



## Video Görüntülerinde Şiddet İçeren Aktivitelerin Lstm Ağı ile Tespiti

### Detection of Violent Activities in Video Images with Lstm Network

Ali Seydi Keçeli <sup>1\*</sup>, Aydın Kaya <sup>1</sup>

<sup>1</sup> Hacettepe Üniversitesi Bilgisayar Mühendisliği Bölümü, Ankara, TÜRKİYE

Sorumlu Yazar / Corresponding Author \*: [aliseydi@cs.hacettepe.edu.tr](mailto:aliseydi@cs.hacettepe.edu.tr)

Geliş Tarihi / Received: 29.08.2018

Kabul Tarihi / Accepted: 14.06.2019

Atıf şekli / How to cite: KEÇELİ, A.S., KAYA A. (2019). Video Görüntülerinde Şiddet İçeren Aktivitelerin Lstm Ağı ile Tespiti. DEUFMD, 21(63), 933-939.

Araştırma Makalesi/Research Article

DOI:10.21205/deufmd.2019216321

#### Öz

Bilgisayarlı görü alanında hareket tanıma gerek RGB videolar, gerekse derinlik haritaları üzerinde fazlasıyla çalışılmış bir konu olmakla beraber; şiddet içeren hareketlerin tespiti görece olarak az çalışılmış bir alandır. Gelişmekte olan teknoloji ve internet ağı sayesinde, büyük miktarlarda video verisine kolayca erişilebilmektedir. Bu sayede, birçok şiddet içerikli video da kolayca erişilebilir hale gelmiştir. Şiddet içeren sahnelere sahip videoların etiketlenmesi, güvenlik ve içerik tabanlı video arama sistemleri için önemlidir. Güvenlik kamera sistemleri genellikle şiddeti ve uygunsuz hareketleri tespit etmek için elverişsizdir. Büyük ölçekli bir güvenlik kamera sistemi için, bir operatörün tüm kameraları aynı anda izlemesi imkânsızdır. Öte yandan, video akışı sitelerine yüklenen videoları kontrol edebilen otomatik video değerlendirme ve etiketleme sistemleri için de giderek artan bir ihtiyaç bulunmaktadır. Bu nedenlerden ötürü şiddet tespiti daha da önem kazanan bir konu haline gelmiştir. Bu çalışmada video görüntüleri üzerinde Transfer Öğrenme ve Long Short Term Memory (LSTM) ağı tabanlı bir yöntem önerilmiştir. Doğrudan RGB görüntülerinden, optik akış değerlerinin ve RGB çerçeve serilerinin türevi hesaplanarak elde edilen hız görüntülerinden GoogleNet kullanılarak derin özellikler elde edilmiştir. Elde edilen derin özellik serileri LSTM ağına girdi olarak verilmiştir. Önerilen yöntem literatürde bu tarz çalışmaların test edilmesinde yaygın olarak kullanılan Hockey Fight ve Violent Flow veri kümeleri ile test edilmiştir. Deney sonuçları literatürdeki çalışmalarla karşılaştırılabilir düzeydedir.

**Anahtar Kelimeler:** Şiddet Eylemi, Derin Öğrenme, CNN, GoogleNet, LSTM

#### Abstract

Although action recognition is a widely studied area on both RGB videos and depth map, violent activity detection is a relatively less studied area. With developing technology and the growing internet, large-scale video data become easily accessible. This also makes videos with violent scenes accessible. Labeling the violent scenes in videos is important for the content-based multimedia retrieval systems. Standard surveillance systems are incapable of detecting violent and improper activities. It is impossible for a human operator to watch all the records for a large-scale surveillance system. On the other hand, there is an increasing demand for automatic labeling systems to check the videos uploaded into video streaming sites. For these reasons, the automatic detection of violent activities is becoming more and more important. In this study, we propose a Transfer Learning and LSTM (Long Short-Term Memory) network-based method. Deep features extracted from GoogleNet by using RGB sequences,

velocity sequences and acceleration sequences computed from the first and second derivative of the pixels values are given into the LSTM network as input. The proposed method is tested with Hockey Fight and Violent Flow datasets that are commonly used in the literature. Experimental results are comparable to those in the literature.

**Keywords:** Violent Activity, Deep Learning, CNN, GoogleNet, LSTM

## 1. Giriş

Şiddet algılama, hareket tanıma probleminin alt sorunu olarak kabul edilmektedir fakat hareket tanımaya nazaran daha az çalışılmıştır [1,2]. Videodan hareket tanıma problemi bazı zorluklar barındırmaktadır. Bu zorluklardan bazıları, aktörlerin hızlı hareketi, örtüşen kişiler veya eylemler, kamera görüntüsünün hareketi ve görsel karışıklıklar olarak sıralanabilir. Literatürdeki birçok hareket tanıma yöntemi belirli eylemlere odaklanmaktadır, fakat şiddet faaliyetleri önceden tahmin edilememektedir ve birçok değişen koşula sahiptir.

Bu alanda yapılan ilk çalışmalar, kan ve ateşli olayları tespit etme ve şiddet olaylarının seslerini algılama üzerine kurulmuştur [1]. Clarin ve diğ. [2] şiddet aktivitelerini tespit etmek için Kohonen'in kendi kendini düzenleyen haritalarını (Self-organizing maps) kullanarak video çerçevelerinde cilt, kan ve hareket tespit etmiştir. Gong ve diğ. [3] düşük seviyeli görsel öznitelikleri ve yüksek seviyeli ses özniteliklerini birleştirerek filmlerdeki şiddetli içeriklerini algılamıştır. Kooij ve diğ. [4] saldırgan insan davranışını saptamak için ses ve video ipuçlarını birleştirmiştir. Ses ve görsel özellikleri bir araya getiren bir başka çalışma Lin ve Wang tarafından önerilmiştir [5]. Araştırmacılar, ses için zayıf denetime tabi tutulan bir sınıflandırıcıyı eğitmişlerdir ve hareket, patlama ve kan algılayan görsel sınıflandırıcıyla birleştirmişlerdir. Kişi üç boyutlu olarak izlenmiş ve aşırı hareketlerle ilgili öznitelikler çıkarılmıştır. Ses bilgileri konuşma çığlık ve şarkı olarak sınıflandırılmıştır. Hassner ve diğ. [6], optik akış bilgilerinden oluşturulmuş şiddet akış tanımlayıcısı (ViF) adlı bir yöntem önermişlerdir. Bu yöntemde, ardışık çerçeveler arasındaki ilk optik akış değerleri hesaplanmaktadır. Daha sonra ikili bir gösterge elde edilmektedir. İkili göstergede, belirli bir eşik değerinden daha büyük optik akış büyüklüğüne sahip pikseller seçilmektedir. İkili göstergeler elde edildiğinde, nicelenmiş değerlerin frekanslarının vektörü olan bir histogram oluşturulmaktadır. En son adım olarak bir destek vektör makinası (SVM) sınıflayıcı, hesaplanan ViF tanımlayıcılarıyla eğitilmektedir. Lloyd ve

diğ. [7] kalabalığı bir doku olarak düşünmüş ve kalabalık dokudaki değişiklikleri kullanarak kalabalık dinamikleri modellemişlerdir. Haralick öznitelikleri her çerçeve için çıkarılmış ve bu özelliklerin zaman içindeki değişimleri hesaplanmıştır. Geliştirilen bu yöntem Şiddetli Kalabalık Dokusu olarak adlandırılmıştır. Boiman ve Irani [8] görüntüdeki aykırı durumları ve şüpheli davranışları tespit etmek için veri odaklı bir yaklaşım önermiştir. Rota ve diğ. [9] öznitelikleri çıkarmak için hareket yörüngelerini kullanmışlardır. Şiddetli etkileşimi tespit etmek için yörünge özniteliklerinden bir kod çizelgesi oluşturulmuştur. Gao ve diğ. [10], Hassner ve diğ. tarafından önerilen ViF yöntemini geliştirmişlerdir. ViF tanımlayıcılarını geliştirmek için optik akışın yönlendirme bilgilerini kullanmışlardır. Bu yöntemde, akış vektör açılarından bir histogram elde edilmektedir. Daha sonra bu histogramlardan ikili göstergeler çıkarılmaktadır. Gerçek sınıflandırma oranını artırmak için önerdikleri yöntemin (OviF) özelliklerini ViF özellikleri ile birleştirmişlerdir. Dai ve diğ. [11] şiddet algılama problemini çözmek için derin öğrenme yöntemlerini kullanmıştır. Önce şiddet tespiti için bir evrimsel sinir ağı (Convolutional Neural Network-CNN) modeli eğitmişlerdir. Ardından, hem istatistiksel hem de optik akış özelliklerini çıkarmak için iki akışlı bir CNN önermişlerdir. Deniz ve diğ. [12] Hızlı Fourier Dönüşümü(FFT) tabanlı bir yöntem önermektedir. Video çerçeveleri üzerinde uygulanan FFT ve Radon dönüşümü ve uzaysal zamansal özellikler çıkarılmıştır ve sınıflandırma için kullanılmıştır. Arceda ve diğ. [13] Horn-Schunck optik akış hesaplama yöntemini kullanarak ViF yönteminin bir varyasyonunu önermektedir. SVM sınıflayıcı kullanarak, farklı veri kümeleri üzerindeki başarımları ölçmüşlerdir. Keçeli ve Kaya [14,15] derin öğrenme ve transfer öğrenmeye dayanan modeller önermişlerdir.

Makalenin ilerleyen kesimlerinde ilk olarak önerilen yönteme ait öznitelik çıkarma adımları anlatılacaktır. Deneyler kesiminde önerilen yöntemin erişime açık veri kümeleri ile yapılan deneylerin sonuçlarından ve kullanılan sınıflayıcı modellerden bahsedilecektir. Son

olarak da sonuç kesiminde deneylerden elde edilen sonuçlar ve gelecek çalışmalar tartışılacaktır.

## 2. Yöntem

Video çerçevelerinden elde edilen hız (velocity) ve ivmelenme (acceleration) serileri ham RGB çerçeveler ile birlikte girdi önceden eğitilmiş GoogleNet ağına girdi olarak verilmiştir. Sınıflama için Uçtan-Uca (end-to-end) bir LSTM (Long Short Term Memory) ağı tasarlanmıştır. Önerilen yöntemin genel akış şeması Şekil-1 de gösterilmektedir.

### 2.1. Hız ve İvme Görüntüleri

Hız görüntüsü ardışık video çerçevelerinin zaman eksenindeki 1. dereceden türevidir. Eğer video serisini  $W(x, y, t)$  fonksiyonu olarak kabul edersek, hız görüntüsünü Eşitlik-1 ile ifade edebiliriz. Bu eşitlikte  $\left(\frac{dw}{dt}\right)$  video çerçevelerinin zaman eksenindeki değişimini göstermektedir. Hız görüntüsü aynı konumda bulunan piksellerin zamana bağlı değişimidir. Bu işlem sonucunda bir çerçeve serisi ya da üç boyutlu bir matris elde edilmektedir.

$$V = \frac{dw}{dt} \quad (1)$$

İvmelenme görüntüsü ise video çerçevelerinin 2. dereceden türevi ya da hız görüntüsü serisinin türevidir. İvmelenme görüntüsü Eşitlik-2'deki gibi ifade edilir ve hesaplanabilir. Bu işlem sonucunda yine bir çerçeve ya da matris serisi elde edilmektedir.

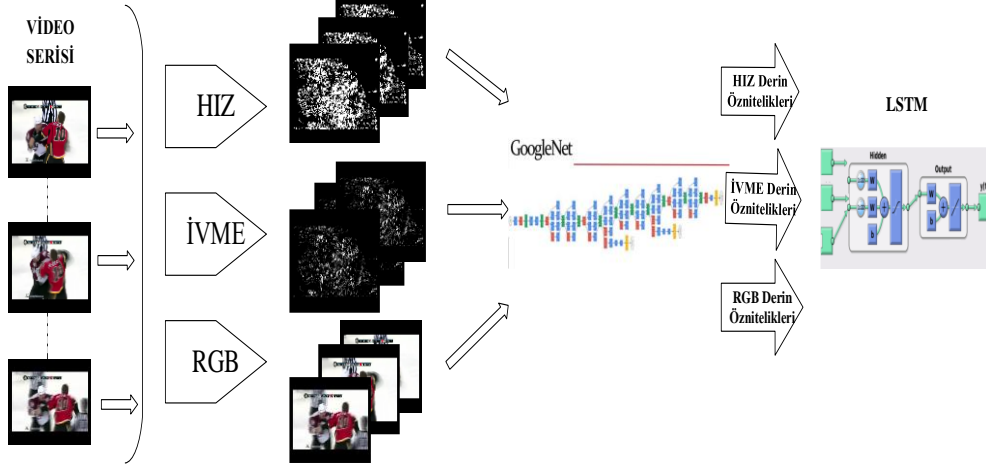
$$a = \frac{dv}{dt} = \frac{d^2w}{dt^2} \quad (2)$$

Hız ve ivme görüntüleri dışında ham video çerçeveleri de öznetelik çıkarılmasında kullanılmıştır. Her üç kaynaktan elde edilen derin öznetelikler birleştirilmiştir.

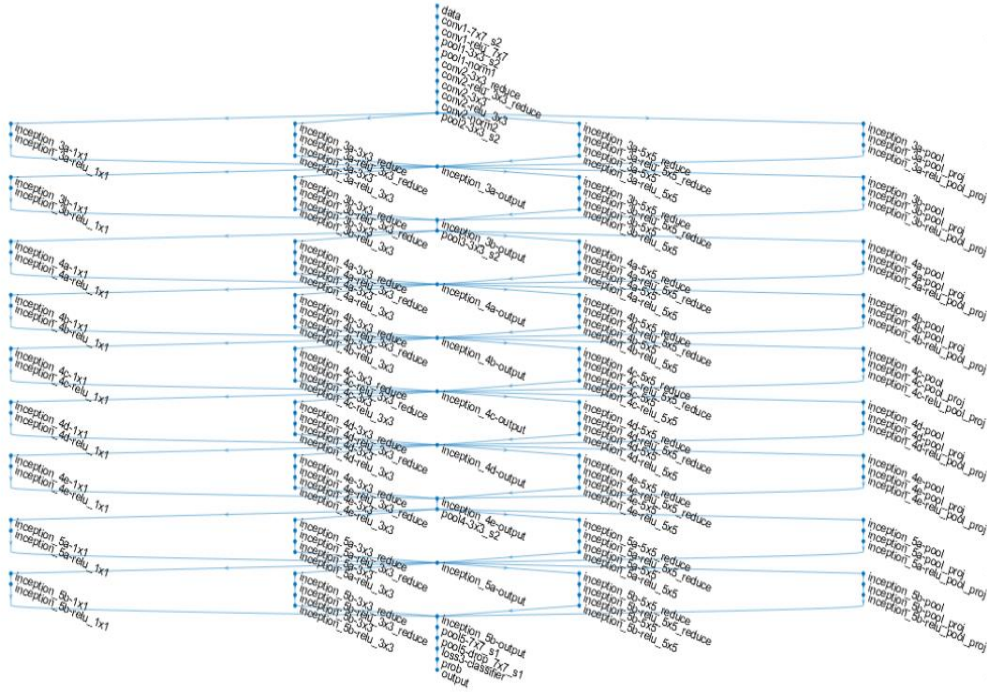
### 2.2. GoogleNet ile Derin Özneteliklerin Çıkarılması

Video serilerinden elde edilen RGB çerçeveleri ile hız ve ivmelenme çerçevelerinden öznetelik çıkarma amacıyla, önceden eğitilmiş GoogleNet derin öğrenme ağı kullanılmıştır. GoogleNet, ImageNet veri kümesinde bulunan milyonlarca görüntü kullanılarak eğitilmiştir. GoogleNet'in literatürde yaygın olarak kullanılan diğer (AlexNet VGG-16, VGG-19) modellerinden farkı, *inception* modüllerine sahip olmasıdır. *Inception* modülleri kendi içerisinde *convolution* ve *pooling* katmanları içermektedir. *Inception* modülleri içerisinde bulunan *convolution* katmanları farklı ölçeklerde evrişim ve birleştirme işlemleri yapabilmektedir. GoogleNet'in diğer bir farklılığı ise koşut *convolution* katmanlarına sahip olmasıdır. GoogleNet ağının mimarisi Şekil-2'de gösterilmektedir.

Transfer öğrenme önceden eğitilmiş bir modelin başka bir görev için tekrar kullanılmasıdır. Eğitilmiş modelin yeni bir problem için uyarlanması olarak da düşünülebilir. Derin öğrenme ağları bağlamında ise, önceden eğitilmiş bir ağdan öğrenilen özneteliklerin başka bir probleme uyarlanmasıdır. Önceden eğitilmiş derin öğrenme ağlarının görsel tanıma problemlerinde etkin öznetelikler çıkarmada üstün bir başarıyı bulunmaktadır. Bu kapsamda video görüntülerini oluşturan çerçeve serilerinden elde edilen ham RGB serileri (video çerçeveleri), hız ve ivme serilerine ait her eleman GoogleNet ağına girdi olarak verilmiştir. GoogleNet ağında bulunan *loss3-classifier* ve *pool5\_7x7\_s1* katmanlarına ait aktivasyon değerleri öznetelik olarak kullanılmıştır. Önceden eğitilmiş CNN'lerde farklı katmanlara ait aktivasyon değerleri farklı soyutlamalara karşılık gelmektedir. *Loss3-classifier* ile her çerçeve için 1000 öznetelik, *pool5\_7x7\_s1* katmanı ile de 1024 öznetelik elde edilmiştir. Kısaca özetlemek gerekirse her biri 2024 özneteliğe sahip seriler elde edilmiştir. Sonuç olarak her harekete ilişkin video serileri derin öznetelik serilerine dönüştürülmüştür.



Şekil 1. Önerilen Sistemin Genel Akış Şeması



Şekil 2. GoogleNet Mimarisi ([16]'nolu kaynaktan uyarlanmıştır)

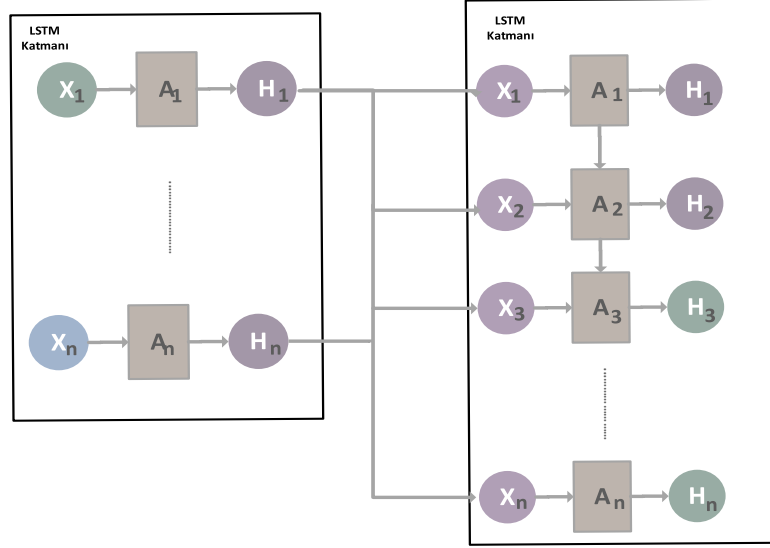
### 2.3 LSTM AĞI

Geleneksel sinir ağları önceki gözlemlere dayanarak tahmin yapamamaktadır. Bu tür gözlem serilerinden tahmin ve sınıflama yapmak için RNN (Recurrent Neural Network) modelleri geliştirilmiştir [17]. LSTM ağlarında bir tür RNN'dir. RNN'lerin asıl amacı ardışık bilgileri kullanmaktır. Geleneksel bir sinir ağında tüm

girişlerin (ve çıktıların) birbirinden bağımsız olduğu varsayılmaktadır. RNN'ler bir serinin her ögesi için aynı görevi gerçekleştirdiğinden tekrarlı (recurrent) olarak adlandırılır ve çıktı önceki hesaplamalara bağlıdır. İleri beslemeli sinir ağlarından farklı olarak, RNN'ler kendi giriş belleğini girdilerin keyfi sıralarını işlemek için kullanabilirler. RNN'leri dizinin gözlemlene değerleri ile ilgili bilgi toplayıp "bellek" taşırlar.

RNN'ler teorik olarak çok uzun diziler için kullanılabilirler, ancak pratikte yalnızca birkaç adım geriye dönmekle sınırlıdır. Örnek bir LSTM Şekil-3'te gösterilmiştir. LSTM katmanları çok sayıda LSTM ünitesinden oluşmaktadır. Bir LSTM'e ünitesinde  $X$ 'ler seri olarak verilen girdileri,  $A$ 'lar kapıları içeren hücreyi,  $H$  ise

çıktıyı belirtmektedir. LSTM katmanları arasındaki bağlantı Şekil-3'te gösterildiği biçimdedir. Bir LSTM ünitesinin çıktısı bir sonraki katmandaki ünitelere girdi olarak verilmektedir.

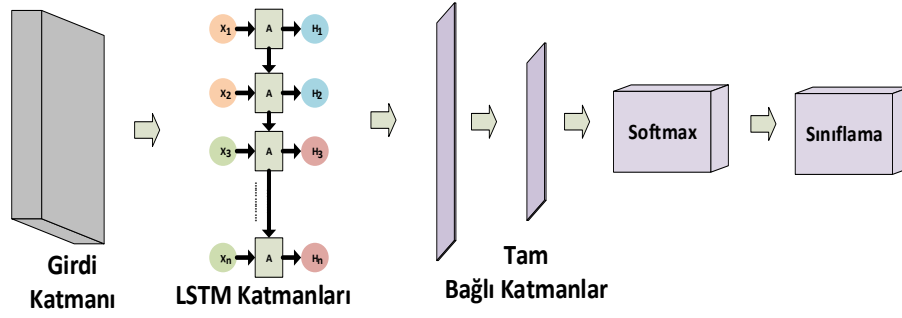


Şekil 3. Örnek bir LSTM gösterimi.

Önerilen yöntem kapsamında altı katmanlı bir LSTM ağı geliştirilmiştir. Geliştirilen ağın 6072 adet girişi olan bir girdi katmanı, 800 ve 400 birimlik 2 LSTM katmanı, çıkış sayıları 40 ve 2 olan 2 adet tam bağlı katman (fully connected layer), bir softmax katmanı ve son olarak bir sınıflama katmanından oluşmaktadır.  $H$  birimlerinin sayısı 1. LSTM biriminde 800 değerinde 400'dür.

Girdi katmanları, gözlem serilerini girdi olarak okuyup LSTM katmanlarına ileten katmandır.

LSTM katmanları, daha önce belirttiğimiz üzere dizi üzerinden hesaplamalarla tahmin yapmaktadır. Tam bağlı katmanlar sınıf puanlarını hesaplamaktadır. Tam bağlı bir katmandaki nöronlar bir önceki katmandaki tüm etkinleşmelere tam bağlantılara sahiptir. Softmax katmanında her sınıf için olasılık değerlerine bir softmax fonksiyonu uygulanır ve tüm girdiler 1'e tamamlanır. Geliştirilen LSTM ağının mimarisi Şekil-4'te gösterilmiştir.



Şekil 4. Geliştirilen LSTM Ağı.

### 3. Deneyler

Önerilen yöntemin test edilmesinde Violent Flow [6] ve Hockey Fight [18] veri kümelerinden faydalanılmıştır. Bu veri kümeleri literatürde yer alan çalışmalarda da kullanılan yaygın ve açık erişimli veri kümeleridir. Hockey Fight veri kümesi buz hokeyi karşılaşmaları sırasında kaydedilmiş 1000 farklı videodan oluşmaktadır. Bu videoların yarısı şiddet içeren videolar iken diğer yarısı şiddet içermemektedir. Violent Flow veri kümesinde ise 246 adet kalabalık aktiviteleri içeren videodan oluşmaktadır. Bu veri kümesinde de yine şiddet içeren ve içermeyen örneklerin oranı yarı yarıdır.

Test aşamasında 5 katlı çapraz geçerlilik yöntemi uygulanmıştır. Bu yaklaşımın uygulanmasının sebebi literatürde yer alan benzer çalışmalarda [6,19] bu test yönteminin kullanılmasıdır. Bu test yönteminde veri kümesi rastgele 5 parçaya ayrılır. Her parçada şiddet içeren ve içermeyen örnek sayısı aynıdır. Bir parça test için seçilir kalan parçalar ise model eğitiminde kullanılır. Bu süreç tüm parçalar için tekrarlanır. Böylelikle tüm parçalar hem test hem eğitim aşamasında kullanılmış olur.

Hockey Fight ve Violent Flow veri kümelerinden elde edilen sonuçlar Tablo-1 de gösterilmiştir.

**Tablo 1.** Başarım sonuçları

Veri Kümesi	Sınıflama Başarımı
Violent Flow	82.2
Hockey Fight	83.7

Önerilen yöntemin Violent Flow verisi üzerinde, literatürde yer alan diğer yaklaşımlar ile sınıflama başarımı açısından karşılaştırılması Tablo-2'de verilmiştir. Önerilen yöntemin Hockey Fight veri kümesi üzerinden karşılaştırması Tablo-3'te verilmiştir. Sonuçlar incelendiğinde, önerilen yöntemin literatürdeki önceki yaklaşımlar ile karşılaştırılabilir başarılı sonuçlar ürettiği gözlemlenmiştir. Violent Flow veri kümesi daha karmaşık görüntüler içermesinden dolayı bu veri kümesi ile yapılan deneylerde daha düşük başarımlar elde edilmiştir.

**Tablo 2.** Yöntemlerin Violent Flow veri kümesi ile elde edilen sonuçları.

Yaklaşım	Sınıflama Başarımı
LTP [20]	71.53
Vif [6]	81.20
Keçeli ve Kaya [14]	80.90
OVif [10]	76.80
Horn-Schunk [13]	73.75
Keçeli ve Kaya [15]	82.60
Önerilen Yöntem	82.20

**Tablo 3.** Yöntemlerin Hockey Fight veri kümesi ile elde edilen sonuçları.

Yaklaşım	Sınıflama Başarımı
LTP [20]	71.90
Vif [6]	81.60
OVif [10]	84.20
Horn-Schunck [13]	79.80
Keçeli ve Kaya [15]	89.90
Önerilen Yöntem	83.70

#### 4. Tartışma ve Sonuç

Bu çalışma kapsamında transfer öğrenme ve LSTM ağları birleştirilerek şiddet içeren videoların sınıflandırılmasında kullanılacak bir yaklaşım önerilmiştir. Video serisinin zaman ekseninde hesaplanan 1. ve 2. türevleri ve video çerçeveleri önceden eğitilmiş bir CNN ağına girdi olarak verilmiş ve yüksek düzeyli öznelikler elde edilmiştir. Video serileri üç kanallı bir derin öznelik serisine dönüştürülmüştür. Bu derin nitelik serilerinden bir LSTM modeli eğitilmiştir. Öznelik çıkarımında önceden eğitilmiş bir ağı kullanmanın sınıflama üzerinde olumlu bir etkisi olmuştur.

Literatürde yer alan temel çalışmalar ile karşılaştırıldığında önerilen yöntemin şiddet içeren hareketlerin tespitinde başarılı sonuçlar ürettiği görülmektedir. Elde edilen sonuçlar literatürde yer alan diğer çalışmalar ile karşılaştırılabilir düzeydedir.

Gelecek çalışma olarak video serilerinden eğitilecek farklı LSTM ağlarını barındıran bir çoklu (ensemble) sınıflayıcı tasarımı planlanmaktadır.

#### Kaynakça

- [1] Nam, J., Alghoniemy, M., Tewfik, A.H., 1998. Audio-visual content-based violent scene characterization, Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on. IEEE, pp. 353-357.
- [2] Clarin, C., Dionisio, J., Echavez, M., Naval, P., 2005. DOVE: Detection of movie violence using motion intensity analysis on skin and blood. PCSC 6, 150-156.
- [3] Gong, Y., Wang, W., Jiang, S., Huang, Q., Gao, W., 2008. Detecting violent scenes in movies by auditory and visual cues, Pacific-Rim Conference on Multimedia. Springer, pp. 317-326.
- [4] Kooij, J.F., Liem, M., Krijnders, J.D., Andringa, T.C., Gavrila, D.M., 2016. Multi-modal human aggression detection. Computer Vision and Image Understanding 144, 106-120.
- [5] Lin, J., Wang, W., 2009. Weakly-supervised violence detection in movies with audio and video based co-training, Pacific-Rim Conference on Multimedia. Springer, pp. 930-935.
- [6] Hassner, T., Itcher, Y., Kliper-Gross, O., 2012. Violent flows: Real-time detection of violent crowd behavior, Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on. IEEE, pp. 1-6.
- [7] Lloyd, K., Marshall, D., Moore, S.C., Rosin, P.L., 2016. Detecting Violent Crowds using Temporal Analysis of GLCM Texture. arXiv preprint arXiv:1605.05106.
- [8] Boiman, O., Irani, M., 2007. Detecting irregularities in images and in video. Int J Comput Vision 74, 17-31.
- [9] Rota, P., Conci, N., Sebe, N., Rehg, J.M., 2015. Real-life violent social interaction detection, Image Processing (ICIP), 2015 IEEE International Conference on. IEEE, pp. 3456-3460.
- [10] Gao, Y., Liu, H., Sun, X., Wang, C., Liu, Y., 2016. Violence detection using Oriented Violent Flows. Image and Vision Computing 48, 37-41.
- [11] Dai, Q., Zhao, R.-W., Wu, Z., Wang, X., Gu, Z., Wu, W., Jiang, Y.-G., 2015. Fudan-Huawei at MediaEval 2015: Detecting Violent Scenes and Affective Impact in Movies with Deep Learning, MediaEval.
- [12] Deniz, O., Serrano, I., Bueno, G., Kim, T.-K., 2014. Fast violence detection in video, Computer Vision Theory and Applications (VISAPP), 2014 International Conference on. IEEE, pp. 478-485.
- [13] Arceda, V.M., Ferna, K., Guti, J., 2016. Real time violence detection in video.
- [14] Keceli, A.S., Kaya, A., 2017. Violent activity detection with transfer learning method. Electron Lett 53, 1047-1048.
- [15] Keceli, A.S., Kaya, A., 2018. Optik Akış Görüntüsü ve Bi-Lstm ile Şiddet İçeren Hareketlerin Sınıflandırılması. Avrupa Bilim ve Teknoloji Dergisi, 204-208.
- [16] GoogleNet katmanları çizgesi. <https://www.mathworks.com/help/deeplearnin g/ref/plot.html> (Erişim Tarihi: 01.09.2018).
- [17] Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput 9, 1735-1780.
- [18] Nievas, E.B., Suarez, O.D., García, G.B., Sukthankar, R., 2011. Violence detection in video using computer vision techniques, International Conference on Computer Analysis of Images and Patterns. Springer, pp. 332-339.
- [19] Bruhn, A., Weickert, J., Schnörr, C., 2005. Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. Int J Comput Vision 61, 211-231.
- [20] Yeffet, L., Wolf, L., 2009. Local trinary patterns for human action recognition, Computer Vision, 2009 IEEE 12th International Conference on. IEEE, pp. 492-497.