Niğde Ömer Halisdemir Üniversitesi Mühendislik Bilimleri Dergisi
Nigde Omer Halisdemir University Journal of Engineering Sciences

ISSN: 2564-6605

Araştırma / Research

ÖN EĞİTİMLİ EVRİŞİMLİ SİNİR AĞLARI DESTEKLİ GÖRÜNTÜ SAHTECİLİK TESPİTİ YÖNTEMİ

Ertuğrul GÜL[1,*], Serkan ÖZTÜRK[2]

[1,2] Erciyes Üniversitesi, Bilgisayar Mühendisliği Bölümü, Kayseri, Türkiye
[1] Niğde Ömer Halisdemir Üniversitesi, Bilgisayar Mühendisliği Bölümü, Niğde Türkiye

ÖZET

İnternet ve bilgisayar teknolojilerinin gelişmesi ile görüntü sahteciliği tespiti önem kazanmıştır. Ayrıca, görüntü iyileştirme uygulamalarında kullanılan tekniklerin iyi başarım göstermesi için görüntülere uygulanan saldırı çeşitlerinin ve bölgelerinin doğru bir şekilde tespit edilmesi gerekmektedir. Bu çalışmada, görüntülere uygulanan saldırı çeşitlerini ve saldırı bölgelerini tespit etmek için ön eğitimli AlexNet ve GoogLeNet evrişimli sinir ağları destekli görüntü sahtecilik tespiti yöntemi önerilmiştir. Öncelikle; MICC-F2000 veri kümesinde bulunan görüntüler kullanılarak orijinal ve saldırılmış görüntülerin olduğu görüntü sahteciliği tespiti veri kümesi oluşturulmuştur. Saldırılmış görüntüleri elde etmek için Gauss bulanıklaştırma, medyan filtreleme, Gauss gürültü ekleme, Poisson gürültü ekleme ve keskinleştirme saldırıları kullanılmıştır. Daha sonra, ön eğitimli AlexNet ve GoogLeNet ağlarının tam bağlantılı katmanları deneysel veri kümesindeki altı veri sınıfı için yeni tam bağlantılı katmanlar ile değiştirilmiştir. Oluşturulan AlexNet ve GoogLeNet destekli ağlar hazırlanan görüntü sahteciliği tespiti veri kümesi ile eğitilerek test edilmiştir. Faklı hiperparametre değerleri için ağların başarımları ölçülmüştür. AlexNet destekli ağlarda en yüksek başarım %99,48'lik doğruluk oranı ile elde edilirken, GoogLeNet destekli ağlarda ise en yüksek başarım %99,92'lik doğruluk oranı ile elde edilmiştir. Ayrıca, geliştirilen AlexNet ve GoogLeNet destekli sahtecilik tespiti yönteminin CoMoFoD veri kümesinden alınan görüntüler üzerindeki saldırıları tespit edebilme başarısı gözlemlenmiştir. Deneysel sonuçlar önerilen yöntemin başarılı bir şekilde görüntü sahteciliği tespiti için kullanılabileceğini göstermiştir.

Anahtar kelimeler: Evrişimli sinir ağı, Sahtecilik tespiti, Transfer öğrenme, AlexNet, GoogLeNet

PRE-TRAINED CONVOLUTIONAL NEURAL NETWORK BASED IMAGE TAMPER DETECTION METHOD

ABSTRACT

With the development of internet and the computer technologies, image forgery detection has become important issue. In addition, in order to obtain successful performance in image enhancement techniques, the types and regions of the attacks applied to the images must be determined correctly. In this study, in order to detect the types and regions of the attacks applied to the images, pre-trained AlexNet and GoogLeNet convolutional neural networks-based forgery detection method has been proposed. Firstly, image forgery detection dataset containing the original and the attacked images has been created using the images in the MICC-F2000 dataset. Gaussian blurring, median filtering, Gaussian noise adding, Poisson noise adding and sharpening attacks have been used to obtain the attacked images. Then, the fully connected layers of the pre-trained AlexNet and GoogLeNet networks have been replaced with the new fully connected layers for the six classes of the created image forgery detection dataset. The modified AlexNet and GoogLeNet based networks have been trained and tested with the created image forgery detection dataset. The networks performances have been evaluated for different hyper parameter values. While the highest accuracy rate of 99.48% has been achieved in AlexNet supported networks, the highest accuracy rate of 99.92% has been achieved in GoogLeNet supported networks. Also, the proposed AlexNet and GoogLeNet-based forgery detection method has been tested on the images from CoMoFoD dataset. Experimental results show that the proposed method can be used successfully for the image forgery detection.

Keywords: Convolutional neural network, Tamper detection, Transfer learning, AlexNet, GoogLeNet

* Sorumlu yazar / Corresponding author, e-posta / e-mail: ertugrulgul@erciyes.edu.tr
Geliş / Recieved: 04.12.2019 Kabul / Accepted: 14.05.2020 doi: 10.28948/ngmuh.654519

1. GİRİŞ

Son yıllarda, sayısal görüntü düzenleme araçları gibi bilgisayar teknolojilerinin hızla gelişmesi, görüntü üzerinde yapılan sahtecilik işlemlerini oldukça kolaylaştırmıştır. Bu sebeple, internet üzerindeki kurcalanmış görüntülerin sayısı gün geçtikçe artmaktadır. Bu durum araştırmacıları görüntü sahteciliği tespiti konusuna yöneltmektedir.

Literatürde görüntü sahteciliği tespiti üzerine yapılan çalışmalar genellikle aktif ve pasif yöntemler olmak üzere ikiye ayrılmaktadır [1]. Görüntü damgalama [2,3,4] ve dijital imza [5] gibi aktif görüntü sahtecilik tespiti yöntemlerinde görüntünün veya depolama alanının içerisine ek bir bilgi gizlenmektedir. Ancak bu işlemin görüntünün elektronik bir ortama kaydedilmesinden önce gerçekleştirilmesi gerekmektedir. Bu yüzden internet üzerinden rastgele elde edilen görüntülerin orijinalliğinin denetlenmesinde aktif yöntemler yetersiz kalmaktadır. Kopyala-yapıştır [6,7] ve ekleme [8] gibi pasif görüntü sahteciliği tespiti yöntemlerinde ise herhangi bir ek bilgiye ihtiyaç duyulmaksızın görüntü özellikleri çıkartılarak sahtecilik tespiti yapılmaktadır.

Son yıllarda, derin öğrenmenin bilgisayar görü ve örüntü tanıma gibi alanlardaki başarısı, araştırmacıları derin öğrenmeyi pasif görüntü sahtecilik tespiti konusu üzerinde kullanmaya yönlendirmiştir [9]. Evrişimli Sinir Ağı (ESA) görüntü sahtecilik tespiti uygulamalarında kullanılan derin öğrenme yöntemlerinin başında gelmektedir. ESA, üç renk kanalında piksel yoğunlukları içeren üç adet iki boyutlu diziden oluşan renkli görüntü gibi çoklu diziler halinde gelen verileri işlemek üzere tasarlanmış yapılardır. ESA yerel bağlantılar, paylaşılan ağırlıklar, ortaklama ve birçok katmanın kullanımı olmak üzere doğal işaretlerin özelliklerinden yararlanan dört temel fikirden oluşmaktadır [10].

Literatürde, ESA mimarisini kullanan birçok görüntü sahteciliği tespiti yöntemi bulunmaktadır. Chen ve arkadaşları [11], küçük boyutlu ve sıkıştırılmış görüntü bloklarında medyan filtrelemeyi tespit etmek için ESA tabanlı bir yöntem önermişlerdir. Önerilen yöntem kullanarak yapılan testlerde, özellikle kesme ve yapıştırma sahteciliği tespitinde önemli başarım iyileştirmeleri sağlandığını belirtmişlerdir. Bayer ve Stamm [12] sahtecilik tespiti özelliklerini eğitim verilerinden otomatik olarak öğrenebilen yeni bir ESA mimarisi önermişlerdir. Bu mimari, önceden seçilmiş özelliklere veya herhangi bir ön işleme gerek duymadan birden fazla sahtecilik çeşidini tespit etmeyi otomatik olarak öğrenebilmektedir. Ayrıca, Bayar ve Stamm [13] yeniden örnekleme algılamasını yeniden sıkıştırılmış görüntülerde gerçekleştirebilen yeni bir ESA mimarisi önermişlerdir. Rao ve Ni [14] ekleme ve kopyala-yapıştır saldırılarını tespit etmek için ESA tabanlı görüntü sahtecilik tespiti yöntemi önermişlerdir. Bu yöntemde ağın ilk katmanındaki ağırlıklar mekânsal zengin modelden (SRM-Spatial Rich Model) alınmaktadır. Amerini ve arkadaşları [15] tek ve çift JPEG sıkıştırma saldırılarını tespit etmek için ESA tabanlı yöntemler önermişlerdir. Çalışmalarında görüntü uzayı, frekans uzayı ve çoklu uzay tabanlı olmak üzere üç çeşit yöntem bulunmaktadır. Wang ve Zhang [16] tek ve çift sıkıştırılmış alanları sınıflandırmak için ESA yapısını kullanan bir yöntem önermişlerdir. Önerilen yöntemin çift sıkıştırılmış bölgelerin tespitinde, özellikle ilk sıkıştırma kalite faktörünün ikincisinden daha yüksek olduğu durumlarda, daha iyi olduğunu öne sürmüşlerdir. Bunk ve arkadaşları [17] görüntü manipülasyonlarının tespit edilmesi ve yerlerinin belirlenmesi için yeniden örnekleme özellikleri ve derin öğrenmenin kombinasyonuna dayalı iki yöntem geliştirmişlerdir. Bondi ve arkadaşları [18] farklı kamera modellerinin görüntülerde bıraktığı karakteristik izleri kullanarak görüntü sahtecilik tespiti ve lokalizasyonu yapan ESA mimarisi tabanlı bir algoritma önermişlerdir.

Bu çalışmada ön eğitimli ESA mimarilerinde transfer öğrenme ile görüntü sahtecilik tespiti gerçekleştirilmektedir. Transfer öğrenme belirli bir görev için eğitilmiş ağın, yeni bir görev ve veri kümesi ile baştan eğitilme işlemidir [19]. Çalışmada, transfer öğrenme için AlexNet ve GoogLeNet ön eğitimli ESA'ları kullanılmıştır. AlexNet [20] 2012 yılında yayınlanan 1,2 milyon yüksek çözünürlüklü görüntü ile eğitilmiş, 1000 farklı sınıfta sınıflandırma yapabilen bir ağdır. GoogLeNet [21] ise 22 katmanlı bir yapıya sahip yüksek doğruluk oranı ile sınıflandırma yapabilen bir ESA'dır. Bu çalışmanın temel katkısı AlexNet ve GoogLeNet destekli ESA mimarisi oluşturup transfer öğrenme ile görüntü sahtecilik tespiti gerçekleştirmektir. Öncelikle, saldırılmış görüntülerin ve orijinal görüntülerin olduğu veri kümesi oluşturulmuştur. Saldırılmış görüntüler; Gauss bulanıklaştırma, medyan filtreleme, Gauss gürültü ekleme, Poisson gürültü ekleme ve keskinleştirme saldırıları kullanılarak elde edilmiştir. Sonrasında, AlexNet ve GoogLeNet ön eğitimli ağlarının tam bağımlı katmanları saldırı tespiti için düzenlenmiştir. Bu iki ön eğitimli ağ destekli ESA mimarileri, oluşturduğumuz veri kümesi kullanılarak yeniden eğitilmiştir. Eğitilen ağların deneysel sonuçları, bu ağların görüntü sahtecilik tespitindeki başarımlarının yüksek olduğunu göstermiştir. Özellikle görüntü iyileştirme uygulamalarında kullanılan tekniklerin iyi başarım göstermesi için görüntülere uygulanan saldırı çeşitlerinin ve bölgelerinin doğru bir şekilde tespit edilmesi önem arz etmektedir. Bu yüzden transfer öğrenme kullanılarak geliştirilen ağlar farklı boyutlardaki görüntülerde sahtecilik çeşidini ve sahtecilik bölgelerini tespit etmek için sistemleştirilmiştir. Genellikle sınıflandırma ve nesne tanıma gibi uygulamalarda kullanılan bu ön eğitimli ağların ağırlıklarının yapılan deneyler sonucunda görüntü sahtecilik tespiti için de iyi sonuçlar verdiği gözlemlenmiştir.

Makalenin geri kalanı aşağıdaki gibi organize edilmektedir: Bölüm 2, ESA mimarilerinden olan AlexNet ve GoogLeNet hakkında genel bilgileri anlatmaktadır. Önerilen yöntem Bölüm 3'te anlatılmaktadır. Bölüm 4'te deneysel sonuçlar gösterilmektedir. Sonuç ve tartışma Bölüm 5'te açıklanmaktadır.

2. EVRİŞİMLİ SİNİR AĞLARI (ESA)

ESA çeşitli bilgisayar görme uygulamalarında yaygın olarak kullanılan bir derin öğrenme yaklaşımıdır [22]. Genellikle ESA mimarileri evrişim katmanı, ortaklama katmanı ve tam bağlantılı katman olmak üzere üç tip katmandan oluşmaktadır. ESA mimarilerini oluşturan bu katmanların birbirinden farklı görevleri ve özellikleri bulunmaktadır [23]. Evrişim katmanında, giriş görüntüsüne bir evrişim filtresi uygulanarak aktivasyon haritası çıkartılmaktadır. Ortaklama katmanında, küçük aktivasyon haritaları oluşturmak için bağımsız olarak her bir aktivasyon haritasının üzerinde çalışan bir aşağı örnekleme işlemi gerçekleştirilmektedir. Tam bağlantılı katmanda ise evrişim katmanları tarafından çıkarılan ve katmanları birleştirerek aşağı örneklenen özellikler üzerinde doğrusal işlemler gerçekleştirilmektedir [24]. ESA uygulamalarından birisi olan transfer öğrenimi ön eğitilmiş ESA mimarilerine yeni bir görev yüklemek için ağın yeniden eğitilmesi işlemidir. Transfer öğrenme kullanmanın temel amacı, yeterli büyüklükte eğitim kümeleri mevcut olmayan görüntü sınıflandırma problemlerinde ESA mimarilerinin kullanılmasını sağlamaktır [25]. Araştırmalar, taban ve hedef veri kümelerinin birbirinden çok farklı olduğu durumlarda bile transfer öğrenme kullanımının ESA ağırlıklarının rastgele başlatılmasına göre daha iyi sonuçlar verdiğini göstermektedir [26]. Farklı mimari yapılara sahip ve transfer öğrenme uygulamalarında kullanılan birçok ESA tabanlı yöntem bulunmaktadır. Transfer öğrenme uygulamalarında kullanılan en yaygın ESA mimarileri AlexNet ve GoogLeNet ağlarıdır.

AlexNet [20] Krizhevsky ve arkadaşları tarafından 2012 yılında önerilen, ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) veri kümesindeki 1000 farklı sınıfı bulunan kabaca 1,2 milyon görüntü ile eğitilmiş bir ESA mimarisidir. AlexNet mimarisi yaklaşık 650.000 nöron ve 60 milyon parametreye sahiptir [27]. Bu mimari beş adet evrişim katmanı, iki adet normalleştirme katmanını, üç adet maksimum ortaklama katmanını, üç adet tam bağlantılı katmanı ve çıkışta softmax aktivasyonu olan doğrusal bir katmanı içermektedir. AlexNet ağı 227×227×3 boyutlardaki bir görüntüyü girdi olarak aldıktan sonra, tekrarlamalı bir şekilde evrişim ve ortaklama işlemlerini uygulayıp, sonuçları tam bağlantılı katmanlara iletmektedir [23].

GoogLeNet [21] Szegedy ve arkadaşları tarafından 2014 yılında önerilen ILSVRC 2014'ün galibi olan 22 katmanlı bir yapıya sahip ESA mimarisidir. GoogLeNet, AlexNet'e göre daha derin bir yapıya sahiptir [23]. GoogLeNet dokuz adet başlangıç modülü, iki adet evrişim katmanı, boyut küçültme için bir adet evrişim katmanı, iki adet normalleştirme katmanı, dört adet maksimum ortaklama katmanı, bir adet ortalama ortaklama katmanı, bir adet tam bağlantılı katman ve çıkışta softmax aktivasyonu olan doğrusal bir katmanını içermektedir [27]. GoogLeNet ağı 224×224×3 boyutundaki görüntüleri girdi olarak alıp işlemektedir.

3. ÖNERİLEN YÖNTEM

Bu çalışmada görüntüler üzerinde yapılan sahteciliklerin tespit ve lokalize edilmesi için ESA mimarileri üzerinde transfer öğrenme işlemi gerçekleştirilmiştir. Ön eğitimli AlexNet ve GoogLeNet mimarileri modifiye edilip transfer öğrenme uygulanarak sahtecilik tespiti görevini gerçekleştirmek üzere yeniden eğitilmiştir. Oluşturulan AlexNet ve GoogLeNet destekli ESA mimarileri kullanılarak farklı boyutlardaki görüntüler üzerinde yapılan sahteciliklerin tespiti ve lokalize edilmesi işlemleri sistemleştirilmiştir.

3.1. Transfer Öğrenme

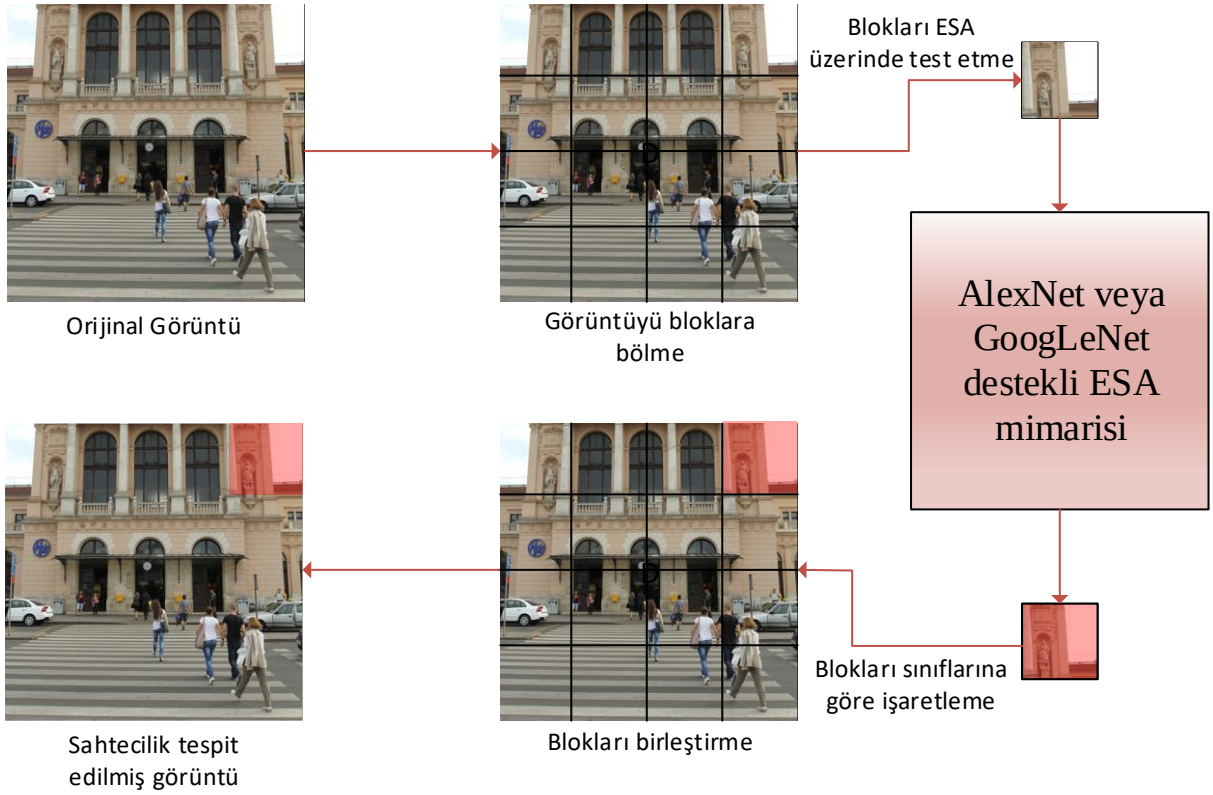
Transfer öğrenme belirli bir görev için eğitilmiş ağın, yeni bir görevi gerçekleştirmesi için bir veri kümesi ile yeniden eğitilme işlemidir. Bu çalışmada, Gauss bulanıklaştırma, medyan filtreleme, Gauss gürültü ekleme, Poisson gürültü ekleme, keskinleştirme saldırılarının ve orijinal görüntülerin olduğu bir veri kümesi hazırlanmıştır. Veri kümesinin hazırlanma süreci çalışmanın deneysel kurulum bölümünde ayrıntılı bir şekilde anlatılmaktadır. Transfer öğrenme işlemini gerçekleştirmek için öncelikle AlexNet ve GoogLeNet ağlarının tüm katmaları kopyalanarak hedef ağlarımız oluşturulmuştur. Daha sonra, 1000 sınıf için tasarlanmış son tam bağlantılı katmanlar veri kümemizdeki 6 sınıf için yeni üç tam bağlantılı katman ile değiştirilmiştir. Böylece, AlexNet ve GoogLeNet mimarileri görüntüler üzerinde yapılan Gauss bulanıklaştırma, medyan filtreleme, Gauss gürültü ekleme, Poisson gürültü ekleme ve keskinleştirme sahteciliklerinin tespit ve lokalize edilmesi görevine uyarlanmıştır. Oluşturulan bu ağlar görüntüler üzerinde saldırı tespiti görevi için, yeni oluşturulmuş veri kümesi ile eğitilip test edilmiştir.

3.2. Sahtecilik Tespiti Sistemi

Önerilen sistemde, sahtecilik tespiti yapılacak görüntü öncelikle kullanılan ESA mimarilerine göre bloklara ayrılmaktadır. Görüntü, AlexNet destekli ESA mimarisi için 227×227×3 boyutlarında, GoogLeNet destekli ESA mimarisi için 224×224×3 boyutlarında bloklara ayrılmaktadır. Blok boyutu seçimi AlexNet ve GoogLeNet ESA mimarilerinin giriş olarak aldığı görüntü boyutlarına göre belirlenmiştir. Daha sonra her bir blok AlexNet ve GoogLeNet destekli ESA mimarileri ile sınıflandırılmaktadır. Bu sınıflandırma ESA eğitiminde kullanılan sınıflara göre gerçekleşmektedir. Son olarak sistem bu sınıflandırılmış blokları

yeniden bir araya getirmektedir. Önerilen sahtecilik tespiti sisteminin ana işlevini AlexNet veya GoogLeNet destekli ESA mimarileri oluşturmaktadır. Bu yüzden ESA'ların eğitimi sırasında elde edilen başarı çok önemlidir. Hata payının çok düşük olması istenmektedir. Önerilen sahtecilik tespiti sistemi; görüntünün hazırlanması, saldırı bölgesi ve türünün tespiti, saldırı bölgesi ve türünün gösterilmesi olmak üzere üç bölümden oluşmaktadır. Önerilen sistemin blok diyagramını Şekil 1'de gösterilmektedir. Sistemin temel adımları aşağıdaki gibidir:

1- **Görüntünün hazırlanması:** Görüntüler kullanılan ESA mimarisine göre bloklara ayrılır (AlexNet destekli ESA mimarisi için 227×227×3 boyutlarında, GoogLeNet destekli ESA mimarisi için 224×224×3 boyutlarında).
2- **Saldırı bölgesi ve türünün tespiti:** Bloklara ayrılmış görüntü parçaları AlexNet veya GoogLeNet destekli ESA mimarileri üzerinde test edilerek sınıflandırılır.
3- **Saldırı bölgesi ve türünün gösterilmesi:** Saldırı yapılmış olarak tespit edilen bölgeler saldırı türü ile işaretlenir ve parçalara ayrılmış bloklar birleştirilir.



Şekil 1. AlexNet veya GoogLeNet destekli ESA mimarisi tabanlı sahtecilik tespiti yöntemi

4. DENEYSEL SONUÇLAR

4.1. Deneysel Kurulum

Ön eğitimli AlexNet ve GoogLeNet ağlarına transfer öğrenme ile görüntüler üzerinde sahtecilik tespiti görevi verilebilmesi için ilk önce orijinal ve manipüle edilmiş görüntülerden oluşan deneysel veri kümesi oluşturulmuştur. Veri kümesinin oluşturulması için MICC-F2000 [28] veri kümesinde bulunan üzerinde herhangi bir değiştirilme ya da oynama yapılmamış 1300 adet işlenmemiş renkli (RGB) görüntü alınmıştır. AlexNet ağlarını eğitmek için 227×227×3 boyutlarında görüntü blokları orijinal görüntüler kullanılarak oluşturulmuştur. Benzer şekilde GoogLeNet ağlarını eğitmek için 224×224×3 boyutlarında görüntü blokları da orijinal görüntülerden elde edilmiştir. Toplamda her bir ağ için kullanılmak üzere 70200 adet görüntü bloğu oluşturulmuştur. Bu görüntülerden, 60000 tanesi eğitim,10200 tanesi ise test görüntüsü olarak rastgele seçilmiştir. Son olarak,

eğitim ve test görüntülerine aşağıda ifade edilen 5 farklı saldırı tipi uygulanarak manipülasyona uğramış görüntüler elde edilmiştir:

- 5×5 pencere boyutu ve "1.1" standart sapma değeri ile Gauss bulanıklaştırma
- 5×5 pencere boyutu ile medyan filtreleme
- "0.01" varyans değeri ile Gauss gürültü ekleme
- Poisson gürültü ekleme
- 5×5 pencere boyutu ile keskinleştirme

Sonuç olarak, saldırı uygulanmış görüntüler dahil toplamda her bir ağın eğitimi için 360000, testi için ise 61200 görüntü elde edilmiştir.

4.2. Deneysel Sonuçlar

AlexNet ve GoogLeNet ön eğitimli ağlarının transfer eğitimi ile görüntü sahtecilik tespiti üzerindeki başarımlarını ölçmek için MICC-F2000 veri kümesi kullanarak içerisinde Gauss bulanıklaştırma, medyan filtreleme, Gauss gürültü ekleme, Poisson gürültü ekleme, keskinleştirme saldırılarının ve orijinal görüntülerin bulunduğu veri kümesi oluşturulmuştur. Oluşturulan eğitim veri kümesi ile eğitilen ağlar test veri kümesi kullanılarak değerlendirilmiştir. Problem karşısında en başarılı sonuçları veren ağları bulmak için farklı hiperparametreler kullanılarak ağlar eğitilmiştir. İlk öğrenme oranının 0,001'den büyük olduğu durumlarda ağların başarılı sonuçlar vermediği yapılan ön çalışmalarda görülmüştür. Bu yüzden ilk öğrenme oranı 0,001 ve 0,0001 olarak seçilmiştir. Maksimum iterasyonun 10 olduğu durumlarda, 16, 32 ve 64 olarak ayarlanan mini yığın büyüklüğü (minibatchsize, mbs) değerlerindeki AlexNet ve GoogLeNet destekli ESA ağlarının eğitim ve test işlemleri gerçekleştirilmiştir. Ayrıca maksimum iterasyonun 5, ilk öğrenme oranının 0,001 seçildiği durumlarda da eğitim ve test işlemleri yapılmıştır. Eğitilen ağların test sonuçlarındaki başarısı Denklem 1 kullanılarak hesaplanmıştır [29].

$$\text{Doğruluk oranı} = \frac{Doğru\ tahmin\ sayısı}{Toplam\ örnek\ sayısı} \tag{1}$$

Eğitilen ağların doğruluk oranları, eğitim ve test süreleri Tablo 1'de gösterilmektedir. Tablo 1'de görüldüğü üzere, farklı mini yığın büyüklüğü değerlerinde eğitilen ağların doğruluk oranında GoogLeNet destekli ESA %99,92 oranında başarı elde ederek AlexNet destekli ESA'dan daha iyi sonuç vermiştir. Ayrıca GoogLeNet destekli ESA'ların eğitim sürelerinin AlexNet destekli ESA'lardan daha düşük olduğu açıkça görülmektedir. Test sürelerindeki başarı ise değişiklik göstermektedir.

Tablo 1. AlexNet ve GoogLeNet destekli ESA mimarilerinin test sonuçlarındaki doğruluk oranları

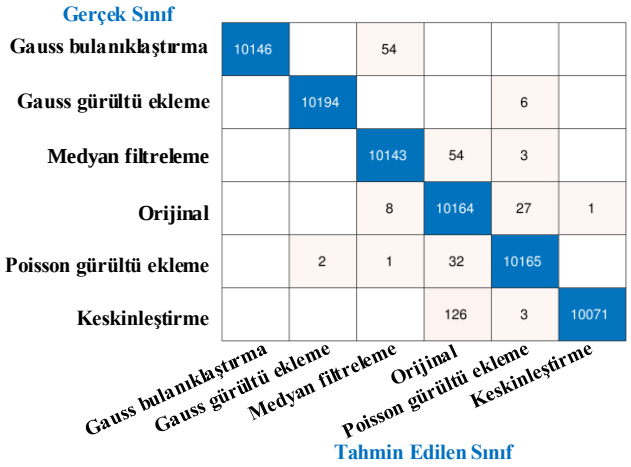
Hiperparametre			**AlexNet**			**GoogLeNet**		
Maksimum iterasyon	**İlk öğrenme oranı**	**Mini yığın büyüklüğü**	**Doğruluk oranı**	**Eğitim süresi (saniye)**	**Test süresi (saniye)**	**Doğruluk oranı**	**Eğitim süresi (saniye)**	**Test süresi (saniye)**
10	0,001	16	%94,89	44419	273,87	%99,88	43513	267,13
		32	**%99,48**	44163	253,83	%99,91	43763	258,09
		64	%98,76	43465	251,41	**%99,92**	42536	256,82
	0,0001	16	%98,80	44505	258,37	%80,39	44277	255,16
		32	%99,46	44168	253,11	%7966	43005	260,05
		64	%99,35	43627	255,30	%82,32	43419	258,16
5	0,001	16	%98,23	22911	249,53	%74,60	22452	260,21
		32	%99,28	23556	249,96	%72,68	22284	257,78
		64	%98,46	22486	250,83	%79,35	22020	257,17

En iyi başarım elde edilen ağlar incelendiğinde, AlexNet destekli ESA'nın test görüntülerinin tamamının %0,52'lik kısmını, GoogLeNet destekli ESA'nın ise %0,08'lik kısmını hatalı bir şekilde sınıflandırdığı gözükmektedir. Ayrıca en başarılı GoogLeNet destekli ESA'nın eğitim süresi AlexNet destekli ESA'dan daha düşüktür. Ancak GoogLeNet destekli ESA'nın test süresi daha yüksektir. Tablo 2'de en başarılı AlexNet ve GoogLeNet destekli ESA'ların sınıflar içerisindeki doğruluk oranları gösterilmektedir. AlexNet destekli ESA %99,94'lik doğruluk oranı ile en yüksek başarıyı Gauss gürültü eklenmiş görüntüleri sınıflandırmakta göstermiştir. GoogLeNet destekli ESA ise Gauss gürültü ekleme, medyan filtreleme, Gauss bulanıklaştırma, Poisson gürültü ekleme saldırıları uygulanmış görüntüleri %100'lük başarılı ile tamamen doğru sınıflandırmıştır. GoogLeNet destekli ESA tüm sınıflarda AlexNet destekli ESA'dan daha iyi sonuç vermiştir.

Tablo 2. En başarılı ağların test sonuçlarının sınıflara göre doğruluk oranları

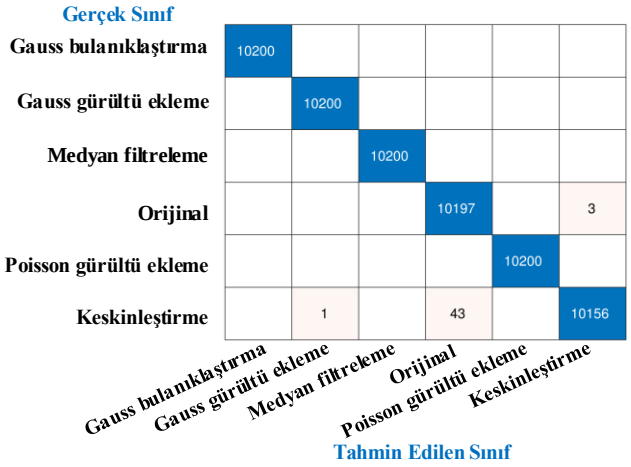
Sınıflar	AlexNet destekli ESA'nın doğruluk oranı	GoogLeNet destekli ESA'nın doğruluk oranı
Orijinal görüntü	%99,64	%99,97
Gauss bulanıklaştırma	%99,47	%100
Medyan filtreleme	%99,44	%100
Gauss gürültü ekleme	%99,94	%100
Poisson gürültü ekleme	%99,65	%100
Keskinleştirme	%98,73	%99,56
Toplam doğruluk oranı	%99,48	%99,92

Şekil 2'de AlexNet destekli ESA'nın karışıklık matrisi gösterilmektedir. Bu matris incelendiğinde keskinleştirme saldırısı sınıfının en düşük başarıya sahip olduğu görülmektedir. Ayrıca bu sınıf içerisinde hatalı olarak sınıflandırılan görüntülerin %97,67'si orijinal olarak sınıflandırılmıştır. AlexNet destekli ESA için en yüksek başarı Gauss gürültü sınıfında elde edilmiştir. Bu sınıfta yalnızca 6 görüntü hatalı olarak sınıflandırılmıştır.



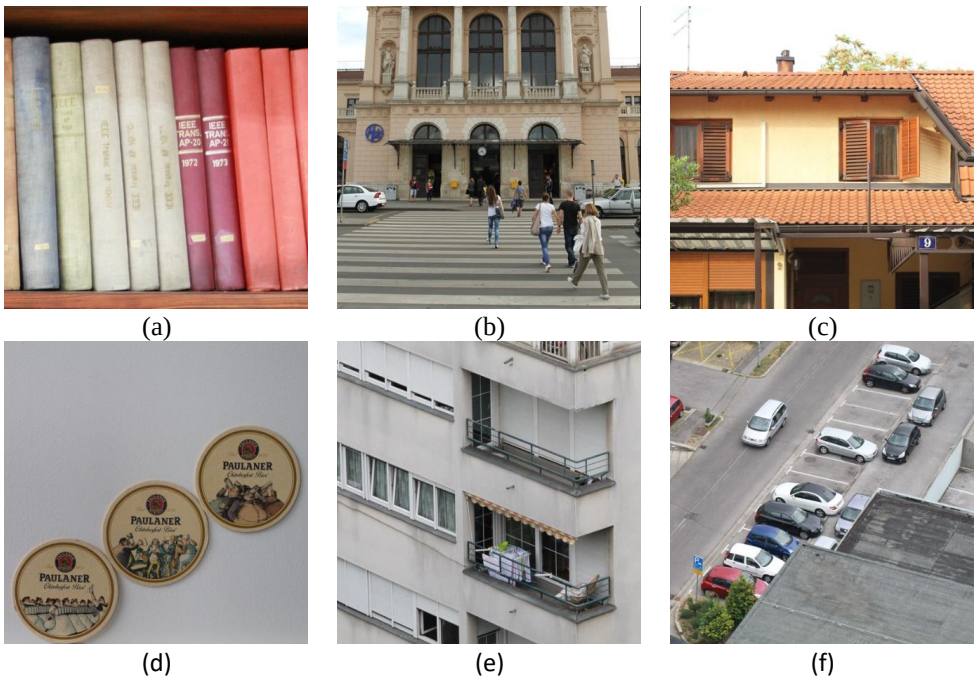
Şekil 2. AlexNet destekli ESA'nın karışıklık matrisi

Şekil 3'te gösterilen GoogLeNet destekli ESA'nın karışıklık matrisi incelendiğinde sadece orijinal ve keskinleştirme saldırısı görüntüleri sınıflandırılırken hataların olduğu gözükmektedir. Orijinal görüntüler sınıfında yalnızca 3 görüntü yanlış sınıflandırılırken keskinleştirme saldırısı sınıfından 44 görüntü yanlış sınıflandırılmıştır.

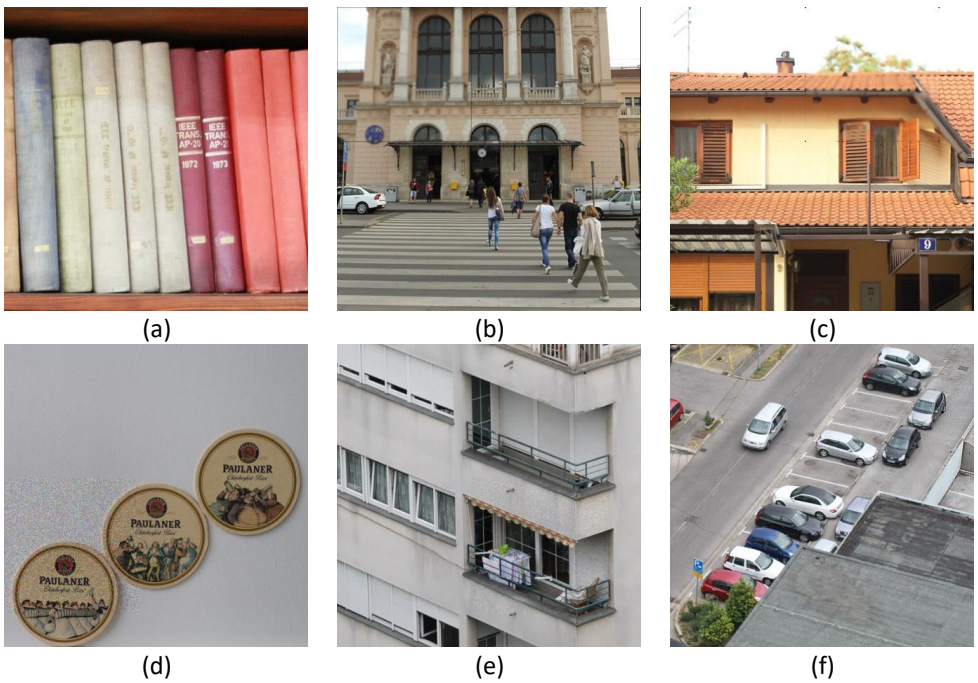


Şekil 3. GoogLeNet destekli ESA'nın karışıklık matrisi

Önerilen AlexNet ve GoogLeNet destekli ESA tabanlı sahtecilik tespiti sistemlerinin başarımlarını değerlendirmek için Şekil 4' de gösterilen CoMoFoD [30] veri tabanından alınmış orijinal görüntüler kullanılmıştır. Bu görüntülerin farklı bölgelerine Gauss bulanıklaştırma, medyan filtreleme, Gauss gürültü ekleme, Poisson gürültü ekleme ve keskinleştirme saldırıları uygulanarak Şekil 5'teki kurcalanmış görüntüler elde edilmiştir.

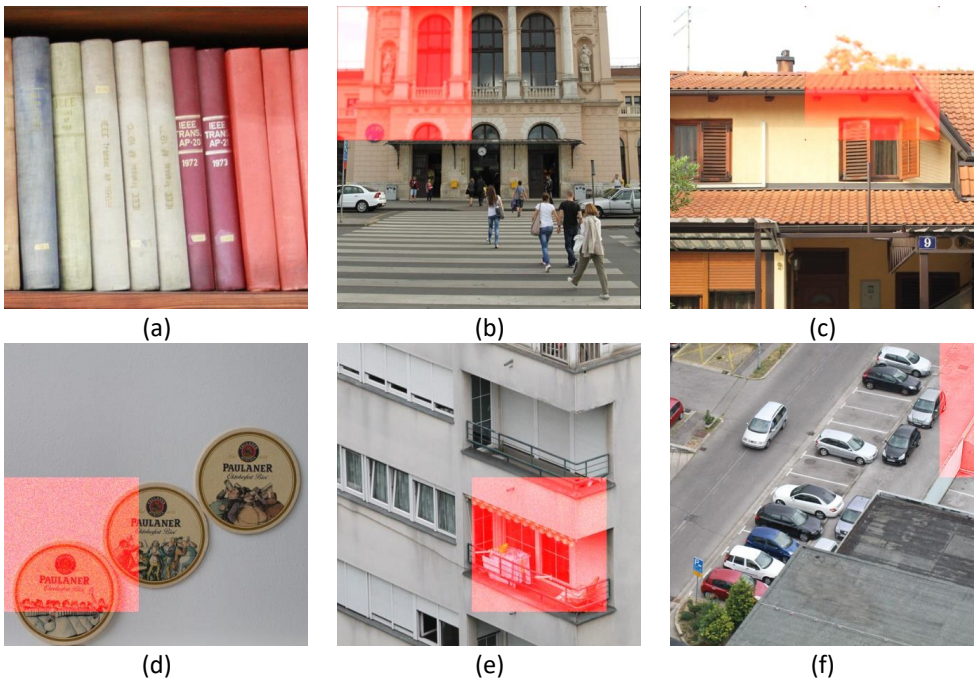


Şekil 4. Önerilen AlexNet ve GoogLeNet destekli ESA tabanlı sahtecilik tespiti sistemlerinin başarımlarının değerlendirilmesi için kullanılan orijinal görüntüler

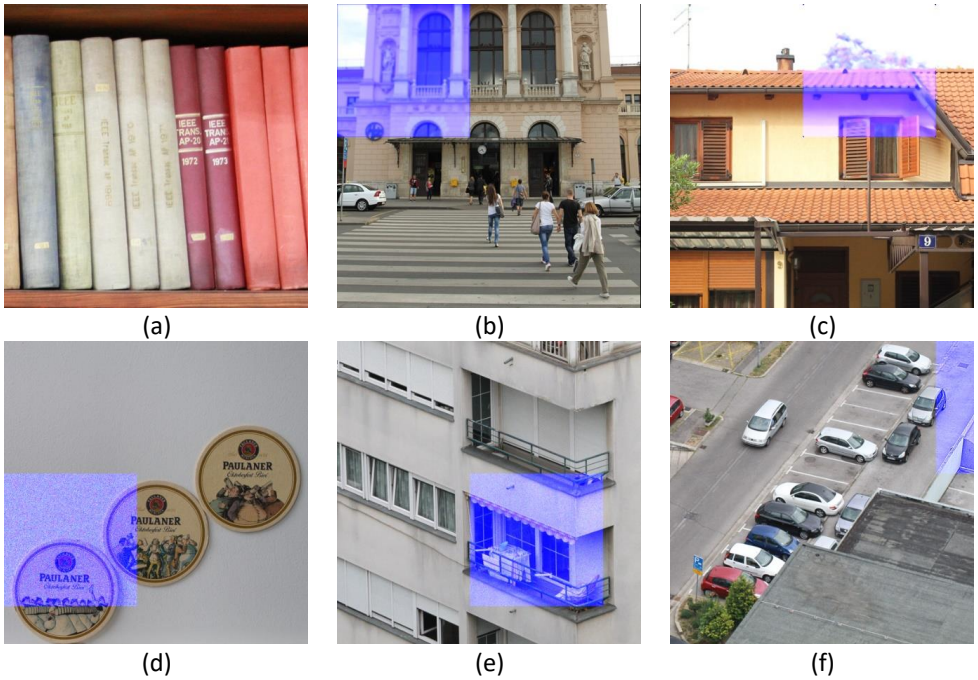


Şekil 5. Görüntülerin farklı bölgelerine faklı saldırıların uygulanması: a) orijinal, b) Gauss bulanıklaştırma, c) medyan filtreleme, d) Gauss gürültü ekleme, e) Poisson gürültü ekleme, f) keskinleştirme

Şekil 5'te gösterilen görüntülerdeki kurcalanmış bölgeler önerilen AlexNet ve GoogLeNet destekli ESA tabanlı sahtecilik tespiti sistemleri kullanılarak işaretlenmiştir. Şekil 6'da AlexNet destekli ESA, Şekil 7'de ise GoogLeNet destekli ESA tabanlı sahtecilik tespiti sistemi kullanılarak işaretlenmiş görüntüler gösterilmektedir. Şekillerdeki işaretlenmiş görüntüler incelendiğinde, kurcalanan bölgelerinin başarılı bir şekilde tespit edildiği görülmektedir.



Şekil 6. AlexNet destekli ESA kullanılarak sahtecilik bölgelerinin tespiti: a) orijinal, b) Gauss bulanıklaştırma, c) medyan filtreleme, d) Gauss gürültü ekleme, e) Poisson gürültü ekleme, f) keskinleştirme



Şekil 7. GoogLeNet destekli ESA kullanılarak sahtecilik bölgelerinin tespiti: a) orijinal, b) Gauss bulanıklaştırma, c) medyan filtreleme, d) Gauss gürültü ekleme, e) Poisson gürültü ekleme, f) keskinleştirme

5. SONUÇLAR

Bu çalışmada görüntüler üzerinde sahtecilik tespitinin gerçekleştirilmesi için ön eğitimli AlexNet ve GoogLeNet mimarilerinde transfer öğrenme gerçekleştirilmiştir. Bu ön eğitimli ESA mimarilerine sahtecilik tespiti görevi verilebilmesi için orijinal ve 5 farklı saldırı uygulanmış görüntülerden oluşturulan veri kümesi hazırlanmıştır. Hazırlanan eğitim veri kümesi ile yeniden eğitilen ağların test veri kümesi kullanılarak başarımları gözlemlenmiştir. Yapılan deneyler sonucunda AlexNet destekli ESA %99,48'lik bir başarı elde ederken, GoogLeNet destekli ESA ise %99,92'lik bir başarı elde etmiştir. GoogLeNet destekli ESA her bir saldırı türünde AlexNet destekli ESA'dan daha başarılı sınıflandırma yapmıştır. Önerilen sahtecilik tespiti sisteminin başarımını değerlendirmek için büyük boyutlardaki görüntüler üzerinde farklı bölgelere ağın eğitiminde kullanılan saldırı türleri ile müdahaleler yapılmıştır. Bu müdahale edilmiş görüntülerdeki müdahale edilmiş tüm bölgelerin tespiti önerilen AlexNet destekli ESA ve GoogLeNet destekli ESA tabanlı sahtecilik tespiti sistemleri kullanılarak başarılı bir şekilde gerçekleştirilmiştir.

Gelecekteki çalışmalarda, AlexNet ve GoogLeNet evrişimli sinir ağları yerine VGGNet [31] ve ResNet [32] gibi farklı mimarilerin kullanılması düşünülebilir. Bununla birlikte, önerilen yöntemin başarısı farklı saldırı türlerini içeren veri kümeleri kullanılarak değerlendirilebilir. Ayrıca önerilen yöntem daha küçük boyutlardaki görüntü bloklarında saldırı tespiti yapması için geliştirilebilir.

KAYNAKLAR

[1] N. B. A. Warif, A. W. A. Wahab, M. Y. I. Idris, R. Ramli, R. Salleh, S. Shamshirband, and K. K. R. Choo, "Copy-move forgery detection: Survey, challenges and future directions", *Journal of Network and Computer Applications*, vol. 75, pp. 259-278, 2016.

[2] W. Ding, W. Yan and D. Qi, "Digital image watermarking based on Discrete Wavelet Transform", *Journal of Computer Science and Technology*, vol. 17, no. 2, pp. 129-139, 2002.

[3] X. Wu, "A new technique for digital image watermarking", *Journal of Computer Science and Technology*, vol. 20, no. 6, pp. 843-848, 2005.

[4] E. Gul and S. Ozturk, "A novel hash function based fragile watermarking method for image integrity.", *Multimedia Tools and Applications*, vol. 78, no. 13, pp. 17701-17718, 2019.

[5] H. Zhang, C. Yang and X. Quan, "Image authentication based on digital signature and semi-fragile watermarking", *Journal of Computer Science and Technology*, vol. 19, no. 6, pp. 752-759, 2004.

[6] M. Alkawaz, G. Sulong, T. Saba and A. Rehman, "Detection of copy-move image forgery based on Discrete Cosine Transform", *Neural Computing and Applications*, vol. 30, no. 1, pp. 183-192, 2018.

[7] Y. Liu, Q. Guan and X. Zhao, "Copy-move forgery detection based on convolutional kernel network", *Multimedia Tools and Applications*, vol. 77, no. 14, pp. 18269-18293, 2018.

[8] J. Han, T. Park, Y. Moon and I. Eom, "Quantization-based Markov feature extraction method for image splicing detection", *Machine Vision and Applications*, vol. 29, no. 3, pp. 543-552, 2018.

[9] D. Cozzolino, G. Poggi and L. Verdoliva, "Recasting Residual-based Local Descriptors as convolutional neural networks", In Proc. 5th ACM Workshop on Information Hiding and Multimedia Securit-IHMMSec '17, 2017, pp. 159–164.

[10] Y. LeCun, Y. Bengio and G. Hinton, "Deep learning", *Nature*, vol. 521, no. 7553, pp. 436-444, 2015.

[11] J. Chen, X. Kang, Y. Liu and Z. J. Wang, "Median filtering forensics based on convolutional neural networks", *IEEE Signal Processing Letters*, vol. 22, no. 11, pp. 1849-1853, 2015.

[12] B. Bayar and M. C. Stamm, "A Deep learning approach to universal image manipulation detection using a new convolutional layer", In Proc. 4th ACM Workshop on Information Hiding and Multimedia Security-IH&MMSec '16, 2016, pp. 5–10.

[13] B. Bayar and M. C. Stamm, "On the robustness of constrained convolutional neural networks to JPEG post-compression for image resampling detection", In Proc. 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2017, pp. 2152-2156.

[14] Y. Rao and J. Ni, "A deep learning approach to detection of splicing and copy-move forgeries in images", In Proc. 2016 IEEE International Workshop on Information Forensics and Security (WIFS), 2016, pp. 1-6.

[15] I. Amerini, T. Uricchio, L. Ballan and R. Caldelli, "Localization of JPEG double compression through multi-domain convolutional neural networks", In Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 1865-1871.

[16] Q. Wang and R. Zhang, "Double JPEG compression forensics based on a convolutional neural network", *EURASIP Journal on Information Security*, vol. 2016, no. 1, 2016.

[17] Bunk, J.; Bappy, J.H.; Mohammed, T.M.; Nataraj, L.; Flenner, A.; Manjunath, B.; Chandrasekaran, S.; Roy-Chowdhury A.K.and Peterson, L. "Detection and localization of image forgeries using resampling features and deep learning", In Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 1881-1889.

[18] L. Bondi, S. Lameri, D. Güera, P. Bestagini, E. J. Delp and S. Tubaro, "Tampering detection and localization through clustering of camera-based CNN features", In Proc. 2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2017, pp. 1855-1864.

[19] D. C. Cireşan, U. Meier and J. Schmidhuber, "Transfer learning for Latin and Chinese characters with deep neural networks", In Proc. The 2012 International Joint Conference on Neural Networks (IJCNN), 2012, pp. 1-6.

[20] A. Krizhevsky, I. Sutskever, G. E. Hinton, "Imagenet classification with deep convolutional neural networks", In Proc. Advances in Neural Information Processing Systems 25 (NIPS 2012), 2012, pp. 1097-1105.

[21] Andrew Rabinovich C Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke and A. Rabinovich, "Going deeper with convolutions", In Proc. IEEE Conference on Computer Vision and Pattern Recognition, 2015, pp. 1-9

[22] B. Zou, Y. Guo, Q. He, P. Ouyang, K. Liu and Z. Chen, "3D Filtering by block matching and convolutional neural network for image denoising", *Journal of Computer Science and Technology*, vol. 33, no. 4, pp. 838-848, 2018.

[23] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu and M. Lew, "Deep learning for visual understanding: A review", *Neurocomputing*, vol. 187, pp. 27-48, 2016.

[24] E. A. Hadhrami, M. A. Mufti, B. Taha and N. Werghi, "Transfer learning with convolutional neural networks for moving target classification with micro-Doppler radar spectrograms,", In Proc. 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), 2018, pp. 148-154.

[25] M. A. Mufti, E. A. Hadhrami, B. Taha and N. Werghi, "Automatic target recognition in SAR images: Comparison between pre-trained CNNs in a tranfer learning based approach,", In Proc. 2018 International Conference on Artificial Intelligence and Big Data (ICAIBD), 2018, pp. 160-164.

[26] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?", In Proc. Advances in Neural Information Processing Systems 27 (NIPS 2014), 2014, pp. 3320-3328.

[27] M. Mehdipour Ghazi, B. Yanikoglu and E. Aptoula, "Plant identification using deep neural networks via optimization of transfer learning parameters", *Neurocomputing*, vol. 235, pp. 228-235, 2017.

[28] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo and G. Serra, "A SIFT-based forensic method for copy–move attack detection and transformation recovery", *IEEE Transactions on Information Forensics and Security*, vol. 6, no. 3, pp. 1099-1110, 2011.

[29] T. Ozcan and A. Basturk, "Transfer learning-based convolutional neural networks with heuristic optimization for hand gesture recognition.", *Neural Computing and Applications*, vol. 31, no. 12, pp. 8955-8970, 2019.

[30] D. Tralic, I. Zupancic, S. Grgic, and M. Grgic, "CoMoFoD-New database for copy-move forgery detection", In Proc. Electronics in Marine ELMAR-2013, 2013, pp. 49-54.

[31] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition.", CoRR, abs/1409.1556, 2014.

[32] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition.", CoRR, abs/1512.03385, 2015.