YAZILIM MÜHENDİSLİĞİ
GÖRÜNTÜ İŞLEME

ÖZGE GÜRBÜZ

MAYIS , 2023

1.Giriş

Dijital manipülasyonlar sonucunda oluşturulan sahte görüntü ve video içeriklerinin sayısı giderek artmaktadır. Dijital manipülasyonlar fotoğraf ve videolarda başlangıçta insanları eğlendirmek amacıyla geleneksel yöntemler kullanılarak yapılmaktaydı. Günümüzde ise dijital manipülasyonlar için makine öğrenmesi algoritmaları kullanılmaya başlanmıştır. Makine öğrenmesi algoritmalarının kullanılması ile gerçek içerikler ile sahte içeriklerin ayırt edilebilmesi oldukça zorlaşmıştır [1,8].

Yüz manipülasyonları, dijital manipülasyonlar arasında en yaygın türdür. Yüz manipülasyonları ile kişiler hiç bulunmadıkları bir yerde gösterilebilir veya hiç yapmadıkları konuşmalar yaptırılabilir. Yüz manipülasyonları tüm yüz sentezi, kimlik değiştirme, nitelik manipülasyonu ve ifade değiştirme olmak üzere dört temel gruba ayrılır [2,8].

Bu proje,transfer öğrenme, evrişimli sinir ağları ve sahte yüz algılama gibi konular üzerine yapılan araştırmaların bir örneğini temsil eder.Böylece dijital manipülasyonlar sonucunda oluşturulan sahte yüz görüntüleri ayrıştırılır.

2.Literatür Özeti

Literatürde tüm yüz sentezi manipülasyon yöntemi için hafif derin öğrenme temelli bir çalışma bulunmamaktadır. Tüm yüz sentezi manipülasyon tespiti için MobileNet, MobileNetV2, EfficientNetB0 ve NASNetMobile evrişimsel sinir ağları kullanılmıştır. Modeller ImageNet veri seti üzerinde ön eğitilmiş olarak transfer öğrenme ile tekrar kullanılmıştır.

3.Materyal ve Metot

Bu bölümde metotlar ve materyaller hakkında bilgiler verilmektedir.VGG16 ve VGG19 CNN ağ modeli import edilmiştir.**VGG16** ve **VGG19** CNN ağ modelinin **ImageNet** veri seti ile eğitildiği ağ ağırlıkları modele yüklemiştir.

Bu çalışma, transfer öğrenme kullanarak önceden eğitilmiş bir evrişimli sinir ağı modelini (Convolutional Neural Network - CNN) eğitmek için oluşturulmuştur. Transfer öğrenme, bir önceden eğitilmiş modelin özellik çıkarmak için kullanılması ve ardından çıkarılan özelliklerin sınıflandırma için yeni bir sinir ağına beslenmesini içerir.

VGG16 ve VGG19 modelleri kullanılarak önceden eğitilmiş bir CNN modeli olan conv_base oluşturulur. Bu model, ImageNet veri kümesi üzerinde eğitilmiştir. Modelin özetini göstermek için summary() fonksiyonu kullanılır.

Daha sonra, hangi katmanların eğitileceği ve dondurulacağı belirlenir. 'block5_conv1' katmanına kadar olan tüm katmanlar dondurulur. Ardından, bir boş model (model) oluşturulur.

Oluşturulan conv_base modeli, model'e eklenir. Ardından, Flatten() katmanı kullanılarak matrisler vektörlere dönüştürülür. Ardından, bir gizli katman (Dense) ve çıkış katmanı eklenir. Gizli katmanda ReLU aktivasyon fonksiyonu kullanılırken, çıkış katmanında softmax

aktivasyon fonksiyonu kullanılır. Bu, modelin iki sınıf olan "fake" ve "real" sınıflarını sınıflandırmasını sağlar.

Model derlenirken binary_crossentropy(ikili çapraz entropi) kayıp fonksiyonu ve RMSprop optimizer kullanılır. Ayrıca, eğitim sırasında doğruluk (accuracy) metriği izlenir.

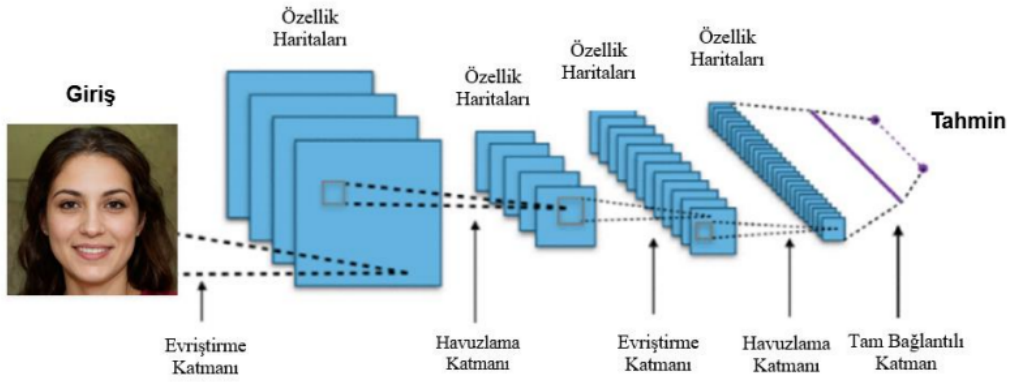
Overfitting'i önlemek için eğitim verilerine veri artırma işlemleri uygulanır. Eğitim verileri, ImageDataGenerator kullanılarak ölçeklendirilir. rescale, rotation_range, width_shift_range, height_shift_range, shear_range, zoom_range, horizontal_flip gibi parametrelerle çeşitli veri artırma teknikleri uygulanır.Benzer şekilde, doğrulama verileri de ölçeklendirilir.

Model eğitimi fit() fonksiyonu kullanılarak gerçekleştirilir. Eğitim ve doğrulama veri akışları ile adım sayıları (steps_per_epoch, validation_steps) ve epoch sayısı belirtilir. Eğitim sırasında elde edilen kayıp ve doğruluk değerleri history değişkenine kaydedilir ,grafiklerle görselleştirilir.

Eğitim tamamlandıktan sonra, eğitilen model 'fakeFaceDetection.h5' adıyla kaydedilir ve test verileri üzerinde değerlendirme yapılır.Test verileri için de ImageDataGenerator kullanılır ve rescale parametresi kullanılarak piksel değerleri 0-1 aralığına ölçeklenir, bir veri akışı (test_generator) oluşturulur.evaluate() fonksiyonu kullanılarak kayıp ve doğruluk değerleri elde edilir.

Aynı zamanda eğitim sürecini izlemek için grafikler oluşturulur. İlk grafik, eğitim doğruluğunu ve doğrulama(geçerleme) doğruluğunu epoch sayısına göre gösterir. İkinci grafik ise eğitim kaybını ve doğrulama kaybını epoch sayısına göre gösterir (Şekil 8 ve 10).

Son olarak, tahmin işlemi gerçekleştirilir. Tahminler, model.predict() fonksiyonu kullanılarak elde edilir ve ardından argmax() fonksiyonu kullanılarak en yüksek olasılığa sahip sınıflar belirlenir. Bu tahminler ve gerçek etiketler arasındaki karşılaştırmalar için Confusion Matrix hesaplanır ve görselleştirilir.



Şekil 1. Evrişimsel sinir ağının çalışması [8]

3.1.Görüntü İşleme

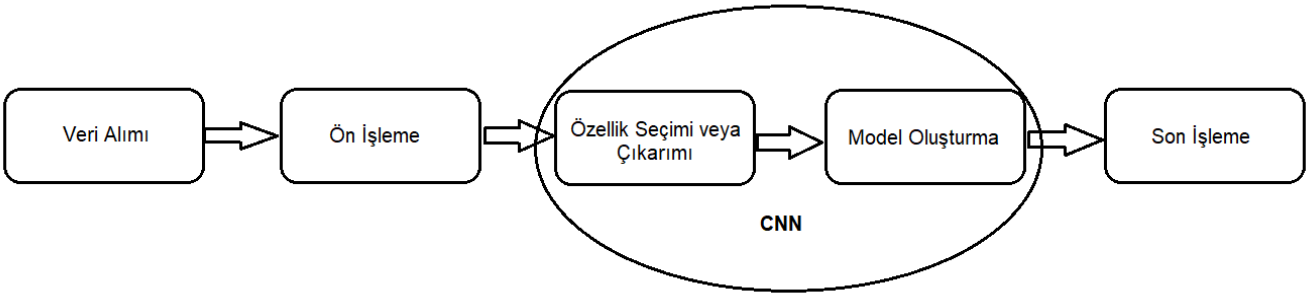
Görüntü işleme, sinyal işlemenin önemli alanlarından biridir. Gerçek yaşamdaki belirli görüntülerin bazı aşamalardan geçerek, bu görüntülerden anlamlı bir bilgi çıkarımı yapılmasıdır [3,4,5].

Bu bağlamda Şekil 2'de gösterildiği gibi, görüntü verileri bazı temel aşamalardan geçer. Bu aşamalar ise; görüntü elde etme, ön işleme, geliştirme ve görüntüleme, bilgi çıkarımıdır [6]. Bu aşamaların ardından ilgili giriş verilerinden ihtiyaca göre sonuçlar elde edilebilmektedir [7].

Şekil 2. Görüntü işleme temel aşamaları [6]

Temel olarak kamera ve algılayıcılar vasıtasıyla elde edilen görüntüde ilgili nesne ve örüntülerin tanınabilmesi için genel bir şablon olarak izlenmesi gereken basamaklar ise şunlardır:

- Ön İşleme
- Özellik Seçimi veya Çıkarımı
- Eğiticili (supervised) sınıflandırma veya eğiticisiz (unsupervised) kümele yöntemleri kullanarak model altyapısının kurulumu
- Son İşleme



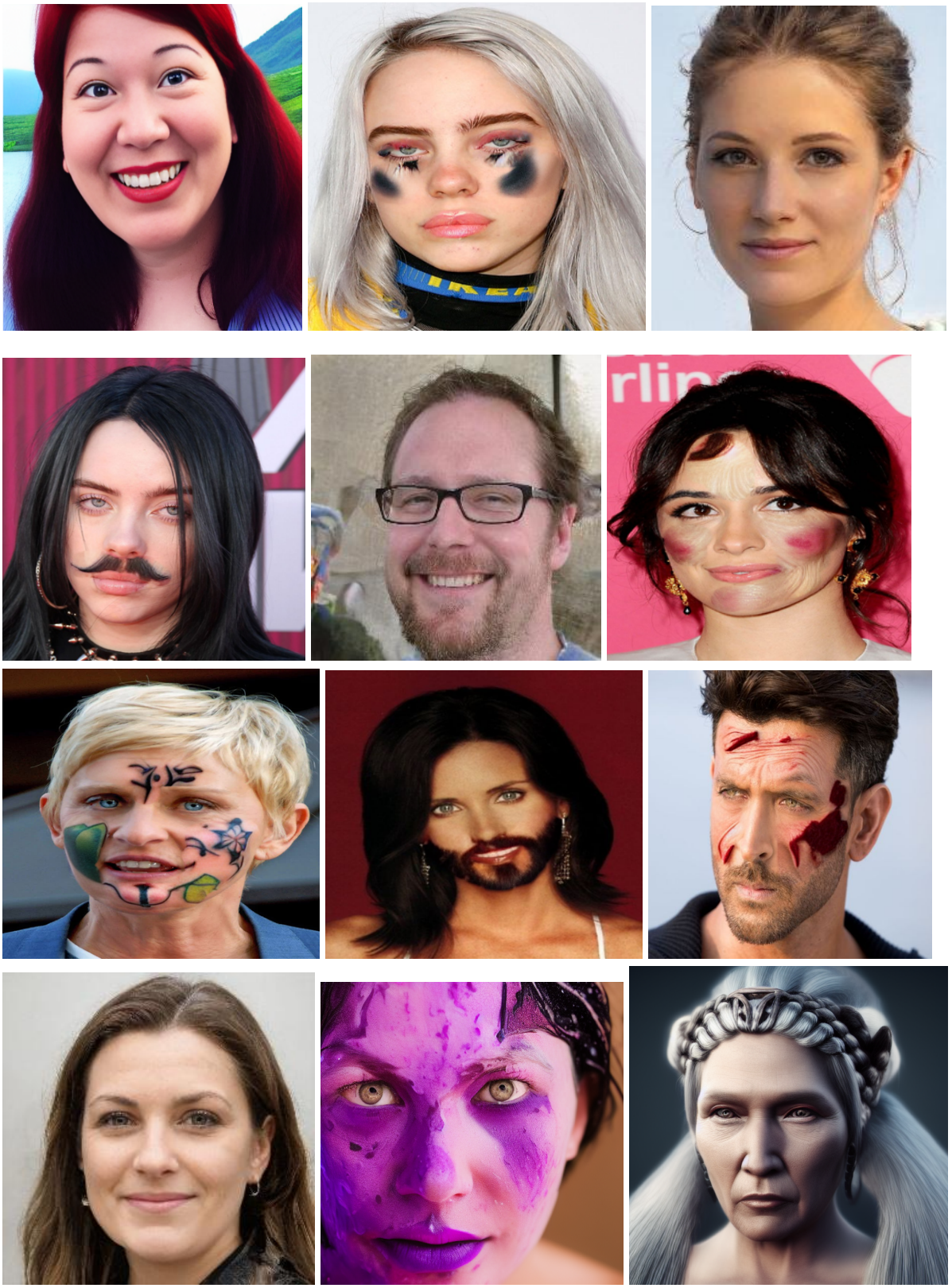
Şekil 3. Nesne ve örüntü tanıma için genel akış şeması

3.1.1. Veri Alma ve Oluşturma

Sahte yüz görüntülerini tespit edebilmek için evrişimsel sinir ağları kullanılmıştır. Evrişimsel sinir ağı modelini eğitmek için **4200** görüntüden oluşan veri seti kullanılmıştır. Kullanılan veri seti hazır Real and Fake Faces veri setlerindeki **2100** gerçek ve **2000** sahte yüz görüntülerinden oluşturulmuştur. **100** sahte yüz görüntüsü ise *Stable Diffusion** ile metinden resim ve resimden resim türünden üretilerek veri setine eklenmiştir. Bu veri seti çekişmeli üretici ağlar için hazırlanmış yaş ve farklı etnik köken çeşitliliğine sahip yüksek çözünürlüklü görüntülerden oluşan veri setidir.

Anita Rácz vd. [7] tarafından yapılan çalışmada %80 eğitim ,%20 test veri seti bölme oranında model performansının maksimuma ulaştığı görülmüştür. Bu nedenle öncelikle veri setindeki görüntülerin %80'i eğitim, %20'si test işlemi için kullanılmıştır [7]. Daha sonra ise ağın Overfitting yapmasını engellemek için veri setinin eğitim,test ve doğrulama(geçerleme) olmak üzere 3'e ayrılmasına karar verilmiştir. Buradaki ağırlıklandırma ise %64 eğitim ,%20 test ve %16 doğrulama olarak ayarlanmıştır.İlgili veri setlerinin bağlantıları kaynakça bölümünde verilmiştir.Veri setindeki bazı örnekler aşağıda Şekil 4 ve 5'te verilmiştir.

**Stable Diffusion, derin öğrenme altyapılı bir metin girişinden görüntüler oluşturan yapay zeka modelidir. Öncelikle metin açıklamalarına bağlı olarak ayrıntılı görüntüler oluşturmak için kullanılır, ancak görüntünün içeriğini değiştirmek ya da dışını genişletmek gibi diğer görevlerde de uygulanabilir. [9]*



Şekil 4. Veri setinde yer alan sahte yüz örnekleri

Şekil 5. Veri setinde yer alan gerçek yüz örnekleri

3.1.2. Ön İşleme

Hem eğitim hem de doğrulama verileri VGG16 ve VGG19 giriş katmanı boyutlarına uygun olacak şekilde 224x224 çözünürlüğünde ölçeklendirilmiştir (gerçek boyutlar 256*256). “batch_size” argümanı her yüklemede 32 resim çekilecek şekilde ayarlanmıştır.Ayrıca veri üreteçleri kullanarak piksel değerlerini 0-255'den 0-1 arasına getirme,verileri kırpma, zoom yapma,girişleri yatay ve dikey olarak rastgele çevirme,girdi sınırları dışındaki noktaları verilen moda göre doldurma,tersine çevirme gibi veri büyütme teknikleri kullanarak veri sayıları arttırılmış ve işlenmiştir.

3.1.3. Özellik Çıkarma

Özellik seçimi ve çıkarımında verinin ayrıştırıcılığını kaybetmeden veya mümkünse arttırarak olabildiğince az sayıda öznitelik kullanılarak sınıflandırma ve kümeleme başarımının arttırılması hedeflenir. Özellik çıkarma işlemi Derin Öğrenmede evrişim ve havuzlama işlemleri ile gerçekleştirilir.Transfer öğrenme modeli ile de bu işlemler direkt olarak alınabilir.Bu çalışmada kullanılan modeller ise VGG16 ve VGG19 ‘dır ve ImageNet ağırlıkları kullanılmıştır.

3.1.4. Model Oluşturma

Öncelikle oluşturulan boş modelin ilk katmanına özellik çıkarılmış katman eklenmiştir.Daha sonra Flatten(düzleştirme) işlemi yapılmıştır. Bu işlemle birlikte çok boyutlu veriler yan yana sıralanarak bir boyutlu vektör haline getirilmiştir.Sonrasında sınıflandırıcı nöronlar eklenir. En sonda ise iki sınıf olduğu ve sınıflandırma yapılacağı için softmax aktivasyon fonksiyonu ile 2 nöron kullanılır.Ve model derlenir(Derleme, kayıp işlevini, optimize ediciyi ve metrikleri tanımlar). Derleme işlemi tamamlandıktan sonra modelin eğitimi yapılmıştır.

Layer (type)	Output Shape	Param #
vgg16 (Functional)	(None, 7, 7, 512)	14714688
flatten_37 (Flatten)	(None, 25088)	0
dense_74 (Dense)	(None, 256)	6422784
dense_75 (Dense)	(None, 2)	514

Layer (type)	Output Shape	Param #
vgg19 (Functional)	(None, 7, 7, 512)	20024384
flatten (Flatten)	(None, 25088)	0
dense (Dense)	(None, 256)	6422784
dense_1 (Dense)	(None, 2)	514

Şekil 6. VGG16 ve VGG19 modelleri dahilinde oluşturulmuş modellerin özeti

3.1.5. Son İşleme

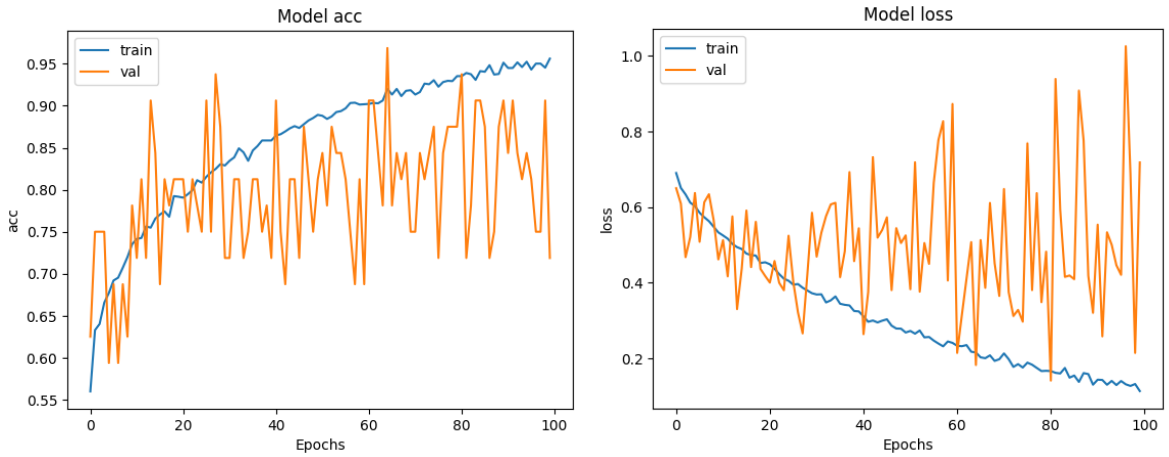
Elde edilen sınıflandırma veya kümeleme sonuçlarının iyileştirilmesi için yapılan her türlü işlem son işleme başlığı altında işlenebilir.

4.Deneyler Ve Bulgular

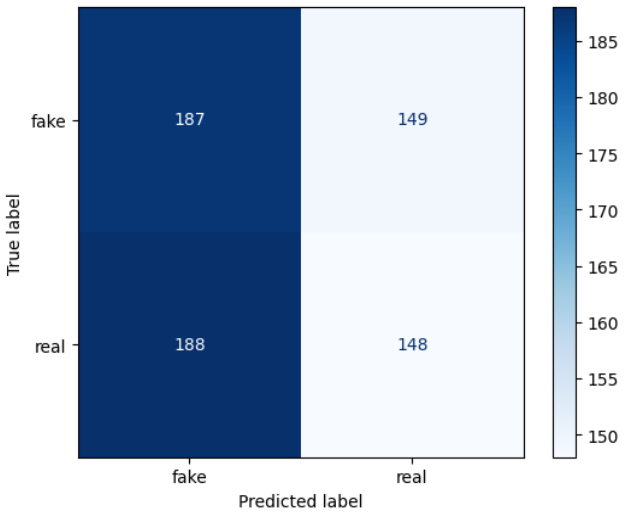
ImageNet ile ağırlıklandırılmış VGG16 ve VGG19 modelleri kullanımı sonucu toplam 4200 görüntüden, 100 epoch ile acc değerinde %89 ve %90 başarı elde edilmiştir. Bununla ilgili bazı grafikler aşağıdaki şekillerde verilmiştir.

	Gerçek Yüz Görüntüleri Sayısı			Sahte Yüz Görüntüleri Sayısı			Başarı (%)	
Model Adı	**Eğitim**	**Doğrulama**	**Test**	**Eğitim**	**Doğrulama**	**Test**	**Test Acc**	**Test Loss**
VGG16	1344	336	420	1344	336	420	78,9	54
VGG19	1344	336	420	1344	336	420	79,5	58

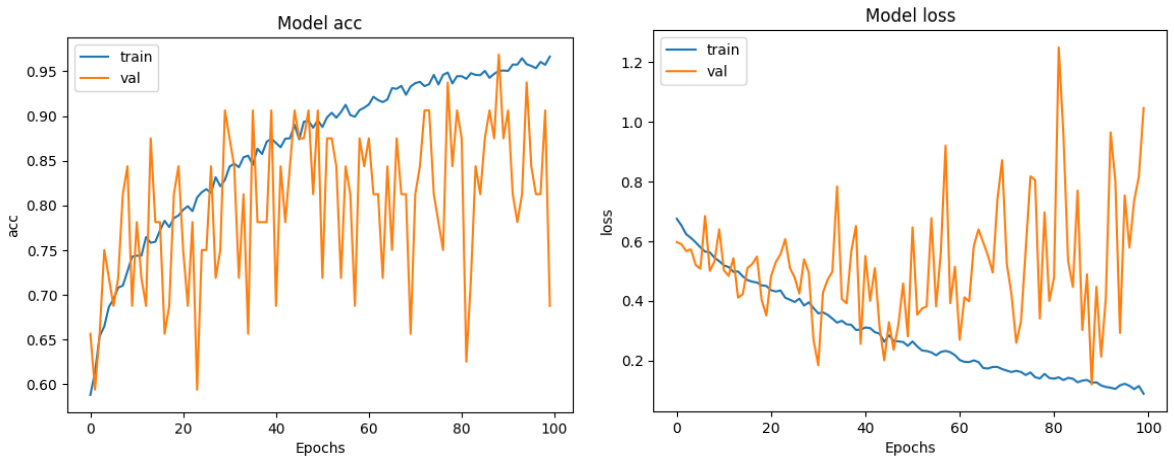
Şekil 7. Deneylerde kullanılan örnek sayıları ve başarı değerleri



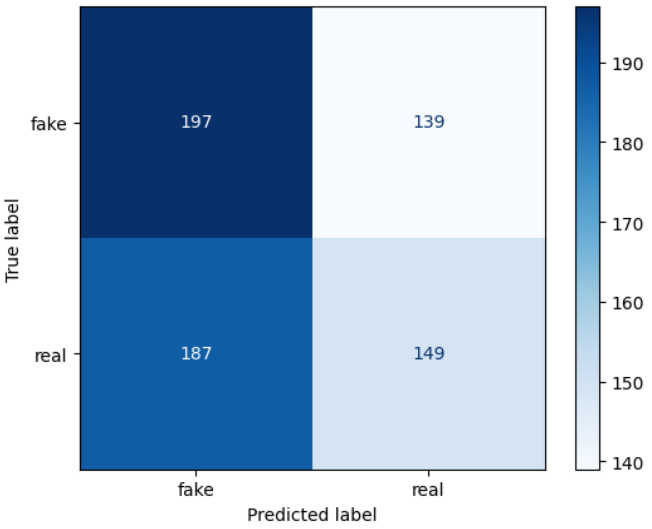
Şekil 8. VGG16 modeli ile eğitim sonucu elde edilen, eğitim ve doğrulama acc ve loss değerlerinin grafikleri



Şekil 9. VGG16 modeli eğitimi sonucu elde edilen karmaşıklık matrisi



Şekil 10. VGG19 modeli ile eğitim sonucu elde edilen, eğitim ve doğrulama acc ve loss değerlerinin grafikleri



Şekil 11. VGG19 modeli eğitimi sonucu elde edilen karmaşıklık matrisi

Bu çalışma i7-13700H 2.40 GHz işlemci ,16GB DDR4 bellek ve 6GB Nvidia RTX4050 ekran kartına sahip bilgisayarda gerçekleştirilmiştir.

5.Sonuçlar

Bu çalışmada gerçek ve sahte yüzleri birbirinden ayıran, iki sınıfta toplanan ve toplamda sadece 2100 adet gerçek ve sahte yüz görüntülerinden oluşan veri seti ile önceden eğitilmiş CNN ağlarını kullanarak bir sınıflandırıcı yapılmıştır.

VGG16 ve VGG19 modellerinin çıkışındaki katmanları değiştirerek yapılan sınıflandırıcı %80 üzerinde doğruluk değerine ulaşmıştır.

Sonuç olarak bu çalışma, sahte yüz algılama gibi uygulamalarda kullanılabilecek bir temel sağlar ve başka veri kümeleriyle veya farklı sınıflarla genişletilebilir.Hiper-parametreler ile optimizasyon sürdürüldüğünde bu doğruluğu arttırmak mümkün olabilir. Ayrıca, farklı önceden eğitilmiş CNN modelleri kullanılarak da deneyler yapılabilir ve performans karşılaştırmaları yapılabilir.

KAYNAKÇA

[1]. Pashine, S., Mandiya, S., Gupta, P. and Sheikh, R., “Deep Fake Detection : Survey of Facial Manipulation Detection Solutions”, arXiv preprint arXiv:2106.126, 2021.

[2]. Wang, R., Juefei-Xu, F., Ma, L., Xie, X., Huang, Y., Wang, J. and Liu, Y., “FakeSpotter: A Simple yet Robust Baseline for Spotting AI-Synthesized Fake Faces”, arXiv preprint arXiv:1909.06122, 2020.

[3]. Yaman, K., Sarucan, A., Atak, M., Aktürk, N., "Dinamik Çizelgeleme İçin Görüntü İşleme Ve Arıma Modelleri Yardımıyla Veri Hazırlama", Gazi Üniv. Müh. Mim. Fak. Der., Cilt:16, No:1, 19-40, Vol 16, 2001.

[4].Karakoç,M.,“Görüntü işleme, Teknolojiler ve Uygulamaları”, https://ab.org.tr/ab12/sunum/21- goruntu_isleme-Karakoc.pdf, Akademik Bilişim’12-XIV, Uşak, 1-3 Şubat 2012.

[5]. Öztaş, O., “Görüntü İşleme Teknikleri-I”, http://www.oguzhanoztas.com/gi/ders1.pdf, 12 Şubat, 2019.

[6]. DUMAN / INTERNATIONAL JOURNAL OF 3D PRINTING TECHNOLOGIES AND DIGITAL INDUSTRY (2019)

[7]. DUMAN, B. GÖRÜNTÜ İŞLEME TEKNİKLERİNİN EKLEMELİ İMALATTA KULLANIMI.

[8]. ŞAFAK, E., & BARIŞÇI, N. (2022). Hafif Evrişimsel Sinir Ağları Kullanılarak Sahte Yüz Görüntülerinin Tespiti. El-Cezeri, 9(4), 1282-1289.

[9] https://tr.wikipedia.org/wiki/Stable_Diffusion

Veri Setlerinin Bağlantıları

1. https://www.kaggle.com/datasets/xhlulu/140k-real-and-fake-faces
2. https://www.kaggle.com/datasets/uditsharma72/real-vs-fake-faces