



Optik Akış Görüntüsü ve Bi-Lstm ile Şiddet İçeren Hareketlerin Sınıflandırılması

Ali Seydi Keçeli^{1*}, Aydın Kaya¹

¹Hacettepe Üniversitesi Bilgisayar Mühendisliği Bölümü, Ankara, Türkiye
*aliseydi@cs.hacettepe.edu.tr

(İlk Geliş Tarihi 15 Eylül 2018 ve Kabul Tarihi 23 Kasım 2018)

(DOI: 10.31590/ejosat.460257)

Öz

Otomatik hareket tanıma sistemlerine ihtiyaç, güvenlik kameralarının sayısındaki hızlı artıştan dolayı giderek artmaktadır. Hareket tanıma, bilgisayarlı görü alanında güncel bir araştırma alanı olmasına karşın şiddet içeren sahnelerin tespiti insan ve toplum güvenliğiyle de ilişkili olması sebebiyle büyük önem taşımaktadır. Optik akış video görüntülerindeki hareketlerin tespit ve modellenmesinde sıklıkla kullanılan bir yaklaşımdır. Bu çalışmada optik akış ve derin öğrenme kullanılarak şiddet içeren aktivitelerin tanınmasında kullanılabilir bir yöntem önerilmiştir. Bir video serisine ait optik akış serisinin bileşenleri birleştirilerek üç kanallı bir görüntü haline getirilmiş ve önceden eğitilmiş VGG-16 evrimsel (convolutional) sinir ağına girdi olarak verilmiştir. VGG-16 ağından elde edilen derin nitelik serileri ile bir Bi-Lstm (Bidirectional long short term memory) sınıflayıcısı eğitilmiştir. Önerilen yöntem literatürde yer alan iki farklı veri kümesi ile test edilmiş ve literatürde yer alan diğer yaklaşımlar ile karşılaştırılabilir ve daha yüksek sınıflama başarımına sahip sonuçlar elde edilmiştir.

Anahtar Kelimeler: Optik Akış, CNN, Derin Öğrenme, Bi-Lstm.

Classification of Violent Activities with Optical Flow Image and Bi-Lstm

Abstract

The need for automated motion recognition systems is increasing due to the rapid increase in the number of security cameras. Although motion recognition is a hot topic in the field of computer vision, the classification of violent scenes is of great importance due to its relation to human and community safety. Optical flow is often used in the detection and modeling of motion in video images. In this study, a method that can be used to recognize violent activities using optical flow and deep learning has been proposed. The components of the optical flow series of a video series were combined into a 3-channel image and pre-trained VGG-16 was input into the convulsive neural network. A Bi-Lstm (Bidirectional long short term memory) classifier has been trained with the deep quality series derived from the VGG-16 network. The proposed method was tested with two different data sets in the literature and comparable and higher classifying results were obtained.

Keywords: Optical Flow, CNN, Deep Learning, Bi-Lstm

1. Giriş

Günümüzde gelişmekte olan teknoloji ve internet ağı sayesinde, büyük miktarlarda video verisine erişim kolaylaşmıştır. Hem video içerik siteleri hem de şehir güvenlik sistemleri bağlamında şiddet içeren hareketlerin tespit ve tanınması giderek önem kazanmaktadır. Şiddet içeren sahnelere sahip videoların etiketlenmesi, şiddet içeren hareketlerin algılanması, bilgisayarlı görü alanının bir alt araştırma alanı olan hareket tanımanın bir problemi olarak kabul edilmektedir.

Benzer şekilde, çözümlenmesi hareket tanımadaki gibi bazı zorluklar içermektedir. Bu zorlukların bazıları aktörlerin değişken hareket hızları, görüntüsü birbiri üzerine örtüşen kişiler ve eylemler, kameranın hareketi, ışık ve çevresel etkenlere bağlı görsel karışıklıklar olarak sıralanabilir [1,2].

Bu makalede iki farklı derin öğrenme yaklaşımı olan transfer öğrenme ve Bi-Lstm ağlarından faydalanılarak şiddet içeren hareketlerin tespitine yönelik bir çalışma gerçekleştirilmiştir. Önceden eğitilmiş VGG derin öğrenme ağından öğrenilen derin öznelik serileri ile Bi-Lstm ağı eğitilmiştir. Öğrenilmiş derin

¹ Sorumlu Yazar: Hacettepe Üniversitesi Bilgisayar Mühendisliği Bölümü, Ankara, Türkiye, aliseydi@cs.hacettepe.edu.tr

öznitelikler el ile çıkarılan özniteliklere göre daha ayırt edici olması sebebiyle tercih edilmiştir. Yöntemin test edilmesinde iki farklı açık erişimli veri kümesinden faydalanılmıştır. Elde edilen sonuçlar incelendiğinde, önerilen yöntemin mevcut pek çok yaklaşımdan daha başarılı sonuçlar ürettiği gözlemlenmiştir.

2. İlgili Çalışmalar

Bu alanda yapılan ilk çalışmalar alev, kan gibi şiddet ile özdeşleştirilebilecek görsel öğeler ve şiddet olayına ait olabilecek seslerin tanınması üzerine kurulmuştur [3]. Clarin ve diğ. [4] video serilerinde cilt, kan ve hareket tespiti ile şiddet içeren videoları sınıflandırmaya çalışmıştır. Gong ve diğ. [5] video serilerinden elde edilen yüksek seviyeli ses özniteliklerini ve düşük seviyeli görüntü özniteliklerini birleştirmiştir. İşitsel ve görsel öznitelikleri birleştiren bir başka yaklaşım da Lin ve Wang [6] tarafından önerilmiştir. Kooij ve diğ. [7] tarafından önerilen yöntemde kişi üç boyutta takip edilmiş ve anormal davranışları tespit edilmeye çalışılmıştır. Hassner ve diğ. [8] tarafından optik akış tabanlı ViF yöntemi önerilmiştir. Bu yöntemde ilk olarak videoya ilişkin optik akış değerleri hesaplanır. Daha sonra hesaplanan optik akış değerlerinden belirli bir eşik değerinin üzerinde olan pikseller seçilerek bir ikili gösterge elde edilir. Seçilen piksellerin frekanslarından oluşturulan histogramlar ile SVM sınıflayıcısı eğitilmiştir.

Gao ve diğ. [9] ViF yöntemine yön bilgisi ekleyerek geliştirmiş ve OVIF yöntemini önermişlerdir. Optik akışın x ve y bileşenleri kullanılarak açı değerleri elde edilmiştir ve ikili gösterge elde edilmesinde bu açı değerleri kullanılmıştır. Bu yaklaşımda başarıyı artırmak için ViF yönteminden elde edilen öznitelikler de kullanılmıştır. Rota ve diğ. [10] öznitelik çıkarmak için hareket yörüngelerini kullanmışlardır. Şiddet içeren etkileşimleri tespit etmek için yörünge özelliklerinden bir kod çizelgesi oluşturmuşlardır. Lloyd ve diğ. [11] kalabalığı bir doku (texture) olarak düşünmüş ve kalabalık dokusundaki değişiklikleri kullanarak kalabalık dinamiklerini modellemişlerdir. Haralick özellikleri tüm çerçeveler için çıkarılmış ve bu özelliklerin zaman içindeki değişimleri hesaplanmıştır. Önerilen bu yöntem çalışmacılar tarafından Şiddetli Kalabalık Dokusu (Violent Crowd Texture) olarak adlandırılmıştır. Arceda ve diğ. [12] ise Horn-Schunck optik akış hesaplama ve ViF yöntemini birleştiren bir yöntem önermişlerdir. Dai ve diğ. [13] şiddet algılama problemi için derin öğrenme yöntemlerini kullanmışlardır. Keçeli ve Kaya [14] transfer öğrenme ile şiddet içeren hareketleri tespit etmişlerdir. Roy ve diğ. [15] HOG değerleri ve SVM ile güvenlik görüntülerindeki şiddet içeren hareketleri tanımayla çalışmışlardır.

3. Yöntem

Önerilen yönteme ilişkin genel akış şeması Şekil-1 de verilmiştir. İlk olarak optik akış bileşenleri hesaplanmıştır. Elde edilen optik akış bilgisi kullanılarak transfer öğrenme ile her çerçeveye ait öznitelik vektörleri oluşturulmuştur. Oluşturulan bu derin öznitelik vektörleri Bi-Lstm ağına girdi olarak verilmiştir. Bu kesimde ilk olarak kullanılan optik akış yöntemi ve Bi-Lstm yaklaşımları açıklanmıştır. Sınıflama alt kesiminde ilk olarak öznitelik çıkarma işlemine ve derin özniteliklerin kullanılmasında kullanılan önceden eğitilmiş VGG-16 ağına ait bilgilere yer verilmiştir. Son olarak da elde edilen öznitelik serilerinin sınıflamasında kullanılan Bi-Lstm ağına ait bilgilere yer verilmiştir.

3.1. Optik Akış

Optik akış (Optical Flow) yöntemi, hareket kestirimi, nesne takibi, video stabilizasyonu (titreşim, hareket giderme), görüntü mozaikleme, çiftli görü (stereo vision) ile derinlik bilgisi elde edilmesi gibi işlemlerde sıklıkla kullanılan temel bir yaklaşımdır [16]. Bu yaklaşımda, piksellerin çerçeveler (frame) arasındaki değişimi gözlenip, değişimin yönü ve hızı hesaplanmaktadır. Algoritmada akışın başta sabit olduğu düşünülmemektedir ve önce yerel (piksel tabanlı) komşulara bakılmaktadır. Daha sonra temel optik akış denklemleri [1] ile bütün piksellerdeki komşular için algoritma işletilmektedir.

İki görüntü arasındaki optik akış Eşitlik-1 ile ifade edilmektedir. I_x, I_y ve I_t görüntüdeki konumsal değişimleri göstermektedir. V_x x eksenindeki akışı, V_y ise y eksenindeki akışı göstermektedir.

$$I_x V_x + I_y V_y + I_t = 0 \quad (1)$$

Bu çalışmada, optik akışı tahmin etmek için Lucas-Kanade yöntemi [1] kullanılmıştır. Bu yöntem için temel algoritma aşağıdaki gibidir.

- Önce gürültü hassasiyetini azaltmak için giriş görüntüsünde Gauss yumuşatması uygulanır.
- Yumuşatmadan sonra, konumsal türevler I_x, I_y ve I_t hesaplanır.
- Görüntüdeki her piksel için türevsel evrişim hesaplanır. p bir pikseli ve p1, ..., pn ise n adet komşu pikseli belirtmektedir. Eşitlik-2 ile t anındaki yerel görüntü akışı tahmin edilir.

$$\begin{aligned} I_x(p_1)V_x + I_y(p_1)V_y + I_t(p_1) &= 0 \\ &\vdots \\ I_x(p_n)V_x + I_y(p_n)V_y + I_t(p_n) &= 0 \end{aligned} \quad (2)$$

Eşitlik-2 matris formunda Eşitlik-3'teki şeklindeki gibi yazılabilir.

$$Av = b$$

$$A = \begin{bmatrix} I_x(p_1) & I_y(p_1) \\ \vdots & \vdots \\ I_x(p_n) & I_y(p_n) \end{bmatrix}, \quad v = \begin{bmatrix} V_x \\ V_y \end{bmatrix}, \quad b = \begin{bmatrix} -I_t(p_1) \\ \vdots \\ -I_t(p_n) \end{bmatrix} \quad (3)$$

Son olarak, iki görüntü arasındaki optik akışı zamanında tahmin etmek için Eşitlik-4 en küçük kareler yaklaşımı ile çözülür. Eşitlik-4'te A^T , A matrisinin devriğidir.

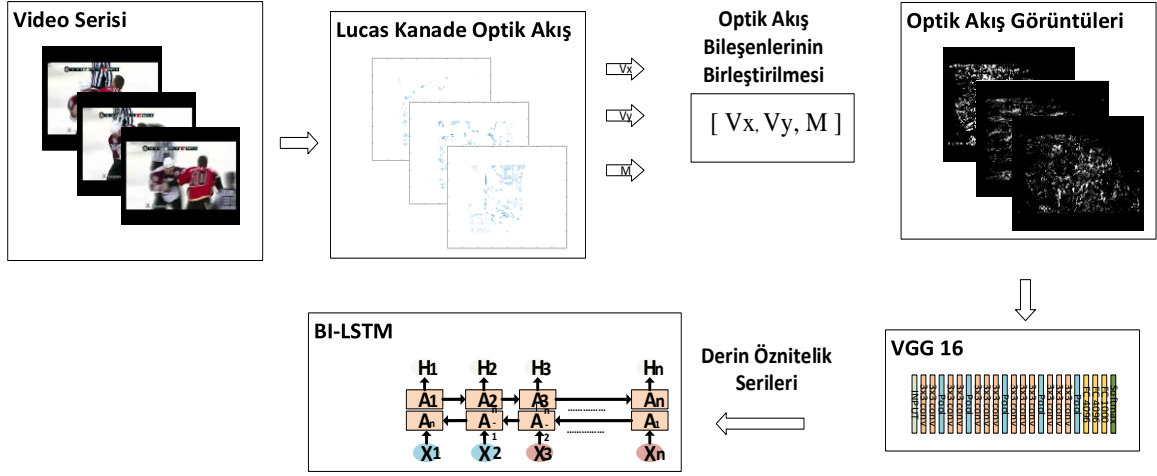
$$v = (A^T A)^{-1} A^T b \quad (4)$$

3.2. Bi-LSTM

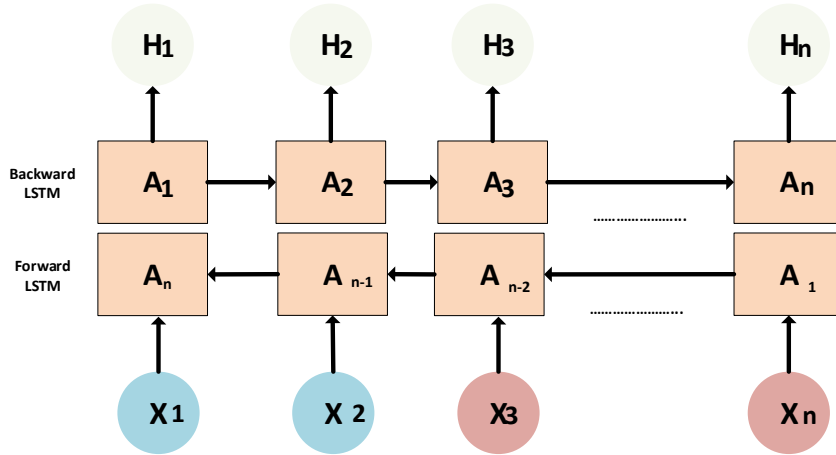
Bi-Lstm geleneksel Lstm ağlarının geliştirilmiş bir versiyonudur [17]. Bi-lstm ağları da Lstm ağlarında olduğu gibi bir çeşit gözlem serileri ile çalışan RNN'dir (Recurrent Neural Network). Klasik sinir ağlarından farklı olarak RNN'ler kendi giriş belleklerini girdinin farklı sıralarını işlemek için kullanabilirler.

Geleneksel Lstm ağlarından farklı olarak bu kez 2 farklı Lstm ağı eğitilmektedir. Birinci Lstm’de girdi olarak verilen seriler oldukları gibi kullanılır. İkinci Lstm ağı ise bu serilerin

ters çevrilmiş versiyonlarını kullanarak eğitilir. Örnek bir Bi-Lstm ağı Şekil-2’de gösterilmiştir.



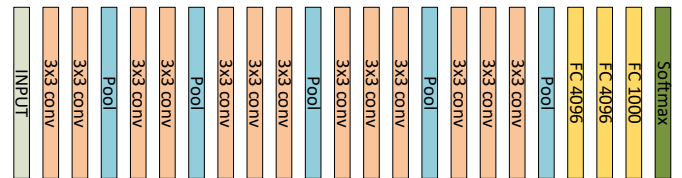
Şekil-1 Yöntemin genel akış şeması



Şekil-2 Örnek bir Bi-Lstm ağı

3.3 Sınıflama

Önerilen yöntemde ilk olarak video serisine ait optik akış değerleri elde edilmiş ve bu değerlerden x ve y eksenlerindeki hız değerleri Kesim 2.1 de belirtilen şekilde elde edilmiştir. Sonrasında büyüklük (magnitüde) değerleri birleştirilerek üç kanallı bir matris elde edilmiştir. Elde edilen üç kanallı matris bir görüntü gibi ele alınıp önceden eğitilmiş olan VGG-16 evrişimsel ağına girdi olarak verilmiştir. VGG-16 ağının katmanları Şekil-3’te verilmiştir [18].

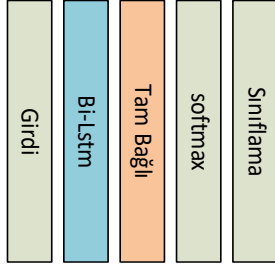


Şekil-3 VGG-16 katmanları [18]

VGG-16 ağına girdi olarak verilen her görüntü için ağın FC-7 katmanı tarafından üretilen aktivasyon değerleri öznetelik olarak kullanılmıştır. Her optik akış görüntüsü için 4096 öznetelik hesaplanmıştır. Transfer öğrenme olarak da bilinen bu işlem [19,20], veri kümelerinin küçük ve yetersiz olduğu durumlarda sıklıkla kullanılan bir yaklaşımdır. Doğası gereği sıfırdan başarılı sonuçlar üretebilen bir CNN ağı eğitmek için yüksek miktarda veri gerekmektedir. Her problem için büyük veri kümeleri oluşturabilmesi mümkün de değildir. Bu gibi durumlarda önceden eğitilmiş bir ağın bilgisinin başka bir problem için uyarlanması

etkili sonuçlar üretmektedir. Bu uyarılma işlemi iki yoldan yapılabilmektedir. Bunlardan ilki önceden eğitilmiş ağı en iyilenmesi (Fine-tuning) ya da mevcut ağı öznetelik çıkarma için kullanımıdır. Bu çalışmada ikinci yöntem kullanılmıştır. Girdiler seri olarak ele alınacağı için en iyileme işlemi yapılamayacaktır. Video serilerinden elde edilen optik akış görüntü serilerinin her elemanı için hesaplanan öznetelikler seri haline getirilerek bir Bi-Lstm sınıflayıcısı eğitilmiştir.

Bi-Lstm sınıflayıcının katmanları Şekil-4'te belirtilmiştir. Bi-Lstm ağında 4096 girişli bir girdi katmanı, 1000 kadar gizli katmanı olan bir Bi-Lstm katmanı ve sınıf sayısı kadar katmanı olan bir tam bağlı katman bulunmaktadır. Bi-Lstm ağına eğitim parametreleri ise aşağıdaki Tablo-1'de verilmiştir.



Şekil-4 Bi-Lstm ağı katmanları
Tablo 1. Bi-Lstm Parametreleri

Parametre	Değer
Gradient Threshold	1
Initial Learn Rate	1e-3
Max Epoch	300
Solver	Sdgm

4. Deneyler

Önerilen yöntem, iki farklı veri kümesi ile test edilmiştir. Bu veri kümelerinden ilki Violent Flow veri kümesidir. Bu veri kümesinde 246 adet örnek bulunmaktadır. Bu örnekler kalabalık toplumsal olaylardan alınmışlardır. Örneklerin yarısı şiddet içeren görüntülerden oluşurken diğer yarısı şiddet içermeyen normal görüntülerdir. Deneyler sırasında kullanılan 2. veri kümesi ise Hockey Fight veri kümesidir. Bu veri kümesi, şiddetin sıkça rastlandığı hokey maçları sırasında çekilmiş şiddet içermeyen 500 ve şiddet içeren 500 olmak üzere toplamda 1000 adet görüntü kaydı içermektedir.

Deneyler sırasında 5 katlı çapraz geçerlilik genelleştirme yöntemi uygulanmıştır. Bu yöntemde veri kümesi rastgele 5 ayrı parçaya bölünmekte ve 1 parça test için ayrılırken kalan parçalar sınıflayıcının eğitimi için kullanılmaktadır. Bu yöntemle elde edilen her iki veri kümesine ait başarımlar Tablo-2'de gösterilmektedir.

Tablo 2. Geliştirilen yöntemin başarımlar sonuçları.

Veri Kümesi	Sınıflama Başarımı
Violent Flow	82.6
Hockey Fight	89.9

Önerilen yöntemin Violent Flow verisi üzerinde, literatürde yer alan diğer yaklaşımlar ile sınıflama başarımları açısından karşılaştırılması Tablo-3'te verilmiştir. Önerilen yöntemin Hockey Fight veri kümesi üzerinden karşılaştırması da Tablo-

4'te verilmiştir. Sonuçlar incelendiğinde, önerilen yöntemin literatürdeki yaklaşımlar ile karşılaştırılabilir sonuçlar ürettiği gözlemlenmiştir.

Tablo 3. Yöntemlerin Violent Flow veri kümesi ile elde edilen sonuçları.

Yaklaşım	Sınıflama Başarımı
LTP [21]	71.53
Vif [8]	81.20
OVif [9]	76.80
Horn-Schunk [12]	73.75
Önerilen Yöntem	82.60

Tablo 4. Yöntemlerin Hockey Fight veri kümesi ile elde edilen sonuçları.

Yaklaşım	Sınıflama Başarımı
LTP [21]	71.90
Vif [8]	81.60
OVif [9]	84.20
Horn-Schunk [12]	79.80
Önerilen Yöntem	89.90

5. Tartışma

Bu çalışmada, derin öznetelik serileri Bi-Lstm sınıflayıcısı ile sınıflandırmıştır. Bi-Lstm kullanmanın en büyük avantajı farklı uzunluktaki girdilerin sınıflandırılabilmesine olanak sağlamasıdır. Derin öznetelik çıkarmak için önceden eğitilmiş VGG ağı kullanılmıştır. Derin öznetelikler görsel tanıma problemlerinde el ile çıkarılan (geleneksel) özneteliklere göre çok daha ayırt edici olmaktadır [22].

Deney sonuçları incelendiğinde, Violent Flow veri kümesi daha karmaşık görüntüler içermesinden dolayı bu veri kümesi ile yapılan deneylerde daha düşük başarımlar elde edilmiştir. Bi-Lstm parametrelerinde öğrenme oranı düşük seçilmiştir. Bu seçim eğitim süresini uzatmasına rağmen daha hassas bir sınıflayıcı elde edilmesini sağlamıştır. Gizli katman sayısı 1000 olarak belirlenmiştir. Bu sayı üzerindeki değerler sınıflama başarımları üzerinde önemsiz düzeyde artış sağlamaktadır.

Önerilen yöntemin en önemli kısıtlarından birisi çevrimiçi (online) olarak, yani sürekli olarak takip edilen bir kamera üzerinde çalışmaya doğrudan elverişli olmamasıdır. Fakat farklı uzunlukta kayan pencereler kullanılarak bu problem aşılabılır. Gelen video çerçeveleri farklı uzunluktaki kayan pencerelerde tutulup önerilen yöntem ile sınıflandırılabilir.

6. Sonuç

Bu çalışmada optik akış ve derin öğrenme tabanlı bir şiddet tanıma yaklaşımı önerilmiştir. Optik akış değerlerinden elde edilen maskeler kullanılarak videoda yer alan her çerçeve için derin öznetelikler çıkarılmıştır. Öznetelik çıkarılmasında VGG-16 ağı kullanılmıştır. Önceden eğitilmiş bir derin ağdan elde edilen nitelikler görsel tanıma problemlerinde başarılı düzeyde ayırt edicidir. Elde edilen çerçeve nitelikleri peş peşe eklenerek bir nitelik serisi haline getirilmiştir. Seri haline getirilen nitelikler Bi-Lstm ağına girdi olarak verilerek sınıflandırılmıştır. Önerilen yöntem iki farklı erişime açık veri kümesi ile test edilmiştir. Derin öğrenme ağlarından çerçeve düzeyinde çıkarılan

özniteliklerin seri haline getirilmesinin sınıflama sonucuna pozitif etki ettiği gözlemlenmiştir. Optik akış değerlerinin de çerçeve düzeyinde öznitelik çıkarmada etkili oldukları gözlemlenmiştir. Elde edilen sonuçlar literatürde yer alan diğer çalışmalar ile karşılaştırılabilir düzeydedir.

Kaynaklar

1. Keçeli AS, Kaya A, Can AB (2017) Depth features to recognise dyadic interactions. *Iet Comput Vis* 12 (3):331-339
2. Herath S, Harandi M, Porikli F (2017) Going deeper into action recognition: A survey. *Image and vision computing* 60:4-21
3. Nam J, Alghoniemy M, Tewfik AH Audio-visual content-based violent scene characterization. In: *Image Processing, 1998. ICIP 98. Proceedings. 1998 International Conference on, 1998. IEEE*, pp 353-357
4. 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
5. Gong Y, Wang W, Jiang S, Huang Q, Gao W Detecting violent scenes in movies by auditory and visual cues. In: *Pacific-Rim Conference on Multimedia, 2008. Springer*, pp 317-326
6. Lin J, Wang W Weakly-supervised violence detection in movies with audio and video based co-training. In: *Pacific-Rim Conference on Multimedia, 2009. Springer*, pp 930-935
7. Kooij JF, Liem M, Krijnders JD, Andringa TC, Gavrilă DM (2016) Multi-modal human aggression detection. *Computer Vision and Image Understanding* 144:106-120
8. Hassner T, Itcher Y, Kliper-Gross O Violent flows: Real-time detection of violent crowd behavior. In: *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on, 2012. IEEE*, pp 1-6
9. Gao Y, Liu H, Sun X, Wang C, Liu Y (2016) Violence detection using Oriented VIolent Flows. *Image and Vision Computing* 48:37-41
10. Rota P, Conci N, Sebe N, Rehg JM Real-life violent social interaction detection. In: *Image Processing (ICIP), 2015 IEEE International Conference on, 2015. IEEE*, pp 3456-3460
11. Lloyd K, Marshall D, Moore SC, Rosin PL (2016) Detecting Violent Crowds using Temporal Analysis of GLCM Texture. *arXiv preprint arXiv:160505106*
12. Arceda VM, Ferna K, Guti J (2016) Real time violence detection in video.
13. Dai Q, Zhao R-W, Wu Z, Wang X, Gu Z, Wu W, Jiang Y-G Fudan-Huawei at MediaEval 2015: Detecting Violent Scenes and Affective Impact in Movies with Deep Learning. In: *MediaEval, 2015.*
14. Keceli AS, Kaya A (2017) Violent activity detection with transfer learning method. *Electron Lett* 53 (15):1047-1048. doi:10.1049/el.2017.0970
15. Roy PK, Om H (2018) Suspicious and Violent Activity Detection of Humans Using HOG Features and SVM Classifier in Surveillance Videos. In: *Advances in Soft Computing and Machine Learning in Image Processing. Springer*, pp 277-294
16. Bruhn A, Weickert J, Schnörr C (2005) Lucas/Kanade meets Horn/Schunck: Combining local and global optic flow methods. *Int J Comput Vision* 61 (3):211-231
17. Schuster M, Paliwal KK (1997) Bidirectional recurrent neural networks. *IEEE Transactions on Signal Processing* 45 (11):2673-2681
18. Keçeli AS, Keçeli SU, Kaya A Classification of radiolarian fossil images with deep learning methods. In: *2018 26th Signal Processing and Communications Applications Conference (SIU), 2018. IEEE*,
19. Shie CK, Chuang CH, Chou CN, Wu MH, Chang EY (2015) Transfer Representation Learning for Medical Image Analysis. *Ieee Eng Med Bio*:711-714
20. Shin HC, Roth HR, Gao MC, Lu L, Xu ZY, Nogues I, Yao JH, Mollura D, Summers RM (2016) Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning. *Ieee T Med Imaging* 35 (5):1285-1298
21. Yeffet L, Wolf L Local trinary patterns for human action recognition. In: *Computer Vision, 2009 IEEE 12th International Conference on, 2009. IEEE*, pp 492-497
22. Keceli AS, Kaya A, Keceli SU (2017) Classification of radiolarian images with hand-crafted and deep features. *Computers & Geosciences* 109:67-74. doi:10.1016/j.cageo.2017.08.011