



DERİN SİNİR AĞLARI YARDIMIYLA FOTOMONTAJ TESPİTİ

Nihat Eren ÖZMEN¹, Ercan BULUŞ¹

¹ Tekirdağ Namık Kemal Üniversitesi, Bilgisayar Mühendisliği Bölümü, Çorlu-Tekirdağ, Türkiye

| Anahtar Kelimeler | Öz |
|--|--|
| Derin Öğrenme, Derin Sinir Ağları, Fotomontaj Tespiti, Mask RCNN, | Son yıllarda hızla gelişen teknoloji ile birlikte verilerin sağlıklı bir şekilde elde edilmesi, elde edilen verilerin korunması ve elde edilen verilerin özgün olması büyük önem taşımaktadır. Özgünlüğün tespiti özellikle görüntüler üzerinde büyük önem teşkil etmektedir. Görüntülerde bozulma ya da değişiklik olup olmadığını tespit etmek ise tıptan, belgede sahteciliğe kadar geniş bir çalışma alanını etkilemektedir. Fotomontaj tespiti için derin öğrenme algoritmaları ile mevcut görüntü işleme metodlarının aynı anda kullanılması verimliliği arttırmaktadır. Yapılan çalışmalar, derin sinir ağları, yüksek boyutlu girdilerden karmaşık istatistiksel özellikleri elde edebildikleri ve hiyerarşik temsillerini etkili bir biçimde öğrenebildiklerini göstermişlerdir. Bu çalışmada görüntü üzerinde değişiklik yapılmış kısım ile yapılmamış kısım arasındaki farkı daha rahat ayırabilmek için geliştirilmiş maske bölgesel evrimsel sinir ağı (Mask R-CNN) ile bu sinir ağına bağlanan sobel filtresi kullanılmaktadır. Sobel filtresi, sinir ağı ile tahmin edilen maskelerin zemin üzerindeki maskeye benzer görüntü gradyanlarına sahip olmasını teşvik etmek için yardımcı bir görev görür. Ağ ile kopyala taşıma ve birleştirme işlemleri algılanabilmektedir. Sinir ağı uygulanırken COCO veri seti kullanılmıştır. Yapılan çalışma ile daha yüksek başarı oranları elde edilmiştir. |

PHOTOMONTAGE DETECTION WITH DEEP NEURAL NETWORKS

| Keywords | Abstract |
|--|--|
| Deep Learning, Deep Neural Networks, Photomontage Detection, Mask RCNN, | With the rapidly developing technology in recent years, obtaining the data properly, protection of the obtained data and it is very important that the obtained data are original. Identification of originality is of great importance, especially on images. Detecting whether there is distortion or change in images affects a wide range of work field from medicine to document forgery. The simultaneous use of deep learning algorithms and existing image processing methods for photomontage detection increases efficiency. Studies have shown that deep neural networks can obtain complex statistical properties from high dimensional inputs and can learn their hierarchical representation effectively. In this study, in order to discriminate the difference between the part that has been changed and the part that has not been changed, we used the improved mask regional convolutional neural network (Mask R-CNN) and the sobel filter connected to this neural network. The Sobel filter acts as an assistant to promote masks to have similar mask image gradients on the ground estimated by the neural network. Copy-move and splicing operations can be detected with the network. The COCO data set was used when applying the neural network. Higher success rates were obtained with the study. |

Alıntı / Cite

Özmen, N.E., Buluş, E., (2020). Derin Sinir Ağları Yardımıyla Fotomontaj Tespiti, Mühendislik Bilimleri ve Tasarım Dergisi, 8(5), 236-240.

| Yazar Kimliği / Author ID (ORCID Number) | Makale Süreci / Article Process |
|--|---|
| N.E. Özmen, 0000-0002-0053-3865 | Başvuru Tarihi / Submission Date 07.12.2020 |
| E. Buluş, 0000-0001-9442-6253 | Revizyon Tarihi / Revision Date 25.12.2020 |
| | Kabul Tarihi / Accepted Date 27.12.2020 |
| | Yayın Tarihi / Published Date 29.12.2020 |

1. Giriş (Introduction)

İlerleyen teknoloji ile birlikte elde edilen verilerin sayısında her geçen gün büyük artışlar gerçekleşmektedir. Bu artışların en önemli sebebi ise internet kullanımı ve bununla birlikte artan sosyal medya kullanımınıdır. Elde edilen verilerin büyük çoğunluğunu ise görüntüler oluşturmaktadır. Büyük sayıda görüntü elde edilmesi ile bir takım sorunlar ile karşılaşmaktadır. Bu sorunların başında ise görüntünün orijinal olup olmadığının tespiti. Görüntü üzerinde herhangi bir değişiklik ya da bozulma olup olmadığının tespiti de bu sayede önem kazanmıştır. Görüntü üzerinde gerçekleştirilen bu değişikliklerde çeşitli yöntemler kullanılmakta ve çeşitli isimler verilmektedir. Bu yöntemlerden biri de fotomontajdır. Fotomontaj, iki ya da daha fazla görüntü ile kesme, yapıştırma, düzenleme ve ya görüntüleri üst üste bindirme yöntemi ile yapılabilmektedir. Fotomontaj yapılmış olan görüntülerin tespiti sadece gündelik hayatta değil bilimsel çalışmalardan, tıp alanına ve adli vakalara kadar geniş bir çalışma alanını etkilemektedir. Bu sebeple fotomontaj yapılmış görüntülerin tespiti ve üzerinde değişiklik yapılmış alanların belirlenmesi için güvenilir yöntemlere ihtiyaç duyulmaktadır (Jarusek vd., 2019).

Görüntü işleme alanında yapılan çalışmalar, elde edilen görüntülerdeki artış sebebi ile son yıllarda hız kazanmıştır. Fotomontaj tespiti de bu çalışmaların arasında yer almaktadır. Fotomontaj tespiti için görüntü işleme algoritmaları ve filtreler mevcuttur.

Derin öğrenme, çoklu katmanlara sahip verilerin temsil öğrenme yöntemi ile bir araya getirilmiş çoklu sayıdaki soyutlama yapısından meydana gelmektedir. Derin öğrenme yöntemleri kullanılarak nesne algılama ve tanıma, ses tanıma gibi birçok alanda başarılı sonuçlar elde edilmiştir. Derin öğrenmede kullanılan derin sinir ağları ile büyük verilerde elde edilmek istenen sonuçlar daha hızlı ve daha başarılı şekilde ortaya konulmuştur. Bu çalışmada fotomontaj tespiti için derin sinir ağları ile klasik görüntü işleme yöntemlerinden biri olan sobel filtresi beraber kullanımı daha başarılı sonuçlar elde edileceği sonucuna varılabilir (LeCun vd., 2015).

2. İlgili Çalışmalar (Related Works)

Son yıllarda derin sinir ağları mimarileri kullanılarak yapılan çalışmalarda önemli bir artış yaşanmıştır. Bu artışın sebebi derin sinir ağları yardımı ile daha hızlı ve daha başarılı sonuçlar elde edilmesidir. Farklı derin sinir ağları mimarilerinin karşılaştırması çalışılacak alan için önem arz etmektedir (Szegedy vd., 2016). Nesne tanıma ve algılama çalışmalarında derin sinir ağları yardımı ile önemli başarılar elde edilmiştir (Zhao vd., 2019). Nesne tanıma alanında elde edilen başarılar beraberinde fotomontaj tespiti üzerine yapılan çalışmalarında gelişmesine önemli katkıda bulunmuştur. Son yıllarda derin sinir ağları kullanılarak fotomontaj tespiti üzerine önemli çalışmalar yapılmıştır (Jarusek vd., 2019; Jaiswal ve Srivastava, 2019; Ahmad ve Khan, 2019).

3. Derin Sinir Ağları (Deep Neural Networks)

Derin öğrenme makine öğrenmesinin alt dalıdır. Derin öğrenme ile kullanılan sinir ağlarına derin sinir ağları denilmektedir. Derin öğrenme ile veriler üzerinde özellik çıkarımı yapılırken temsil öğrenme yöntemini kullanılmaktadır. Temsil öğrenme veriler üzerinde yapılması planlanan algılama ve sonrasında yapılacak olan sınıflandırma işlemleri gibi sonuçları otomatik bir şekilde öğretilen makinenin yapmasını sağlayan bir takım yöntemler tümüdür. Bu işlem yapılırken öncelikle basit parçalar oluşturulur ve basit parçalar yardımı ile daha yüksek düzeydeki temsillerin elde edilmesi ile gerçekleştirilir (LeCun vd., 2015).

3.1. Evrişimsel Sinir Ağları (Convolution Neural Network)

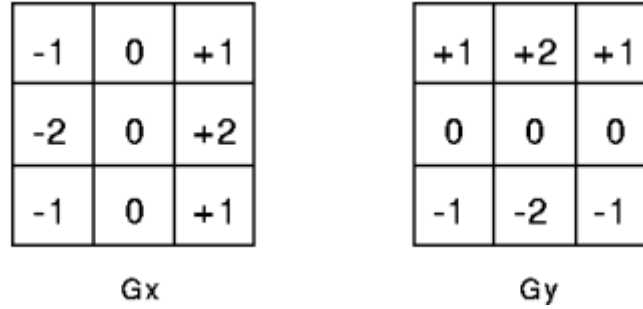
Derin öğrenme mimarilerinde biri olan evrişimsel sinir ağları (Convolution Neural Network-CNN) çok katmanlı algılayıcıların kullanıldığı bir derin öğrenme türüdür. Evrişimsel sinir ağları mimarisi birden çok katmandan oluşmakta, her bir katmanda ve katmanlar arasında ise bir dizi aşamadan meydana gelmektedir. İlk birkaç aşamada iki farklı katman kullanılmaktadır. Bunlar: Evrişimsel katmanlar ve havuz katmanlarıdır. Evrişimsel katmanlarda ağırlık seti ile önceki katmanlardan çıkan sonuçlar ile özellik haritaları oluşturulur. Bu ağırlıkların toplamının sonucu ReLu fonksiyonun geçirilir. ReLu doğrusal olmayan bir fonksiyondur. $f(z)$ ile gösterilmektedir.

$$f(z) = \max(z, 0) \quad (1)$$

Evrişimsel katmanlar ile katmanlardaki özelliklerin yerel bağlantıları ortaya çıkarılmaktadır. Bir başka deyişle görüntünün bir bölümünde görülen bir desen aynı görüntünün başka bir yerinde de görülebilir bu sebeple farklı konumlarda yer alan desen aynı ağırlığı paylaşmaktadır. Evrişimsel katmanlar ile ortaya konulan katman özelliklerinin bağlantıları havuz katmanı ile bir araya getirilmektedir. Havuz katmanında özellik haritalarında fazla yer kaplayan alanlar ele alınır ve komşu havuz birimleri, birden fazla satır veya birden fazla sütun alanlarını girdi olarak alır. Böylelikle temsilin boyutu düşürülür. Evrişimsel katman, ReLu fonksiyonu ve havuz katmanları birden çok sayıda arka arkaya dizilir ve bunları daha çok evrişimsel katman ve tam bağlantılı katmanlar takip eder. Evrişimsel sinir ağları ile görüntü ve dil işleme alanları başta olmak üzere birçok alanda başarılı sonuçlar elde edilmiştir (LeCun vd., 2015; He vd., 2015; Şeker vd., 2017).

3. 2. Sobel Filtresi (Sobel Filter)

Görüntü işleme alanında birçok algoritma ve yöntem kullanılmaktadır. Bunlardan biri de sobel filtresidir. Sobel filtresi, kenar algılama alanında kullanılan basit bir filtredir. Sobel filtresi görüntüde yer alan kenar bölgelerini ortaya çıkarmak için kullanılır. Filtre 3x3 boyutunda konvolüsyon matrisinden meydana gelmektedir.



Şekil 1. Konvolüsyon Matrisi

Bu matrisler dikey ve yatay kenarlar için ayrı ayrı olacak biçimde oluşturulmuştur. Matrisler görüntü üzerinde piksel piksel hareket eder ve kenarları bulma işlemi gerçekleştirilir. Bunun için

$$|G| = \sqrt{Gx^2 + Gy^2} \quad (2)$$

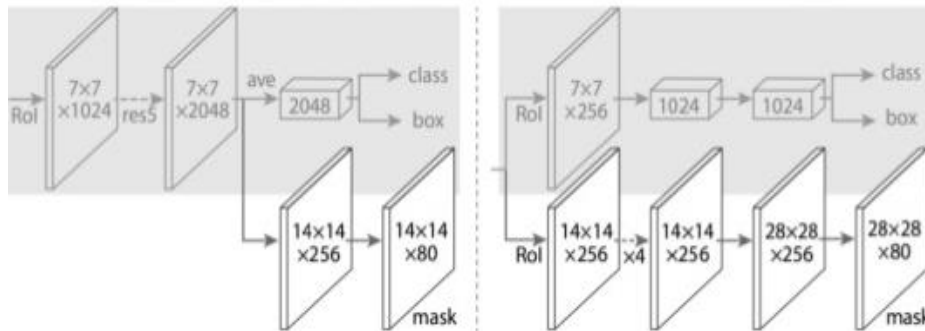
formülü ile piksel mutlak değeri hesaplanır.

$$\theta = \arctan\left(\frac{Gy}{Gx}\right) \quad (3)$$

formülü ile kenarın yön açısı hesaplanmaktadır. θ yardımı ile 0 derece yatay çizgileri 90 derecesi ise dikey çizgileri temsil etmektedir. Eğik çizgiler içinde aradaki açılar kullanılmaktadır. Görüntüdeki tüm kenarları ortaya çıkarmak için her iki matrisi kullanarak çıkan sonuçları toplayarak elde edilebilmektedir (Gonzalez ve Woods, 1992).

3.2. Mask R-CNN (Mask R-CNN)

Mask R-CNN, daha hızlı R-CNN (Faster R-CNN) mimarisine yeni bir adım eklenerek oluşturulmuş bir mimaridir. Mask R-CNN iki aşamalı bir prosedürü ele almaktadır. Bu iki aşama, daha hızlı R-CNN mimarisinde de kullanılmaktadır. Bunların ilki Bölge Teklif Ağı (Region Proposal Network-RPN) bu aşamada aday nesnelere için sınırlama kutuları önerilir. İkinci aşamada ise aday olarak belirlenen her bir kutudan RoIPool kullanılarak özellikler çıkarılır. Bu çıkarım ile sınıflandırma ve sınırlama kutusu regresyonu gerçekleştirilir. Mask R-CNN mimarisinde ise ikinci aşamada gerçekleştirilen işlemlere paralel olarak her bölge için ikili bir maske oluşturur. Mask R-CNN mimarisi oluşturulurken ResNet ağının 50 veya 101 katmanlı ağ mimarisi kullanılmaktadır. ResNet-50 ağ mimarisi daha hızlı R-CNN için yaygın bir kullanımdır. Bu sebeple ağ mimarisi seçiminde ResNet ağ mimarilerinin kullanımı daha başarılı sonuçlar sunmaktadır (He vd., 2017).

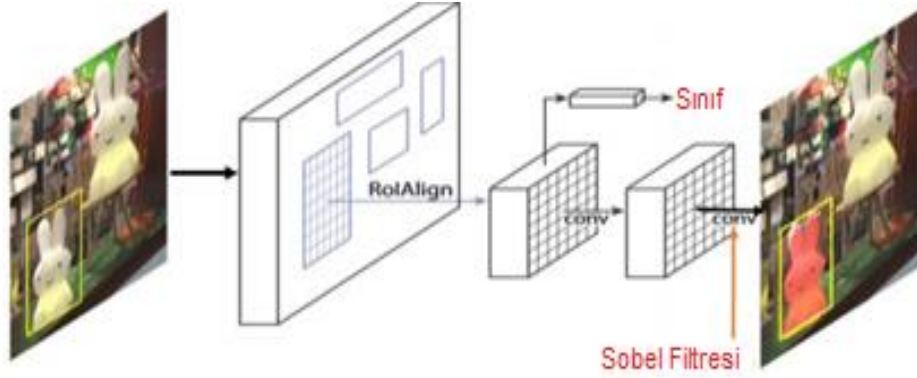


Şekil 2. Ağ Mimarisi

4. Mask R-CNN ile Tespit (Detection with Mask R-CNN)

Bu çalışmada manipüle edilmiş bölgeleri başarılı bir şekilde tespit etmek ve bulmak için Mask R-CNN mimarisi uygulanmaktadır. Manipüle edilmiş bölgelerin piksel düzeyindeki bilgileri Mask R-CNN için eğitimi denetleyici bilgi sağlamaktadır. Manipüle edilmiş bölgedeki sınırlara odaklanmak için sobel filtresi eklenmektedir. Bu filtre tahmin edilen manipüle edilmiş maskelerin algılama doğruluğunu arttırmaya yardımcı olmaktadır.

Oluşturulan ağ mimarisi ile kopyala-taşıma ve birleştirme işlemleri algılanabilmektedir. Bu mimari kullanılırken COCO veri seti kullanılmaktadır(Wang vd., 2019).



Şekil 3. Metodun gösterimi

5. Uygulama Çıktıları (Experimental Results)

Program gerçekleştirilirken python programlama dili ve tensorflow kütüphanesi kullanılmıştır. Model oluşturulurken şu adımlar izlenmiştir.

```
keras.layers.ZeroPadding2D((3, 3))
keras.layers.Conv2D(64, (7, 7))
BatchNormalization
keras.layers.Activation('relu')
keras.layers.MaxPooling2D
```

Yukarıda yer alan ZeroPadding2D fonksiyonu, görüntü tensörünün üst, alt, sol ve sağ taraflarına sıfır satırları ve ve sıfır sütunları eklemektedir. Conv2D fonksiyonu, çıkış tensörü oluşturmak için katman girdisini 2 boyutlu bir matrise düzleştirmektedir. BatchNormalization fonksiyonu bir katmana 0'a ortalanmış ve 0 ile 1 arasında değerlere sahip veriler vermemizi sağlamaktadır. Activation('relu') fonksiyonu yerel ağırlıklı toplamların sonuçların girdi olarak alır ve max(özellik,0) doğrusal olmayan fonksiyonun sonucunu verir. MaxPooling2D fonksiyonu, özelliklerin yer aldığı eksen boyunca mevcut olan her boyut için pool_size ile tanımlanmış olan pencere üzerinden maksimum değeri girdi olarak alt örnekler oluşturmaktadır.

ve sonrasında aynı adımları aşağıda verilen blok büyüklükleri için tekrarlanmıştır.

```
conv_block(, , [64, 64, 256])
conv_block(, , [128, 128, 512])
conv_block(, , [256, 256, 1024])
conv_block(, , [512, 512, 2048])
```

Model oluşturulduktan sonra epoch sayısı 500 olarak belirlenmiştir. Önceden eğitilmiş COCO ağırlık verileri kullanılmıştır sonrasında gerekli kod düzenlemeleri yapılmıştır(Chollet vd., 2017).



Şekil 4. İlk görüntü fotomontaj yapılmış görüntü ve İkinci görüntü ise üzerinde fotomontaj yapılmış alanın belirlenmiş hali (Wang vd., 2019)

Sobel filtresi eklenerek yapılan uygulama klasik Mask R-CNN uygulamasına kıyasla daha başarılı sonuçlar ortaya çıkarmıştır.

Tablo 1. Yöntem Sonuçları

| YÖNTEM | TEST ORANI |
|------------------------------|------------|
| Mask R-CNN | 0,69 |
| Mask R-CNN ve Sobel Filtresi | 0.73 |

6. Sonuç ve Tartışma (Result and Discussion)

Fotomontaj tespitinde birçok yöntem mevcuttur. Derin sinir ağları, klasik görüntü işleme yöntemlerinden farklı olarak bu alanda son yıllarda yaygın olarak kullanılan bir mimaridir. Derin sinir ağları ile daha büyük ve daha karmaşık verilerde çok daha hızlı ve başarılı sonuçlar elde edilmiştir. Sonuç olarak derin sinir ağları ile klasik görüntü işleme metotlarının beraber kullanımı bu alanda daha başarılı sonuçlar ortaya koyulacağı varsayılabilir. Mask RCNN görüntü bölütlemek için kullanılan derin sinir ağıdır. İleriki çalışmalarda Mask RCNN ile klasik bilgisayar ile görü yöntemlerinin hangi durumlarda kullanmanın daha yararlı olacağı araştırılacaktır. Ayrıca Mask RCNN ile farklı metotların kullanımı amaçlanmaktadır.

Çıkar Çatışması (Conflict of Interest)

Yazarlar tarafından herhangi bir çıkar çatışması beyan edilmemiştir.

Kaynaklar (References)

- Ahmad MF., Khan Z.,2019, Fake Image Detection Using Convolutional Neural Network.
- Chollet F., Yue-Hei Ng J., (ResNet50), 2017, GitHub repository, <https://github.com/fchollet/deep-learning-models/blob/master/resnet50.py>
- Gonzalez R. ve Woods R., 1992, Digital Image Processing, Addison Wesley, 414 - 428.
- He K., Gkioxari G., Dollár P. ve diğerleri, 2017, Mask R-CNN.
- He K., Zhang X., Ren S., ve Sun J.,2015, Deep residual learning for image recognition. 7.
- Jaiswal AK., Srivastava R., 2019, Image Splicing Detection using Deep Residual Network. SSRN Electronic Journal.
- Jarusek R., Volna E.,Kotyrba M., 2018, Robust steganographic method based on unconventional approach of neural networks. Neural Computing and Applications, 26, 111-116.
- Jarusek R., Volna E.,Kotyrba M., 2019, Photomontage detection using steganography technique based on a neural network. Neural Networks, 116, 150-165.
- LeCun Y., Bengio Y., Hinton G., 2015, Deep Learning. Nature, 521, 436-44.
- Szegedy C.,Ioffe S., Vanhoucke V.,Alemi A., 2016, Inception- v4, Inception-ResNet and the Impact of Residual Connections Learning. pp, 4278-4284.
- Şeker A., Diri B., Balık H. H., 2017, Derin Öğrenme Yöntemleri ve Uygulamaları Hakkında Bir İnceleme. Gazi Mühendislik Bilimleri Dergisi, 3.
- Wang X., Wang H., Niu S. ve Zhang J., 2019, Detection and localization of image forgeries using improved mask regional convolutional neural network. Mathematical Biosciences and Engineering, 16, 4581-4593.
- Zhao ZQ., Zheng P., Xu S.S, Wu X., 2019, Object detection with deep learning: A review. IEEE Transactions on Neural Networks and Learning Systems, 30, 3212-3232.