



GAZİ ÜNİVERSİTESİ
MÜHENDİSLİK FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

BM495 BİTİRME PROJESİ I
DÖNEM SONU RAPORU

181180030- İsmail ERTAYLAN
181180006- Büşra ARIK

Dr. Öğr. Üyesi Çağrı ŞAHİN

2022

Kelime Sayısı : 5073

İNTİHAL BEYANI

Bu çalışmadaki tüm bilgilerin akademik kurallara ve etik davranışa uygun olarak alındığını ve sunulduğunu ve bu belgede alıntı yaptığımı belirttiğim yerler dışında sunduğum çalışmanın kendi çalışmam olduğunu, Yükseköğretim Kurumları Bilimsel Araştırma Ve Yayın Etiği Yönergesinde belirtilen bilimsel araştırma ve yayın etiği ilkelerine uygun olduğunu beyan ederim.

Numara : 181180006

Ad Soyad : Büşra Arık


Tarih : 08.01.2023

İmza : 

Numara : 181180030

Ad Soyad : İsmail Ertaylan

Tarih : 08.01.2023

İmza : 

İçindekiler

SAHTE FOTOĞRAF ANALİZİ	1
ÖZET	1
1. GİRİŞ	1
1.1. Kapsam	2
1.2. Amaç.....	2
1.3. Hedef Kitle	2
2. LİTERATÜR TARAMASI	2
2.1. Forged Face Detection using ELA and Deep Learning Techniques	2
2.2. Methods of Deepfake Detection Based on Machine Learning.....	2
2.3. Exposing AI Generated Fake Face Videos by Detecting Eye Blinking	3
2.4. A Detection Method of Operated Fake-Images Using Robust Hashing	3
2.5. Detecting Fake Images on Social Media using Machine Learning	4
2.6. Comparative Analysis of Deepfake Image Detection Method Using Convolutional Neural Network.....	4
2.7. Fake Image Detection Using Machine Learning	5
2.8. Image Forgery Detection Using Deep Learning by Recompressing Images	6
2.9. A Landscape View Of Deepfake Techniques And Detection Methods	6
2.10. Deep Fake Image Detection Based on Pairwise Learning.....	6
3. GELİŞTİRİLEN YAKLAŞIM VE BULGULAR	7
3.1. Ürün Perspektifi.....	7
3.2. Sistem Arayüzü.....	7
3.3. Ürün İşlevleri.....	7
3.4. Gerekli Durum ve Modlar	8
3.5. Fonksiyonel Gereksinimleri	8
3.6. Arayüz Gereksinimleri	8
3.7. Veri Seti Gereksinimleri.....	8
3.8. Tasarım ve Uygulama Kısıtlamaları.....	8

3.8.1.	Yazılım Kısıtları.....	9
3.8.2.	Donanım Kısıtları.....	9
3.9.	Yazılım Kalite Faktörleri.....	9
3.9.1.	Sayısal Gereksinimler	9
3.9.2.	Kullanılabilirlik.....	9
3.9.3.	Güvenilirlik	9
3.9.4.	Erişilebilirlik	9
3.9.5.	Esneklik.....	9
3.9.6.	Test Edilebilirlik	9
3.9.7.	Taşınabilirlik	9
3.10.	Tasarım Görünümleri.....	10
3.11.	Tasarım Bakış Açıları	10
3.12.	Tasarım Endişeleri	10
3.13.	Tasarım Öğeleri	10
3.13.1.	Tasarım Varlıkları	11
3.13.2.	Tasarım Özellikleri.....	11
3.13.3.	Tasarım Kısıtlamaları.....	11
3.14.	Tasarım Katmanları	12
3.15.	Tasarım Gerekçesi	12
3.16.	Tasarım Dilleri	12
3.17.	Tasarım Bakış Açıları	12
3.18.	Bağlam Bakış Açısı.....	12
3.19.	Kompozisyon Bakış Açısı.....	13
3.20.	Mantıksal Bakış Açısı	13
3.21.	Bağımlılık Bakış Açısı.....	14
3.22.	Bilgi Bakış Açısı	14
3.23.	Arayüz Bakış Açısı	14

3.24. Etkileşim Bakış Açısı.....	15
4. SONUÇ VE ÖNERİLER.....	16
Referanslar	17

SAHTE FOTOĞRAF ANALİZİ

ÖZET

Günümüzde teknolojinin ilerlemesi ve yapay zekanın yaygınlaşması sürmektedir. Derin sahte fotoğraflar yapay zekanın konularındandır ve bu görsellerin oluşturulması için gerekli modeller yapay zeka ile yapılmaktadır. Her geçen gün algoritmaların gelişimi ile gerçekçiliği artan derin sahte içerikler tehlikeli bir hal almaktadır. Bilinen kişilerin sahte içeriklerinin yayınlanabilmesi oldukça riskli ihtimallerdir. Bu teknoloji ileride ciddi sorunlar oluşturabilir hale gelmektedir. Bu da teknolojinin yarattığı problemleri teknolojinin çözmesine yol açmaktadır. Derin sahte içeriklerin tespiti yapılabilmektedir. Bu tespitlerin yapılabilmesi hukuki anlamda siber güvenlik konularında önemli bir rol oynar. Bu konu, güncel bir teknoloji olmakla birlikte gelişmeye hali hazırda devam etmektedir. Sahte içeriklerin üretiminde çeşitli algoritmalar kullanılırken tespitinde de durum aynıdır. Tanımlamak gerekirse sahte fotoğraf analizi, yapay zekaya dayalı yöntemler ve derin öğrenme teknikleri ile üzerinde değişiklik yapılan fotoğraf ve görsellerin tespitini sağlamaktır. Sahte fotoğraf analizini gerçekleştirecek ilgili sistemde derin öğrenme teknikleri, hata seviye analizi, meta veri analizi ve evrişimli sinir ağı algoritmaları kullanılacaktır.

1. GİRİŞ

Günümüzde yapay zekanın yaygınlaşması ile teknoloji ciddi boyutlarda gelişmiştir. Yapay zeka konularından biri de derin sahte fotoğraf ve videolardır. Bu medya içerikleri için gerekli modeller yapay zeka ile yapılmaktadır. Fonksiyonel açıdan derin sahte, bireyin yüzü veya bedeninin tamamı ile yapay zeka teknolojisinden faydalanılarak hareket ve konuşmaların değiştirilmesidir.

Yapay zekanın ve makine öğrenmesi algoritmalarının gelişmesi ile daha gerçekçi hale gelen derin sahte içerikler tehlikeli hale gelmektedir. Örneğin ünlü bir siyasinin sahte bir suç videosu yayınlanabilir. Bu teknoloji ileride ciddi sorunlar oluşturabilir hale gelmektedir. Bu durumda ise teknolojinin kötüye kullanımını yine bir başka teknoloji tarafından durdurulup, derin sahte içeriklerin tespiti yapılabilmektedir. Büyük öneme sahip olan bu teknoloji aynı zamanda hukuki anlamda siber güvenlik sorunlarında da ciddi bir role sahiptir. Derin sahte içeriklerin tespiti de oldukça günceldir ve gelişmeye devam etmektedir. Bu içerikler üretilirken çeşitli algoritmalar kullanılmaktadır ve tespitinde de benzer durum söz konusudur.

Derin sahte içeriklerin tespitinde yüz ifadeleri ve kafa hareketleri önemli ipuçları vermektedir. Renk tutarsızlıkları, bozuk dokular, optik hatalar, artefaktlar ve izler de derin sahte tespitinde öne çıkmaktadır. Yüz değişimleri için dişler ve göz yansımaları, tutarsız ağız ve dudak hareketleri analiz edilmelidir. Göz kırpma anormallikleri algoritmalar tarafından yeterince geliştirilmemiş ve tespiti kolaylaştıran unsurlardandır. Göz kırpmanın frekansı olağandışı olabilmektedir. Araştırmacılar yöntemlerde artefaktlara, tutarsız renklere, doku bozulması ve parmak izlerine odaklanmışlardır. Baş pozisyonundaki farklılıklar, bahsedilen yapay göz kırpmaları ve yüz çarpıklıkları da araştırmacıların konularından olmuştur.

Yüz artefaktları da tespit yöntemlerinde kullanılır ve bunlar yüzde bulanıklık ve ışık farkları, ton değişiklikleri, kaydırma sonucu çift bölgelerin oluşması, titreme olmasıdır. En sık kullanılan yöntemlerden biri yüz manipülasyonlarıdır. Görüntü işlemleri yapılırken yeniden

ölçeklendirme, döndürme ve sentezleme işlemleri yapılmaktadır. Bu işlemler yapılırken bazı çarpıtmalar oluşur ve bu manipülasyonlar tespit edilebilmektedir.

1.1. Kapsam

Sahte fotoğraf analizi projesi, kullanıcıların programa yüklediği görsellerin sahte olup olmadığı analizini gerçekleştiren bir uygulama olarak geliştirilecektir. Uygulamada programlama dili olarak Python kullanılacaktır. Proje 2 öğretim dönemini kapsayan bir projedir. Projenin ilk döneminde araştırmalar, literatür taramaları, gereksinim belirlemeleri ve planlamalar gerçekleştirilmiştir. Ayrıca kullanılacak veri kümesi olan CASIA veri kümesi belirlenmiştir. Bu veri seti üzerinde araştırmalar gerçekleştirilmiştir. Veri ön işleme operasyonları gerçekleştirilmiştir. Hata seviyesinde analiz(ELA) işlemi tamamlanmıştır. Diğer dönemde ise meta veri analizi ve CNN aracılığıyla tahmin işlemleri yapılacaktır. Ayrıca projenin platformlara uyumlu hale getirilmesi de planlanmaktadır.

1.2. Amaç

Projede amaç kullanıcının girdi olarak kullandığı görsellerin derin öğrenme yöntemleri aracılığıyla sahtelik analizinin gerçekleştirilmesi ve sonucun kullanıcıya basit bir şekilde yansıtılmasıdır.

1.3. Hedef Kitle

Proje genel olarak bir görselin sahteliğini analiz etmek isteyen tüm kitlelere hitap etmektedir.

2. LİTERATÜR TARAMASI

2.1. Forged Face Detection using ELA and Deep Learning Techniques

Qurat-ul-ain[1], sahte fotoğraf analizi için CNN'i kullanan bir teknik önermiştir. Başlangıçta tüm veri kümesinin (128*128) piksel olarak yeniden boyutlandırıldığı ve normalleştirildiği pre-processing işlemi yapılmaktadır. Daha sonra ELA kullanılarak extraction işlemi yapılmıştır. İşlemden geçen görüntüler training ve test setlerine bölünüp, gerçek ve sahte görüntüleri tanımak için Deep-CNN modellerine iletilir. Bu modeller VGG-16, ResNet-50, InceptionV3 ve VGG-19'dir. VGG-16 ve 19 modelleri %91,97 ve %92,09 oranında training doğruluğu verirken, VGG-16'nın aynı veri setlerinde diğer önceden eğitilmiş modellere kıyasla daha iyi olan %64,49 test seti doğruluğu verdiği gözlemlenmiştir [1].

2.2. Methods of Deepfake Detection Based on Machine Learning

Makalede face swapping AI tabanlı algoritmalarla videonun/fotoğrafın değiştirilip değiştirilmediğine karar vermek için kullanılacak indikatörler açıklanmıştır. Bunlar; çok pürüzsüz bir cilt, sentezlenen yüz ile orijinal yüz arasındaki renk uyumsuzluğu, baş pozisyonu, göz kırpma oranı, küçük hareketli parçalardaki artefaktlar ve yüz çarpıtma artefaktlarıdır. Yüz çarpıtma artefaktları, düşük çözünürlüklü yüz çıktısına sahip algoritmalar (64x64 veya 128x128) tarafından oluşturulan sahte videoların en iyi indikatörlerindendir. Model olarak DenseNet169 ile yüz çarpıtma artefakt indikatörü kullanılmıştır. Modeli değerlendirmek için Celeb-DF veri seti kullanılmıştır. Bu veri setindeki içeriklerin test sonucunda doğru çıktılar verip vermediği anlaşılması için üzerinde değişiklikler yapılmıştır. İçeriklere gürültü eklenip Gauss bulanıklığı, exponential

bulanıklığı ve Rayleigh bulanıklığı test edilmiştir. En yüksek AUC değerine sahip model %60.1 ile DenseNet169 + Rayleigh blur modeli olmuştur [2].

2.3. Exposing AI Generated Fake Face Videos by Detecting Eye Blinking

Bu çalışmada, sinir ağı ile oluşturulan sahte yüz videolarının tespiti için göz kırpmaya dayalı bir yöntem anlatılmıştır. Spontane göz kırpma, refleks olarak göz kırpma ve istemli göz kırpma olmak üzere 3 göz kırpma türü vardır. Burada kullanılan yöntem, göz kırpma sürecindeki fenomenolojik ve zamansal düzenlilikleri yakalamak için CNN'i, RNN ile birleştiren bir derin öğrenme modeline dayanmaktadır. Deneyde yapılan işlemler pre-processing, LRCN ve model eğitimi olmak üzere 3 başlıkta toplanmaktadır. Pre-processingte, yer işareti tabanlı yüz hizalama algoritmaları kullanılarak yüz bölgeleri hizalanır ve yüz dedektörü kullanılarak yüz işaretleri çıkarılmaktadır. LRCN aşamasında model özellik çıkartma, dizi öğrenme ve durum tahmini işlemlerinden geçirilip eğitim aşamasına gönderilmektedir. LRCN modeli, gözün açık halinin görüntü veri kümelerine göre eğitilmiştir. Daha sonra Deep Fake algoritması ile oluşturulan gerçek ve sahte videolarda göz kırpmayı algılayan algoritma test edilmiştir. Deneyde CEW veri seti kullanılmıştır. LRCN metodu, EAR ve CNN metotları ile karşılaştırılarak değerlendirildiğinde CNN görüntü sınıflandırıcısı, farklı sınıfları ayırt etmek için görüntü alanında eğitilmiştir. Göz durumunu ayırt etmek için CNN modeli olarak VGG16 kullanılmıştır. EAR metodu üst ve alt kapak mesafesi ile sol ve sağ köşe noktası arasındaki mesafe arasındaki oran açısından göz durumunu analiz etmek için göz işaretlerine yanıt vermektedir. En büyük dezavantajı tamamen göz işaretlerine bağlı olmasıdır. Hesaplamalara bakıldığında LRCN %99 başarı oranıyla en iyi sonucu verirken CNN %98 ve EAR %79 oranında başarılıdır. Ama göz durumuna göre tespit konusunda CNN olağanüstü başarılı bir sonuç vermektedir. Mevcut çalışmada eksiklik olarak dinamik göz kırpma modeli dikkate alınmamaktadır. Göz kırpma, sahte yüz videolarını tespit etmede kolay bir ipucudur ve daha gelişmiş modeller, daha fazla eğitim verisi ile hala gerçekçi yanıp sönme efektleri oluşturabilmektedir. Bu nedenle, önerilen metot eksiklikler barındırmaktadır [3].

2.4. A Detection Method of Operated Fake-Images Using Robust Hashing

Bu makalede, görüntü işlemlerinden kaynaklanan bozulmalar da dahil olmak üzere sahte görüntüleri tespit etmek için bir yöntem önerilmiştir. Makalede, Robust hash yöntemi kullanılarak referans görüntülerden robust hash değerleri hesaplanır ve değerler veri tabanında saklanır. Referans görüntüleri benzer şekilde robust hash yöntemi kullanılarak bir sorgu görüntüsünden robust hash değeri hesaplanmaktadır. Sorgunun hash değeri, veritabanında depolananlarla karşılaştırılıp, hash değerleri arasındaki mesafeye göre sorgu görüntüsünün gerçekliğine karar verilmektedir. Sahte görüntü algılamaya yönelik hash değerlerinin, sıkıştırma ve yeniden boyutlandırma gibi bir dizi görüntü işlemi türüne karşı yeterince sağlam olması gerekir çünkü bu tür bir işlem, görüntülerin kalitesini düşürmesine rağmen görüntülerin içeriğini değiştirmez. Bu nedenle, sorgu görüntülerine benzer görüntüleri sağlam bir şekilde almayı amaçlayan robust hashing yöntemi kullanılmıştır. Buna karşılık, robust hash yöntemi kullanılarak oluşturulan hash değerlerin, kopyala-taşı ve GAN'lar gibi sahte görüntüler oluşturmak için kullanılan manipülasyonun etkisine duyarlı olması gerekmektedir. Bu gereksinimler altında, Li et al.'s yönteminin sahte görüntü tespiti için uygun bir performansla sahip olduğu anlaşılmıştır. Deneyde Görüntü Manipülasyon Veri Kümesi, UADFV, CycleGAN ve StarGAN veri setleri kullanılmıştır. Orijinal görüntüler referans olarak kullanılmış, her deney için farklı sahte görüntü veri seti ile oluşturulan ayrı bir referans veri seti hazırlanmıştır. Sorgu görüntüleri olarak hem orijinal görüntüler hem de

sahte görüntüler kullanılmıştır. Bu deneyde, önerilen yöntem, gerçek sorgu görüntüleri olmasına rağmen, veri kümelerinden gelen sorgu görüntülerinin herhangi bir ek işlem yapılmadan doğrudan kullanıldığı Wang'ın yöntemi ve Xu'nun yöntemiyle karşılaştırılmıştır. Wang'ın yöntemi, sınıflandırıcının ProGAN kullanılarak eğitildiği GAN modelleriyle birlikte çeşitli CNN'ler tarafından oluşturulan görüntüleri tespit etmek için önerilmiştir. Önerilen yöntemin neredeyse tüm kriterler açısından daha yüksek bir doğruluğa sahip olduğu gösterilmektedir. Ayrıca, Görüntü Manipülasyonu ve UADFV veri kümeleri kullanıldığında geleneksel yöntemlerin doğruluğu oldukça azalmıştır. Bunun nedeni, geleneksel olanların CNN'ler kullanılarak oluşturulan sahte görüntüleri tespit etmeye odaklanmasıdır. Görüntü Manipülasyonu Veri Kümesi, GAN'larla oluşturulan görüntülerden oluşmaz. Ayrıca UADFV derin sahte videolardan oluşsa da veri setindeki videolar zaten video sıkıştırma etkisine sahiptir. Sıkıştırma için orijinal bir hash kod olduğunda, önerilen yöntem, geleneksel yöntemlere göre sahte görüntüleri daha iyi bir şekilde tespit edebilmektedir, eğer bir hash kodu yoksa görüntü tespiti yapılamamaktadır. Deneyde, önerilen yöntemin diğerlerinden daha iyi performans gösterdiği ve aynı zamanda birden fazla işlemi birleştirirken de iyi bir sonuç verdiği gözlemlenmiştir [4].

2.5. Detecting Fake Images on Social Media using Machine Learning

Bu makalede sosyal medya üzerindeki sahte fotoğrafların makine öğrenmesiyle tespiti incelenmiştir. Araştırmacı, makine öğrenimi algoritmalarını kullanan ve CNN aracılığıyla bu tespiti sağlayan sınıflandırıcı bir model önermiştir. İlgili modelde normal görsel ve sahte görsel olmak üzere iki sınıf vardır. Araştırmacı, CNN aracılığıyla derin öğrenme tekniğini kullanmıştır. Yönteme göre önce Instagram'daki IJACSA veri setinden tespiti yapılacak görüntüler elde edilir. CNN aracılığıyla geleneksel matematiksel işlemler kullanılır ve görüntü özellikleri çıkartılıp aktivasyon fonksiyonu oluşturulur. Görüntü verilerinde doğrusallık olmadığından RELU işlevi kullanılmaktadır. Dizi boyutunu küçültmek amacıyla max pooling algoritması kullanılmaktadır. Bu aşamalardan sonra tahmin gerçekleştirilir ve bir sinir ağı ile görüntünün eşleşip eşleşmediğine karar verilir. SoftMax ile çıktının olasılıklar halinde görünmesi sağlanır. Sinir ağı eğitimi tamamlandığında da veri seti test edilir ve doğruluğun hesaplandığı değişkenleri içeren karışıklık matrisi çıkarılır. Araştırmada performans metrikleri 3 ağı göre incelenir: Alexnet, klasik CNN ve AlexnetTL. Eğitim verilerine göre ortalama sonuçlar değişse de her seferinde sıralama Alexnet-AlexnetTL-klasik CNN şeklinde olmuştur. Eğitim verilerinden yola çıkıldığında Alexnet %99.3, AlexnetTL %94 ve klasik CNN ise %83.9 oranında başarı sağlamıştır. Sonuçlardan da anlaşıldığı üzere klasik CNN modeline göre Alexnet/AlexnetTL'in kullanımı, daha başarılı çıktılar elde etmektedir [5].

2.6. Comparative Analysis of Deepfake Image Detection Method Using Convolutional Neural Network

Bu çalışmanın amacı, derin sahte görüntüleri tespit etmenin güvenilir bir yolunu bulmak ve CNN mimarisiyle başarılı sonuçlar elde etmektir. Çalışmada, büyük bir veri setinden derin sahte görüntüleri tespit etmek için 8 CNN mimarisi kullanılmaktadır. Bunlardan üçü DenseNet mimarisi (DenseNet121, DenseNet169 ve DenseNet201), ikisi VGGNet mimarisi (VGG16, VGG19), biri ResNet50 mimarisi, biri VGGFace mimarisi ve sonuncusu ise özel bir CNN mimarisidir. CNN, özellik çıkartma ve sınıflandırma kısımlarından oluşmaktadır. Deneyde veriler, Kaggle üzerinden toplanan bir veri kümesinden elde edilir ve daha sonra evrişim katmanına gönderilir. Bu katman, girdi olarak alınan fotoğraflardan çok sayıda özellik çıkartır. Daha sonra havuzlama katmanına geçilir. Bu katmanın amacı, evrişimli

özellik katmanının boyutunu en aza indirmektir. Önceki seviyelerden gelen girdiler düzleştirilir ve girdi FC katmanına gönderilir. Düzleştirilmiş vektör üzerinde matematiksel fonksiyonel işlemleri yürütmek için başka FC katmanları kullanılır. Bu aşama fotoğrafların sınıflandırma sürecini başlatmaktadır. Deney sonunda VGGFace, doğruluk, kesinlik, F1 puanı ve ROC eğrisi altındaki alan gibi ölçümlerde en iyi performansı göstermiştir. En kötü performansı ise VGG16 göstermiştir, %92 doğruluk elde etmiştir. ResNet50 de %97 doğruluk elde etmiştir. DenseNet201 ve DenseNet169, sırasıyla %96 ve %95 doğruluk elde etmiştir. En yüksek AUC puanı, %99.8 ile VGGFace mimarisi tarafından elde edilirken, en düşük AUC puan DenseNet121 mimarisi tarafından elde edilmiştir. Yazarlar tarafından önerilen model,%90 doğruluk elde etmiştir.Genele bakıldığında VGGFace en iyi performansı göstermiştir [6].

2.7. Fake Image Detection Using Machine Learning

Çoğu görüntü dosyası resim hakkında bilgi veren meta verilerini de barındırmaktadır. Meta veriler, dosyanın nasıl oluşturulduğu ve işlendiği ile ilgili bilgiler vermektedir. Meta veride aranması gereken bilgiler şunlardır;

1. Model ve yazılım: Bunlar, resmi oluşturan cihazı veya uygulamayı tanımlar. Kameralar, EXIF bilgisi olarak marka ve model içermektedir.
2. Görüntü boyutu: Meta veriler genellikle resmin boyutlarını kaydeder. İşlenen görüntü boyutu, meta verilerdeki diğer boyutlarla eşleşiyor mu diye kontrol edilir.
3. Zaman bilgisi: Bunlar genellikle fotoğrafın çekim ve değişim tarihlerini içerir. Zaman bilgilerinin beklenen zaman dilimine uyumu kontrol edilir.
4. Meta veri türü: Meta veri türlerinin bazıları sadece kameralar tarafından üretilirken, diğerleri yalnızca uygulamalar tarafından üretilir.
5. Açıklamalar: Gömülü ek açıklamalar içerir.
6. Eksik meta veri: Belirli meta verilerin olmaması genellikle orijinal bir fotoğrafın değil, resmin kaydedildiğini göstermektedir.
7. Değiştirilmiş meta veriler: Kasıtlı olarak meta verileri değiştirilebilmektedir.

Makalede ELA ve Meta veri analizi yöntemleri birlikte kullanılmıştır. Bir JPEG'in kalitesi kayıt edildikçe düşmektedir. ELA, fotoğraftaki görüntü kalite farklarına bakarak manipülasyonu anlayabilmektedir. ELA, ImageJ kütüphanesi aracılığıyla yapılmaktadır. ImageJ, görüntüyü belirli bir sıkıştırma yüzdesiyle JPEG formatında kaydetme seçeneği sunmaktadır. Sistem önce görüntüyü kayıpsız kaydeder. Daha sonra aynı görüntü ImageJ kullanılarak %90 kaliteli görüntüye dönüştürülür. Aradaki fark, fark yöntemiyle bulunmaktadır. Elde edilen görüntü, giriş görüntüsünün gerekli ELA görüntüsüdür. Bu görüntü, arabelleğe alınmış bir görüntü olarak kaydedilir ve daha sonraki işlemler için sinir ağına gönderilir. Eğitim sırasında dizi, çok katmanlı algılayıcı ağına girdi olarak verilir ve çıktı nöronları ayarlanır. MLP, tamamen bağlı bir sinir ağıdır ve 2 çıkış nöronu vardır. İlki sahte, ikincisi gerçek görüntüyü temsil etmektedir. Verilen görüntü sahte ise, sahte nöron bire, gerçek ise sifıra ayarlanır. Testte, görüntü dizisi giriş nöronlarına beslenir ve çıkış nöronlarının değerleri alınır. Meta veri analizinde ise önce meta verilerin çıkarılma işlemi yapılmaktadır. Sonra meta veri metni, meta veri analizi modülüne gönderilir. Bu analiz temelde bir etiket arama algoritmasıdır. Metinde Photoshop, Gimp, Adobe vb. kelimeleri arar. Sahtelik ve gerçeklik olarak adlandırılan ve gerçek ve sahteyi temsil eden iki değişken oluşturulur. Bir etiket alındığında, analiz edilir ve karşılık gelen değişken önceden tanımlanmış belirli bir ağırlıkla artırılır ve buradan alınan sonuçlarla ELA yönteminden alınan sonuçlar birleştirilir. Bu analiz, çok küçük bir işlem altında dahi tüm 'photoshopped' veya 'gimped' görüntülerde sahteliği tespit edebilmektedir ama WhatsApp, Google+ vb.

üzerinden paylaşılan görsellerde hata vermektedir. Sinir ağı CASIA veri seti ile eğitilmiştir. Eğitimli sinir ağı, görüntüyü %83 başarı oranıyla tanıyabilmiştir [7].

2.8. Image Forgery Detection Using Deep Learning by Recompressing Images

Bu yazıda, çift görüntü sıkıştırma bağlamında görüntü sahteliğini belirlemek için robust derin öğrenme tabanlı bir sistem anlatılmıştır. Bir görüntünün orijinal ve yeniden sıkıştırılmış sürümleri arasındaki fark, modeli eğitmek için kullanılmıştır. Görüntü yeniden sıkıştırıldığında, sahtelik içeriyorsa, orijinal görüntünün kaynağı ile sahte bölümün kaynağı arasındaki fark nedeniyle görüntünün sahte kısmı görüntünün geri kalanından farklı şekilde sıkıştırılmaktadır. Orijinal görüntü ile yeniden sıkıştırılmış versiyonu analiz edilir. Makalede, CNN mimarisi yaklaşımını vurgulayan, sinir ağları ve derin öğrenmeye dayalı bir görüntü sahteciliği tespit sistemi sunulmuştur. Bu yöntem, görüntü sıkıştırmasındaki varyasyonları içeren CNN mimarisini kullanmaktadır. Modeli eğitmek için orijinal ve yeniden sıkıştırılmış görüntüler arasındaki fark kullanılmıştır. Önerilen teknik, birleştirme ve kopyala-taşı ile değişiklik yapılmış fotoğraflardaki sahteliği saptayabilmektedir. Deney sonuçları, %92,23 genel doğrulama oranı göstermektedir. Mevcut teknik, minimum 128x128 çözünürlük gerektirmektedir [8].

2.9. A Landscape View Of Deepfake Techniques And Detection Methods

Bu makalede derin sahte çalışma ve kavramları, teknikleri ve algoritmaları incelenmiştir. Derin sahte içerikler manipülasyon derecesine göre tüm yüzün sentezi, kimlik değişikliği, özellik manipülasyonu ve ifade değişimi olarak 4 sınıfta toplanmaktadır. Tüm yüz sentezi, StyleGAN kullanarak aslında var olmayan tam yüz görsellerini üretmektedir. Bu sentez video oyunları, 3 boyutlu modelleme, fotoğrafçılık gibi çeşitli alanlarda avantajlar sağlamaktadır. Kimlik değiştirme yöntemi, FaceSwap6 ve DeepFakes7 gibi yöntemler ile yüz değişimini sağlar. Özellik manipülasyonu, bir GAN ve StarGAN yöntemi kullanılarak yüzde saç/ten rengi, sakal, bıyık, yaş, cinsiyet değişiklikleri gibi rötüslara olanak sağlamaktadır. İfade değişimine bakıldığında ise, standart GAN mimarileri aracılığıyla bir kişinin mimiklerinin değişimi sağlanmaktadır. Çalışmalarda odaklanılan çeşitli noktalarla farklı sonuçlar elde edilebilmektedir. NIST MFC2018 veri setinde renk farklılıklarına odaklanılarak yapılan bir çalışma ile %70 AUC elde edilmiştir. Sinirsel davranışı izleyen başka bir çalışmada ise FakeSpotter yöntemleri kullanarak bir SVM eğitilmiştir. Önerilen teknikler CelebA-HQ, FFHQ veri setlerinden gerçek yüz görselleri; InterFaceGAN ve styleGAN'ın ürettiği sentetik yüz görselleri kullanarak test edilmiş ve %84.7 oranında doğruluk elde edilmiştir [9].

2.10. Deep Fake Image Detection Based on Pairwise Learning

Bu makalede, kontrast kaybı kullanarak sahte görüntülerin tespiti için derin öğrenmeye dayalı bir yaklaşım önerilmektedir. Sahte-gerçek görüntü çiftlerini oluşturmak için son teknoloji GAN'lar kullanılmıştır. İndirgenmiş DenseNet, girdi olarak ikili bilgiye izin vermek için iki akışlı bir ağ yapısına dönüştürülmüştür. Ardından, önerilen ortak sahte özellik ağı, görüntüler arasındaki özellikleri ayırt etmek için ikili öğrenme kullanılarak eğitilmiştir. Son olarak, sahteliğini algılamak için önerilen ortak sahte özellik ağına bir sınıflandırma katmanı eklenmiştir. Yöntemi doğrulamak için, sahte yüz ve genel görüntüleri tanımlamak için önerilen DeepFD uygulanmıştır. Deneysel sonuçlar, yöntemin yöntemlerden daha iyi performans sağladığını göstermiştir. Önerilen ikili öğrenme stratejisi, eğitilmiş sahte görüntü dedektörünün, eğitim aşamasına dahil edilmemiş olsa bile, yeni bir GAN tarafından oluşturulan sahte görüntüyü algılama yeteneğine sahip olmasını sağlayan

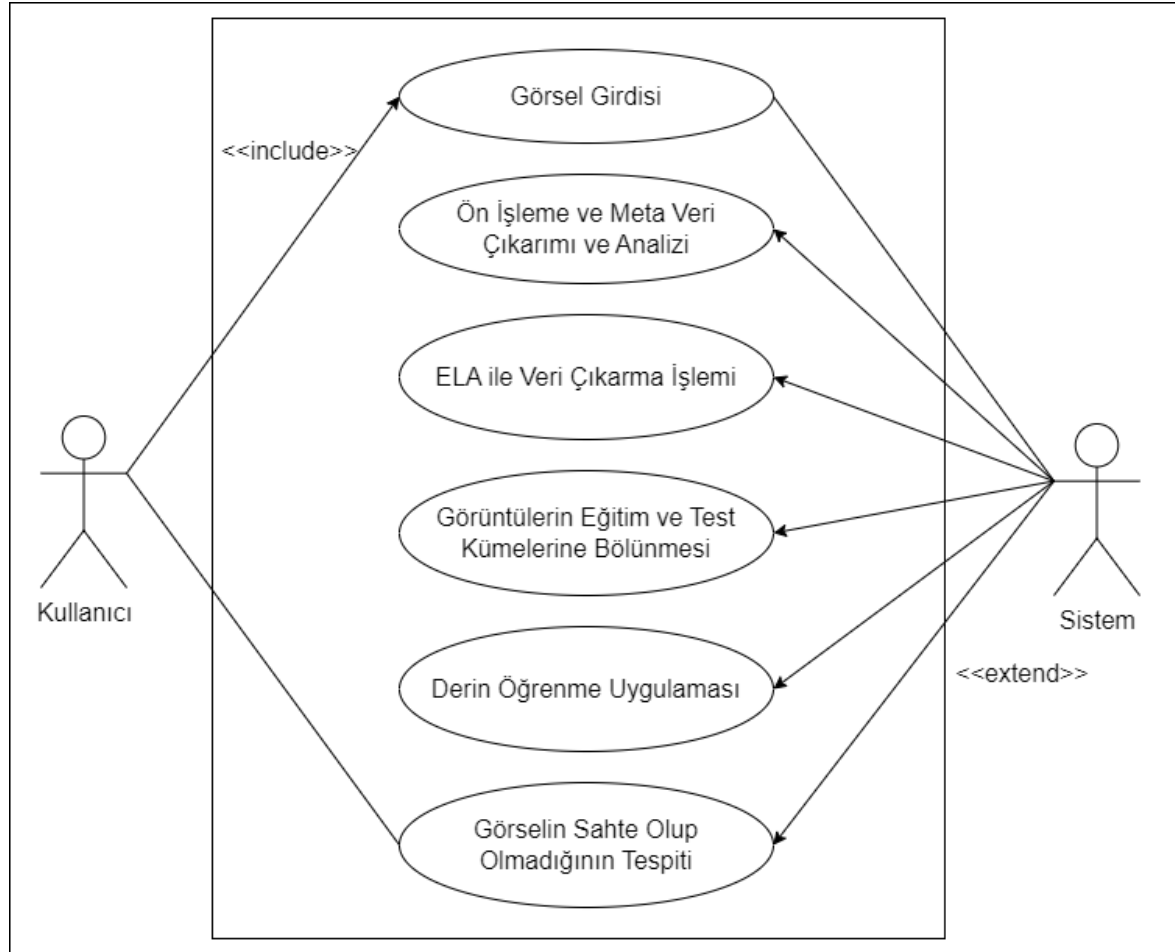
sahte özellik öğrenmesini sağlamaktadır. Deneysel sonuçlar, yöntemin kesinlik ve geri çağırma oranı açısından daha başarılı olduğunu göstermektedir. Yöntemin dezavantajı, eğitim örneklerinin toplanması ile ilgilidir. Bazı sahte görüntü oluşturunucuların teknik detayları açıklanmadığından eğitim örneklerinin toplanması zor olabilmektedir. Bunu aşabilmek için CFF küçük bir eğitim setinden birkaç aşamalı olarak öğrenilmelidir [10].

3. GELİŞTİRİLEN YAKLAŞIM VE BULGULAR

3.1. Ürün Perspektifi

Sahte fotoğraf analizinde sistem tek bir kullanıcı tipine sahiptir. Bu sebeple ilgili aşamalar değişkenlik göstermez. Akış kullanıcıdan sisteme ve sistemden kullanıcıya doğrudur. Akış kullanıcının sonucu elde etmesiyle son bulur.

3.2. Sistem Arayüzü



Şekil 1 - Sistem Use-Case Diyagramı

3.3. Ürün İşlevleri

- Uygulama açıldığında kullanıcı görsel yükleme butonuna tıklar.
- Kullanıcı tespit yapmak istediği görseli bilgisayarından seçer.
- Kullanıcı görseli seçtikten sonra “tespit et” butonuna tıklar ve görseli sisteme yükler.
- Sistem, ilgili algoritmaları yaptıktan sonra sonuçları ekrana yansıtır.

- Kullanıcı fotoğrafın sahte olup olmadığı bilgisine erişir.
- Kullanıcı sonucun doğruluk analizini ekranda pasta grafiği biçiminde görüntüler.

3.4. Gerekli Durum ve Modlar

Sistemin değişken herhangi bir modu bulunmamaktadır. Sistem her an tek bir çalışma modundadır ve yüklenen görselin tespitini yapmaya hazır haldedir. Her tespit sonrasında tekrar tespit yapılabilir duruma gelmektedir.

3.5. Fonksiyonel Gereksinimleri

Fonksiyonel olarak sistem, yüklenen görselin belirli aşamalardan geçmesiyle hata seviyelerini analiz etmelidir. Bu analiz sonuçlarında görseller eğitim ve test aşamalarından geçmelidir. Ardından görseller, sahteliğin tespit edilebilmesi için derin öğrenme modellerine iletilmelidir. Sistem başarı oranlarını sonuç olarak vermelidir.

- 3.5.1. Sistem kullanıcının fotoğrafları veya görselleri JPG ve PNG formatında yüklemesine izin vermelidir.
- 3.5.2. Sistem veri kümesini ön işleme(pre-processing) aşamasından geçirmelidir.
- 3.5.3. Sistem hata seviye analizi(ELA) tekniğiyle çıkarma işlemi gerçekleştirmelidir.
- 3.5.4. Sistem hata seviye analizinden geçen görüntüleri eğitim ve test kümelerine bölmelidir.
- 3.5.5. Sistem görüntülerin derin öğrenme teknikleri aracılığıyla sahte olup olmadığını analiz edebilmelidir.
- 3.5.6. Sistem sonuç olarak başarı oranını ve sahte olup olmadığını kullanıcıya çıktı halinde sunmalıdır.

3.6. Arayüz Gereksinimleri

- 3.6.1. Program arayüzü basit ve tek tasarımlı sayfadan oluşmalıdır. Bu ekranda tespiti yapılacak görselin programa yüklenmesi için gerekli bölümler ve sonuçların ekrana yansıtacağı bölümler yer almalıdır.
- 3.6.2. Arayüz her kullanıcının kolayca kullanabileceği şekilde tasarlanmalıdır. Basit bir görsel yükleme aracı ve anlaşılır bir sonuç ekranı içermelidir.

3.7. Veri Seti Gereksinimleri

- 3.7.1. Sistem, eğitim ve test aşamalarında kullanılmak üzere uygun veri setini barındırmalıdır.
- 3.7.2. Kullanıcının yüklediği görseller, sistem tarafından tek seferlik kullanılacak olup veri setine dahil edilmeyecektir.
- 3.7.3. Sistem, veri seti elemanlarını önceden belirlenmiş ölçüde yeniden boyutlandırarak algoritmalarda kullanılmak üzere sabit boyutlu hale getirmelidir.

3.8. Tasarım ve Uygulama Kısıtlamaları

Sistemin kullanılacağı bilgisayar, yazılım ve donanım olmak üzere bazı gereksinimleri içerir.

3.8.1. Yazılım Kısıtları

3.8.1.1. Sistem geliştirme ortamı olarak Jupyter Notebook çatısı altında Python programlama dili ile gerçekleştirilmelidir. Veri seti üzerinde değişiklik yapılmayacağından bir veri tabanı yönetim sistemine ihtiyaç duyulmamaktadır.

3.8.1.2. Sistemin çıktıları pasta grafiği üzerinde gösterilmelidir.

3.8.2. Donanım Kısıtları

3.8.2.1. Sistemde gerekli işlemlerin yapılacağı bir bilgisayar ve müşterinin çıktıları görebileceği bir monitör olmalıdır.

3.8.2.2. Bilgisayarın çalışacağı sistemde 1.8 GHz ve üzeri işlemci hızına sahip olmalıdır.

3.8.2.3. Bilgisayarın çalışacağı sistem, en az 2 GB RAM'e sahip olmalıdır.

3.8.2.4. Bilgisayarın çalışacağı sistem, ilgili görselleri de barındıracağından en az 10 GB depolama alanına sahip olmalıdır.

3.9. Yazılım Kalite Faktörleri

Sistem aşağıdaki yazılım kalite faktörlerini gerçekleştirecek şekilde geliştirilmelidir.

3.9.1. Sayısal Gereksinimler

3.9.1.1. Sistem, fotoğrafın yüklenmesi işlemini kullanıcının da internet hızına ve yüklediği fotoğrafın boyutuna göre 15 saniyeden kısa bir süre içerisinde tamamlamalıdır.

3.9.1.2. Sistem, yüklenen fotoğrafın analizini 15 saniyeden kısa bir süre içerisinde gerçekleştirmelidir.

3.9.1.3. Sistem, 3 megabayttan daha fazla boyutta bir fotoğrafı kabul etmemelidir.

3.9.2. Kullanılabilirlik

Sistem, her kullanıcı tarafından kolayca kullanılma yeteneğine sahip olmalıdır.

3.9.3. Güvenilirlik

Sistem tarafından üretilen çıktı algoritmanın sonuçlarına göre doğru ve tutarlı olmalıdır.

3.9.4. Erişilebilirlik

Sistem, her an kullanıcı tarafından kullanıma açık halde olmalıdır.

3.9.5. Esneklik

Sistem, olası güncellemelere kolayca adapte olmalıdır.

3.9.6. Test Edilebilirlik

Sistem, veri setinden elde ettiği sonuçlara dayanarak çıktıyı test edebilmelidir.

3.9.7. Taşınabilirlik

Sistem, gereksinimleri karşılayan her mobil cihazda kullanılabilir durumda olmalıdır.

3.10. Tasarım Görünümleri

Projede yapısal olarak modüler bir tasarım uygulanacaktır. Nesne yönelimli programlama ilkelerine dayanarak birbirinden etkilenmeyen ve birbirleriyle uyumlu çalışabilen yapılar inşa edilecektir. Bu sayede olası hataların çözümü kolaylaşacaktır.

3.11. Tasarım Bakış Açıları

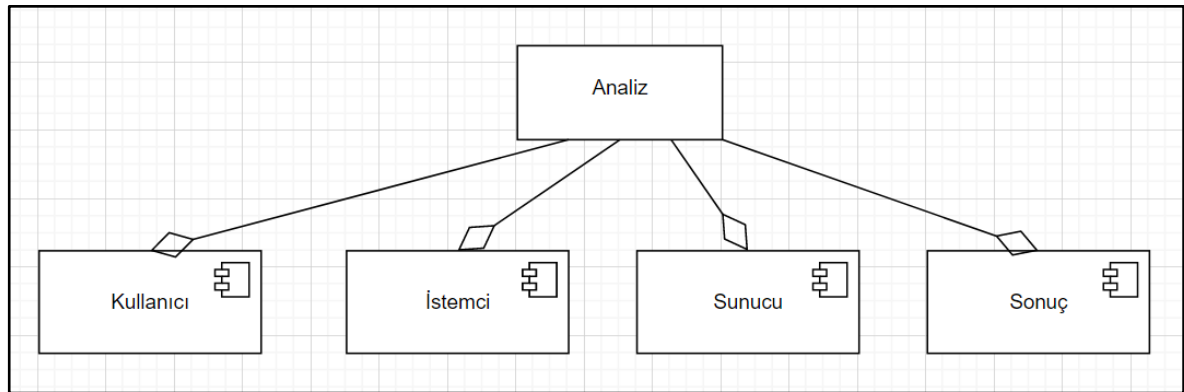
Bu bölümde, her tasarım bakış açısı kısaca açıklanmaktadır.

- Bağlam bakış açısı, kullanıcılar ve sistem arasındaki ilişkileri ve etkileşimleri tanımlar. Her işlevi temsil etmek için use-case diyagramları kullanılır.
- Kompozisyon bakış açısı, uygulamanın ana yapısını açıklar. Sistemin bileşenleri arasındaki etkileşimleri gösterir. Genel sistem mimarisi, bileşen diyagramı kullanılarak gösterilmiştir.
- Bağımlılık bakış açısı, sistemdeki bileşenlerin birbirleri arasındaki kapsamlarını belirtir.
- Mantıksal bakış açısı, temel olarak majör arabirimleri ve bu arabirimlerin arasındaki iletişimi temsil eder. Bu bakış açısı sayesinde projenin büyük parçalarına genel bir bakış yapılmış olur.
- Arayüz bakış açısı, sistemde kullanılan arayüzler açıklanır.
- Etkileşim bakış açısı, her kullanıcı işlemi için etkileşimleri ve ilişkileri tanımlar. Bu ilişkileri temsil etmek için sıra diyagramı kullanılır.

3.12. Tasarım Endişeleri

Sistem üzerinde geliştiricilerin çeşitli tasarım kaygıları vardır. Bunların başında teslim tarihine dayalı olarak projenin çalışır versiyonunun ilgili tarihten önce yayınlanmış olması yer alır. Bir diğer endişe fotoğraf analiz algoritmalarının performansıdır. Bu algoritmaların doğruluk oranlarının olabildiğince yüksek olması hedeflenmektedir. Bu da iyi bir önışleme ve eğitim aşamaları ile sağlanmaya çalışılacaktır.

3.13. Tasarım Öğeleri



Şekil 2. Tasarım Öğeleri Bileşen Diyagramı

Analiz

- Tür: Sistem
- Açıklama: Bu diyagramda analiz bileşeni, sistemin temel amacını temsil eder. 5 ana bileşenden oluşmaktadır. Bunlar kullanıcı, istemci, sunucu ve sonuç bileşenleridir.

Kullanıcı

- Tür: Bileşen

- Açıklama: Sistemin hizmet edeceği temel bileşendir. Kullanıcı istemci aracılığıyla sunucudan analiz hizmeti ister. Analizin sonucunu istemci üzerinden alır.

İstemci

- Tür: Bileşen
- Açıklama: Sistemde kullanıcı ile sunucu arasında iletişimi sağlayan köprü niteliğindeki bileşendir. Kullanıcının isteğini alıp sunucuya, sunucudan gelen sonucu ise kullanıcıya iletir.

Sunucu

- Tür: Bileşen
- Açıklama: Sistemde en komplike bileşendir. Görselin analizini sağlayan yapay zeka unsurlarının tamamı bu bileşenin içerisindedir. Geliştiriciler en çok bu bileşenin arka planındaki algoritmalar üzerinde çalışır.

Sonuç

- Tür: Bileşen
- Açıklama: Sunucuda algoritmaların gerçekleştirdiği analizlerin sonucudur. Kullanıcının ulaşmak istediği bileşendir.

3.13.1. Tasarım Varlıkları

- Kullanıcı
- Görsel
- Mesaj

3.13.2. Tasarım Özellikleri

Varlık ismi	Açıklama
Kullanıcı	Görseli yükleyen kullanıcının bilgilerini kapsar.
Görsel	Analizi yapılacak görselin bilgilerini kapsar.
Mesaj	Analiz sonucunda kullanıcıya verilecek çıktının bilgilerini kapsar.

3.13.3. Tasarım Kısıtlamaları

Sahte fotoğraf analiz uygulaması kullanıcının yüklediği görsellerin herhangi bir sistemde saklanmadığını bildirir. Kullanıcının yüklediği görseller, 3.bir şahıs ya da kurumla paylaşılmayacağı bildirilir. Ayrıca analiz sonuçları yalnızca kullanıcıya ilgili ekranda paylaşılır ve bu sonuçlar da görseller gibi uygulama kapsamında saklanmayacaktır. Ek olarak kullanıcıdan herhangi bir bilgi alınmamakta ve cihazındaki etkinlikleri de herhangi bir amaçla takip edilmemektedir. Özetle sahte fotoğraf analiz uygulaması, sadece amacına yönelik işlemleri gerçekleştireceğini ve güvenilir bir şekilde kullanıma uygun olduğunu beyan eder.

3.14. Tasarım Katmanları

Tasarım Bakış Açıları bölümünde, mevcut tüm bilgiler açıklanır ve bu bölümde sunulacak ek bir bilgi yoktur.

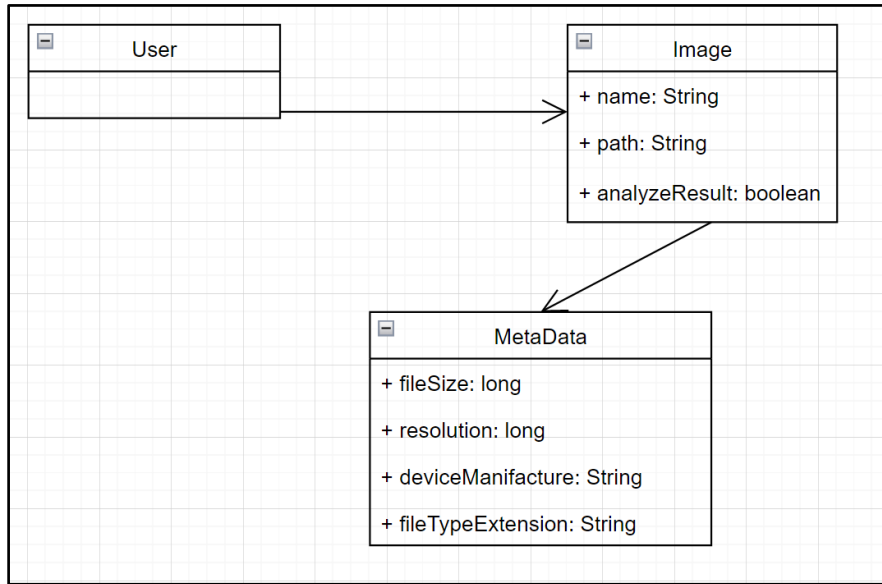
3.15. Tasarım Gerekçesi

Nesne yönelimli tasarım ile nesnelerin sınıflandırılması gerçekleştirilir. Bu sayede sisteme sonradan ekleme ve çıkarmalar yapılabilir. Bu da programda uygulanabilirliği kolaylaştırır. Sistemin uygulanması için test odaklı geliştirme seçilmiştir. Yazılan test senaryolarıyla kodun netliği sağlanmıştır. Olası hata durumlarında yapılacaklar planlanmıştır. Değişken ve sınıf isimleri ileride yapılacak değişikliklerde kolaylık açısından özenle seçilmiştir.

3.16. Tasarım Dilleri

Sistemin tasarlanmasında UML (Unified Modeling Language) dili kullanılmıştır. Yazılım mühendisliği alanında geniş çapta kabul görmeleri ve geliştiriciler arasında yazılım tasarım kavramlarını iletmedeki etkinlikleri nedeniyle UML seçilmiştir.

3.17. Tasarım Bakış Açıları

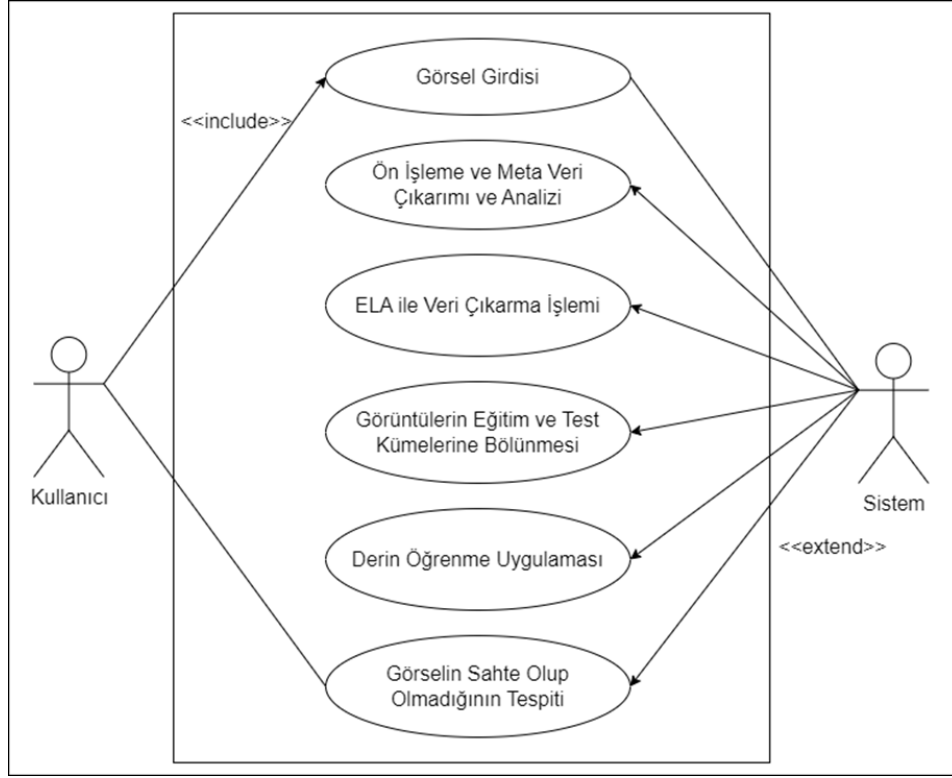


Şekil 3. Sistem Sınıf Diyagramı

Yukarıdaki sınıf diyagramı sistemin temel nesnelerini ve aralarındaki ilişkileri temsil etmektedir. Genel olarak işleyiş kullanıcının sahip olduğu bir görseli sistem aracılığıyla analiz etmesi üzerine gerçekleşmektedir. Görsel sınıfının isim, yol ve analiz sonucu gibi temel özellikleri bulunmaktadır. Aynı zamanda görsele yapılan ilk analizde metadata bilgileri ortaya çıkar. Bunlardan bazıları dosya boyutu, çözünürlük, cihaz üreticisi ve dosya uzantısıdır. Bunlar metadata nesnesinin özellikleridir.

3.18. Bağlam Bakış Açısı

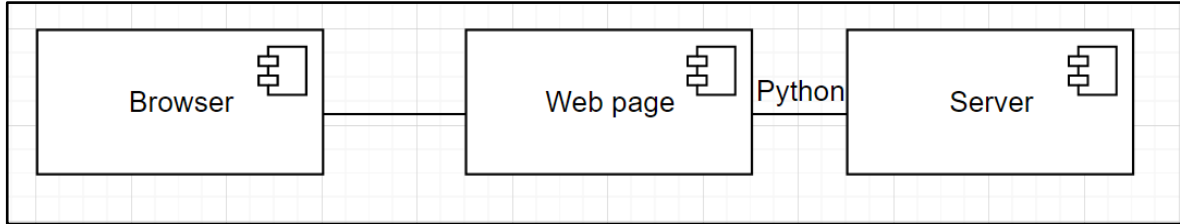
Sistemde tek tip kullanıcı bulunmaktadır. Sistem, kullanıcılara sahte fotoğraf analizi hizmeti sunar. Kullanıcı bu hizmeti aynı anda tek bir görsel için kullanabilir. Kullanım durumlarıyla ilgili daha ayrıntılı bilgi, SRS dokümanında belirtilmiştir.



Şekil 4. Use-Case Diyagramı

3.19. Kompozisyon Bakış Açısı

Bileşenler arasındaki mantıksal ilişki, aşağıdaki bileşen şemasında gösterilmiştir.



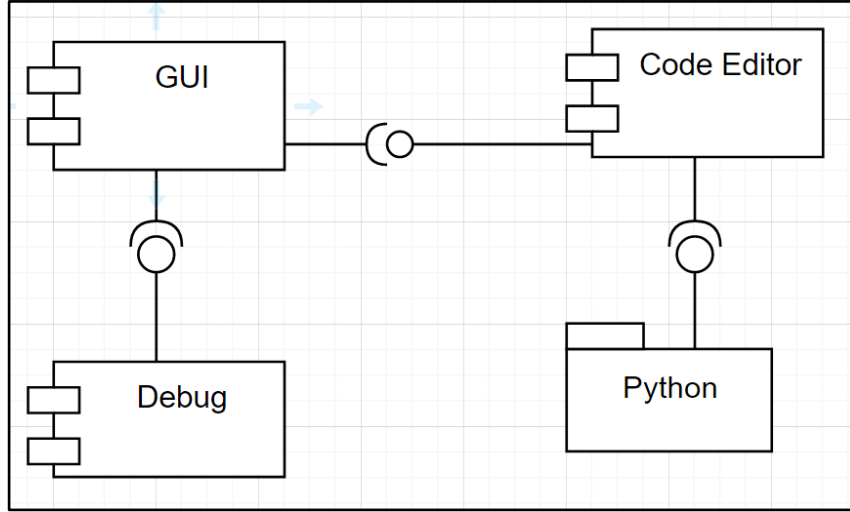
Şekil 5. Kompozisyon Bileşen Diyagramı

Şemaya göre tarayıcı bileşeni istenilen mesajı web sayfasına iletir. Web sayfası Python programlama dilini kullanarak mesajı sunucuya aktarır. Analiz sonucunda elde edilen mesaj aynı yol üzerinden kullanıcıya dönüş yapar.

3.20. Mantıksal Bakış Açısı

Sistem, alt sistemlere sınıflar aracılığıyla parçalanmıştır. Her sınıfın yerine getirdiği işlev aracılığıyla sistem işlevini yerine getirir.

3.21. Bağımlılık Bakış Açısı



Şekil 6. Bağımlılık Bileşen Diyagramı

Kod Editörü, programın GUI' sinin bir parçasıdır, bu nedenle bileşen şemasında GUI' ye arayüz olarak temsil edilir. Kod Editöründe gerçekleştirilen komut dosyasını çalıştırmak gibi eylemlerin sonuçlarının GUI' nin diğer bölümlerini etkileyeceğini belirtmek gerekir. Python, programlama dili kullanılarak kod editöründe programın kodları yazılmıştır.

3.22. Bilgi Bakış Açısı

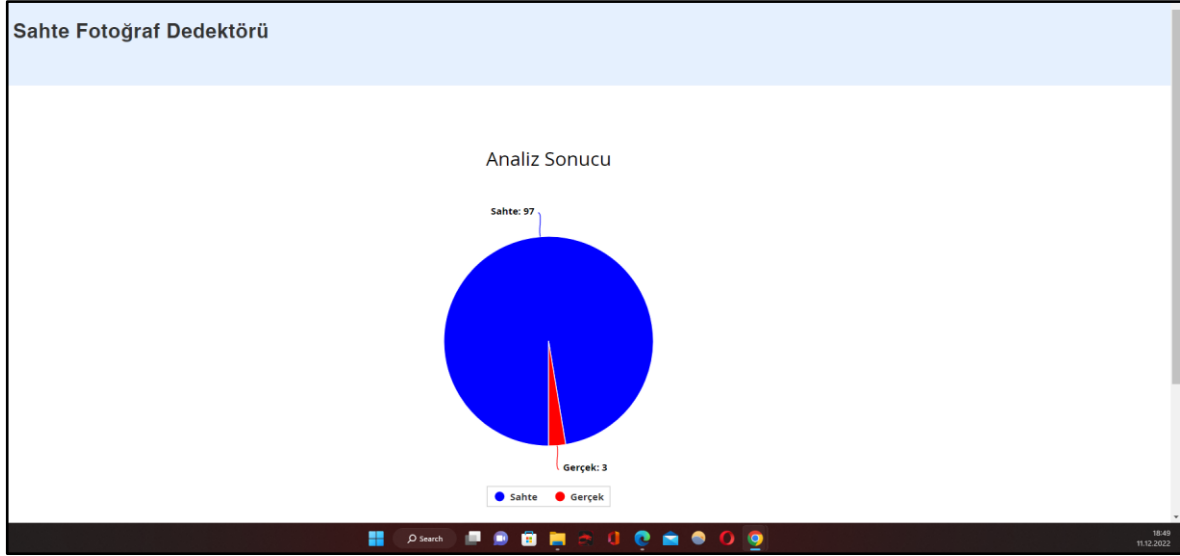
Herhangi bir bilgi sisteminin temel amacı verileri manipüle etmektir. Sistemde, kullanıcıdan herhangi bir bilgi alınmamakla birlikte kullanıcının bilgilerine erişim sağlanmamaktadır. Bu yüzden, sistemde bir bilgi depolama sistemi bulunmamaktadır. Sistemde kullanıcıdan alınan görsel, anlık olarak işlenip sonuç ekrana yansıtılır.

3.23. Arayüz Bakış Açısı



Şekil 7. Temsili Arayüz 1

Yukarıdaki temsili görselde sahte fotoğraf dedektörü web sayfasının giriş penceresinin bir görüntüsü bulunmaktadır. Kullanıcı yükle butonuna tıklayarak seçtiği görseli görüntüler ve yükler. Ardından analiz işlemine geçilir.

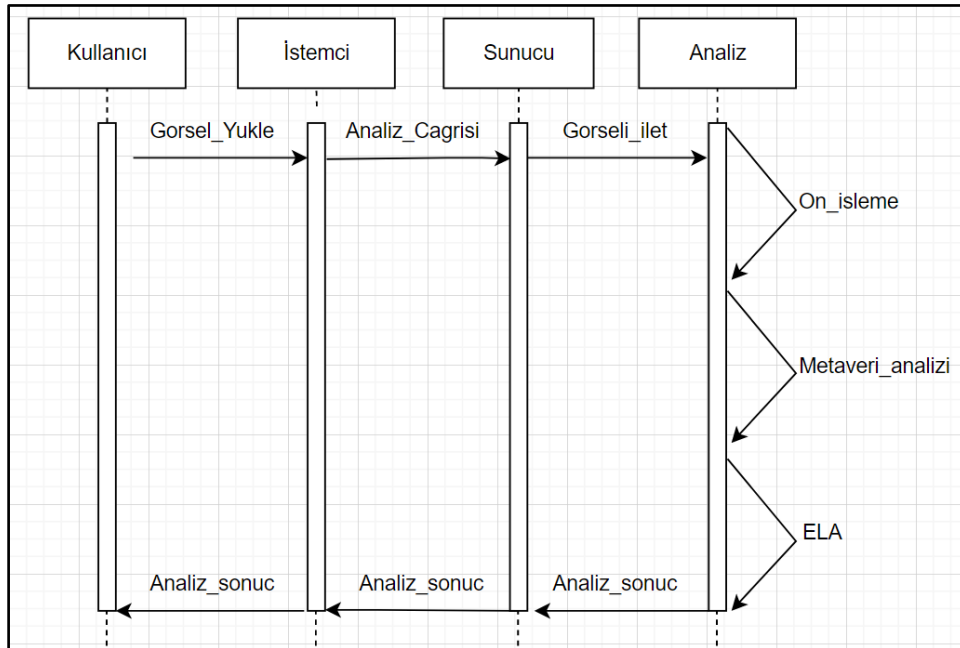


Şekil 8. Temsili Arayüz 2

Sunucuya iletilen görselin arka planda analiz işlemlerini gerçekleştirilir. Gerçekleştirilen bu işlemler sonucunda yukarıdaki ekrana görselin analiz sonucu ve doğruluk oranının pasta dilimi grafiği yansıtılır. Kullanıcı tekrar analiz işlemi yapmak istersen önceki sayfaya dönüş yapabilir.

3.24. Etkileşim Bakış Açısı

Bu bölümde, sistem kullanıcılarının sistem ve sistem nesneleri ile etkileşimi bulunur. Sistemin işlevlerinin yerine getirilmesinde nesneler arasındaki mesajlar görev alır.



Şekil 9. Sıra Diyagramı

4. SONUÇ VE ÖNERİLER

Proje 2 öğretim dönemini kapsayan bir projedir. Projenin ilk döneminde araştırmalar, literatür taramaları, gereksinim belirlemeleri ve planlamalar gerçekleştirilmiştir. Ayrıca kullanılacak veri kümesi olan CASIA veri kümesi belirlenmiştir. Bu veri seti üzerinde araştırmalar gerçekleştirilmiştir. Veri ön işleme operasyonları gerçekleştirilmiştir. Hata seviyesinde analiz(ELA) işlemi tamamlanmıştır.



Şekil 10 ve 11-CASIA2 Datasetinden gerçek bir fotoğrafın ELA öncesi ve sonrası görüntüler



Şekil 12 ve 13-CASIA2 Datasetinden sahte bir fotoğrafın ELA öncesi ve sonrası görüntüler



Şekil 14 ve 15-İnternette alınmış gerçek bir görselin ELA öncesi ve sonrası görüntüler



Şekil 16 ve 17-Photoshop üzerinden değiştirilmiş bir görselin ELA öncesi ve sonrası görüntüler

Diğer proje döneminde ise meta veri analizi ve CNN aracılığıyla tahmin işlemleri yapılacaktır. Ayrıca projenin platformlara uyumlu hale getirilmesi de planlanmaktadır. Derin sahte tespiti üzerine daha önceden gerçekleştirilen çalışmalar araştırılmıştır. Bu araştırmalardan elde edilen sonuçlar aşağıdaki gibi özetlenmiştir:

Derin Sahte Tespit Yöntemi - Başarı Oranı Tablosu			
Yöntem	Başarı Oranı % (AUC)	Yöntem	Başarı Oranı % (AUC)
VGG-16	%91.97	Alexnet(RELU)	%99.3
VGG-19	%92.09	Alexnet-TL(RELU)	%94
DenseNet169 + Rayleigh blur	%60.1	CNN(RELU)	%83.9
LRCN(Eye blanking)	%99	DenseNet121 & ResNet50	%97
CNN(Eye blanking)	%98	VGGFace	%99.8
EAR(Eye blanking)	%79	DenseNet201	%96

Şekil 10. Derin sahte yöntemi modellerinin başarı oranı tablosu

CNN temelli bir önermeye göre VGGFace mimarisi %99.8 gibi zirve başarı oranlarını elde etmektedir. Resim dosyalarının Meta verilerine odaklanan bir önerme ise ELA aracılığıyla görselin her kaydedilişinde oluşan kalite farkına dayanır. Sinir ağı CASIA veri setiyle eğitilir ve %83 başarı oranı yakalar.

NIST MFC2018 veri setiyle renk odaklı çalışmada %70 başarı oranı sağlanırken, yüz değişimlerine bakıldığında %85'e yakın bir oran elde edilmiştir. Kontrast kaybına odaklanan bir çalışmada ise GAN mimarisi kullanılıp DenseNet aracılığıyla ağ yapıları oluşturulmuştur. Katmanlı aşamalardan sonra DeepFD uygulanıp çoğu yönetime kıyasla yüksek başarı oranları elde edilmiştir. Fakat eğitim örneklerinin toplanmasının zorluğu da bu yöntemi pratikte dezavantajlı hale getirmektedir.

Referanslar

1. Qurat-ul-ain, Nida, N., Irtaza, A., & Ilyas, N. (2021). Forged Face Detection using ELA and Deep Learning Techniques. 2021 International Bhurban Conference on Applied Sciences and Technologies (IBCAST). <https://doi.org/10.1109/ibcast51254.2021.9393234>

2. Maksutov, A. A., Morozov, V. O., Lavrenov, A. A., & Smirnov, A. S. (2020). Methods of Deepfake Detection Based on Machine Learning. *2020 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (EIConRus)*. <https://doi.org/10.1109/eiconrus49466.2020.9039057>
3. Yuezun Li, Ming-Ching Chang, & Siwei Lyu. (2018). In Ictu Oculi: Exposing AI Generated Fake Face Videos by Detecting Eye Blinking. *ArXiv: Computer Vision and Pattern Recognition*. <http://export.arxiv.org/pdf/1806.02877>
4. Tanaka, M., Shiota, S., & Kiya, H. (2021). A Detection Method of Operated Fake-Images Using Robust Hashing. *Journal of Imaging*, 7(8), 134. <https://doi.org/10.3390/jimaging7080134>
5. AlShariah, N. M., & Khader, A. (2019). Detecting Fake Images on Social Media using Machine Learning. *International Journal of Advanced Computer Science and Applications*, 10(12). <https://doi.org/10.14569/ijacsa.2019.0101224>
6. Shad, H. S., Rizvee, M. M., Roza, N. T., Hoq, S. M. A., Monirujjaman Khan, M., Singh, A., Zaguia, A., & Bourouis, S. (2021). Comparative Analysis of Deepfake Image Detection Method Using Convolutional Neural Network. *Computational Intelligence and Neuroscience*, 2021, 1–18. <https://doi.org/10.1155/2021/3111676>
7. Villan, M., Kuruvilla, A., Paul, J., & Elias, E. (2017). Fake Image Detection Using Machine Learning. *IRACST-International Journal of Computer Science and Information Technology & Security (IJCSITS)*.
8. Ali, S. S., Ganapathi, I. I., Vu, N. S., Ali, S. D., Saxena, N., & Werghi, N. (2022). Image Forgery Detection Using Deep Learning by Recompressing Images. *Electronics*, 11(3), 403. <https://doi.org/10.3390/electronics11030403>
9. Ahmed S Abdulreda, & Ahmed J. Obaid. (2022). A landscape view of deepfake techniques and detection methods. *International Journal of Nonlinear Analysis and Applications*, 13(1), 745–755. <https://doi.org/10.22075/ijnaa.2022.5580>
10. Hsu, C. C., Zhuang, Y. X., & Lee, C. Y. (2020). Deep Fake Image Detection Based on Pairwise Learning. *Applied Sciences*, 10(1), 370. <https://doi.org/10.3390/app10010370>