanılan Modeller	29
3. BULGULAR.....	30
4. TARTIŞMA.....	36
4.1. Stres Değerlendirmesindeki Zorluklar	36
4.2. Stres Analizinde Yüz Görüntülemenin Etkinliği	36
4.3. Çok Yönlü Stres Değerlendirmesinin Geliştirilmesi	37
4.4. Stres Yönetimi Müdahalelerinin Uygulanması.....	37
4.5. İş Optimizasyonu için İnsan-Makine İş Birliği.....	37
5. SONUÇ VE ÖNERİLER	39
KAYNAKLAR.....	40

ŞEKİLLER DİZİNİ

Şekil	Sayfa
No	No
Şekil 2.1 : Sınıflandırma ve otomatik özellik çıkarımı	8
Şekil 2.2 : Evrişimli sinir ağı modeli	8
Şekil 2.3 : Derin öğrenme	9
Şekil 2.4 : Sigmoid aktivasyon fonksiyonu.....	10
Şekil 2.5 : Tanh aktivasyon fonksiyonu	11
Şekil 2.6 : ReLU aktivasyon fonksiyonu	12
Şekil 2.7 : Softmax aktivasyon fonksiyonu.....	13
Şekil 2.8 : Maksimum havuzlama ve ortalama havuzlama.....	14
Şekil 2.9 : Tam bağlı tabaka.....	15
Şekil 2.10 : VGG16.....	19
Şekil 2.11 : VGG16 katman yapısı	20
Şekil 2.12 : MobileNetV2 mimari.....	24
Şekil 2.13 : Dlib'in 68 modeli ile yüz dönüm noktası tespiti	26
Şekil 2.14 : Çözüm akışı	27
Şekil 3.1 : Uygulama test görüntüsü	35

TABLULAR DİZİNİ

Tablo	Sayfa
No	No
Tablo 2.1 : Optimizasyon algoritmaları	16
Tablo 2.2 : FER2013 Veri seti dağılımı	28
Tablo 2.3 : Çalışmada kullanılan kriterler	29
Tablo 3.1 : MobileNetV2 - 256	30
Tablo 3.2 : MobileNetV2 - 32	31
Tablo 3.3 : VGG16 - 32	31
Tablo 3.4 : Resnet50 - 32	32
Tablo 3.5 : Resnet50 - 32	33

GİRİŞ

Bireyler, çalışanlar ve personelden oluşan işgücünün sağlık ve güvenliği, endüstri ve tehlikeli çalışma ortamları alanında büyük önem taşımaktadır. Bireyler, profesyonel çevrelerinde sürekli olarak çok çeşitli koşullara maruz kalmakta ve çalışma koşulları onları önemli ölçüde etkilemektedir. Bununla birlikte, bu faktörlerin insanların stres düzeyleri üzerindeki etkisini ölçmek, sayısal değerler kullanılarak tam olarak ölçülemediği için çok yönlü bir zorluk olmaya devam etmektedir. Dahası, bir bireyin stres seviyesi içinde bulunduğu koşullara bağlı olarak değişmekte ve çeşitli psikolojik sonuçlara yol açmaktadır. Aşırı stresin çalışanların fiziksel ve zihinsel refahı üzerindeki zararlı etkileri, üretkenliğin azalması, iş performansının düşmesi, karar verme becerilerinin zayıflaması, odaklanmanın azalması ve işe bağlılık ve iş tatmini düzeylerinin düşmesi gibi geniş kapsamlı sonuçlar doğurmaktadır. Bu nedenle, daha sağlıklı ve daha üretken çalışma ortamlarını teşvik etmek için çalışanların benzersiz iş sorumluluklarıyla karşılaştıkları fiziksel ve psikolojik zorlukları ölçmek çok önemlidir.

Duygusal sıkıntı durumlarını ve streslerini ayırt etmek için 68 yüz noktasından hesaplanan 7 yüz özelliğini (Ağız dışı ($F1$), ağız köşesi ($F2$), göz çevresi ($F3$), kaş aktivitesi ($F4$) ve yüz alanı ($F5$), yüz hareketi ($F6$) ve tüm yüz işaretlerinin ortalama değeri ($F7$)) otomatik olarak alınmıştır. Sunulan yöntem gözün algılayabildiği ve video tabanlı olduğundan nesnelere eklenecek fiziksel sensör ihtiyacını ortadan kaldırdığı gözlemlenmiştir (Bevilacqua vd., 2018).

Yüksek doğruluk için tasarlanmış üç katman üzerinde benzersiz bir noninvaziv mimari ile Yüz Hareketleri Kodlama Sistemi (YHKS) kullanılarak yüz ifadelerini analiz ederek Depresyon Kaygı Stres Ölçeği (DKSÖ) düzeylerini belirlemek amaçlanmıştır. İlk katmanda, Eylem Birimi (EB) sınıflandırması için Aktif Görünüm Modelleri (AGM) ve bir dizi çok sınıflı Destek Vektör Makinesi (DVM) kullanılmıştır. İkinci katmanda ise, EB'lerin yoğunluk seviyelerini içeren bir matris oluşturulur ve üçüncü katmanda ise, optimal bir İleri Beslemeli Sinir Ağı (İBSA), bir örüntü tanıma görevinde ikinci katmandaki matrisi analiz ederek DKSÖ seviyelerini tahmin etmesi amaçlanmıştır. Depresyon için %87,2, anksiyete için %77,9 ve stres için %90,2 doğruluk elde edilmiştir. Sunulan yöntem diğer son teknoloji yöntemlere kıyasla %5 daha yüksek doğruluk, hassasiyet ve özgüllük sunmuştur (Gavrilescu ve Vizireanu, 2019).

Göze çarpmayan akıllı giyilebilir cihazlardan elde edilen fizyolojik sinyalleri kullanarak bireylerin günlük yaşam rutinleri sırasında taşıyabileceği otomatik stres algılama sistemi geliştirilmiştir. Bu sistem, gerçek yaşam koşulları için modaliteye özgü artefakt kaldırma ve özellik çıkarma yöntemlerine sahiptir. Kalp aktivitesi, cilt iletkenliği ve ivmeölçer sinyallerini kullanarak, farklı makine öğrenme yöntemlerini kullanarak farklı durumlardaki stres tespit edilmiştir (Can vd., 2019).

Bireylerin yüzlerinden alınan kayıtlı görüntülerden bireyin duygu analizini yapabileceğimiz “Mutlu”, “Üzgün” ve “Doğal” gibi ifadeleri göz ve ağızların anlık durumlarına bakarak tespiti yapılabilmektedir. Bu nedenle, bireylerin yüz bölgelerinde bulunan göz ve ağız durumları ile psikolojik analizleri yapılmaktadır (Tenekeci vd., 2014). Göz video sekanslarının otomatik analizi yoluyla zihinsel stres tanıma için yeni bir yaklaşım önerilmiştir.

Pupil Çapı (PÇ) ve Pupil Genişleme Hızlandırması (PGH) dahil olmak üzere göz bebeği parametreleri, göz bölgesi Genetik Algoritma (GA) kullanılarak tespit edilmiştir. Çıkarılan özellikler, duyuşsal durumları “stres” ve “rahatlamış” olarak sınıflandırmak için öğrenme sistemine (bulanık DVM) aktarılmıştır. Gözbebeği parametrelerinden elde edilen sonuçlar, Elektrokardiyogram (EKG) ve Fotopletismogram (PPG) dahil olmak üzere diğer iki fizyolojik sinyal ile karşılaştırılmıştır. Deneysel sonuçlar, gözbebeği parametrelerinin diğer iki fizyolojik sinyale kıyasla stres tanıma için büyük bir potansiyele sahip olduğunu ve önerilen stres tanıma sisteminin umut verici olduğunu göstermektedir (Mokhayeri ve Akbarzadeh, 2011).

Bir zamanlama grafiği, bir dizide kullanıcının ardışık fotoğraf yakalamalarının başlatılmasını belirleyen bir fonksiyonun zaman içindeki bir sonucudur. Yüzün ifade edebileceği başlıca 7 duygu olasılıksal bir şekilde ve ağırlık tablosunda belirtilen öncelik ağırlıklarıyla, getirilen her duyguya atanarak formüle edilebilir (Mohan, 2021).

Yüz ifadesini analiz ederek bir kişinin duygusal durumunu tespit eden, gerçek zamanlı, videolar çekilir. Videonun her karesinde ayrı bir duygu tanımlanır ve ardından kaydedilen çekimin ilerleyen saatlerinde stres seviyesinin belirlenmesi sağlanır. Önerilen sistem, stresi tespit etmek için görüntü işleme ve derin öğrenmeyi entegre etmektedir. Uygun girdilerle görüntü işleme mimarisinden elde edilen sonuçlar, doğrusal regresyon modelini eğitmek ve bu modeli test veri seti ile test etmek için kullanılmıştır. Gelecekteki sağlık risklerini en aza

indirmek için son kullanıcıların devam eden stresini başarılı bir şekilde tanımasını sağlayarak elde edilir (Raichur vd., 2017).

Yüz ifadelerini analiz ederek sürücünün duygusal durumlarını tespit eden, gerçek zamanlı, bir izleme sistemi geliştirilmiştir. Sistem, iki olumsuz temel duyguyu, öfke ve iğrenmeyi stresle ilgili duygular olarak kabul eder. Daha iyi performans elde etmek için kafa hareketi ve akustik sinyaller gibi diğer ipuçları da entegre edilebileceği öngörülmüştür (Gao vd., 2014).

Günlük yaşam senaryolarındaki zihinsel iş yükü düzeylerinin bireysel kalibrasyon önlemleri dahil edilerek mobil bir EKG kaydediciden alınan verilerle nasıl ayırt edilebileceği sunulmuştur. Ofis çalışması sırasında zihinsel iş yükü düzeyini belirlemek için üç sınıflandırma yönteminin performansı karşılaştırılmıştır. En iyi sonuçlar, yedi denekten altısı için doğru bir sınıflandırma sağlayan Lineer Diskriminant Analizi (LDA) ile elde edildiği gözlemlenmiştir. K-en yakın komşu algoritması (k-NN) ve DVM, yedi denekten beşi için ofis çalışması sırasında zihinsel iş yükü seviyesinin doğru bir şekilde sınıflandırılmasıyla sonuçlanmıştır (Cinaz vd., 2013).

Bilgisayar ile sürekli olarak ilgilenen bir bireyin anlık durumunda rahatlamadan stresli bir duygusal yapıya dönüşümünü yüz bölgelerinde bulunan göz bebeğindeki boyut değişimiyle ve bireyin yüzündeki periorbital bölgedeki ısı ölçümündeki değişimden çıkarılabilecek öznitelikler ile makine öğrenme modellerinin birleşimiyle sağlanmıştır. Özniteliksel olarak ölçülebilen sinyal değerleri ve entropi bazlı yapıların birlikte ele alınması doğru bir şekilde duygu analizinin yapılmasını sağlamıştır (Baltacı ve Gökçay, 2014).

Hem nesnel (yani, cilt iletkenliği) hem de öznel (yani, değerlik-uyarılma VA derecelendirmeleri) ölçümler için insan bilgisayar etkileşiminde stres tanımadaki cinsiyet farklılıkları araştırılmıştır. Deri iletkenlik sinyalleri, yedi popüler makine öğrenimi sınıflandırıcısı kullanılarak analiz edilmiştir. Her iki grupta da tüm görevler için en iyi stres tanıma doğruluğu LDA ile elde edilmiştir (Liapis vd., 2015).

Stresin tanımlanabileceği temel parametreler, bir kişinin zihinsel durumu hakkında ayrıntılı bilgi sağlayan kalp hızı, galvanik cilt tepkisi, vücut ısısı, kan basıncıdır. Önerilen sistemin temel amacı, farklı pozisyonlarda ve ruh hallerinde elektrokardiyografi kullanarak fizyolojik veriler aracılığıyla zihinsel stresi analiz etmektir. Stres tespiti için farklı ön işleme teknikleri kullanılabilir. Özellik çıkarmada ayrık dalgacık dönüşümü uygulanabilir. Yapay sinir ağı,

destek vektör makinesi, bayes ağı ve karar ağacı gibi birçok sınıflandırıcı, doğruluk temelinde daha doğru sonuçlar elde etmek için kullanılmaktadır. Sensör destekli taşınabilir, giyilebilir ve implante edilebilir cihazların mevcudiyeti, büyüyen Nesnelerin İnterneti'nde (IoT) her yerde yaygınlaştıkça, fizyolojik sensör analitiği giderek daha önemli hale gelmektedir. Stresi tespit etmek için başarılı bir şekilde fizyolojik çoklu sensör çalışmaları yapılmıştır (Shelke vd., 2021).

ESA derin öğrenme teknikleri ile yüz ifadelerinin belirli sınıflara ayrılarak anlık olarak bir bireyin yüz ifadesinin duygusal olarak sınıflandırılmasına bağlı olarak stres durumunun analizini yapmaktadır. Önceden eğitilmiş VGG16, VGG19 ve Inception-ResNet V2 ağları kullanılmıştır (Almeida ve Rodrigues, 2021).

Bireylerin yüz ifadelerinde stres tepkilerini gözlemleyebilmekteyiz. İnsanlar arasındaki uyum ve istikrarı zenginleştirecek şekilde stresi analiz etmek ve tespit etmek için çeşitli tekniklerin serbest bırakılmasına yardımcı olacaktır, böylece zamanla tespit edilerek ciddi sorunlara yol açan depresyona neden olmaz (Singh, 2019).

Çeşitli dış ve iç stres etkenler yoluyla duygusal durumlarda (nötr, rahat ve stresli/endişeli) sistematik değişkenliği indüklemek için kapsamlı bir deneysel protokol oluşturulmuştur. İncelenen özellikler arasında gözle ilgili olaylar, ağız aktivitesi, baş hareketi parametreleri ve kamera tabanlı fotopletismografi ile tahmin edilen kalp hızı yer almaktadır. Özellikleri seçmek için bir öznitelik seçim prosedürü, ardından her deney aşamasında rahat bir duruma referansla stres/kaygı ve nötr durumlar arasında ayırım yapan sınıflandırma şemaları kullanılmıştır. Ek olarak, yüz parametrelerinin katılımcının algıladığı stres/kaygı miktarı ile korelasyonunu araştırmak için öz raporlar kullanılarak bir sıralama dönüşümü önerilmiştir. Göz aktivitesi, ağız aktivitesi, kafa hareketleri ve kameraya dayalı kalp aktivitesinden elde edilen belirli yüz ipuçlarının iyi bir doğruluk sağladığı ve stres ile kaygının ayırt edici göstergeleri olarak uygun olduğunu gösterilmiştir (Giannakakis, 2017).

İçsel duygu durumunun gerçek zamanlı izlenmesi için yeni bir metodoloji sunmayı amaçlayan araştırmalar yapılmaktadır. Kullanıcı stresinin, periorbital, supraorbital ve maksiller olan sempatik öneme sahip üç yüz bölgesinde artan kan akışı ile ilişkili olduğu bulunulmuştur. Bu artan kan akışı, termal görüntüleme yoluyla izlenebilen konvektif ısıyı dağıtır. Gerçekleştirilen stres deneyinde, termal görüntüleme yoluyla birçok denek için gerçek zamanlı olarak kan damarı da tespit edilmiştir. Hem termal hem de görsel olarak yüz

özelliğinin tespit edilmesi için yeni bir yöntem olan burun deliği maskesi uygulanarak, kişinin yüz özelliği olan burun bölgesinin termal ve görsel olarak bulunmasını sağlayan yeni bir yöntem sunulmuştur.

Son zamanlarda, yakın kızılötesi (YKÖ) görüntüler, elde edilen görüntülerdeki aydınlatma değişikliklerine karşı yüksek sağlamlığı nedeniyle birçok yüz tanıma sisteminde kullanılmıştır. Bu kızılötesi alanda bazı araştırmalar yapılmış olsa da YKÖ yöntemlerinden çok termal kızılötesi yöntemlere odaklanmışlardır (Farokhi vd., 2016).

Derin öğrenme, birden çok özellik çıkarma düzeyiyle verilerin temsillerini öğrenmek için birden çok işlem katmanı uygular. Ortaya çıkan bu teknik, DeepFace ve DeepID'nin çığır açan buluşları tarafından başlatılan 2014'ten bu yana yüz tanıma araştırma ortamını yeniden şekillendirmektedir (Wang ve Deng, 2021). Görsel dikkati başarılı ESA duygu sınıflandırma çerçevesine uçtan uca eğiterek entegre eden yeni bir mimari olan görsel dikkatli duygu ağlarını içermektedir. Görsel dikkati modellemek için, görüntünün bölgeleri üzerindeki dikkat dağılımını oluşturmak için çoklu katmanlar geliştirilir. Ayrıca, görüntünün belirginlik haritası, duygu tahmini için dikkat dağılımını bütünsel olarak iyileştirmek için önsel bilgi ve düzenleyici olarak kullanılmıştır (Song vd., 2018).

Yüz ifadelerinden yedi temel duyguyu (mutlu, üzgün, kızgın, korkmuş, şaşırılmış, iğrenmiş ve tarafsız) tanıyabilen birkaç model oluşturulmuştur. FER2013 etiketli vesikalık görüntü veri setini kullanarak, DVM kullanarak %45,95 ve ESA kullanarak %66,67 test doğruluğu elde edilmiştir; CK+ veri setinde %98,4 doğruluk elde edilmiştir. Daha sonra, statik görüntülerde öğrenilen becerileri, bir video beslemesinden sürekli olarak yüzleri algılayan ve bireyin baskın duygusunu sınıflandıran gerçek zamanlı bir duygu tanıma sistemine aktarılmıştır (Quinn vd., 2017).

Bu çalışmada sadece yüz resimlerine bakarak ne kadar başarılı stresin ölçülebildiğini görmek amaçlanmaktadır. Alınan yüz görüntüleri üzerinden analizler yapılarak çalışanın ne kadar stresli olduğu tespit edilmek istenmektedir. FER2013 veri seti ile 7 farklı duygu tespit için derin öğrenme algoritma mimarisi kullanılmaktadır. Bu araştırmanın amacı, sadece yüz resimlerine bakarak derin öğrenme tekniklerinden başarı oranı yüksek seviyede olan özelleştirilmiş Evrişimli Sinir Ağları (ESA) modeli kullanılarak stresin ne kadar başarılı ölçülebildiğini görmekle birlikte, alınan yüz görüntüleri üzerinden analizler yapılarak çalışanın ne kadar stresli olduğu tespit edilmek istenmektedir.

MATERYAL VE METOT

2.1. Yapay Zekâ

Yapay zekâ; bilgisayarların veya bilgisayar kontrollü makinelerin, genellikle bir özellik olarak kabul edilen insanlara özgü, önceki deneyimlerden etkilenme, muhakeme, genelleme ve genellikle bilgi edinimi de dahil olmak üzere bilişsel işlevlerle ilgili faaliyetlerde bulunma yeteneği olarak tanımlanır (Nabiyev, 2012).

Yapay zekâ, akıllı programlara yönelik bir bilimdir. Bu programlar şunları yapabilmelidir (Copeland, 1993):

- İnsan düşüncesini taklit ederek karmaşık sorunları çözme becerisi,
- Yorumlarını açıklayabilme, yani belirli bir durumda karşınızdaki kişiye yanıt verebilme,
- Öğrenerek uzmanlığınızı geliştiren ve eski bilgileri yeni bilgilerle uyum içinde kullanarak bilgi tabanını büyütme.

2.2. Makine Öğrenmesi

Makine öğrenimi, mantıksal veya ikili işlemlere dayalı otomatikleştirilmiş hesaplama prosedürlerini kapsayan ve bir dizi örnek aracılığıyla görevlerde yeterlilik kazanan bir sorgulama alanıdır (Donald vd., 2009). Bu bağlamda, makine öğrenimi, yapay zekâ ile ilgili görevleri gerçekleştirmek için sistemleri değiştirmekle ilgilidir (Nilsson, 2017) ve yapay zekâ alanında bir alt bölüm oluşturmaktadır. Bilgi keşfinde, bilgi keşfi sürecindeki adımlardan biri olan çıkarım algoritmalarını uygulamak için makine öğrenimi sıklıkla kullanılmaktadır (Kohavi ve Provost, 1998).

Çıkarımsal makine öğrenimi algoritmaları, etiketli verilerden, yani bilinen çıktılara sahip verilerden kalıpları öğrenebilir (Duda vd., 2001). Makine öğrenimi alanında, tüm gerçek dünya sorunları için eğitim verilerinden kusursuz bilgi ihsan edebilen tek bir öğrenme algoritması yoktur. Bu nedenle, algoritmalar deneysel olarak seçilir (Alpaydın, 2014). Aynı zamanda sınıflandırıcının eğitim verilerine dayalı olarak oluşturduğu model, eğitim verilerine göre değişmektedir ve en iyi modelin bir seçimi görüntülenir. Optimal bir öğrenme algoritması yoktur ve her algoritma bazı verilerde çok başarılı olurken bazılarında başarısız olabilmektedir. Oluşan bu teoreme “no free lunch” teoremi denilir (Wolpert, 1996).

2.3. Görüntü İşleme

Görüntü işleme, ölçülen veya kaydedilen elektronik (sayısal) görüntü verilerinin elektronik ortamda (bilgisayarlar ve yazılımlar yardımıyla) amaçlar doğrultusunda değiştirilmesine yönelik hesaplamalı çalışmadır.

Görüntü işleme, sinyal işlemeden farklı bir işlemdir. Sinyal işleme, verilerin başka bir cihaz tarafından okunabilir bir formata dönüştürülmesini, verilerin bir elektronik ortamdan diğerine aktarılmadan önce yakalanmasını, ölçülmesini ve değerlendirilmesini içerir.

Görüntü düzenleme genellikle önceden çekilmiş görüntüleri düzenlemek için kullanılır. Bu araçlar, önceden var olan görüntüleri ve grafikleri değiştirmek veya daha belirgin hale yükseltmek konusunda uzmandır.

Elektronik veri işleme alanı, kırk yılı kapsayan şaşırtıcı derecede hızlı bir evrim geçirdi. Bu gelişme bilgisayar teknolojisinin gelişimine paralel olarak gerçekleşmiştir. Bilgisayarların küçülmesi, bellek kapasitesinin artması ve veri işleme hızının artması nedeniyle görüntü işleme teknolojisinin gelişimi hızlanmaktadır.

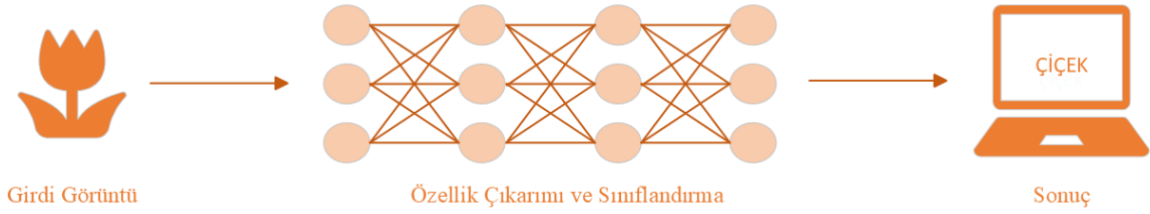
2.4. Derin Öğrenme

Derin öğrenme; segmentasyon, sınıflandırma, konuşma tanıma ve doğal dil işleme gibi birçok araştırma alanında kullanılmaktadır. İnsan benzeri kararlar veren ve kavramsal bir hiyerarşiye dayalı bir öğrenme sürecini takip eden bir makine öğrenimi tekniği olarak tanımlanır (Goodfellow vd., 2016).

Bu yöntemin temel avantajı, bilgisayarın öğrenme amacıyla kullandığı özellikleri/öznitelikleri bağımsız olarak elde edebilmesidir (Patterson ve Gibson, 2017). Klasik yöntemlerde bu işlem bir operatör tarafından gerçekleştirilir ve kullanılan veri setlerinin analiz edilmesi uzun bir süreç olabilir. Fakat derin öğrenme Şekil 2.1'de belirtildiği gibi, otomatik öznitelik/özellik çıkarma yetenekleriyle bu süreci büyük ölçüde hızlandırır. Derin öğrenme mimarileri denetimli ve denetimsiz yöntemlere ayrılabilir. Segmentasyon, sınıflandırma ve nesne tanıma için kullanılır.

Denetimsiz derin öğrenme tekniklerine örnek olarak AE (Otomatik Kodlayıcılar) ve SSAE (Sparse Stacked Auto Encoders) gibi teknikler verilebilir.

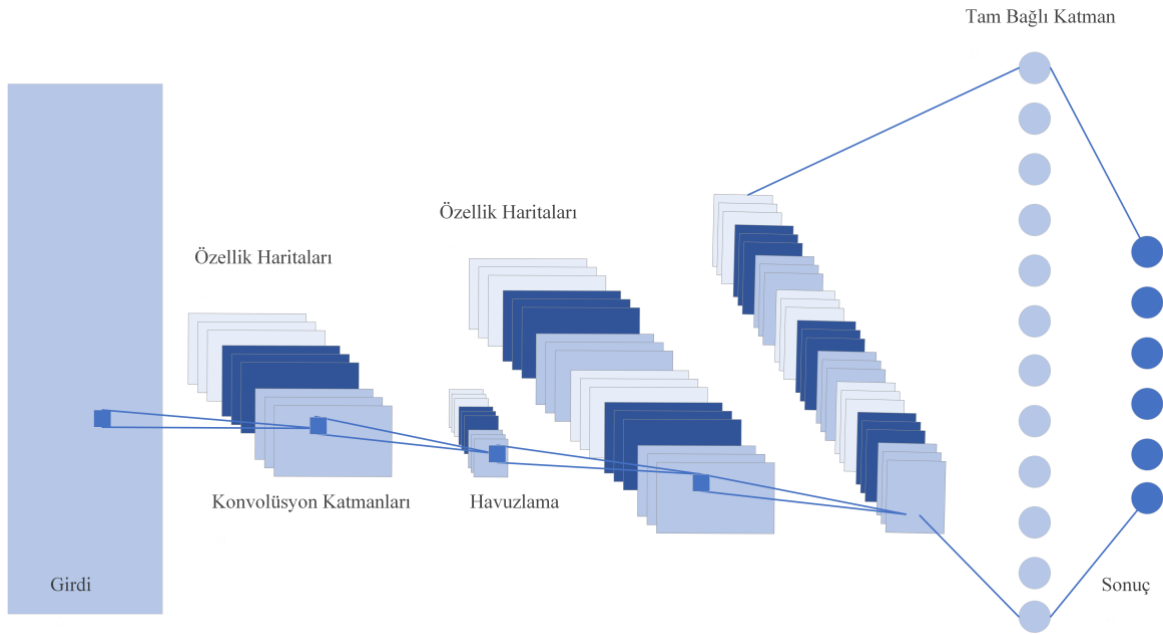
V-Net, U-Net, FractalNet, ResNet ve DenseNet gibi mimariler de denetimli derin öğrenme mimarileri alanına girer. Mimari türüne göre ESA, R-CNN, FCN vb. dayanmaktadır.



Şekil 0.1 : Sınıflandırma ve otomatik özellik çıkarımı

2.5. Evrişimli Sinir Ağı

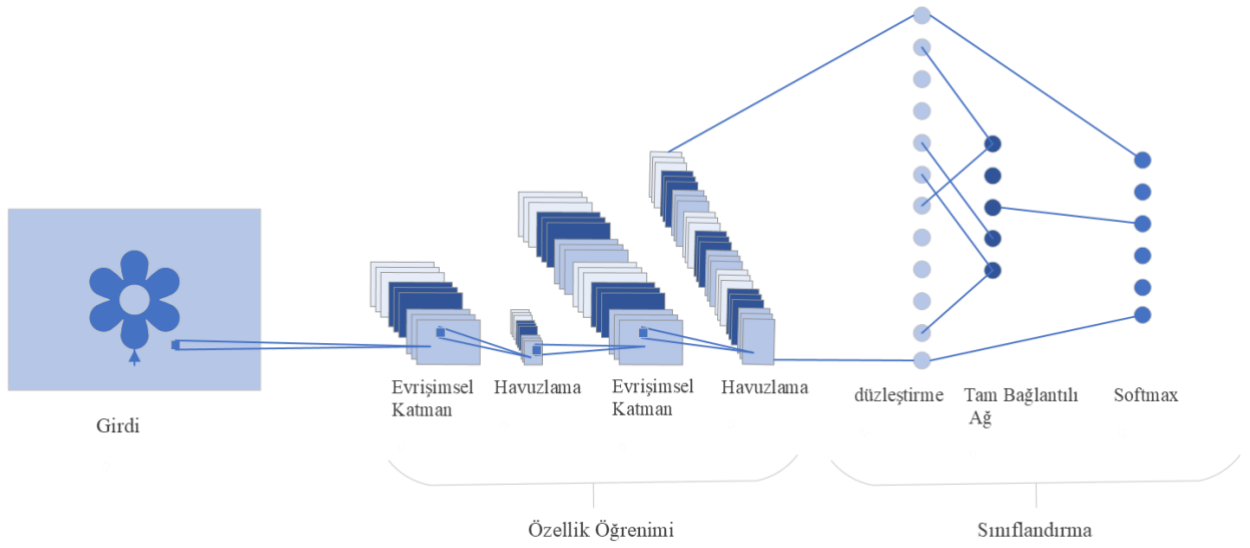
Bir Evrişimli Sinir Ağı (ESA), girdiler, evrişimli katmanlar, aktivasyon fonksiyonları, havuzlama, tamamen bağlı katmanlar ve segmentler gibi temel bileşenlerden oluşur. Şekil 2.2.' de bu bileşenler arasındaki ilişki, bileşen sayısı ve kullanılan işlevsellik mimari seçiminin mimari seçimine değişebilir.



Şekil 0.2 : Evrişimli sinir ağı modeli

Evrişimli katmanlar, derin öğrenmenin anahtarıdır. Hedef sınıfın nesnelere bölünmesi için gereken özellik bilgisi, evrişimli katmanın çekirdek matrisinde bulunur. Bu matristeki değerlere ağırlıklar denir. Bu filtreler, giriş görüntüsündeki piksellere düzenli yatay ve dikey aralıklarla uygulanır ve giriş verilerini farklı şekilde yorumlar. Şekil 2.3 de olduğu gibi evrişimli katmanların sayısı ve filtre matrisinin boyutu, tasarlanan ağın yapısına bağlıdır.

Evrişimli katmanların sayısını artırmak, ağıın öğrenmesi gereken özellikleri çeşitlendirerek girdi görüntüsünden daha ayrıntılı bilgi çıkarmamızı sağlar.



Şekil 0.3 : Derin öğrenme

2.5.1. Aktivasyon Fonksiyonları

Aktivasyon fonksiyonları, sinir ağlarındaki evrişimden kaynaklanan özelliklerin sayısal dağılımını değiştirerek, bu değerlerin doğrusallığını ortadan kaldırır ve değerlerin dağılımının gerçekçi bir şekilde temsil edilmesini sağlar. Verileri istenilen dağılımlara dönüştürmek, sinir ağlarından uyumsuz/aykırı değerleri çıkarmak ve farklı ölçeklerde evrişimli katmanlardan elde edilen değerleri anlamak için kullanılır (Ketkar, 2017).

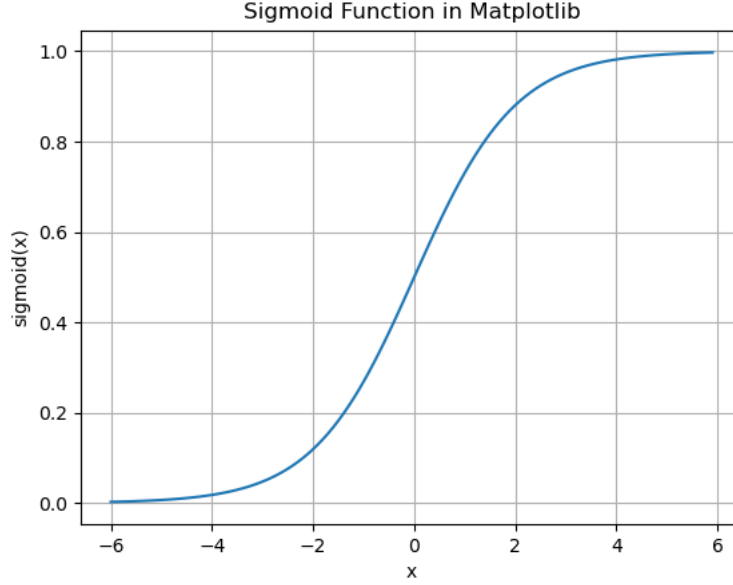
Sigmoid, hiperbolik tanjant, softmax ve ReLU (Rectified Linear Unit), sıklıkla kullanılan baskın aktivasyon fonksiyonlarıdır.

Evrişim işleminin ortaya çıkan değeri; sigmoid fonksiyonu kullanılarak [0-1] arasında ve hiperbolik tanjant fonksiyonu kullanılarak [-1,1] arasında ifade edilebilir. ReLU, giriş özellik haritasındaki negatif değişkenleri sıfıra ve pozitif değişkenleri değişmeden sonraki seviyeye aktarır.

İyi bir çıktı performansı elde etmek için probleme göre doğru fonksiyon seçilmelidir. Yanlış özelliği seçmek tüm sistemi çökertebilir. Çeşitli fonksiyonlar kullanılarak birbirleriyle karşılaştırılabilirler ve sisteme uygun fonksiyon seçilir.

2.5.1.1. Sigmoid

Yaygın olarak kullanılan bir etkinleştirme işlevi olan sigmoid, 0 ila 1 aralığında çıktılar üretir. Karakteristik düz eğrileri, giriş değerlerindeki küçük değişikliklere bile karşılık gelen çıkış değerlerinin ayırt edilmesini sağlar. Türevlenebilir bir fonksiyon olan sigmoid fonksiyon şekil 2.4 de grafikte gösterilmiş olup, yinelemeli öğrenmeye uygundur.



Şekil 0.4 : Sigmoid aktivasyon fonksiyonu

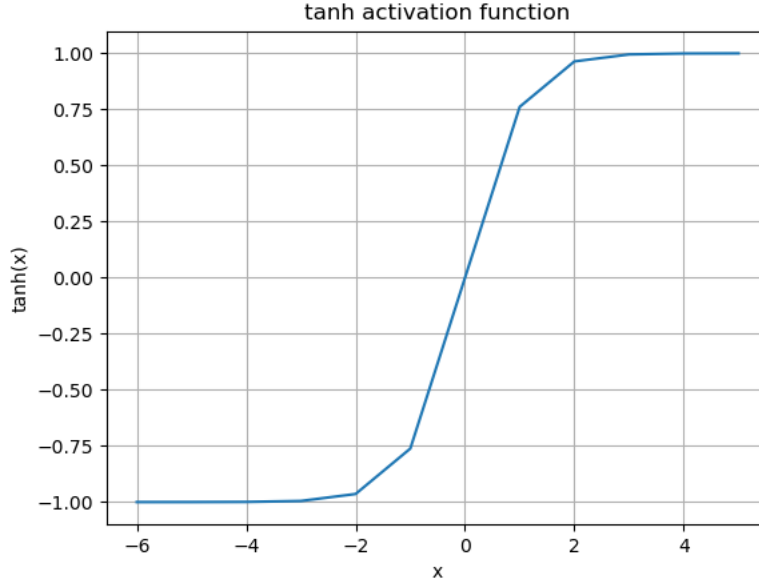
$$Sigmoid(x) = \frac{1}{1 + e^{-x}} \quad (1)$$

Sigmoid fonksiyonunun dezavantajları bulunmaktadır. En önemli dezavantajı "kaybolan gradyan problemi" dir. Bu problem, yapay sinir ağlarında veri eğitilirken ortaya çıkar. Grafikte görüldüğü gibi x değerleri giderken ekstremelerde değerler 0'a yakınsıyor.

2.5.1.2. Tanh

Hiperbolik tanjant fonksiyonunun sürekli olması, çok yönlülüğünü sergileyerek her x değeri için bir çıktı üretmesine olanak tanır. -1 ila +1 arasında değişen çıkışlarla, bu aktivasyon fonksiyonu sigmoid fonksiyonuna benzerlik gösterir ve özellikle ikili sınıflandırma problemlerinde fayda sağlar. Şekil 2.5 de Tanh aktivasyon fonksiyonu görülmektedir. Özellikle, daha geniş bir değer yelpazesini barındırma yeteneği, sinir

ağlarının daha hızlı ve daha verimli bir şekilde eğitilmesine olanak tanır.

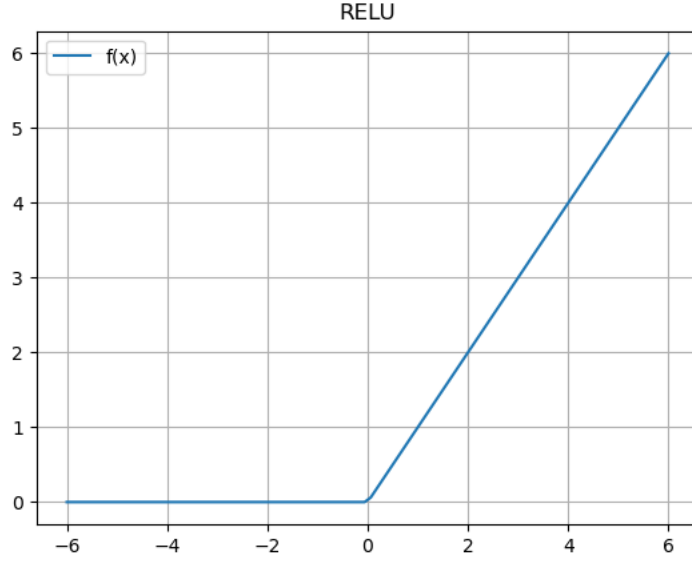


Şekil 0.5 : Tanh aktivasyon fonksiyonu

$$f(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})} \quad (2)$$

2.5.1.3. ReLU (Doğrultulmuş Doğrular Birimler)

Son zamanlarda ReLU, muadili olan sigmoide göre hatırı sayılır bir beğeni topladı. Doğrudan yaklaşımı ile Doğrultulmuş Doğrusal Birim, şekil 2.6'daki grafikte sıfırdan sonsuza kadar değişen çıkış değerlerinin sınırları içinde bulunur.



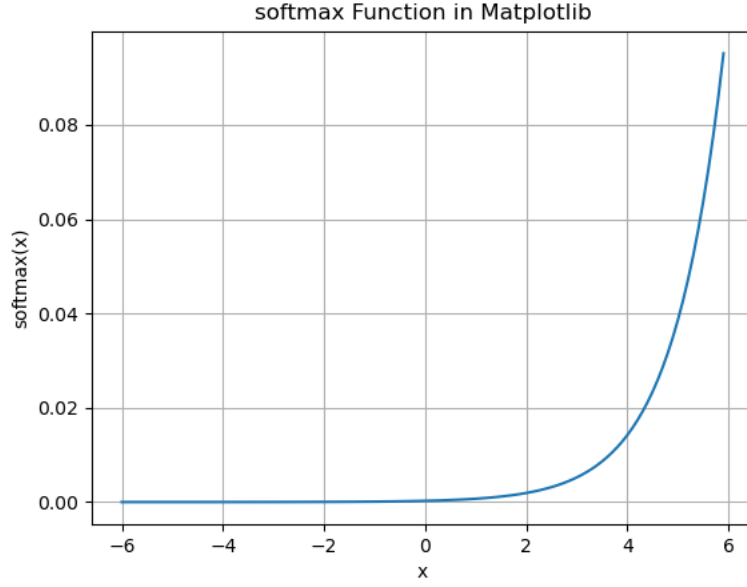
Şekil 0.6 : ReLU aktivasyon fonksiyonu

$$F(x) = \text{MAX}(X, 0) \quad (3)$$

X fonksiyonunun çıktısı sıfırdan büyükse ($X > 0$) sonuç X olur, X işlevi sıfırdan küçük veya sıfıra eşitse, sonuç 0'dır. ReLU aktivasyon fonksiyonu sadece pozitif değerler için aktiftir, sınır ağının gizli katmanlarının negatif çıkışları aktif değildir. Bu nedenle, tüm nöronları aynı anda aktif olacak şekilde atamak mümkün değildir. Bu performans için iyidir. Sonuç olarak, sistem daha hızlı ve daha verimli çalışır.

2.5.1.4. Softmax

Softmax fonksiyonlarının kullanımı, sınıflandırma ikilemleri alanında yaygın bir uygulamadır. Softmax, sistemdeki her bir girdinin kategorizasyonunu gösteren bir çıktı üretir. Bu çıkışlar, 0 ile 1 arasındaki değerler aralığında yer alır. Sigmoid fonksiyon benzeri yapısı ile sigmoid fonksiyon arasındaki tek fark, elde edilen çıktılarının toplamının bire normalleştirilmesidir. Girdinin tipine bakılmaksızın ister pozitif, ister negatif veya sıfır olsun, softmax girişi 0 ile 1 arasında etkin bir şekilde dönüştürür. Olasılıkların ortaya çıktığı yer burasıdır. Şekil 2.7'deki grafikte Softmax aktivasyon fonksiyonu çıkış değerleri görülmektedir. Giriş değeri küçük veya negatifse, softmax işlevi bu değeri küçük bir olasılık olarak atar. Giriş büyükse, yüksek olasılık olarak atılır. Ancak bu olasılıklar her zaman 0 ile 1 arasındadır.



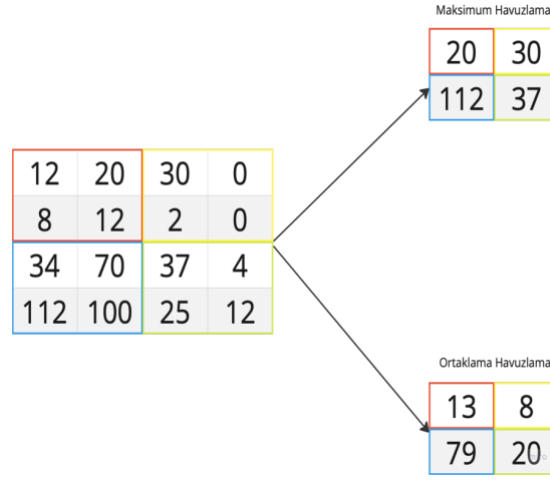
Şekil 0.7 : Softmax aktivasyon fonksiyonu

Sınıflandırma problemleri alanında, sigmoid ve softmax fonksiyonlarını kullanmak yaygındır. Sigmoid, özellikle istenen sonucun ikili olduğu durumlar için uygundur, softmax işlevi ise çok katmanlı sınıflandırma problemlerinin karmaşıklıklarını ele almada faydasını bulur.

2.5.2. Havuzlama

Havuzlama katmanları, girdi verilerinin ayırt edici özelliklerini çıkararak hesaplamayı azaltma amacına hizmet eder, böylece ağırlık hesaplaması karmaşıklığını azaltır (Goodfellow vd., 2016).

Etkinleştirme işlevi tarafından üretilen verileri örnekleme için bir havuzlama katmanı kullanılır. Girdi verilerinin boyutu, bu kayan veri penceresi tarafından küçültülür. Giriş verileri, seçilen havuzlama operatörüne bağlı olarak farklı yaklaşımlar kullanılarak sonraki katmana aktarılır. Şekil 2.8, bir 2x2 pencere üzerinde ustaca uygulanan maksimum ve ortalama havuzlama işlemlerinin özü ifade edilmiştir.



Şekil 0.8 : Maksimum havuzlama ve ortalama havuzlama

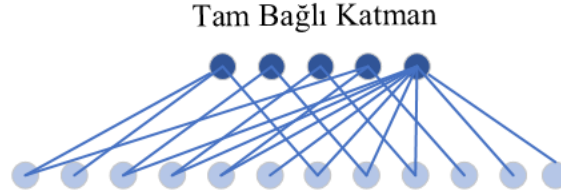
2.5.3. Tam Bağlı Tabaka ve Softmax Tabakası

Tabaka (Layer) kavramı, derin öğrenme ve yapay sinir ağları gibi makine öğrenimi alanlarında sıkça kullanılan bir terimdir. Tam bağlı katman (fully connected layer) ve Softmax katmanı, bu tabakalardan ikisidir.

1. **Tam Bağlı Katman (Fully Connected Layer):** Tam bağlı katman, bir yapay sinir ağının temel yapı taşlarından biridir. Girdi verilerini (input) alır ve her bir giriş nöronunu tüm çıkış nöronlarına bağlayarak tam bir bağlantı sağlar. Her giriş nöronu, ağırlıklar ve bias değerleri ile çıkış nöronlarına katkıda bulunur. Bu işlem, matris çarpımı ve aktivasyon fonksiyonu kullanılarak gerçekleştirilir. Tam bağlı katmanlar, özellik çıkarımı ve sınıflandırma gibi birçok görevde yaygın olarak kullanılır.
2. **Softmax Katmanı:** Softmax katmanı, özellikle çok sınıflı sınıflandırma problemlerinde kullanılan bir çıkış katmanı türüdür. Bu katman, nöronların çıkış değerlerini olasılık dağılımına dönüştürmek için kullanılır. Softmax fonksiyonu, her bir çıkış nöronunun değerini, toplam çıkış değerlerinin etkisiyle normalize eder. Böylece, çıkışların sınıflar arasındaki olasılık tabanlı bir dağılımını elde ederiz. Bu, sınıflandırma problemlerinde hangi sınıfa ait olasılığın daha yüksek olduğunu tahmin etmemize yardımcı olur. Softmax katmanı genellikle çok sınıflı doğrusal olmayan (non-linear) problemlerde kullanılır ve çıkışlar toplamda 1'e eşit olacak şekilde normalize edilir.

Katmanlı sinir ağının son katmanı, evrişim yoluyla çıkarılan düşük, orta ve yüksek düzey özellikleri kullanarak sınıf puanlarını hesaplayan bu çerçeveyi içerir.

Tam bağı katmandan Şekil 2.9 ile oluşan değerleri ait oldukları sınıf ile ilişkilendirmek için bir sınıflandırıcıya ihtiyaç duyulur. Rastgele Ormanlar, Destek Vektör Makineleri ve Softmax gibi sınıflandırma algoritmalarının kullanılması, etiketlerin tahmininde yardımcı olur.



Şekil 0.9 : Tam bağı tabaka

2.5.4. ESA Parametreleri

Evrişimli sinir ağları alanında, eğitilmiş ağın etkinliği, operatör tarafından ayarlanan parametrelerden derinden etkilenir. Bu nedenle, parametre seçimine dikkatli bir şekilde karar verilmelidir. Bu parametreler, öğrenme oranını, dönem sayısını, parti boyutunu, başlangıç ağırlıklarını ve düzenlemeyi kapsar.

2.5.4.1. Öğrenme Oranı

Bu metrik, ağ parametrelerinin geri yayılım düzeltmesinin büyüklüğü ile uyumlu olarak güncellendiği hızı belirtir. Bu metriğin aşırı derecede yüksek bir değeri, eğitim için gereken süreyi azaltır, ancak yetersiz ve istikrarsız sonuçlara yol açabilir. Bu değerinin çok düşük seçilmesi, global minimuma ulaşamama riskiyle ağın eğitim süresini artırabilir.

2.5.4.2. Epok Sayısı

Tüm veri kümesinin ağıdaki eğitime kaç kez katıldığını gösterir. Epok sayısı düşük tutulursa yetersiz uyum oluşabilir. Bu sayı büyükse, belirli bir seviyeden sonra doğruluğun değişmediğini, yani işlem yükünün arttığını biliyoruz.

2.5.4.3. Küme Boyutu

Ağ derinliği ve veri boyutu göz önüne alındığında, eğitim için tüm veri setini aynı anda kullanmak hem donanım hem de maliyet sorunlarını gündeme getirir. Bu nedenle eğitim sırasında veri setini birkaç parçaya ayırmamız gerekiyor. Parçadaki veri sayısı, kümenin boyutunu gösterir. Ayrıca, GPU üzerinde lineer cebir işlemlerinin vektör hesaplamasını

sağlar (Kriesel, 2007).

2.5.4.4. Başlangıç Ağırlıkları

Bir sinir ağındaki nöronlar arasındaki ağırlıkların ilk tahsisi, ileri besleme yoluyla görülen ayarlama seviyesini doğrudan etkilediği için ağ eğitimi sürecinde önemli bir öneme sahiptir. Bu değerler, normal dağılıma uymak için rastgele başlatılma potansiyeline sahiptir (Bengio, 2012).

2.5.4.5. Düzenleştirme (Regularization)

Fazla uydurma/ezberleme (overfitting), makine öğrenimi tekniklerinde en yaygın sorunlardan biridir. Bu eğitim veri seti için yüksek doğruluk, diğer veri türleri için düşük kesinlik olarak tanımlanabilir. Bu durumun temel nedeni, eğitim verilerinde çok sayıda benzer verinin bulunmasıdır. Tek bir baskın türdeki verileri kullanarak bir ağı eğitmek, diğer türleri temsil etmeyen ağırlıklar üretir.

2.6. Optimizasyon Algoritmaları

Ağın çıktı değeri ile girdi değeri arasındaki fark edilebilir ayrım, ağ hatası için bir gösterge görevi görür. Optimizasyonun birincil amacı, hata fonksiyonunun net minimum değerini ortaya çıkarmaktır. Derin öğrenme topluluğu içinde en yaygın optimizasyon yaklaşımları, Gradient Descent ile Adam, RMSProp, AdaGrad ve Adadelta algoritmaları gibi çeşitli yinelemelerini kapsar. Tablo 2.1 de optimizasyon algoritmaları avantajları ve dezavantajları belirtilmiştir.

Tez çalışması kapsamında küçük hafızaya ihtiyaç duyduğu için optimizasyon algoritması olarak Adam kullanılmıştır.

Tablo 0.1 : Optimizasyon algoritmaları

Optimasyon Algoritması	Nedir?	Avantajlar	Dezavantajlar
Gradient Descent	Amaç, rastgele seçilen parametre değerleri ile global bir minimuma ulaşmaktır.	➤ Minimum noktaya kararlı yakınsama yapar,	➤ Yörünge, tüm veri setinin kullanılması nedeniyle

		<ul style="list-style-type: none"> ➤ Vektörleştirme yararları kullanılabilir, ➤ Hesaplama açısından verimlidir. 	<p>öğrenme hızı azalırken, yerel minimuma doğru inebilir.</p>
RMSprop (Root Mean Square Propagation)	<p>Yukarı ve aşağı hareketin kapsamını azaltır. Dolayısıyla öğrenme hızını artırarak yatay yönde ilerlemesini hızlandırır ve minimuma daha hızlı ulaşır.</p>	<ul style="list-style-type: none"> ➤ Çok geniş kapsamlı veri setlerinde iyi şekilde çalışır. ➤ Daha büyük öğrenme oranları kullanmak mümkündür. 	
Adam (Adaptive Moment Estimation)	<p>Momentum ve RMSprop algoritmalarını birleştirip kullanılan algoritmadır.</p>	<ul style="list-style-type: none"> ➤ Algoritmanın uygulanması basittir, ➤ Minimum bellek kullanımı ile gürültülü veri kümelerini işlemek için idealdir. ➤ Adam algoritması, hesaplama verimliliği ile öne çıkıyor. 	<ul style="list-style-type: none"> ➤ Moment ve RmsProp algoritmaları, boş bir sayfadan başlama ve hız söz konusu olduğunda dezavantajlar sergiler.
Adagrad	<p>Bu, gradyan tabanlı tekniklerle optimize etmek için tasarlanmış dikkate değer bir algoritmadır. Öğrenme oranı, önceki gözlemler dikkate alınarak parametrelere göre akıllıca ayarlanır. Her parametre, genel performansı etkili bir</p>	<ul style="list-style-type: none"> ➤ Yöntem, öğrenme oranının ayarlanmasını otomatikleştirerek manuel müdahale gerekliliğini ortadan kaldırır. ➤ Ek olarak, ağırlık ölçeklemenin eşit olmadığı senaryolarda 	<ul style="list-style-type: none"> ➤ Öğrenme oranını yüksek oranda küçültebilir.

	şekilde artıran benzersiz öğrenme oranına sahiptir.	doğrudan stokastik gradyan inişine kıyasla daha hızlı ve daha güvenilir yakınsama sunarken adım boyutuna minimum hassasiyet gösterir.
Adadelta	Adagrad'ın bir uzantısıdır, Adagrad'ın öğrenme oranı yüksekliğini azaltmaya çalışır.	➤ Varsayılan öğrenme oranını ayarlamaya gerek yoktur.

2.7. Model Özellikleri

2.7.1. VGG16

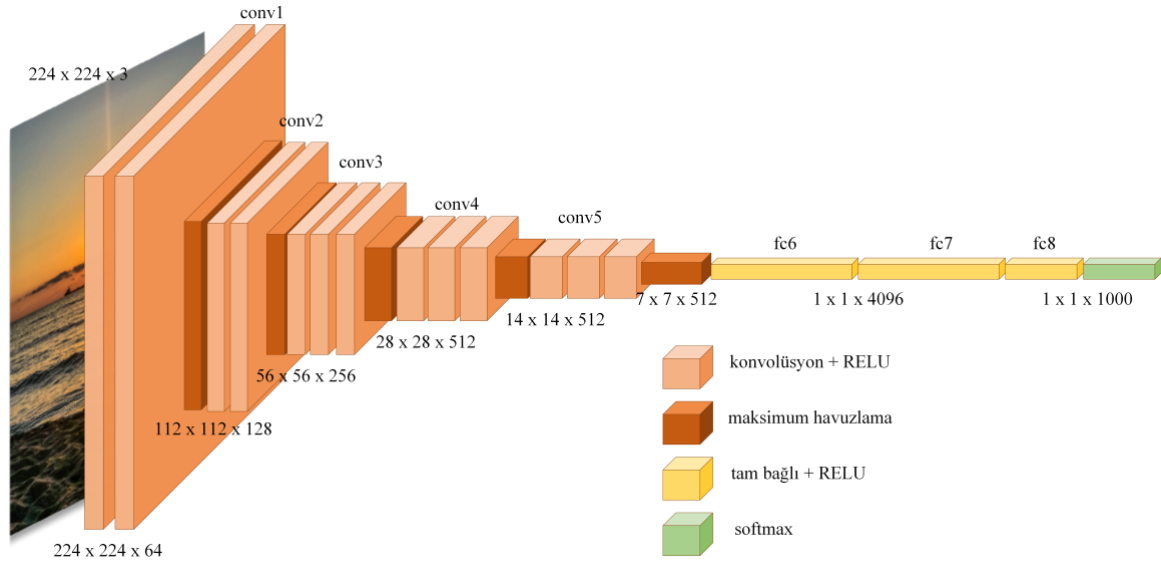
ImageNet Large-Scale Visual Recognition Challenge (ILSVRC), bilgisayarla görme modellerini sergilemek ve meydan okumak için yıllık bir etkinliktir. 2014 ImageNet yarışmasında, Oxford Üniversitesi mühendislik bilimi bölümü görsel geometri grubundan Karen Simonyan ve Andrew Zisserman, modellerini nesne algılama ve sınıflandırmada 1. ve 2. sırada yer alan "Büyük Ölçekli Görüntü Tanıma İçin Çok Derin Kavramlı Ağlar" başlıklı makalede sergilediler.

VGG16, yapay sinir ağları çatısı altına giren, yaygın olarak ConvNet olarak adlandırılan evrişimli bir sinir ağını temsil eder. Böyle bir ağ yapısı, bir girdi katmanı, bir çıktı katmanı ve çok sayıda gizli katman içerir. Bilgisayar görüşündeki mükemmelliği ile tanınan bir Evrişimli Sinir Ağı (ESA) olan VGG16, bu alandaki dikkate değer ilerlemelerin bir kanıtıdır. Mimarları, çeşitli ağ yapılarını titizlikle incelediler ve sonuçta son derece küçük (3×3) evrişim filtrelerinin uygulanması yoluyla derinliği artıran bir tasarımı seçtiler. Bu dahiyane konfigürasyon, önceki tüm teknolojik düzenlemeleri geride bırakan muazzam bir ileri atılımı temsil ediyor. Derinliği 16-19 ağırlık katmanına ittiler ve yaklaşık 138 eğitilebilir parametre yaptı.

VGG16, 1000 farklı kategorideki 1000 görüntüyü %92,7 doğrulukla sınıflandırabilen nesne algılama ve sınıflandırma algoritmasıdır. Görüntü sınıflandırması için popüler algoritmalarından biridir ve transfer öğrenimi ile kullanımı kolaydır.

VGG16 modelindeki sayısal atama "16", her biri kendi önemli ağırlık setiyle doldurulmuş on altı katmanın varlığına karşılık gelir. VGG16'da on üç evrişimli katman, beş Maksimum Havuzlama katmanı ve 21 katmana kadar toplamı olan üç Yoğun katman vardır, ancak yalnızca on altı ağırlık katmanına, yani öğrenilebilir parametreler katmanına sahiptir. VGG16 katman giriş modeli şekil 2.10'da gösterilmiştir.

VGG16, giriş tensör boyutunu 3 RGB kanallı 224x224 olarak alır.



Şekil 0.10 : VGG16

VGG16, çok sayıda hiperparametreden kaçınarak bunun yerine 3x3 filtrelerle süslenmiş evrişim katmanlarını ve aynı dolgunun tutarlı bir şekilde kullanılmasını tercih eden olağanüstü özelliği ile öne çıkıyor. Ayrıca, maksimum havuz katmanı için sürekli olarak 2x2 filtre adımı uygular.

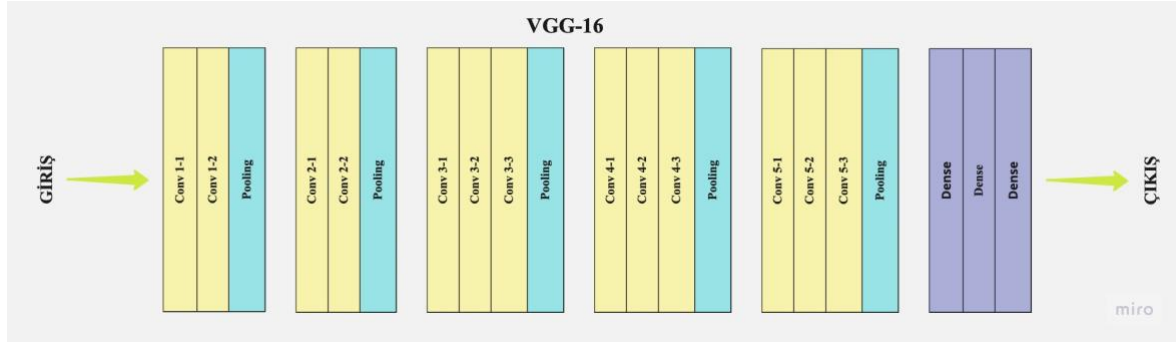
Evrişim ve maksimum havuz katmanlarının düzeni, mimarinin tamamında tutarlı kalır.

Conv-1 Katmanı 64 filtreye, Conv-2'de 128 filtreye, Conv-3'te 256 filtreye, Conv 4 ve Conv 5'te 512 filtre var. Şekil 2.11 de VGG16 modeline ait katman yapısı yer almaktadır.

Bir dizi evrişimli katmanın ardından, üç Tam Bağlantılı (FC) katman ortaya çıkar. İlk iki

katmanın her biri etkileyici 4096 kanala sahiptir, üçüncü katman ise her sınıf için bir tane olmak üzere 1000 kanalı kapsayan 1000 yollu ILSVRC sınıflandırmasını yürütme gibi önemli bir görevi üstlenir. Son olarak, sonuç katmanı soft-max katmanı şeklini alır.

VGG-16'nın eğitilmesi çok yavaştır. VGG-16 eğitilmiş imageNet ağırlıklarının boyutu 528 MB'dir. Bu nedenle, onu verimsiz kılan oldukça fazla disk alanı ve bant genişliği gerektirir. (URL-1, 2023)



Şekil 0.11 : VGG16 katman yapısı

2.7.2. VGG19

VGG19, bilgisayar görüşü alanında sık kullanılan bir derin öğrenme modelidir. Oxford Üniversitesi'nden Visual Geometry Group (VGG) tarafından geliştirilmiştir. "VGG" adı, grubun adından gelir.

VGG19, toplam 19 katmandan oluşan derin bir evrişimli sinir ağıdır. Bu 19 katmandan 16 tanesi evrişimli katmanlardan (convolutional layer) ve 3 tanesi tam bağlı katmanlardan (fully connected layer) oluşur.

VGG19, evrişimli sinir ağlarının önemli bir kilometre taşı olarak kabul edilir ve bilgisayar görüşü alanında başarılarına sahiptir. Ağın temel mimarisi, küçük boyutlu (3x3) filtrelerin ardışık olarak kullanılmasıyla oluşturulan derin bir yapıya sahiptir. Bu yapı sayesinde, VGG19 daha iyi özellik çıkarımı yeteneklerine sahip olmuş ve birçok görsel tanıma ve sınıflandırma görevinde üstün performans sergilemiştir.

Ancak, VGG19'nun daha yeni ve daha verimli modellerin (örneğin, ResNet, Inception, ve EfficientNet gibi) ortaya çıkmasıyla beraber popüleritesi zamanla azalmıştır. Fakat hala eğitim için bir başlangıç noktası olarak veya bazı özel durumlarda kullanılabilecek güçlü bir derin öğrenme modeli olarak kabul edilmektedir.

2.7.3. ResNet

ResNet (Residual Neural Network), Microsoft Research tarafından 2015 yılında Kaiming He ve ekibi tarafından geliştirilen bir derin öğrenme modelidir. ResNet, bilgisayar görüşü alanında ve diğer pek çok görevde büyük başarı elde etmiş ve derin öğrenme modellerinin gelişiminde önemli bir dönüm noktası olarak kabul edilmiştir.

ResNet'in temel fikri, derin ağların daha kolay eğitilmesine yardımcı olacak şekilde tasarlanmış "artıklık bağlantıları" (residual connections) kullanmaktır. Geleneksel derin ağ yapılarında, her katmandaki çıkış, bir sonraki katmana aktarılmaktadır. Ancak, bu durum derin ağların derinleştikçe "degradasyon" sorununa yol açar. Yani, ağın daha fazla katman eklemesi performansın kötüleşmesine neden olabilir.

ResNet, bu degradasyon sorununu artıklık bağlantılarıyla aşmayı amaçlar. Artıklık bağlantıları, katmanlardan geçen çıktıyı doğrudan sonraki katmanın çıktısı ile birleştirir. Böylece, katmanlar arasındaki hata azalır ve ağın daha kolay ve etkili bir şekilde eğitilebilmesi sağlanır. Bu yaklaşım sayesinde ResNet, 152 ve daha fazla katman içeren derin modellerin bile eğitilebilmesini mümkün kılmıştır.

ResNet'in en popüler versiyonları ResNet-50, ResNet-101, ve ResNet-152 gibi derinliklere sahip olanlarıdır. Bu modeller, çeşitli bilgisayar görüşü görevlerinde (nesne tanıma, nesne dedeleme, yüz tanıma gibi) ve diğer görevlerde state-of-the-art sonuçlar vermiştir. Ayrıca, ResNet'in farklı alanlarda transfer öğrenme için temel bir mimari olarak da kullanılması yaygındır.

ResNet'in başarıları şu şekildedir:

- ImageNet veri kümesi alanında, ünlü ILSVRC 2015 sınıflandırma görevinde yalnızca %3,57'lik bir hata oranıyla en üst sırayı alarak başarının zirvesine ulaştı. Bu tür bir zafer, kalan ağların titiz bir şekilde incelenmesiyle elde edildi ve hepsi övgüye değer bir karmaşıklık düzeyini korurken etkileyici 152 katmanla VGG ağlarının derinliğini aştı.
- Faster R-CNN çerçevesinde VGG-16 katmanları ResNet-101 ile değiştirilerek, COCO nesne algılama veri setinin performansında %28'lik kayda değer bir artış gözlemlendi.

Görüntü sınıflandırma alanında, derin evrişimli sinir ağları bir dizi dikkate değer ilerlemeye

öncülük etti. Bu ağlar, görüntü tanıma ve sınıflandırma gibi görevlerde olağanüstü yetenekler sergileyerek teknolojik ilerlemenin ön saflarında yer alan sonuçlar verdi. Sonuç olarak, araştırmacılar, giderek karmaşıklaşan zorlukların üstesinden gelmek için ek katmanlar dahil ederek derin öğrenme mimarilerinin derinliğini sürekli olarak genişletti. Bu uyumlu çaba, sınıflandırma ve tanıma çabalarının etkinliğini önemli ölçüde artırdı ve süreçteki dayanıklılıklarını güçlendirdi (URL-2, 2023).

Ünlü VGG-19 çerçevesinden ilham alan ResNet modeli, VGG muadillerine kıyasla azaltılmış filtreler ve gelişmiş basitliğe sahip aerodinamik 34 katmanlı yapısıyla kendisini farklı kılar. Etkinliğini artırmak için bu mimari, sıçrama bağlantılarının veya artık blokların orijinal ağ dokusuna stratejik olarak dahil edilmesi yoluyla ustaca bir artık kafese dönüştürülür.

Model, azaltılmış filtrelere ve VGG ağlarına kıyasla daha basit bir tasarıma sahiptir. 34 katmanlı bir yapıyla yalnızca 3,6 milyar FLOP (çarpma ve toplama işlemleri) kullanır, bu da VGG-19'un etkileyici 19,6 milyar FLOP'unun yalnızca %18'ine tekabül eder (He vd., 2015).

2.7.4. MobilNetV2

MobileNetV2, Google tarafından geliştirilen bir derin öğrenme modelidir. MobileNetV2, bilgisayar görüşü ve diğer görevlerde yüksek performans sağlamak için özellikle mobil ve gömülü cihazlar gibi kaynak kısıtlı ortamlarda kullanılmak üzere tasarlanmış bir yapay sinir ağı mimarisidir.

MobileNetV2, önceki MobileNet modelinin (MobileNetV1) geliştirilmiş bir versiyonudur. MobileNetV1, hesaplama verimliliği ve düşük bellek gereksinimleri sağlamak için 1x1 evrişim katmanları ve depthwise separable convolution (derinlik tabanlı ayrıştırılmış evrişim) yöntemini kullanarak başarılı olmuştur. Ancak MobileNetV2, bu özelliklere ek olarak bazı yeni tekniklerle performansı daha da artırır.

MobileNetV2'nin temel özellikleri:

1. Inverted Residuals: MobileNetV2'de "ters çözüm" (inverted residual) adı verilen bir yapı kullanılır. Bu yapı, düşük boyutlu girişlerdeki (feature map) bilgi kaybını azaltmaya ve yüksek boyutlu girişlerdeki hesaplama maliyetini düşürmeye yardımcı olur.

2. Linear Bottlenecks: MobileNetV2'de doğrusal şişirme (linear bottleneck) kullanılır. Bu, modelin daha fazla parametresiz bir şekilde genişletilmesine olanak tanırken, yine de yüksek doğruluk elde etmeye yardımcı olur.

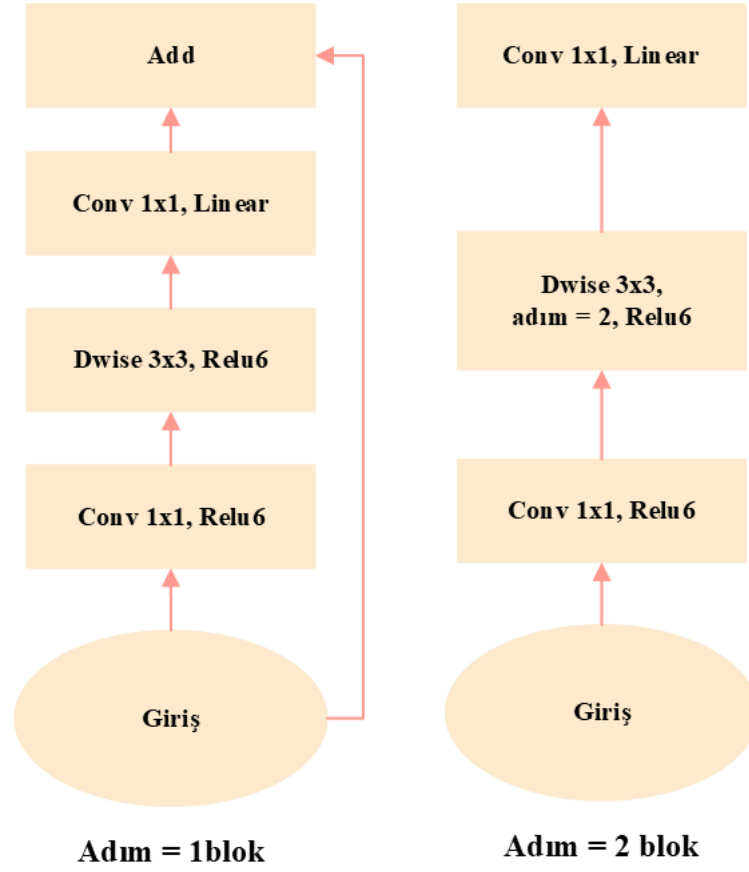
3. Expansion Layer: MobileNetV2'de genişletme (expansion) katmanı kullanılır. Bu katman, 1x1 evrişim katmanlarıyla özellikleri daha yüksek boyutlara genişletir, ardından daha az sayıda özellik haritası üzerinde daha az hesaplama yapılmasını sağlar.

MobileNetV2, görüntü sınıflandırma, nesne tespiti, yüz tanıma gibi birçok görsel tanıma görevinde ve mobil cihazlarda çalışan uygulamalarda tercih edilen bir modeldir. Düşük bellek gereksinimleri ve yüksek hızlı hesaplama yetenekleri sayesinde, kaynak sınırlı cihazlarda etkin bir şekilde kullanılabilir.

Geleneksel evrişim işlemi yerine, evrişim filtreleri kullanan görüntülerden özniteliklerin çıkarılması amacıyla Derin Ayrılabilir Evrişim tekniğinin kullanılması kullanılır (Zhang vd., 2021).

Bu metodolojinin uygulanmasıyla, karmaşık özellik çıkarma işlemi, geleneksel evrişimli prosedürde kullanılanlardan yaklaşık sekiz veya dokuz kat daha az, önemli ölçüde azaltılmış parametre sayısı ile yürütülebilir.

MobileNetV2'de ters çevrilmiş artık yapıya sahip daha iyi bir modüldür. Dar katmanlardaki doğrusal olmayan yapı kaldırılmıştır. Özellik çıkarma için omurga olarak MobileNetV2 ile nesne algılama ve anlamsal segmentasyon için de son teknoloji performanslar elde edilmiştir.



Şekil 0.12 : MobileNetV2 mimari

MobileNetV2'de iki tür blok vardır. Şekil 2.12 de bloklarda görüldüğü gibi, biri 1 adımlı artık bloktur. Bir diğeri, küçülme için 2 adımlı bloktur. Her iki blok türü için de 3 katman vardır. Bu kez ilk katman ReLU6 ile 1×1 evrişimdir. İkinci katman, derinliksel evrişimdir. Üçüncü katman başka bir 1×1 evrişimdir, ancak herhangi bir doğrusal olmayanlık yoktur. ReLU tekrar kullanılırsa, derin ağların yalnızca çıktı alanının sıfır olmayan hacim kısmı üzerinde doğrusal bir sınıflandırıcının gücüne sahip olduğu iddia edilmektedir.

Her darboğaz modülünün çıkışında ReLU6'nın çıkarılmasıyla doğruluk artırılır.

Darboğazlar arasındaki kısayolla, genişletmeler ile artık bağlantısı olmayan kısayoldan daha iyi performans gösterir (URL-3, 2023).

2.8. Stres

Stres kelimesi Latince "estritia" kelimesinden gelir. 17. yüzyılda stres, felaket, bela, ıstırap, keder ve gönül yarası anlamlarında kullanılmıştır. 18. ve 19. yüzyıllarda terim anlam değiştirdi ve kuvvet, baskı ve kuvvet anlamında nesnelere, insanlara, organlara ve zihinsel

yapılara atıfta bulunmak için kullanıldı. Bu nedenle, stres, bu tür kuvvetlerin etkisi nedeniyle nesnelere ve insanların deformasyonuna ve deformasyonuna karşı bir direnç olarak giderek daha fazla kullanılmaktadır.

Stres, vücudun ve zihnin günlük problemlere, zorluklara ve çevreye nasıl tepki verdiğini ifade eder. Durumsal talep ile bireyin yeteneği arasındaki tutarsızlık ve uyumsuzluk nedeniyle oluşur. Kişinin değişen iç ve/veya dış parametrelere ve psikolojik taleplere tepkisini içeren birbirine bağlı ve birbirine bağlı birçok parametreden (kapsamlı, ayırt etme, zihinsel ve fizyolojik) oluşur.

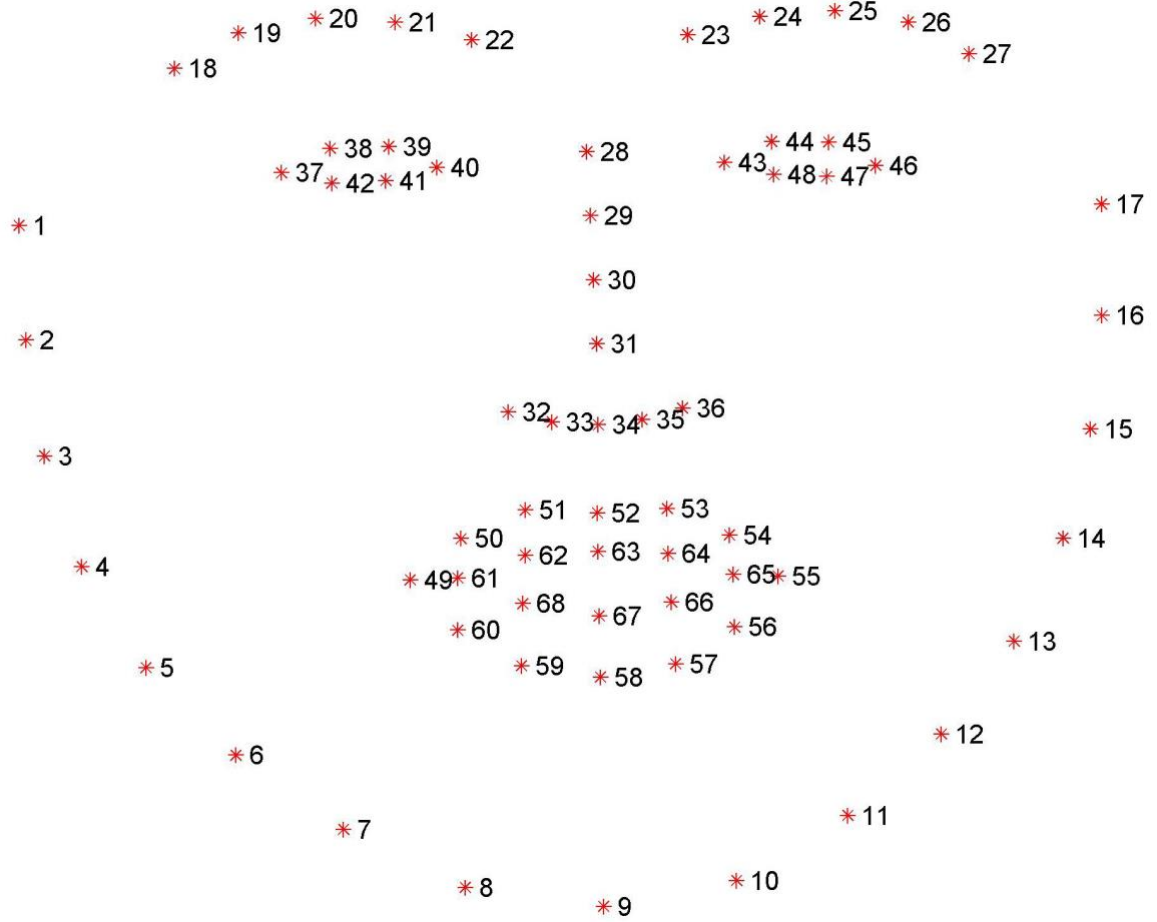
2.9. Dlib Kütüphanesi (68 nokta)

Dlib, C++ programlama dilinde yazılmış evrensel, açık kaynaklı, platformlar arası bir yazılım kütüphanesidir. Tasarımı, sözleşme tasarımından ve bileşen tabanlı yazılım geliştirmeden alınan fikirlerden büyük ölçüde etkilenmiştir. Her şeyden önce, her biri kapsamlı belgelere ve kapsamlı hata ayıklama modlarına sahip bağımsız yazılım bileşenlerinden oluşan bir koleksiyondur.

Davis King, geliştirilmeye başladığı 2002 yılından bu yana dlib'in baş yazarıdır. Bu süre zarfında dlib çok sayıda araçla büyüdü. Spesifik olarak artık ağları, iş parçacıklarını, grafik ara yüzleri, karmaşık veri yapılarını, lineer cebir, istatistiksel makine öğrenimini, görüntü işlemeyi, veri madenciliğini, XML ve metin analizini, sayısal optimizasyonu, Bayes ağlarını ve diğer birçok görevi işlemek için yazılım bileşenleri içerir. Son yıllarda, çeşitli gelişmeler çok çeşitli istatistiksel makine öğrenimi araçları oluşturmaya odaklandı. Bununla birlikte, Dlib genel amaçlı bir kitaplıktır ve her alanda yararlı olan yüksek kaliteli yazılım bileşenlerinden gelen katkıları memnuniyetle karşılar.

Dlib'in geliştirme felsefesinin temelinde, taşınabilirlik ve kullanım kolaylığı taahhüdü yer alır. Bu nedenle, dlib'deki tüm kodlar, kullanıcının herhangi bir şey yapılandırması veya yüklemesi gerekmeden mümkün olduğunca taşınabilir olacak şekilde tasarlanmıştır. Bunu başarmak için, platforma özgü tüm kodlar API sarmalayıcıları ile sınırlandırılmıştır. Geri kalan her şey ya bu paketleyicilerin üzerine katmanlaşmıştır ya da saf ISO standardı C++ ile yazılmıştır.

Şekilde bir Dlib'in 68 puanlık modeline bir örnektir. Orada 1'den 68'e kadar olan noktaları görülebilmektedir (URL-4, 2023).

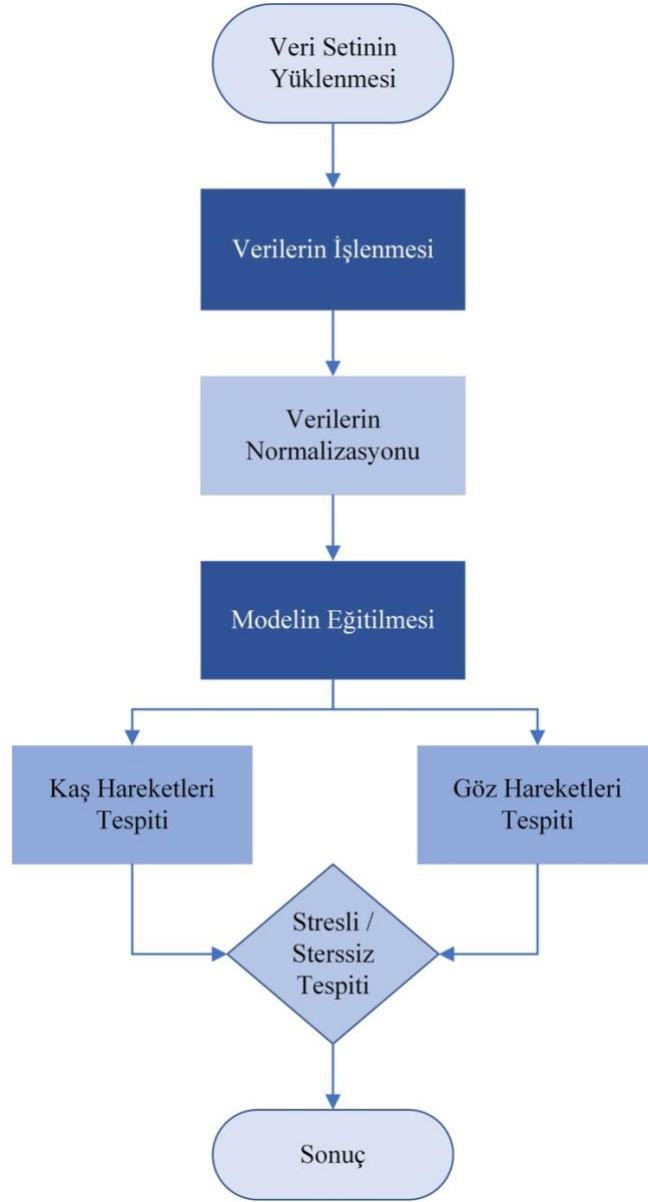


Şekil 0.13 : Dlib'in 68 modeli ile yüz dönüm noktası tespiti

2.10. Çözüm Akışı

Uygulama yöntemi olarak derin öğrenme algoritmaları kullanılmaktadır. Üzerinde çalışma yapılacak veri internet üzerindeki resimler, açık kaynak kütüphanelerindeki veri setleri kullanılmaktadır. Çalışma kapsamında en yaygın kullanılan FER2013 veri seti kullanılarak yüz görüntüleri üzerinden keras uygulama modelleri ile 7 farklı duygu (kızgın, iğrenme, korku, mutlu, nötr, üzgün, şaşkın) analizi tespiti sağlanmıştır. FER2013 veri seti ile 7 farklı duygu tespiti ve yüz üzerindeki 68 noktadan göz, ağız, kaş aktiviteleri referans alınarak çalışanların duygu durumları ile stres analizi sonuçları başarılı bir şekilde elde edilmektedir.

Şekil 2.14'teki akış diyagramında ifade edildiği üzere veri seti sisteme yüklenir, veriler ilk olarak bir ön işlemde geçirilir. Ön işlemde geçirilen veriler gürültüden arındırılmak ve standart yapıya gelebilmesi için normalizasyon işlemi uygulanır, veri çeşitli modeller ile eğitilir ve eğitilen veriler işlem sonucunda etiketlenerek stresli ya da stressiz olarak ayrılır.



Şekil 0.14 : Çözüm akışı

Görüntüler üzerinden keras uygulama modelleri üzerindeki VGG16, ResNet50, MobileNetV2, VGG19 modelleri ile işlemler gerçekleştirilmiştir.









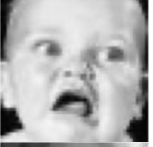
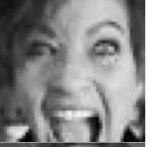






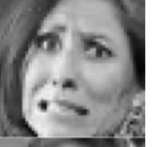




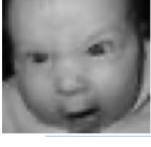

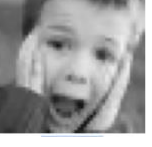




Şekil 2.13 ile ifade edilen yüz üzerindeki 68 farklı nokta referans alınarak kaş, göz ve dudak tespiti yapılmıştır. Kaş hareketi, göz hareketi belli oranlarda ölçülerek stresli ya da stressiz olduğu belirlenmiştir.

2.11. Kullanılan Veri Seti

Bu tez çalışmasında FER2013 veri setinden faydalandım. Kaggle'ın yüz ifadesi analizi yarışması için özel olarak tasarlanan bu yüz ifadeleri derlemesi, 48x48 piksel boyutlarında

35.887 gri tonlamalı etkileyici bir görüntüden oluşuyor. Bunların arasında 28.709 görüntü eğitim amaçlı, 3.589 görüntü doğrulama ve 3.589 görüntü test amaçlı olarak ayrıldı. Tablo 2.2 de spesifik olarak, veri seti 4953 öfkeli görüntü, 547 tiksinti görüntüsü, 5121 korku görüntüsü, 8989 mutlu görüntü, 6077 üzgün görüntü, 4002 şaşkın görüntü ve 6198 nötr görüntüden oluşuyor (Söylemez ve Ergen, 2020).

Tablo 0.2 : FER2013 Veri seti dağılımı

Kızgın	İğrenme	Korku	Mutlu	Üzgün	Şaşkın	Nötr
						
						
						
						
4953	547	5121	8989	6077	4002	6198
% 13.8	% 1.5	% 14.2	% 25.0	% 16.9	% 11.1	% 17.2

2.11.1. Veri Setinin Hazırlanması

Ağı besleyen veri kümesindeki 48x48 görüntüleri yeniden boyutlandırmak için en yakın komşu enterpolasyonunu kullanılmıştır. Bu enterpolasyon, çok az işlem yükü gerektirir ve yeni oluşturulan piksellerin değerini en yakın pikselin değerine ayarlayarak gerçekleştirilmiştir.

Görüntünün hazırlanma sürecinde yeniden boyutlandırma haricinde görüntünün yoğunluk değeri 255'e bölünerek ayarlanır. Bu nedenle giriş görüntüsünün 0 ile 1 arasında değişen yoğunluk değerlerine sahip olması beklenir. Yığındaki piksellerin ortalama yoğunluğu yığındaki piksel değerlerinden çıkarılır ve bu yeni piksel değerleri daha sonra yığının

standart sapmasına bölünür. Sonuç olarak, ortaya çıkan piksel yoğunluğu değerleri -1 ila 1 aralığında bir dağılım sergiler. Mevcut verileri çoğaltmak için veri artırmayı içeren hiçbir tekniğin kullanılmadığına dikkat edilmelidir.

Eğitim ağı için optimizasyon yöntemi “Adam” kullanıldı. Adam, birinci dereceden gradyan tabanlı teknikler kullanarak skolastik amaç fonksiyonlarının optimizasyonu için tasarlanmış devrim niteliğinde bir algoritmadır. Bu algoritmanın uygulanması kolaydır ve hesaplama açısından verimlidir, ancak çok az bellek gerektirir ve büyük miktarda veri veya parametre olarak sınıflandırılabilen problemlerde başarılı olur.

Çalışma, veri kümesi geliştiricisi tarafından sağlanan eğitim ve doğrulama verilerini, herhangi bir dış veriyi dahil etmeden veya veri kümesinden negatif örnekleri ortadan kaldırmadan kullanmıştır.

2.11.2. Kullanılan Modeller

Görüntüler üzerinden keras uygulama modelleri üzerindeki VGG16, ResNet50, MobileNetV2, VGG19 modelleri ile işlemler gerçekleştirilmiştir. Çalışmada kullanılan kriterler tablo 2.3 de belirtilmiştir.

Tablo 0.3 : Çalışmada kullanılan kriterler

Eğitim / Test Verileri (224×224)	3890 / 898
Düzeltilme Fonksiyonu	Adam
Epok Sayısı	5
Batch Size	Varsayılan Değer (32)
Kayıp Fonksiyonu	Sparse categorical crossentropy
Doğruluk Ölçütü	Accuracy

BULGULAR

Görüntüler üzerinden keras uygulama modelleri üzerindeki VGG16, ResNet50, MobileNetV2, VGG19 modelleri ile işlemler gerçekleştirilmiştir.

Tablo 3.1. de belirtildiği gibi, veri seti MobilNetV2 modeli 256 batch boyutu ve 5 epoch değeri ile eğitildiğinde doğruluk ölçütü 0.9092 olmaktadır. Epoch sayısı arttıkça istenilen doğruluk oranına yaklaşılmaktadır.

Tablo 0.1: MobileNetV2 - 256

Model	MobilNetV2
Kayıp Fonksiyonu	Sparse categorical crossentropy
Metrik	Doğruluk (Accuracy)
Batch Boyutu	256
Epochs	5

Dönüm 1/5 15/15 [=] - 1291s 83s/adım - kayıp: 2.4750 - doğruluk: 0.3780 - val_loss: 5.5431 - val_accuracy: 0.1069

Dönüm 2/5 15/15 [=] - 1190s 79s/adım - kayıp: 1.1221 - doğruluk: 0.5947 - val_loss: 22.5800 - val_accuracy: 0.0156

Dönüm 3/5 15/15 [=] - 1185s 79s/adım - kayıp: 0.8553 - doğruluk: 0.6844 - val_loss: 23.0896 - val_accuracy: 0.0156

Dönüm 4/5 15/15 [=] - 1192s 79s/adım - kayıp: 0.5776 - doğruluk: 0.7939 - val_loss: 21.8307 - val_accuracy: 0.0156

Dönüm 5/5 15/15 [=] - 1381s 93s/adım - kayıp: 0.2626 - doğruluk: **0.9092** - val_loss: 21.5248 - val_accuracy: 0.0156

Tablo 3.2 de veri seti MobileNetV2 modeli 32 batch boyutu ve 5 epoch değeri ile eğitildiğinde doğruluk ölçütü **0.7889** olmaktadır. Batch size küçük seçildiğinde verilerin

eđitilmesinde istenilen dođruluk deđerinden uzaklařılmaktadır.

Tablo 0.2: MobileNetV2 - 32

Model	MobilNetV2
Kayıp Fonksiyonu	Sparse categorical crossentropy
Metrik	Dođruluk (Accuracy)
Batch Boyutu	32
Epochs	5

Dönüm 1/5 113/113 [=] - 539s 5s/adım - kayıp: 1.2626 - dođruluk: **0.5359** - val_loss: 4.1199
- val_accuracy: 0.1893

Dönüm 2/5 113/113 [=] - 511s 5s/adım - kayıp: 1.0037 - dođruluk: **0.6315** - val_loss:
17.5854 - val_accuracy: 0.1336

Dönüm 3/5 113/113 [=] - 495s 4s/adım - kayıp: 0.8236 - dođruluk: **0.6953** - val_loss:
10.8244 - val_accuracy: 0.1370

Dönüm 4/5 113/113 [=] - 508s 4s/adım - kayıp: 0.6798 - dođruluk: **0.7610** - val_loss:
11.8678 - val_accuracy: 0.1370

Dönüm 5/5 113/113 [=] - 605s 5s/adım - kayıp: 0.5994 - dođruluk: **0.7889** - val_loss:
12.0401 - val_accuracy: 0.1359

Tablo 3.3. de veri seti VGG16 modeli 32 batch boyutu ve 5 epoch deđerı ile eđitildiđinde dođruluk ölçütü **0.2468** olmaktadır. Kullanılan model veri setini eđitme konusunda istenilen deđerden uzak bir sonuç üretmiştir.

Tablo 0.3: VGG16 - 32

Model	VGG16
Kayıp Fonksiyonu	Sparse categorical crossentropy

Metrik	Doğruluk (Accuracy)
Batch Boyutu	32
Epochs	5

Dönüm 1/5 113/113 [=] - 4853s 43s/adım - kayıp 6.4118 - doğruluk: **0.2067** - val_loss: 2.0012 - val_accuracy: 0.1715

Dönüm 2/5 113/113 [=] - 4662s 41s/adım - kayıp: 1.8444 - doğruluk: **0.2265** - val_loss: 1.8228 - val_accuracy: 0.2472

Dönüm 3/5 113/113 [=] - 4762s 42s/adım - kayıp: 1.8260 - doğruluk: **0.2479** - val_loss: 1.8143 - val_accuracy: 0.2461

Dönüm 4/5 113/113 [=] - 4079s 36s/adım - kayıp: 1.8200 - doğruluk: **0.2384** - val_loss: 1.8123 - val_accuracy: 0.2472

Dönüm 5/5 113/113 [=] - 3945s 35s/adım - kayıp: 1.8225 - doğruluk: **0.2468** - val_loss: 1.8272 - val_accuracy: 0.2461

Tablo 3.4 veri seti RESNET50 modeli 32 batch boyutu ve 5 epoch değeri ile eğitildiğinde doğruluk ölçütü **0.7111** olmaktadır. Kullanılan model veri setini eğitme konusunda istenilen değerden uzak bir sonuç üretmiştir.

Tablo 0.4: Resnet50 - 32

Model	RESNET50
Kayıp Fonksiyonu	Sparse categorical crossentropy
Metrik	Doğruluk (Accuracy)
Batch Boyutu	32
Epochs	5

Dönüm 1/5 113/113 [=] - 1601s 14s/adım - kayıp: 1.8117 - doğruluk: 0.3660 - val_loss:

2.5078 - val_accuracy: 0.2439

Dönüm 2/5 113/113 [=] - 1577s 14s/adım - kayıp: 1.2992 - doğruluk: 0.5092 - val_loss: 2.2324 - val_accuracy: 0.2472

Dönüm 3/5 113/113 [=] - 1774s 16s/adım - kayıp: 1.1357 - doğruluk: 0.5657 - val_loss: 3.4348 - val_accuracy: 0.1158

Dönüm 4/5 113/113 [=] - 2160s 19s/adım - kayıp: 0.9355 - doğruluk: 0.6557 - val_loss: 2.0884 - val_accuracy: 0.1737

Dönüm 5/5 113/113 [=] - 1867s 17s/adım - kayıp: 0.7760 - doğruluk: **0.7111** - val_loss: 2.2446 - val_accuracy: 0.1526

Tablo 3.5 de veri seti VGG19 modeli 32 batch boyutu ve 5 epoch değeri ile eğitildiğinde doğruluk ölçütü **0.2513** olmaktadır. Kullanılan model veri setini eğitirken çok uzun zamanda işlemini tamamlamıştır ve istenilen doğruluktan uzak bir sonuç elde edilmiştir.

Tablo 0.5: Resnet50 - 32

Model	VGG19
Kayıp Fonksiyonu	Sparse categorical crossentropy
Metrik	Doğruluk (Accuracy)
Batch Boyutu	32
Epochs	5

Dönüm 1/5 113/113 [=] - 5234s 46s/adım - kayıp: 113.7592 - doğruluk: 0.2159 - val_loss: 1.9759 - val_accuracy: 0.1737

Dönüm 2/5 113/113 [=] - 10060s 89s/adım - kayıp: 1.8472 - doğruluk: 0.2201 - val_loss: 1.8472 - val_accuracy: 0.2472

Dönüm 3/5 113/113 [=] - 6177s 55s/adım - kayıp: 1.8442 - doğruluk: 0.2276 - val_loss: 1.8179 - val_accuracy: 0.2472

Dönüm 4/5 113/113 [=] - 6738s 60s/adım - kayıp: 1.8269 - doğruluk: 0.2443 - val_loss: 1.8467 - val_accuracy: 0.2472

Dönüm 5/5 113/113 [=] - 6823s 60s/adım - kayıp: 1.8213 - doğruluk: **0.2513** - val_loss: 1.8180 - val_accuracy: 0.2472

Yukarıdaki modellerden farklı olarak tez çalışması için evrimsel katmanları içeren yeni model oluşturuldu ve eğitilen model yüz görüntüleri üzerinde kullanıldı. Oluşturulan modele ait kod parçacığı aşağıdaki gibidir.

İlk olarak giriş verisi ve düzenleme katsayısı alınır. Keras kütüphanesi içindeki düzenleme metodu kullanılarak bir düzenleme değeri belirlenir.

```
regularization = l2(0.07)
img_input = Input(input_shape)
```

Sisteme verilen resim temel katman ile işlemlerden geçirilir. İki boyutlu evrimsel katmana düzenleme değerleri verilir ve girdi verisi işlemi başlatılır. Veri üzerinde işlemler tutarlı yapılabilmesi için normalize edilir. Aktivasyon fonksiyonu olarak ReLU eklenir. İkinci adımda ilk adımda işlem yapılan değerlere tekrar normalizasyon işlemleri yapılır.

```
# ilk iki boyutlu konvolüsyonel katman
x = Conv2D(8, (3, 3), strides=(1, 1), kernel_regularizer=regularization,
use_bias=False)(img_input)
x = BatchNormalization()(x)
x = Activation('relu')(x)

# ikinci iki boyutlu konvolüsyonel katman
x = Conv2D(8, (3, 3), strides=(1, 1), kernel_regularizer=regularization,
use_bias=False)(x)
x = BatchNormalization()(x)
x = Activation('relu')(x)
```

Ön işleminden geçirilen veri ilk adım olan modül bir sürecine verilir. Burada 1x1 konvolüsyon ardından normalizasyon yapılır.

```
# giriş verisine 1x1 konvolüsyonel işlem yapılır
residual = Conv2D(filters, (1, 1), strides=(2, 2), padding='same',
use_bias=False)(input_tensor)
residual = BatchNormalization()(residual)
```

Giriş verisi üzerinde 3x3 konvolüsyon işlemi uygulanır ardından normalizasyon yapılır ve aktivasyon fonksiyonu olarak ReLU sisteme verilir.

```

# Giriş tensörüne 3x3 konvolüsyon yapılır
x = SeparableConv2D(filters, kernel_size, padding='same',
kernel_regularizer=regularization, use_bias=False)(input_tensor)
x = BatchNormalization()(x)
x = Activation('relu')(x)
x = SeparableConv2D(filters, kernel_size, padding='same',
kernel_regularizer=regularization, use_bias=False)(x)
x = BatchNormalization()(x)

```

Azami havuzlama yapılır ve kalan tensör sisteme eklenir.

```

x = MaxPooling2D((3, 3), strides=(2, 2), padding='same')(x)
x = add([x, residual])

```

Konvolüsyon katmanına giriş değeri 8'den 256 ya kadar artırılarak uygulanır. Oluşturulan modül 5 defa artarda olacak şekilde işleme alınır.

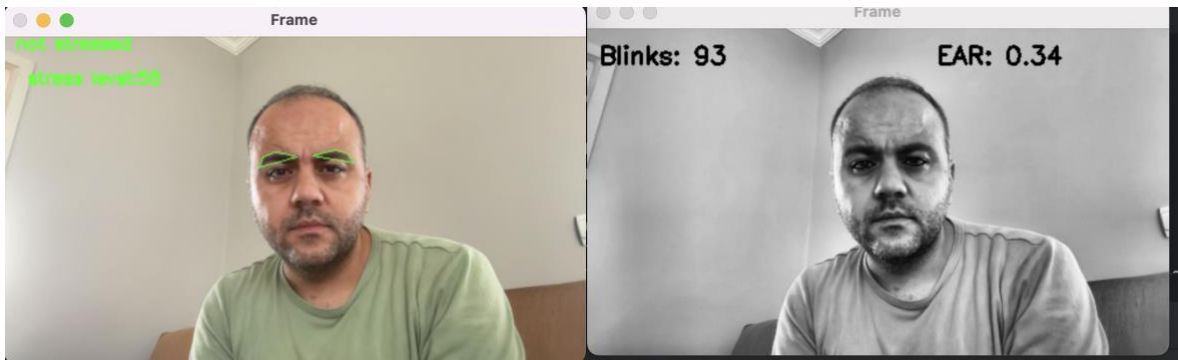
5 modül üzerinden geçen giriş verisi son adım olarak softmax aktivasyon fonksiyonundan geçirilir.

```

x = GlobalAveragePooling2D()(x)
output = Activation('softmax', name='predictions')(x)

```

Eğitilen model yüz görüntülerini ayırt etmede kullanıldı, bilgisayar kamerası ile gerçek zamanlı görüntüler üzerinde kaş ve göz hareketleri tespit edildi. Stres düzeyi, kaşların kasılması ve ortalama konumundan yer değiştirmesi yardımıyla hesaplanır. Sol ve sağ kaş arasındaki mesafe hesaplanır ve ardından üstel fonksiyon kullanılarak stres seviyesi hesaplanır ve 1 ile 100 arasında normalleştirilir. Belirlenen eşik değerinin üstündeki görüntüler stresli olarak etiketlenir.



Şekil 0.1 : Uygulama test görüntüsü

TARTIŞMA

Çalışmamızın bu bölümünde stres analizi altyapısını oluşturma sürecindeki bazı zorluklar ve bunlara bağlı olarak geliştirme yapılabilecek eksikliklerden bahsedilmektedir.

4.1. Stres Değerlendirmesindeki Zorluklar

Bu çalışmanın birincil amacı, belirli bir Evrişimli Sinir Ağları (ESA) modeli kullanarak yüksek riskli iş ortamlarında çalışanlar arasındaki stres seviyelerini değerlendirmenin karmaşıklığını çözümlenerek stres tespit sonuçlarını araştırmaktır. Sonuçlar, stresin karmaşık ve sübjektif karakterinden dolayı sayısal ölçümler kullanarak stresi doğru bir şekilde ölçmeyle ilgili zorlukları vurgulamaktadır. Bireylerin stres düzeyleri, farklı çevresel faktörler nedeniyle önemli ölçüde değişkenlik göstermektedir ve bir dizi psikolojik sonuca yol açabilmektedir. Artan stres düzeylerinin yaygınlığı, iş gücündeki bireylerin bütünsel refahı üzerinde önemli bir etki yaratan önemli bir sorun haline gelmektedir. Bu, bilişsel yetileri, iş doyumunu ve çalışanların genel yeterliliğini içermektedir. ESA modeli kullanılarak yüz görüntüleme teknolojisinin etkili bir şekilde uygulanması, gerçek zamanlı stres değerlendirmesinin olanaklarını gözler önüne sermiştir. Çalışma, yüzdeki 68 referans bölgesi üzerinden özellikle göz, dudak ve kaş hareketlerine odaklanmayı hedefleyip yüz ifadelerini inceleyerek duygusal durumları ve stres derecelerini başarılı bir şekilde tanımladı. Yukarıda belirtilen metodoloji, çalışanların stres deneyimlerine ilişkin önemli bilgiler sunarak, yüksek riskli iş ortamları için özelleştirilmiş stres değerlendirme tekniklerinin önemini vurgulamaktadır ve bu sistemin hali hazırda uygulama ortamı üzerinde de test edilerek sonuçlarının analiz edilmesi kullanılmakta olan modellerin sistemin sonuçlarını belirlemedeki doğruluklarının tespitini bizlere sağlayabilecektir.

4.2. Stres Analizinde Yüz Görüntülemenin Etkinliği

Stres seviyelerinin değerlendirilmesi için yüz görüntüleme teknolojisinin kullanılması, potansiyel olarak etkili ve müdahaleci olmayan bir yöntem olarak görülmektedir. Evrişimli sinir ağları modelinin yüz ifadelerine dayalı stres seviyelerini ayırt etmedeki doğruluğu, stresi değerlendirmek için bir araç olarak potansiyel etkinliğini göstermektedir. Bununla birlikte, yüz görüntülemenin stresle ilişkili incelikleri kapsamlı bir şekilde yakalayamayacağını kabul etmek çok önemlidir. Bu nedenle, stres reaksiyonlarının daha kapsamlı bir şekilde anlaşılmasını sağlamak için kalp atış hızı veya bireylerin böbrek üstü

bezlerinde doğal olarak üretilen ve vücudun strese verdiği tepkiyi düzenleyen kortizol isimli steroid hormonu seviyeleri gibi ek fizyolojik belirteçlerin dahil edilmesini gelecekteki araştırmaların derinlemesine incelemeleri önerilmektedir. Ayrıca, bu çalışmada, örneklerinin çeşitliliği açısından belirli sınırlamaları olabilecek FER2013 veri setinden yararlanıldığına dikkat edilmelidir. Sonuçların uygulanabilirliğini artırmak için, gelecekteki çalışmaların, bir dizi tehlikeli iş ortamını ve işçi demografisini kapsayan daha büyük ve daha çeşitli veri kümeleri elde etmeye çalışması önerilir.

4.3. Çok Yönlü Stres Değerlendirmesinin Geliştirilmesi

Potansiyel bir araştırma alanı olarak, yüz görüntüleme teknolojilerinin ek fizyolojik endikasyonlarla, yani bir ilacın hangi hastalıklara ve hangi biçimde uygun olacağını veya bir operasyonun hangi koşullarda gerekli olduğunun entegrasyonu, stresi değerlendirmek için kapsamlı bir metodoloji sunmaktadır. Kalp atış hızı izleme, cilt iletkenliği ve diğer biyometrik belirteçlerin kullanımı, çalışanların stres tepkilerini değerlendirmenin kapsamlılığını ve kesinliğini artırma potansiyeline sahip olabilmektedir. Çok yönlü bir yaklaşımın kullanılması, işverenlerin stres seviyelerini gerçek zamanlı olarak aktif bir şekilde izlemeleri ve çalışanların refahını teşvik etmek ve sürdürmek için müdahaleleri hızlı bir şekilde gerçekleştirmeleri için değerli bir araç olarak görülebilir.

4.4. Stres Yönetimi Müdahalelerinin Uygulanması

Mevcut çalışmada yapılan keşiflerin ışığında, gelecekteki araştırma çabalarının, stresin gerçek zamanlı olarak ölçülmesine dayanan stres yönetimi terapilerinin değerlendirilmesini ele alması önerilebilir. Yüz görüntüleme ve fizyolojik belirteçlerin kullanılmasıyla, işletmeler stres yaşayan çalışanları belirleme yeteneğine sahiptir. Bu, stresi azaltmak ve çalışan dayanıklılığını artırmak amacıyla gevşeme teknikleri, farkındalık eğitimi ve esnek çalışma düzenlemeleri gibi özel müdahaleler geliştirmelerini sağlayabilir.

4.5. İş Optimizasyonu için İnsan-Makine İş Birliği

Evrimsel sinir ağı modelleri gibi gelişmiş makine öğrenimi tekniklerinin profesyonel ortamlara entegrasyonu, insanlar ve makineler arasındaki iş birliğini geliştirme, stresi azaltma ve iş verimliliğini artırma potansiyeline sahip olabilmektedir. Çalışan stres düzeylerini izlemek ve uygun iş pozisyonları veya iş yükleri için öneriler sağlamak üzere tasarlanmış akıllı teknolojilerin uygulanması hem daha güvenli hem de daha destekleyici bir

alıřma ortamı geliřtirme potansiyeline sahip olabilmektedir.

Bahsedilen geliřmelere ek olarak, stres analizi konusuyla alakalı olarak etik hususları ve mahremiyet endiřeleri de gz nnde bulundurulmalıdır.

Stres deęerlendirme teknolojisinin srekli geliřmesiyle, alıřanların mahremiyeti ve verilerinin korunmasıyla ilgili etik hususları kabul etmek ve bunlarla mcadele etmek zorunlu hale gelmektedir. Stres deęerlendirmesi iin yz grntleme ve fizyolojik verilerin sorumlu bir Őekilde kullanılmasını saęlamak iin, gelecekteki arařtırmaların veri iřleme iin Őeffaf ve hesap verebilir yntemlerin geliřtirilmesine vurgu yapması zorunlu hal almaktadır.

zet olarak, mevcut arařtırma, yksek riskli mesleki ortamlarda stres seviyelerinin deęerlendirilmesiyle ilgili karmařık yapıların ve anlık stres analizi iin yz grntleme teknolojilerinin potansiyel etkinlięinin altını izmektedir. Sonular, farklı mesleki ortamlar iin zel olarak tasarlanmış zelleřtirilmiş stres deęerlendirme teknikleri kullanmanın nemini vurgulamaktadır. Gelecekteki arařtırmalar, ok ynl tekniklerin keřfi, stres ynetimi terapilerinin uygulanması, insan-makine iř birlięinin teřvik edilmesi ve etik sonuların deęerlendirilmesi yoluyla stres deęerlendirmesini geliřtirme potansiyeline sahip olmaktadır. Bu arařtırmalar, alıřanların refahını iyileřtirmeyi ve iřyeri verimlilięini arttırmayı amalamaktadır.

SONUÇ VE ÖNERİLER

Üzerinde test işlemi yapılan modellerden MobileNetV2 modelinin diğer modellerden doğruluk oranının yüksek olduğu gözlemlenmiştir. MobilnetV2 kullanıldığında batch boyutu artırılarak elde edilen sonuçların doğruluğu artmaktadır. MobilNetV2 modeli boyut olarak diğer modellerden küçük olduğu için donanım gereksinimi açısından biz kullanıcılara avantaj sağlamaktadır. Çalışmanın bu aşamasına kadar duygu analizi işlemleri yapılarak literatür incelemesi doğrultusunda korku, iğrenme ve kızgınlık duygularını stresli olarak etiketleyip model üzerinde işlemler ile sürecin devamı olarak stres analizi oluşturulmaktadır. Ayrıca yüz üzerindeki 68 nokta referans alınarak göz bebeği hareketleri ve ağız hareketleri tespit edilmiştir. Ağız hareketleri ve göz bebeği hareketlerinin konumuna göre kişinin stres durumu tespit edilmiştir. Geçmiş araştırmalar doğrultusunda bu çalışmanın geliştirilmesi için stres tespiti yapacak sensörler ya da kızılötesi kameralar da kullanılabilir.

KAYNAKLAR

- Almeida, J. ve Rodrigues, F. (2021). Facial Expression Recognition System for Stress Detection with Deep Learning.
- Alpaydın E. (2014), Introduction to Machine Learning, 3rd ed., The MIT Press
- Baltacı, S. ve Gökçay, D. (2014), Role Of Pupil Dilation And Facial Temperature Features İn Stress Detection, 22nd Signal Processing and Communications Applications Conference (SIU), Trabzon, Türkiye
- Bengio, Y. (2012). Practical recommendations for gradient-based training of deep architectures. Springer, Berlin, Heidelberg.
- Bevilacqua, F., Engström, H. ve Backlund, P. (2018). Automated Analysis of Facial Cues from Videos as a Potential Method for Differentiating Stress and Boredom of Players in Games. International Journal of Computer Games Technology.
- Can, Y. S., Chalabianloo, D., Ekiz, D. ve Ersoy, C., (2019). Continuous stress detection using wearable sensors in real life: Algorithmic programming contest case study. Sensors (Switzerland).
- Cinaz, B., Arnrich, B., Marca, R. ve Tröster, G. (2013). Monitoring Of Mental Workload Levels During An Everyday Life Office-Work Scenario, Personal and Ubiquitous Computing.
- Copeland, J., (1993). Artificial Intelligence: A Philosophical Introduction, Blackwel, Oxfrod.
- Donald M., Spiegelhalter D.J. ve Taylor C.C. (2009) Machine Learning: Neural and Statistical Classification, Overseas Press
- Duan, K., Keerthi, S. S., Chu, W., Shevade, S. K., ve Poo, A. N. (2003). Multi-category classification by soft-max combination of binary classifiers. In International Workshop on Multiple Classifier Systems. Springer, Berlin, Heidelberg.
- Duda R.O., Hart P.E., Stork D.G. (2001), Pattern Classification, 2nd ed., Wiley-Interscience, New York.
- Farokhi, S., Flusser, J. ve Ullah Sheikh, U. (2016), Near infrared face recognition: A literature survey, Computer Science Review
- Gao, H., Yüce, A., Thiran, J.P., (2014). Detecting emotional stress from facial expressions for driving safety, IEEE International Conference on Image Processing (ICIP), Paris, France.
- Gavrilescu, M. ve Vizireanu, N. (2019). Predicting depression, anxiety, and stress levels from videos using the facial action coding system. Sensors (Switzerland).

- Giannakakis, G. (2017). Stress and anxiety detection using facial cues from videos, Biomedical Signal Processing and Control.
- Goodfellow, I., Bengio, Y., ve Courville, A. (2016). Deep learning. MIT press.
- He, K., Zhang, X., Ren, S. ve Sun, J. (2015), Deep Residual Learning for Image Recognition, Microsoft Research
- Ketkar, N. (2017). Deep Learning with Python, Apress.
- Kohavi R., Provost F. (1998), Glossary of Terms, Machine Learning
- Kriesel, D. (2007). A brief introduction on neural networks, <http://www.dkriesel.com>
- Liapis, A., Katsanos, C., Sotiropoulos, D., Xenos, M. ve Karousos, N. (2015). Stress recognition in human-computer interaction using physiological and self-reported data: A study of gender differences, ACM International Conference Proceeding Series.
- Mohan, A. (2021). Face Emotions based Stress Index Measurement using Machine Learning, ScienceOpen Preprints.
- Mokhayeri, F. ve Akbarzadeh-T, M. R. (2011). Mental stress detection based on soft computing techniques, IEEE International Conference on Bioinformatics and Biomedicine.
- Nabiyev, V. (2012). Yapay Zeka. 4. Baskı, Seçkin Yayıncılık
- Nilsson N.J. (2017), Introduction to Machine Learning: An Early Draft of a Proposed Textbook, Robotics Laboratory, Department of Computer Science, Stanford University.
- Patterson, J., ve Gibson, A. (2017). Deep learning: A practitioner's approach. O'Reilly Media, Inc.
- Quinn, M. A., Sivesind, G. ve Reis, G. (2017). Real-time Emotion Recognition From Facial Expressions., Stanford University.
- Raichur, N., Lonakadi, N. ve Mural, P. (2017). Detection of Stress Using Image Processing and Machine Learning Techniques, International Journal of Engineering and Technology.
- Shelke, S., Kor, S., Bavaskar, S., Rajadnya, K. ve Jondhale, S. S. (2021). Stress Detection Using Machine Learning., Iconic Research And Engineering Journals.
- Singh, T. G., Chaudhari, N. S., Barbosa, J. L. V. ve Aghwariya, M. K. (2019). International Conference on Intelligent Computing and Smart Communication, Proceedings of ICSC.

Song, K., Yao, T., Ling, Q. ve Mei, T. (2018). Boosting image sentiment analysis with visual attention, Neurocomputing.

Söylemez, Ö. F. Ve Ergen, B. (2020). Farklı Evrişimsel Sinir Ağı Mimarilerinin Yüz İfade Analizi Alanındaki Başarımlarının İncelenmesi, Dicle University Journal of Engineering (DUJE).

Tenekeci, M. E., Gümüşçü, A. ve Aslan, E., (2014). Görüntüden OpenCV ile Duygu Analizi. XVI. Akademik Bilişim Konferansı Bildirileri, Mersin.

Wang, M. ve Deng, W. (2021). Deep face recognition: A survey, Neurocomputing.

Wolpert, D. H. (1996). The Lack of A Priori Distinctions Between Learning Algorithms. Neural Comput

Zhang, A., Lipton, Z. C., Li. M. ve Smola, A. J. (2021). Dive into Deep Learning

URL-1 (2023). <https://medium.com/@mygreatlearning/everything-you-need-to-know-about-vgg16-7315defb5918>, Everything you need to know about VGG16, Medium., (26.07.2023)

URL-2 (2023). <https://suhedacilek.medium.com/resnet-residual-network-nedir-49105e642566> , Resnet (Residual Network) Nedir, Medium, (26.07.2023)

URL-3 (2023). <https://ai.googleblog.com/2018/04/mobilenetv2-next-generation-of-on.html> , MobileNetV2: The Next Generation of On-Device Computer Vision Networks, googleblog, (26.07.2023)

URL-4 (2023). <https://www.studytonight.com/post/dlib-68-points-face-landmark-detection-with-opencv-and-python> , Dlib 68 points Face landmark Detection with OpenCV and Python, studytonight, (26.07.2023)

