Özellik Çıkarma Yöntemlerinin Zararlı Yazılım Sınıflandırmasına Etkileri

Effects of Feature Extraction Methods to Malware Classification

Yusuf Kartal[a], Feyza Yılmaz[b], Kemal Özkan[a]
[a]*Eskişehir Osmangazi Üniversitesi, Bilgisayar Mühendisliği Bölümü, Eskişehir, Türkiye*
[b]*Eskişehir Osmangazi Üniversitesi, Elektrik-Elektronik Mühendisliği Bölümü, Eskişehir, Türkiye*
{ykartal,f.yilmaz,kozkan}@ogu.edu.tr

***Özetçe* — Teknolojik ilerlemeler zararlı yazılımların yayılmasını da beraberinde getirmektedir. Zararlı yazılımlara karşı siber güvenlik önemli bir rol oynamaktadır. Araştırmalar zararlı yazılımla mücadele ve tespit edilmesinde çeşitli alternatif yaklaşımlar geliştirmeye çalışmaktadır. Bu çalışmada önerilen bir yaklaşım, kötü amaçlı yazılımın tespitinde makine öğrenmesi teknikleri kullanmaktır. Bu çalışma kapsamında derin öğrenme teknikleri ile özellik çıkarma işlemi yapılarak farklı özellik çıkarma yöntemlerinin başarıları ölçülmüştür.**

Anahtar Kelimeler* — *derin öğrenme, zararlı yazılım sınıflandırma, özellik çıkarma

***Abstract* — Technological advances bring about the spread of malware. Cyber security plays an important role against malware. Studies are trying to develop various alternative approaches to combating and identifying malware. In this study, a recommended approach is to use machine learning techniques to detect malware. Within the scope of this study, the success of different feature extraction methods was measured by performing feature extraction with deep learning techniques.**

Keywords* — *deep learning, malware classification, feature extraction

I. GİRİŞ

Bilgisayar teknolojisinin hızlı bir şekilde gelişmesiyle birlikte kullanılan cihazların güvenliği ile ilgili konular önem kazanmaya başlamıştır. Günümüzdeki teknolojik ilerlemeler zararlı yazılımların yayılmasını da beraberinde getirmektedir. Zararlı yazılım, sahibinin rızası olmadan bir bilgisayar sistemine sızmak veya zarar vermek için tasarlanmış herhangi bir kod olarak tanımlanmaktadır [1].

Günlük hayatımızın önemli bir parçası haline gelen internet kullanımı gün geçtikçe artmaktadır. Zararlı yazılımlar tüm dünyaya internet üzerinden yayılırken bilgisayar ve internet kullanıcılarının zarar görmesine neden olan kötü amaçlı yazılımlardan kaynaklanan tehditlere karşı siber güvenlik önemli bir rol oynamaktadır. Siber güvenlik kötü amaçlı kişiler ya da yazılımlardan verilerin güvenliği, bütünlüğü ve gizliliğinin sağlanmasını amaçlamaktadır.

Özellikle, internet üzerinden dış dünyaya bağlı bilgisayarlar daha savunmasız olmakla birlikte zararlı yazılım tehditleriyle daha sık karşı karşıya kalmaktadırlar. Bu tehditler, yalnızca bilgisayarlara zarar vermekle kalmayıp, aynı zamanda gizlilik haklarını ihlal ederek bilgi hizmetlerinin erişilebilirliğine ve değerli kişisel bilgilerin toplanmasına zarar veren Virüsler, Solucanlar, Truva Atları, Kök Setleri, Botnetler, Arka Kapılar ve diğer kötü amaçlı yazılım saldırıları olarak da gözlenmektedir. Bu tür zararlı yazılımlar ve saldırılar bilgi güvenliği için en tehlikeli tehditler olarak kabul edilebilir. Son yıllarda, giderek artan zararlı yazılım örnekleri ve verdikleri zararlar, kötü amaçlı yazılım algılama teknolojisinde çarpıcı bir iyileştirme gerektiğini ortaya koymaktadır. Zararlı yazılımlara karşı koruma teknolojisi, 2003'ten bu yana dünya bilgi ve iletişim teknolojisi geliştirme eğilimlerinin en önemli 30 başlığından biridir [2]. Bununla birlikte, araştırmalar zararlı yazılımla mücadele ve tespit edilmesinde çeşitli alternatif yaklaşımlar geliştirmeye çalışmaktadır. Önerilen bir yaklaşım, kötü amaçlı yazılımın tespitinde etkinlik ve verimlilik elde etmek için makine öğrenmesi (sınıflandırma) teknikleri gibi veri madenciliği görevleri ile birlikte otomatik dinamik (davranış) zararlı yazılım analizi kullanmaktır. Paketlenmiş çalıştırılabilir bir ikili(binary) dosyayı 2 boyutlu görüntü biçimine dönüştürdükten sonra, benzer tanıdık üyelerle benzer özellikleri ortaya koyma eğiliminde olduğu görülmüştür. Bu tür çalışmaların ilham noktası, [3]'te önerilen test edilmiş ve onaylanmış hipoteze atfedilebilir. Metin editörleri veya ikili editörler gibi birkaç araç ikili veriyi hem görselleştirip hem de değiştirebilir [3]. Görselleştirme ve oldukça etkileşimli görsel analiz sistemlerinin kullanılması, çalıştırılabilir dosyaların araştırılması, karşılaştırılması ve özetlenmesi ile ilgili analiz sürecini desteklemeye yardımcı olabilir [4].

II. YÖNTEM

Bu bölüm, çalışmada izlenen yaklaşımı adım adım göstermektedir. İlk olarak zararlı yazılımların görselleştirilmesini ve yeniden düzenlenmesi gösterilmektedir. Ardından zararlı yazılımlar için VGG16,VGG19 ve Densenet ile özellik çıkarımı yapılmıştır. Eğitim ve sınıflandırma için k-NN ve SVM yöntemleri uygulanmıştır.

A. Önişlem

Hem sistemin eğitilmesi hem de eğitilen sistemin zararlı yazılımı sınıflandırabilmesi için eğitim için kullanılacak zararlı yazılımların ve sınıflandırılacak zararlı yazılıma ait ikili dosyanın ayrı ayrı görüntü dosyasına çevrilmesi gerekmektedir. Bu işlem için öncelikle zararlı yazılım 8 bitlik işaretsiz tamsayı değerler içeren vektöre, elde edilen vektörler 0 değeri siyah, 255 değeri beyaz olacak şekilde gri tonlamalı bir görüntü dosyasına dönüştürülür (Şekil 1). Oluşan görüntü dosyasında genişlik bir değere sabitlenirken boy dosya boyutuna bağlı olarak değişkenlik göstermektedir. Dosya boyutuna göre genişlik için sabitlenecek değerler Tablo 1'de verilmiştir [3].

B. Öznitelik Çıkarımı

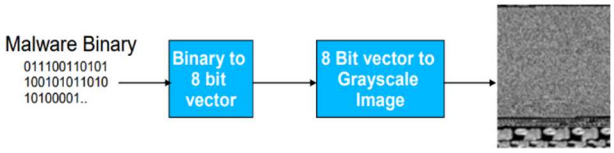
1) GIST

Görüntü işleme için sahne ve nesne sınıflandırmasında başarılı olmuş bir yöntemdir. Her görüntünün konumu farklı yönlere ve ölçeklere ayarlanmış filtrelerin çıktısıyla gösterilir. Daha detaylı bilgi [5]'da yer almaktadır.

2) VGG16 ve VGG19

VGG ağ mimarisi, Simonyan ve Zisserman tarafından 2014 yılında "Büyük Ölçekli Görüntü Tanıma için Çok Derin Konvolüsyon Ağları" çalışmasında yayınlanmıştır. VGG16 derin öğrenme algoritması 13 konvolüsyon ve 3 tam bağlı katmanından oluşan ağ mimarisidir (Şekil 2). Mimaride toplam 41 katman yer almaktadır. Bunlar Maxpool, Fullconnectedlayer, Relulayer, Dropoutlayer ve Softmaxlayer katmanlarıdır. Girdi katmanındaki görüntü 224x224x3 boyutundadır. Son katman ise sınıflandırma katmanıdır [6].

VGG19 derin öğrenme algoritması ise 16 konvolüsyon ve 3 tam bağlı katmanlı bir ağ mimarisine sahiptir. Bu mimaride toplam 47 katman yer almaktadır. Bunlar Maxpool, Fullconnectedlayer, Relulayer, Dropoutlayer ve Softmaxlayer katmanlarıdır. VGG16'da olduğu gibi giriş katmanı 224x224x3 boyutundadır ve son katman sınıflandırma katmanıdır [6].



Şekil 1. Zararlı Yazılımın Görüntüye Dönüştürülmesi [3]

TABLO I. DOSYA BOYUTU - GÖRÜNTÜ GENİŞLİĞİ TABLOSU [3]

Dosya Boyutu Aralığı (kB)	Görüntü Genişliği
<10	32
10-30	64
30-60	128
60-100	256
100-200	384
200-500	512
500-1000	768
>1000	1024

3) DenseNet

Her katmanı diğer katmanlarla ileri beslemeli olarak birbirine bağlayan ağ mimarisidir (Şekil 3). Her katman önceki tüm katmanların özellik haritaları giriş olarak kullanırken, kendi özellik haritalarını ise kendinden sonraki tüm katmanlara girdi olarak kullanır [7].

C. Sınıflandırma

1) SVM

SVM denetimli bir öğrenme yöntemidir ve veri sınıflandırma için iyi bir araçtır. SVM sınıflandırıcı, büyük verilerde bile hızlı bir öğrenmeye sahiptir. Veri sınıflandırma süreci temel olarak eğitim ve test veri setlerini içerir. SVM, iki veya daha fazla sınıflandırma problemi için kullanılır [8].

2) k-NN

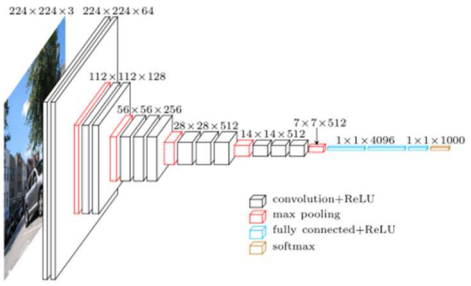
k-NN sınıflandırma yöntemi, çıkarılan özelliklere göre yeni verinin daha önceki verilerden k tanesinin yakınlığına bakılarak sınıflandırılması işlemidir [9].

III. VERİ SETİ

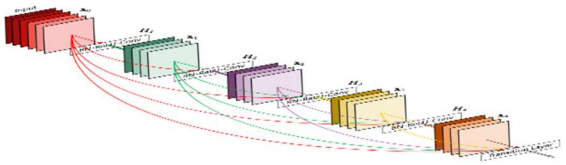
Zararlı yazılım sınıflandırma çalışmaları sırasında eğitim ve test aşamaları MalImg veri seti üzerinde yapılmıştır. MalImg veri seti 25 farklı aileye ait 9339 zararlı yazılım içermektedir. Tablo 2'de 25 farklı aileye ait bilgiler yer almaktadır [3].

IV. DENEY ÇIKTILARI

GIST, VGG16, VGG19 ve Densenet kullanılarak çıkarılan özellikler üzerinde k-NN ve SVM modellerinin eğitilmesi ile MalImg veri setinin sınıflandırılması sağlanmıştır. 4 farklı yönteme ilişkin çıktılar Şekil 4, Şekil 5, Şekil 6 ve Şekil 7'de karışıklık matrisi olarak verilmiştir. Karışıklık matrisinin köşegen elemanları doğru sınıflamaları, diğerleri ise yanlış sınıflamaları temsil eder [10]. C2LOP.gen!g ve C2LOP.P gibi bazı türler aynı sınıfın farklı tipleri olduğundan benzerlikleri ölçüsünde GIST ve k-NN kullanımı sonucunda hatalı



Şekil 2. VGG Ağ Mimarisi [6]

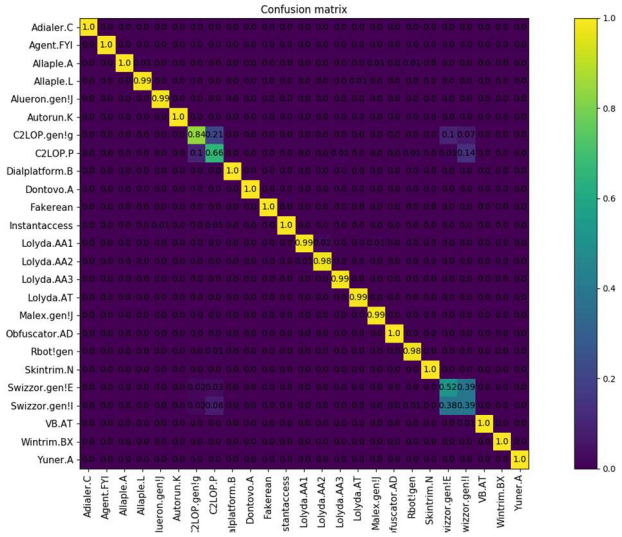


Şekil 3. DenseNet Mimarisi [7]

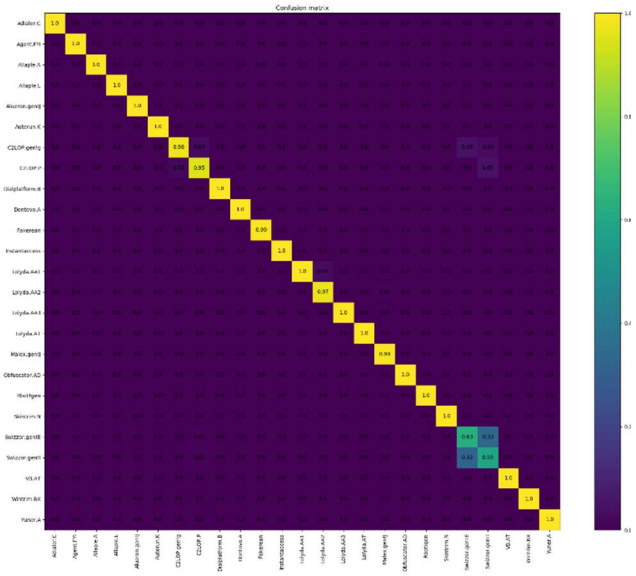
saptanmıştır. C2LOP için VGG16, VGG19 ve densenet ile SVM kullanımı sonucunda C2LOP örnekleri doğru saptanmış olmakla birlikte Swizzor.gen!E ve Swizzor.gen!I örneklerinde benzer bir durumun oluştuğu görülmektedir. Tablo 3'te verilen tabloda ortalama doğruluk değerleri açısından VGG16 ile özellik çıkarma sonucunda elde edilen değer bu çalışmada ulaşılan en iyi doğruluk oranını vermektedir

TABLO II. MALIMG VERİ SETİ ZARARLI YAZILIM SINIFLARI [3]

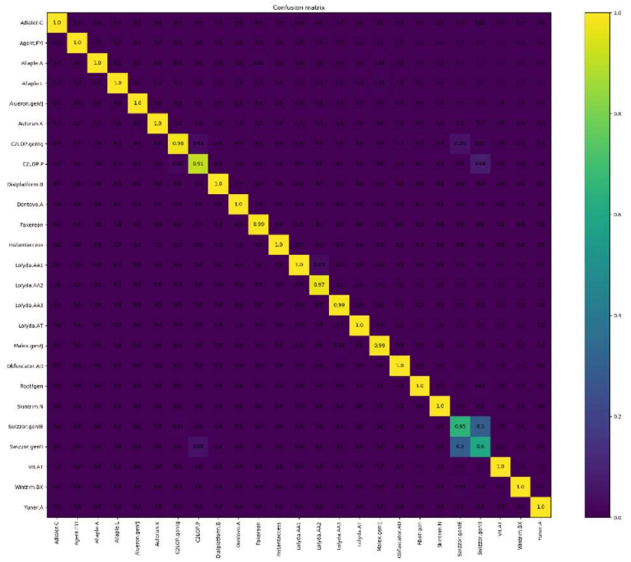
No	Sınıf	Sınıf Adı	Adet
1	Dialer	Adialer.C	122
2	Backdoor	Agent.FYI	116
3	Worm	Allaple.A	2949
4	Worm	Allaple.L	1591
5	Trojan	Alueron.gen!J	198
6	Worm:AutoIT	Autorun.K	106
7	Trojan	C2Lop.P	146
8	Trojan	C2Lop.gen!G	200
9	Dialer	Dialplatform.B	177
10	Trojan Downloader	Dontovo.A	162
11	Rogue	Fakerean	381
12	Dialer	Instantaccess	431
13	PWS	Lolyda.AA 1	213
14	PWS	Lolyda.AA 2	184
15	PWS	Lolyda.AA 3	123
16	PWS	Lolyda.AT	159
17	Trojan	Malex.gen!J	136
18	Trojan Downloader	Obfuscator.AD	142
19	Backdoor	Rbot!gen	158
20	Trojan	Skintrim.N	80
21	Trojan Downloader	Swizzor.gen!E	128
22	Trojan Downloader	Swizzor.gen!I	132
23	Worm	VB.AT	408
24	Trojan Downloader	Wintrim.BX	97
25	Worm	Yuner.A	800



Şekil 4. GIST ve k-NN Eğitimi Sonucu Karışıklık Matrisi



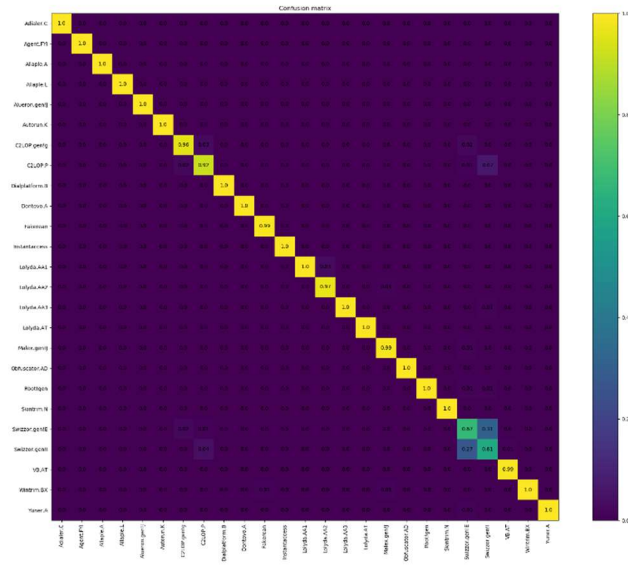
Şekil 5. VGG16 ve SVM Eğitimi Sonucu Karışıklık Matrisi



Şekil 6. VGG19 ve SVM Eğitimi Sonucu Karışıklık Matrisi

TABLO III. DOĞRULUK DEĞERLERİ

Özellik Çıkarma	Sınıflandırıcı Model	Ort. Doğruluk	Maks. Doğruluk
GIST	k-NN	0.9708	-
VGG16	SVM_Linear	0.9869	0.9894
VGG19	SVM_Linear	0.9864	0.9894
densenet	SVM_Linear	0.9864	0.9873



Şekil 7. Densenet ve SVM Eğitimi Sonucu Karışıklık Matrisi

V. SONUÇ

Bu çalışmada GIST, VGG16, VGG19 ve densenet modellerinin zararlı yazılım tespiti ve sınıflandırılmasında özellik çıkarma işlemi sonucundaki başarımları üzerinde durulmuştur. Veri seti her bir model için python dili kullanılarak analiz edilmiştir. Doğruluk oranlarının yüksek olarak görüldüğü deneysel çıktılar incelendiğinde aynı sınıfa ait farklı tipteki zararlı yazılımların birbirlerinin yerine tespit edilebildiği görülmektedir. GIST ve k-NN ile benzerliklerinden dolayı hatalı sınıflandırılan örneklerin VGG16, VGG19 ve densenet ile doğru sınıflandırıldığı tespit edilmiştir. Bu durum özellik çıkarma yöntemlerinin iyileştirilmesi ile elde edilen doğruluk oranlarının daha yüksek olabileceği anlamı taşımaktadır. Benzerlik gösteren bu zararlı yazılımlar için assembly kodları üzerinde metin işleme tekniklerinin kullanılarak daha başarılı sonuçlar alınması ileride yapılacak çalışmalar arasındadır.

KAYNAKLAR

[1] Sanjaa, B., & Chuluun, E. (2013, Haziran). Malware detection using linear SVM. In Strategic Technology (IFOST), 2013 8th International Forum on (Vol. 2, pp. 136-138). IEEE.

[2] Sukhbaatar, B. (2007, Ekim). Introduction of new information and Communication technologies in mongolia. In Strategic Technology, 2007. IFOST 2007. International Forum on (pp. 24-27). IEEE.

[3] Nataraj, L., Karthikeyan, S., Jacob, G., & Manjunath, B. S. (2011, Temmuz). Malware images: visualization and automatic classification. In Proceedings of the 8th international symposium on visualization for cyber security (p. 4). ACM.

[4] Wagner, M., Fischer, F., Luh, R., Haberson, A., Rind, A., Keim, D. A., ... & Viola, I. (2015, Mayıs). A survey of visualization systems for malware analysis. In EG Conference on Visualization (EuroVis)-STARs (pp. 105-125).

[5] Oliva, A., & Torralba, A. (2001). Modeling the shape of the scene: A holistic representation of the spatial envelope. International journal of computer vision, 42(3), 145-175.

[6] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.

[7] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). Densely connected convolutional networks. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 4700-4708).

[8] Vapnik, V. N. (1999). An overview of statistical learning theory. IEEE transactions on neural networks, 10(5), 988-999.

[9] Soucy, P., & Mineau, G. W. (2001). A simple KNN algorithm for text categorization. In Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on (pp. 647-648). IEEE.

[10] Uyar, M., Kaya, Y., & Ataş, M. (2013, Nisan). Classification of power quality disturbances based on S-transform and image processing techniques. In Signal Processing and Communications Applications Conference (SIU), 2013 21st (pp. 1-4). IEEE.