All Sciences Proceedings

1st International Conference on Pioneer and Innovative Studies

June 5-7, 2023 : Konya, Turkey

ICPIS

All Sciences Proceedings
http://as-proceeding.com/

https://as-proceeding.com/index.php/icpis

Bilinen CNN mimarilerinin görsel Captcha sınıflandırması açısından değerlendirilmesi

Abdulmuttalip DURAN[1*] , Halit BAKIR [2 I], Hacı Mehmet GUZEY[3]

[1] *Savunma Teknolojileri ABD, Lisans Üstü Eğitim Enstüsü, Sivas Bilim ve Teknoloji Üniversitesi, Türkiye*
[2]*Bilgisayar Mühendislik Bölümü, Mühendislik Ve Doğa Bilimleri Fakültesi, Sivas Bilim ve Teknoloji Üniversitesi, Türkiye*
[3]*Elektrik Elektronik Mühendislik Bölümü, Mühendislik Ve Doğa Bilimleri Fakültesi, Sivas Bilim ve Teknoloji Üniversitesi, Türkiye*

[*]*(Abdulmuttalipduran58@gmail.com) Başlıca yazarın mail adresi*

Özet – İnternet siteleri erişim sağlayan kullanıcıların gerçek kişi mi yoksa robot mu olduğunu tespit etmek için captcha güvenlik sistemleri kullanmaktadır. Bu çalışmada derin öğrenme teknikleri kullanılarak görsel captcha'ların görsellerinin sınıflandırılması yapılmıştır. Projede Keras kütüphanesi altında bulunan tüm iyi bilinen CNN derin öğrenme modelleri test edilmiştir. İlk olarak tüm modeller kullanılarak analizler yapılmıştır. 95%'nin üzerinde başarı gösteren algoritmaların epoch değerleri artırılarak tekrar testler yapılmıştır. Testler sonucunda en başarılı sonuç 98.89% doğruluk değeriyle EfficientNetB1 ve EfficientNetB3 modellerinde elde edilmiştir.

Anahtar Kelimeler – Derin öğrenme- Güvenlik duvarı- Captcha- CNN mimarileri

I. GİRİŞ

Günümüzde gelişen teknoloji ile beraber internet kullanımı da hızla artmaktadır. İnternet kullanımı arttıkça faydalı içerik üreten siteler de artmaktadır. Bunun ile beraber bu siteleri suistimal etmek isteyen kesimler de ortaya çıkmaktadır [1]. İnternet sitelerine siber saldırılar, otonom bot saldırıları, dos saldırıları artmaktadır. Bu saldırıların artması ile birlikte günümüzde internet sitelerinin güvenliğini sağlamak da önemli bir konu haline gelmiştir [2]. İnternet sitelerinin güvenliğini sağlamak için captcha olarak adlandırılan kullanıcı doğrulama sistemleri ortaya çıkmıştır. Bu sistemler internet sitelerine erişim sağlayan cihazların gerçek kişi kontrolünde mi erişim sağladığını yoksa otonom olarak botlar tarafından mı erişim sağlandığını tespit ederek bot erişimlerini engellemektedir [3]. İnternet sitelerinin güvenliğini sağlamak için ilk başlarda metin ve ses tabanlı captcha doğrulama sistemleri kullanılmıştır. Metin tabanlı doğrulama sistemleri ilk olarak 1996 da çıkmıştır [3]. Bu sistemlerde kullanıcıya karışık şekilde yazılmış metinler veya sesler verilerek belirtilen kelimenin ne olduğunu kullanıcının yazması isteniyor [4]. Bu sistemler başlangıçta güvenli olmasına karşın ilerleyen süreçte kolaylıkla kırılabilir hale gelmiştir ve yetersiz kalmıştır. Bunun üzerine görüntü tabanlı captcha lar geliştirilimiştir bu captchalarda kullanıcıya bazı görseller verilerek bu görsellerden istenilen objelerin seçilmesi istenilmektedir [5]. Bu çalışmada captcha nın ne kadar güvenli olduğu denenecektir ve captcha görsellerindeki nesnelerin insanların yanı sıra makineler tarafından da görüntü analizi yapılarak tanınıp tanınmayacağı tespit edilecektir. Görüntü analizinde klasik makine öğrenmesi algoritmaları kullanmak için özellik vektörünün çıkartılması gerekmektedir. Bu işlem için ise alanında uzman kişilere ihtiyaç olmaktadır. Bu sebepten görüntü analizinde makine öğrenmesi modelleri kullanmak oldukça maliyetlidir. Bunun yanında derin öğrenme teknikleri görüntü analizi konusunda çok daha gelişmiştir. Derin öğrenme ilk defa 2012 yılında yapılan görsel tanıma yarışmasında elde ettiği başarı ile dikkatleri üzerine

[I] Khaled Bakour or Halit Bakır: Due to the author's dual citizenship, his name can be written in two different ways.

çekmiştir. Derin öğrenme modelinde makine öğrenmesinde olan özellik vektörünün çıkartılması gibi ön işlemlere ihtiyaç olmadan doğrudan ham veri üzerinde analizler yapılabilmektedir [6]. Bu çalışmada derin öğrenme teknikleri kullanılarak görsel captcha larda kullanılan nesne görsellerinin sınıflandırması yapılacaktır. Bu görsellerde kullanılan nesnelerin makineler tarafından ne düzeyde doğru sınıflandırılabileceği tespit edilecektir ve captcha nın ne düzeyde güvenli olduğu analiz edilecektir.

Literetür özeti

[3] Sivakorn ve ark. göre internet başlangıcından beri captchalar dolandırıcıların yasa dışı faaliyetlerini önlemek için kullanılmaktadır. Ancak günümüzde dolandırıcılar captchaları kırmak için yeni sistemler geliştirmeye başlamıştır. Sivakorn ve ark. Yaptıkları çalışmada Deep Learning kullanarak Image CAPTCHAs larını yenmeye çalışmışlardır ve Google nin sunduğu reCAPTCHA üzerinde %70.78, Facebook image CAPTCHA üzerinde ise %83.5 doğruluk değerinde başarı elde etmişlerdir.

[7] Akrout ve ark. Google recaptcha v3 güvenlik duvarını kırmak için bir takviyeli öğrenme yöntemini (RL) kullanmışlardır. Yaptıkları sistemde ekran çözünürlüğünü değiştirerek böl ve fethet stratejisi kullanmışlardır. 100×100 ekran çözünürlüğünde 97.7%, 1000×1000 ekran çözünürlüğünde 96.7% başarı elde etmişlerdir.

[8] L. Mookdarsanit ve P. Mookdarsanit' e göre derin öğrenme 2012'den beri nesne tanıma doğrultusunda büyük bir başarı katetmiştir. Bir bilgisayar korsanı tarafından Konvolüsyonel sinir ağı (CNN) yöntemi kullanılarak yapay zeka tabanlı botlar oluşturulmuştur ve bu botlar ile internet sitelerinde güvenlik duvarı olarak kullanılan recaptcha kırılmıştır. L. Mookdarsanit ve P. Mookdarsanit Yaptıkları çalışmada recapthca görüntü saldırıları için savunma mekanizması önermişlerdir. Önerdikleri modelde recaptcha görüntüsüne pertürbasyon (gürültü) eklemişlerdir. Daha sonra farklı CNN algoritmaları (R-CNN, SPPNet, Fast RCNN, Faster RCNN, FPN, YOLO, SSD) ve farklı mimariler (AlexNet, VGGNet, GoogLeNet, and ResNet) kullanılarak testler yapmışlardır. Yaptıkları test sonuçlarına göre geliştirdikleri model reCaptcha saldırısını %95'in üzerinde savunabilmiştir.

[9] Mittal ve ak. Göre güvenlik duvarı olan captcha tasarımcıları ve kırıcıları her zaman birbirlerini alt etmek için yeni yöntemler geliştirmeye çalışıyorlar. Mittal ve ak. Nın yapmış oldukları çalışmada captcha güvenlik duvarını kırmak için imagenet veri tabanı üzerinde eğitilmiş inception v3 modelini kullanmışlardır. Çalışma sonucunda 91% in üzerinde başarı elde etmişlerdir.

[1] Demirel ve kılınca göre insanlar tarih boyunca yeni şeyler ürettikçe bunları suistimal eden kesimler de ortaya çıkmaktadır. Aynı şekilde internette faydalı içerikler ve faydalı siteler oluştukça bu siteleri kırmaya çalışan bu sitelere karşı siber saldırı yapan kişiler de çoğalmaktadır. Captcha lar internet ortamında siber saldırılarını önlemek, otonom bot saldırılarını önlemek ve insan bot ayrımını yapmak için geliştirilmiştir.

[2] moradi ve keyvanpour'a göre günümüz internet dünyasında internet sitelerinin kabul edilebilir bir korunma düzeyine ulaşması gerekmektedir. Bunun için de internet sitelerine giren kişinin bilgisayar mı yoksa insan mı olduğunu anlamak için tamamen otomatik olarak çalışan captcha lar tasarlanmıştır. Bu captchaların temel görevi botların internet sitelerine girerek belirli eylemleri gerçekleştirilmesi önlemektir. Bu görevlere kaydolma, üye girişi, dosya indirme, dosya yüklemek gibi eylemler örnek verilebilir. Moradi ve keyvanpour yaptıkları çalışmada captchaların çeşitli yönleri, teknolojileri ve alternatifleri hakkında araştırma yapmışlardır.

[4] Alqahtani ve alsulaiman'a göre captchalar web sitelerinde kullanılan en yaygın kimlik doğrulama yöntemlerinden biridir. İlk olarak metin tabanlı ve sesli captcha lar kullanılmıştır bu yöntemlerin yetersiz olduğundan ötürü bu yöntemler yerini görüntü tabanlı captcha sistemlerine bırakmıştır. Alqahtani ve alsulaiman yaptıkları çalışmada görüntü tabanlı bir captcha kırma sistemi önererek capctha güvenlik duvarını kırmayı denemişlerdir. Yaptıkları çalışmada including random forest, CART, bagging with CART ve naive bayes gibi derin öğrenme tekniklerini ve makine öğrenimi

algoritmalarını kullanmışlardır. Yaptıkları proje sonunda recaptcha zorluklarının 56.29% unu başaryıla çözerken ortalama 85.32% başarı elde etmişlerdir.

[5] sukhani ve ark. Göre internet sitelerini Dos saldırılarından korumak için captchalar kullanılmaktadır. Metin tabanlı capchaları kırmak için çok sayıda çalışma yapılmıştır bu sebepten metin tabanlı capchalar daha geleneksel olup güvenliği düşük hale gelmiştir o sebepten metin tabanlı kapchalar yerini görüntü tabanlı captchalara bırakmıştır. Sukhani ve ark. Yaptıkları çalışmada konvolüsyonel sinir ağı (CNN) kullanarakile bu modeli otomatikleştirerek recaptcha v2 yi kırmayı denemişlerdir. Çalışma sonucunda 92.98% başarı elde etmişlerdir.

[10] Hossen ve ark. yaptıkları çalışmada Google Cloud Vision, Microsoft Azure, Computer Vision ve Amazon Rekognition object detection servislerini kullanarak reCAPTCHA v2 veri setinde nesne tespiti yapmışlardır. Çalışma sonucunda en yüksek başarıyı Amazon Rekognition servisi ile 47% oranında elde etmişlerdir.

[6] İnik ve Ülker'in yaptıkları çalışmaya göre görüntü analizinde klasik makine öğrenme tekniklerini kullanmak için alanında uzman kişiler tarafından bir takım ön işlemler yapılması gerekmektedir ve bu sebepten ötürü görüntü analizinde makine öğrenmesi kullanmak oldukça maliyetli ve zaman alıcı olmaktadır. Derin öğrenme ise bunun aksine görüntü analizini doğrudan ham veri üzerinden yapmaktadır. İnik ve Ülker yaptıkları çalışmada derin öğrenme hakkında detaylı bilgi vermişlerdir. Evrişimsel sinir ağı (ESA) katmanları olan konvolüsyon, havuzlama, relu, dropout hakkında açıklamalar yapmışlardır.

Ayrıca, iyi bilinen CNN derin öğrenme mimarileri, sağlık alanı başta olmak üzere farklı alanlarda kullanılmış olup yüksek performans elde etmiştir. Örneğin, [11] Bakır ve ark. Yaptıkları çalışmada, akciğer röntgen görüntülerinin derin öğrenme teknikleri kullanılarak analiz edildiğini ve akciğer hastalıklarının tespitinde etkili olduğunu göstermiştir. Çalışmada, yapay sinir ağı (YSA) modelleri ve evrişimli sinir ağı (CNN) mimarileri kullanılmıştır. En iyi performansı gösteren model, ResNet özellik çıkarma aşamasını kullanan önerilen YSA modelidir. Bu model, çok sınıflı sınıflandırmada %81,67 doğruluk oranı ve ikili sınıflandırmada %95,67 doğruluk oranı elde etmiştir. [12] Bakır ve ark. Yaptıkları çalışmada, retinal fundus görüntülerinden katarakt hastalığının tespiti için bir derin öğrenme modeli önerilmiştir. Önerilen model, ResNet kullanılarak eğitilmiş ve test edilmiştir. Sonuçlar, %95,51 gibi yüksek bir tespit doğruluğu elde edildiğini göstermektedir. Bu çalışma, katarakt hastalığının teşhisinde etkili ve başarılı bir yöntem olduğunu ortaya koymaktadır.

II. MATERYAL VE YÖNTEM

A. Materyal

İnternet siteleri, giriş yapan cihazların gerçek kişi tarafından mı yoksa otonom botlar tarafından mı sağlandığını tespit etmek için güvenlik önlemi olarak giriş yapılan cihazda insanların bilebileceği şekilde resimler göstermektedirler. Bu resimler araba, motor gibi çeşitli araçlar olabileceği gibi yangın musluğu, merdiven gibi çeşitli nesneler de olabilmektedir. Bu çalışmada veri seti olarak Kaggle den alınan 8 sınıf ve 4068 görselden oluşan recaptcha veri seti kullanılmıştır. Kullanılan görseller jpg formatındadır, çözünürlükleri ise 64x64 piksel ölçülerindedir.

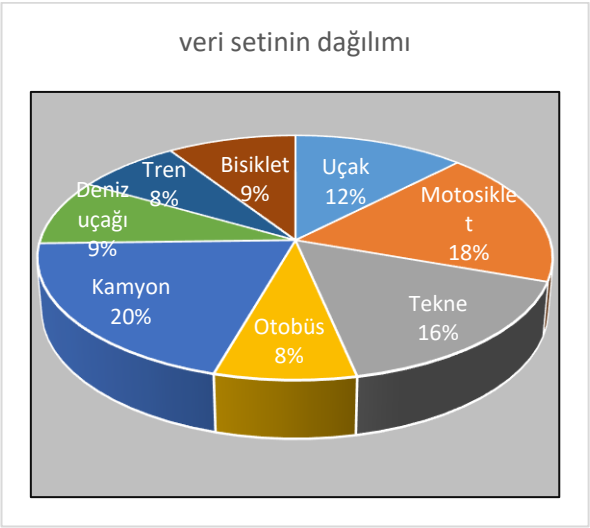
Kullanılan veri setinin sınıf isimleri ve görsellerin sayıca dağılımları *Tablo 1* de gösterilmiştir.

Tablo 1. Veri setinin sınıf sayıları ve örnek görselleri

Sınıf adı	Görsel sayısı	Örnek resim
Uçak	503	
Motosiklet	733	

Tekne	662	
Otobüs	317	
Kamyon	819	
Deniz uçağı	355	
Tren	304	
Bisiklet	375	

Veri setlerinin dağılımı ise *Şekil 1* de verilen pasta grafiğinde gösterilmiştir.

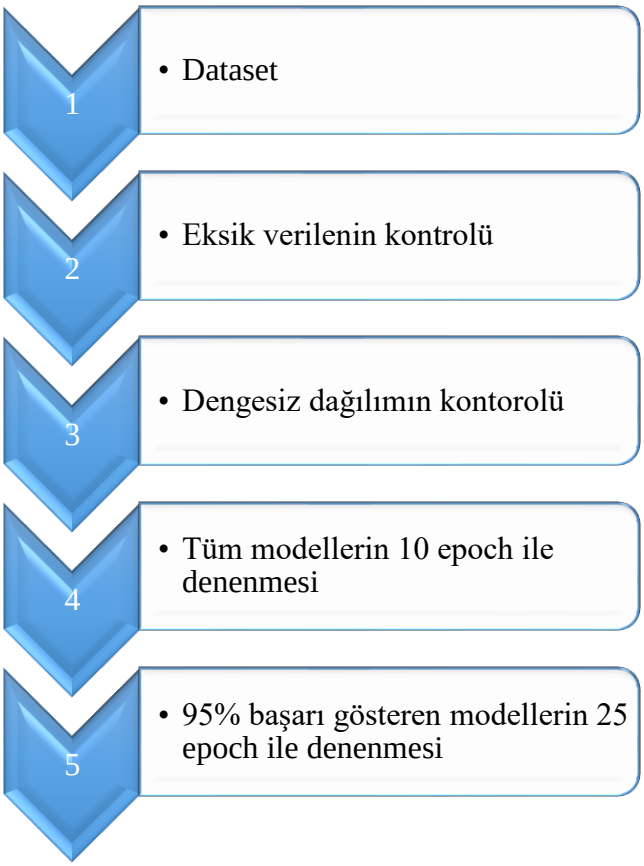


Şekil 1 Veri setinin sınıf dağılımları

Pasta grafiği dağılımından da görüldüğü üzere veri setlerinin sınıf dağılımı homojen olduğu için veri görsellere, veri çoğaltma vb. ön işlemler uygulanmamıştır.

B. Yönterm

Kaggle den alınan veri seti ilk olarak Google drive yüklenmiştir ve işlemler Google colab üzerinden gerçekleştirilmiştir. Yazılım geliştirme aşamasında başlangıç olarak veri setinin dağılımı ve veri setinde herhangi bir eksiklik olup olmadığı incelenmiştir. Sonrasında veri setinin sınıfları ve görselleri ayrıştırılarak eğitim ve test veri setleri oluşturulmuştur. Veri setinin 80%'i eğitim için 20%'si ise test için ayrılmıştır. Bu çalışmada veri setinin sınıflandırılması için Keras kütüphanesindeki iyi bilinen CNN mimarileri kullanılmıştır. İlk olarak kütüphane içerisindeki tüm modeller 10 epoch değerinde çalıştırılarak sonuçlar listelenmiştir. Kullanılan mimariler şu şekildedir: Xception, Vgg16, Vgg19, Resnet50, Resnet50v2, Resnet101, Resnet101v2, Resnet152, Resnet152v2, Inceptionv3, Inceptionresnetv2, Mobilenet, Mobilenetv2, Densenet121, Densenet169, Densenet201, Nasnetmobile, Nasnetlarge, Efficientnetb0, Efficientnetb1, Efficientnetb2, Efficientnetb3, Efficientnetb4, Efficientnetb5, Efficientnetb6, Efficientnetb7, Efficientnetv2b0, Efficientnetv2b1, Efficientnetv2b2, Efficientnetv2b3, Efficientnetv2s, Efficientnetv2m, Efficientnetv2l, Convnexttiny, Convnextsmall, Convnextbase, Convnextlarge, Convnextxlarge. Bu analiz sonucunda %95 üzerinde başarı gösteren mimariler 25 epoch değeriyle tekrar analiz edilmiştir. Veri setine uygulanan işlemler *Şekil 2*. de gösterilmiştir.



Şekil 2. İşlem adımları

III. BULGULAR VE TARTIŞMA

Bu çalışma Google colab üzerinden 12.7 GB Ram'e sahip sanal sunucuda Python dili kullanılarak gerçekleştirilmiştir. Projede colab, os, matplotlib, opencv, numpy, keras, tensorflow ve gc kütüphaneleri ile keras kütüphanesi içerisinde bulunan Xception, VGG16, VGG19, Resnet50, Resnet50v2, Resnet101, Resnet101v2, Resnet152, Resnet152v2, Inceptionv3, Inceptionresnetv2, Mobilenet, Mobilenetv2, Densenet121, Densenet169, Densenet201, Nasnetmobile, Nasnetlarge, Efficientnetb0, Efficientnetb1, Efficientnetb2, Efficientnetb3, Efficientnetb4, Efficientnetb5, Efficientnetb6, Efficientnetb7, Efficientnetv2b0, Efficientnetv2b1, Efficientnetv2b2, Efficientnetv2b3, Efficientnetv2s, Efficientnetv2m, Efficientnetv2l, Convnexttiny, Convnextsmall, Convnextbase, Convnextlarge, Convnextxlarge modelleri kullanılmıştır.

Çalışmada ilk olarak tüm keras modelleri kullanılarak analizler gerçekleştirilmiştir. Analizlerde modeller 10 epoch kullanılarak yapılmıştır. Sonuçlar *Tablo 2* de gösterilmiştir.

Tablo 2. Tüm Modellerin Sonucu

CNN mimarisi	Doğruluk değeri	Epoch
Xception	88.08%	10
VGG16	53.44%	10
VGG19	53.81%	10
ResNet50	92.87%	10
ResNet50V2	92.01%	10
ResNet101	85.38%	10
ResNet101V2	89.68%	10
ResNet152	42.26%	10
ResNet152V2	84.89%	10
InceptionV3	95.33%	10
InceptionResNetV2	96.31%	10
MobileNet	96.07%	10
MobileNetV2	64.13%	10
DenseNet121	94.23%	10
DenseNet169	69.29%	10
DenseNet201	92.26%	10
NASNetMobile	14.99%	10
NASNetLarge	7.25%	10
EfficientNetB0	97.79%	10
EfficientNetB1	98.16%	10
EfficientNetB2	98.77%	10
EfficientNetB3	94.84%	10
EfficientNetB4	98.16%	10
EfficientNetB5	98.40%	10
EfficientNetB6	85.01%	10
EfficientNetB7	91.65%	10
EfficientNetV2B0	98.16%	10
EfficientNetV2B1	96.93%	10
EfficientNetV2B2	97.67%	10
EfficientNetV2B3	98.40%	10
EfficientNetV2S	96.68%	10
EfficientNetV2M	98.03%	10
EfficientNetV2L	97.91%	10
ConvNeXtTiny	95.58%	10
ConvNeXtSmall	95.58%	10
ConvNeXtBase	94.10%	10
ConvNeXtLarge	44.35%	10
ConvNeXtXLarge	65.60%	10

Daha sonra doğruluk değeri %95 olan modeller 25 epoch ile denenmiştir. Sonuçlar *Tablo 3* de

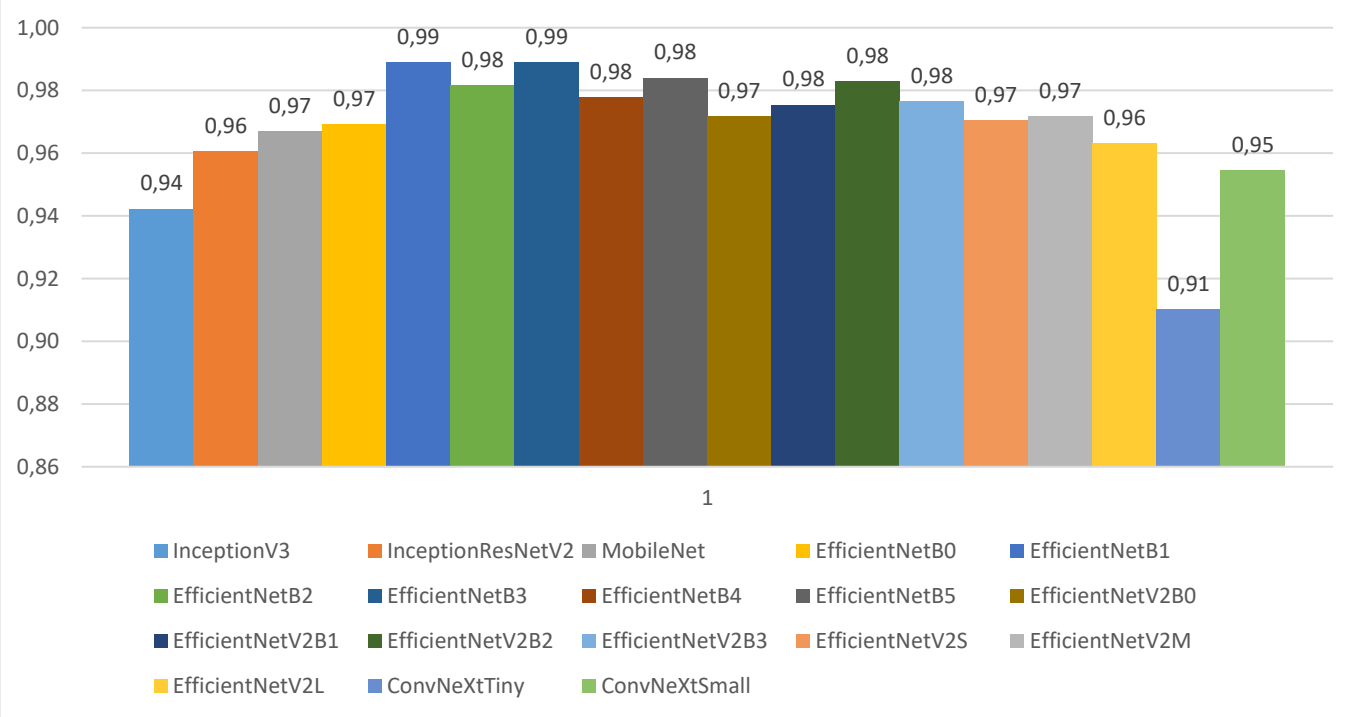
gösterilmiştir. Sonuçların sütun grafiği de *Şekil 3* te gösterilmiştir.

Tablo 3. İlk aşamada %95 üstü başarı gösteren modeller

CNN mimarisi	Doğruluk değeri	Epoch
InceptionV3	94.23%	25
InceptionResNetV2	96.07%	25
MobileNet	96.83%	25
EfficientNetB0	96.93%	25
EfficientNetB1	98.89%	25
EfficientNetB2	98.16%	25
EfficientNetB3	98.89%	25
EfficientNetB4	97.79%	25
EfficientNetB5	98.40%	25
EfficientNetV2B0	97.18%	25
EfficientNetV2B1	97.54%	25
EfficientNetV2B2	98.28%	25
EfficientNetV2B3	97.67%	25
EfficientNetV2S	97.05%	25
EfficientNetV2M	97.15%	25
EfficientNetV2L	96.32%	25
ConvNeXtTiny	91.03%	25
ConvNeXtSmall	95.46%	25

IV. SONUÇLAR

Bu çalışmada internet sitelerinin güvenlik önlemlerinden birisi olan captcha görsellerinin derin öğrenme algoritmaları kullanılarak sınıflandırılması gerçekleştirildi. Proje kapsamında Keras kütüphanesinin tüm derin öğrenme modelleri kullanılmıştır. Proje sonucunda en başarılı yöntemler 98.89% doğruluk değeriyle EfficientNetB1 ve EfficientNetB3 modelleridir. Gelecekte bu sistem kullanılarak sitelerin captcha güvenlik duvarını otomatik kıran bot yazılımı geliştirilecektir.



Şekil 3 Modellerin doğruluk değerleri

SON NOT

Bu çalışma “Derin öğrenme teknikleri” dersi kapsamında yapılmıştır.

KAYNAKLAR

[1] C. Demİrel and D. Kilinç, “Captcha ile Güvenliğe Genel Bakış”.

[2] Özkaya, U., Öztürk, Ş., & Barstugan, M. (2020). Coronavirus (COVID-19) classification using deep features fusion and ranking technique. Big Data Analytics and Artificial Intelligence Against COVID-19: Innovation Vision and Approach, 281-295.

[3] S. Sivakorn, I. Polakis, and A. D. Keromytis, “I am Robot: (Deep) learning to break semantic image CAPTCHAs,” *Proc. - 2016 IEEE Eur. Symp. Secur. Privacy, EURO S P 2016*, pp. 388–403, 2016, doi: 10.1109/EuroSP.2016.37.

[4] F. H. Alqahtani and F. A. Alsulaiman, “Is image-based CAPTCHA secure against attacks based on machine learning? An experimental study,” *Comput. Secur.*, vol. 88, p. 101635, 2020, doi: 10.1016/j.cose.2019.101635.

[5] K. Sukhani, S. Sawant, S. Maniar, and R. Pawar, “Automating the Bypass of Image-based CAPTCHA and Assessing Security,” *2021 12th Int. Conf. Comput. Commun. Netw. Technol. ICCCNT 2021*, 2021, doi: 10.1109/ICCCNT51525.2021.9580020.

[6] Ö. İnik and E. Ülker, “Derin Öğrenme ve Görüntü Analizinde Kullanılan Derin Öğrenme Modelleri,” *Gaziosmanpasa J. Sci. Res.*, vol. 6, no. 3, pp. 85–104, 2017.

[7] I. Akrout, A. Feriani, and M. Akrout, “Hacking Google reCAPTCHA v3 using Reinforcement Learning,” 2019, [Online]. Available: http://arxiv.org/abs/1903.01003

[8] L. Mookdarsanit and P. Mookdarsanit, “An Adversarial Perturbation Technique against reCaptcha Image Attacks”.

[9] S. Mittal, P. Kaushik, S. Hashmi, and K. Kumar, “Robust Real Time Breaking of Image CAPTCHAs Using Inception v3 Model,” *2018 11th Int. Conf. Contemp. Comput. IC3 2018*, pp. 2–4, 2018, doi: 10.1109/IC3.2018.8530607.

[10] M. I. Hossen, Y. Tu, M. F. Rabby, M. N. Islam, H. Cao, and X. Hei, “An object detection based solver for Google’s image reCAPTCHA v2,” *RAID 2020 Proc. - 23rd Int. Symp. Res. Attacks, Intrusions Defenses*, pp. 269–284, 2020.

[11] D. Of, P. From, X. I. Using, and D. Learning, “DETECTION OF PNEUMONIA FROM X-RAY IMAGES USING DEEP LEARNING,” pp. 419–440, 2023.

[12] H. Bakir and G. Tarihi, “Using Transfer Learning Technique as a Feature Extraction Phase for Diagnosis of Cataract Disease in the Eye,” *Usbtu*, vol. 1, no. 1, p. 2022, 2022.