

Facial Limb Detection for the Protection of Face in Occupational Safety

*¹M. Fatih ADAK and ²Nimetullah NECMETTİN
^{1,2} Department of Computer Engineering Sakarya University, Turkey

Abstract

A little carelessness may endanger life safety or loss of life may occur in heavy industry, factories and sectors where occupational safety is very important. In order to prevent or minimize them, it is necessary to take advantage of informatics facilities. Today, the fact that artificial intelligence and the internet of things have come to a very advanced position can provide life safety by developing autonomous systems and building modules that will give warnings, warning the employee working in these sectors in a wrong position immediately. In this study, a YOLO-based system was developed and a proposal was made that aims to protect human facial parts. Tests show that results can be obtained quickly and to prevent harm to life.

Key words: deep learning, protection of face, yolov4, occupational safety

İş Güvenliğinde Yüzün Korunması İçin Yüz Uzun Tespiti

Özet

Ağır sanayide fabrikalarda ve iş güvenliğinin çok önemli olduğu sektörlerde çalışanlarda ufak bir dikkatsizlik can güvenliğini tehlikeye atabilir veya can kaybı yaşanabilir. Bunların önlenmesi veya en minimal seviyeye indirmek için bilişim olanaklarından faydalanmak gereklidir. Günümüzde yapay zekanın, nesnelere interneti alanın çok ileri bir konuma gelmesi otonom sistemler geliştirerek uyarı verecek modüller inşa ederek bu sektörlerde çalışan işçiyi yanlış bir pozisyonda anında uyarı vererek can güvenliğini sağlayabilir. Bu çalışmada YOLO tabanlı bir sistem geliştirilerek insan yüz uzunlarını korumayı amaçlayan bir öneride bulunulmuştur. Yapılan testler hızlı ve cana gelecek zararı önleyebilecek sonuçlar alınabildiğini göstermektedir.

Anahtar Kelimeler: derin öğrenme, yüzün korunması, yolov4, iş güvenliği

1. Giriş

Sanayi sektöründe çalışanlarda beden korunması ne kadar önemli ise insan yüzünün uzunlarının korunması çok daha önemlidir. Örneğin bir kaynak atölyesinde gözlerin korunması bir an bile olsa koruyucu gözlüğün çıkarılmaması ya da yüksek ses ile çalışılan bir fabrikada kulaklıkların sürekli takılı olması ya da tehlikeli gazın bulunduğu bir alanda gaz maskesinin doğru bir şekilde takılıp takılmadığını kontrolü otonom bir sistem ile sağlanabilir. İşçilerin çalıştığı alanı birkaç farklı açıdan izleyecek kameralar insan yüz uzun tespiti ve korumanın doğru bir şekilde takıldığını tespit edecek adımlar gereklidir. Bunlar sağlandığı durumda insan sağlığı önemli ölçüde iş kazalarından korunmuş olacaktır. Türkiye’de her yıl yüzlerce çalışan iş kazalarında zarar görmektedir [1]. Derin öğrenme teknikleri, nesnelere interneti uygulamaları ile bu sayı aşağı çekilebilir.

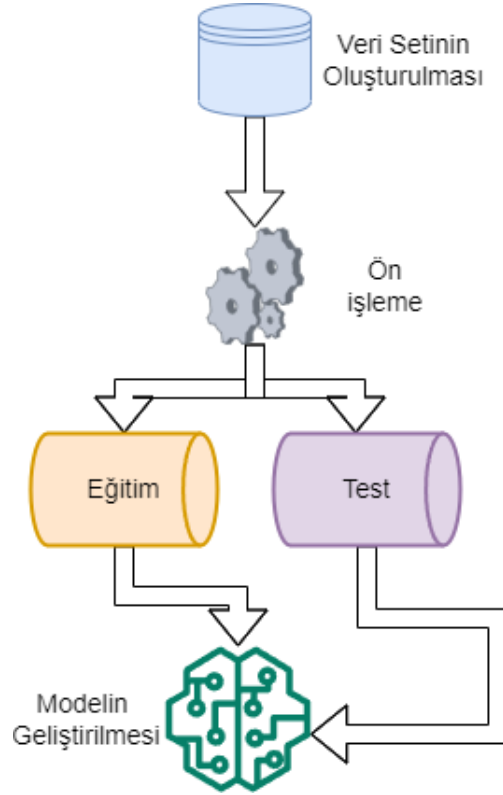
Dünya çapında verilere erişim kolaylığı, bu verileri işlemek için gerekli yazılımların geliştirilmesi, daha yaygın kullanımı ve donanımdaki hızlı genişlemeler nedeniyle derin öğrenme daha yaygın olarak kullanılmaya başlanmıştır. Derin öğrenme, günlük hayatımızdaki küçük problemler ve önemli uzay bilimleri için kullanılmaktadır ve karşılaştığımız sorunları hızlı ve yüksek başarı oranları ile çözmektedir. Günümüzde üniversitelerden özel şirketlere kadar görüntü sınıflandırma, video analizi, doğal dil işleme, ses tanıma, borsa fiyat tahminleri, medikal tedavi sanal gerçeklik (VR), artırılmış gerçeklik ve daha birçok konuda yaşanan sorunlara derin öğrenme yöntemleri ile çözüm sağlamaktadır [2], [3]. Yüz ve yüz uzuv tespiti, nesne tanımadaki zorlu problemlerdir. Farklı derin öğrenme ve gerçek zamanlı nesne algılama yöntemleri kullanılarak yüzün ve yüzün diğer bölümlerinin tespit edilmesine yönelik çalışmalar bulunmaktadır. Görüntülerdeki yüz algılama yöntemleri, poz, yüz ifadesi, konum ve yön, ten rengi, gözlük veya sakal, aydınlatma koşulları ve görüntü çözünürlüğü gibi değişkenler nedeniyle zorluk düzeyi artmaktadır. İki aşamalı hızlı CNN, Hızlı_CNN gibi nesne algılama algoritmaları ve one-step YOLO gibi gerçek zamanlı nesne algılama algoritmaları, daha hızlı ve çok yüksek performans sergileyen önerilen çalışma ve uygulamalardır [4], [5]. Yüz uzuv tespit çalışmaları incelendiğinde örneğin göz bölgesinin tespiti için SVM tabanlı bir algoritma önerilmiş ve başarılı sonuçlar alınmıştır [6]. Karahan ve Ark. Caffe kütüphanesi [7] ile çok katmanlı evrimsel sinirsel ağ kullanarak HAAR algoritması Fddb [8] ve CACD [9], veri setlerindeki yüz görüntüleri ile test edilerek kesinlik ve geri çağırma değerleri hesaplanmıştır [2]. Göz ve ağız tespiti için yapılan diğer bir çalışmada göz algılama için 674 görüntü ve ağız algılama modeli için 623 görüntü kullanılmıştır. Sistemin doğruluğu %90 üzeri olduğu görülmüştür [10]. Yolo tabanlı bir çalışmada yüz tespiti yapılmış yüz algılama sistemleri için bir iyileştirme önermek için VGG16 modeli kullanılmıştır hassasiyet %95 olarak ölçülmüştür [11]. Yine bir yüz algılama çalışmasında NMS yöntemi kullanılmıştır [12]. Elde edilen bu sonuçlar, YOLO ve benzeri görüntü işleme modellerinin yüz uzuvlarının tespitinde başarılı sonuçlar verdiğini göstermektedir.

2. Materyal ve Metot

Bu çalışmada, temel iş sağlığı ve güvenliği önlemlerinden biri olan insan uzuvlarının korunması bağlamında yüz uzuvlarının tespiti ve koruyucu bulunmuyorsa bunun için anında uyarı veren bir sistem önerilmiştir. Tek aşamalı nesne algılama algoritması olan YOLO ile koruyucu tespiti ve gerçek zamanlı bir sistemle işletmelerde iş güvenliği ve sağlığı önlemleri açısından yüz uzuvlarının korunması amaçlanmıştır. Çalışmada derin öğrenme yöntemleri kullanılmış ve gerçek zamanlı koruyucu algılama sistemi geliştirilmiştir. Şekil 1'de görüldüğü gibi çalışma 3 adımdan oluşmaktadır. İlk aşamada veri seti ön işleme ile hazırlanmış, ikinci aşamada Derin öğrenme yaklaşımı için gerekli yapılandırma işlemleri yapılmış son olarak sistemin eğitim ve testleri gerçekleştirilmiştir.

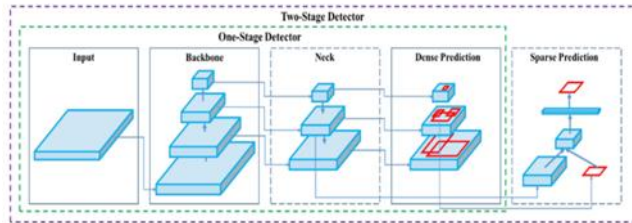
2.1. Verinin Hazırlanması

Veri seti internet üzerinden toplanan çeşitli insan yüzü uzuvları içeren 800 görüntüden oluşmaktadır. Veri seti ön işleme olarak insan yüz uzuvları etiketlenmiş daha sonra %80 eğitim %20 test olacak şekilde bölünmüştür. Boyutlandırma işlemi sonrasında model eğitilmiştir.



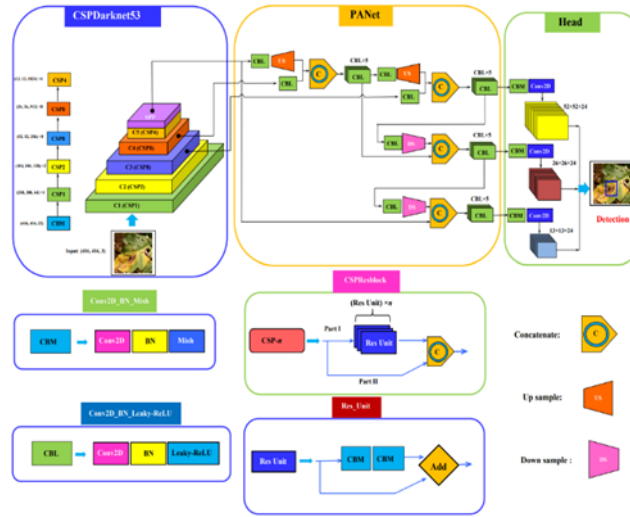
2.2. YOLOv4'ün Mimarisi

YOLOv4'ün genel ağ mimarisi (Şekil 2) Omurga, Boyun ve Baş'tan oluşur (Şekil 3). Omurga, farklı görüntülerden ayrıntılı bir düzeyde görüntü özelliklerini toplayan ve oluşturan evrimsel bir sinir ağıdır. Öte yandan, boyun, görüntü özelliklerini işleyen ve bunları tahmin katmanına ileten bir dizi ağ katmanıdır. Baş, kısmi görüntü özelliklerini tahmin eden ve sınırlayıcı kutular ve bunların kategorilerini oluşturan katmandır [13].



Omurga

YOLOv4, omurga olarak CSPDarknet'i kullanır, omurga, özellik çıkarımının yapıldığı katmandır. CSPDarknet giriş görüntülerinden öznitelikler çıkarmak için kullanılır. CSPNet [14] modülü, Omurgalarda tekrarlayan gradyan problemlerini çözer ve gradyan değişikliklerini özellik haritasına entegre ederek model parametrelerini azaltır [15]. Bu işlem modelin parametre miktarını ve boyutunu azaltırken ağın çıkarım hızını ve doğruluğunu artırır [13], [14].



Şekil 3. YOLOv4 ağ mimarisinin şeması [14]

YOLOv4'ün Omurga bölümü, Bag-of-Freebie ve Bag-Of-Specials adlı optimizasyon yöntemlerini içerir. Bu optimizasyon yöntemleri, model eğitimi ve çıkarımı sırasında Class Label Smoothing, DropBlock Regularization ve Mish Activation fonksiyonları gibi farklı teknikleri kullanır. Eğitim sürecinde Bag-of-Freebie (BoF) kullanılmıştır. Görüntüleri zenginleştirmek için veri büyütme, sınıf dengesizliğini yapmak için Sınıf Etiketleri Yumuşatma ve modelin aşırı uyum sorunlarını önlemek için DropBlock düzenleme gibi iyileştirme tekniklerini kullanır. Ağın çıkarım performansını iyileştirmek için çıkarım aşamasında Bag-Of-Specials (BoS) kullanıldı. Bu iyileştirmeler, doğruluğu artırmak için Mish etkinleştirme fonksiyonu gibi gelişmiş kayıp etkinleştirme fonksiyonlarının kullanılmasını ve hesaplama karmaşıklığını azaltmak için Aşamalar Arası Kısmi Bağlantıların kullanılmasını içerir [16].

Boyun

Omurga ile baş arasındaki katmandır ve bilgi akışını artırmak için PANet [17] kullanır. Nesneleri tahmin ederken daha fazla bilgi elde etmek için bir ara katman olarak boyun eklenmiştir. Burada aşağıdan yukarıya veya yukarıdan aşağıya akıştan gelen bilgilerle komşu özellik haritalarından ayrıntılı bilgi çıkarımları yapılabilir. Boyun esas olarak özellik piramitleri oluşturmak için kullanılır. Özellik piramidi, modelin aynı nesneyi çeşitli boyutlarda tanıyabilmesi için farklı ölçeklerdeki nesnelerin algılanmasını güçlendirir [15].

Baş

Baş kısmında sınırlayıcı kutular vardır ve her kutunun sınıfı tahmin edilir ve aynı zamanda nesne tespit edici olarak da bilinir. Burada YOLOv3'te uygulanan modelleri kullanır ve Görüntü özelliklerini tahmin eder, sınırlayıcı kutular oluşturur ve sınıflarını ve koordinatlarını tahmin eder.

3. Bulgular

Derin öğrenme problemlerinde geliştirilen modelin başarısını ifade etmek için çeşitli kriterler bulunmaktadır. Hassasiyet, geri çağırma, F1 puanı ve ortalama hassasiyet gibi standartlar modelin performansını ölçer. Modelin performans değerlendirmesinde kullanılabilirler.

Geri Çağırma: görüntüdeki kaç pozitif örneğin doğru tahmin edildiğini gösterir (1).

$$Recall = \frac{TP}{TP+FN} \quad (1)$$

Kesinlik: Kesinlik oranı, tahmin sonucu içindir ve tahmin edilen pozitif numunelerin kaçının gerçek pozitif numuneler olduğunu gösterir (2).

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

Ortalama Kesinlik (AP): Nesne tanıma çalışmalarının doğruluğunu ölçmek için popüler bir ölçüm olan kesinlik-hatırlama eğrisinin altındaki alandır. Ayrıca, ağırlığın hassasiyetteki artış olduğu her bir eşikteki hassasiyetlerin ağırlıklı toplamı olarak da ifade edilebilir. R_n ve P_n , denklemdaki (3) güven eşliğinde kesinliği ve geri çağırmaı ifade etmektedirler.

$$AP = \sum_0^n (R_n - R_{n-1})P_n \quad (3)$$

F1-skoru: Harmoniğin ortalamasını alarak hatırlama ve kesinliği birleştiren bir ölçüm. P, denklemden kesinlik anlamına gelir ve R, hatırlama anlamına gelir (4).

$$F1 - score = 2 \left(\frac{P \cdot R}{P+R} \right) \quad (4)$$

Ortalama Hassasiyet Ortalaması (mAP): AP'nin ortalamasıdır. Bazı bağlamlarda, AP her sınıf için hesaplanır ve mAP elde etmek için ortalaması alınır. Ortalama kesinlik (mAP) puanı, mevcut farklı algılama zorluklarına dayalı olarak ortalama AP kategorileri ve genel IoU eşikleri alınarak hesaplanır (5).

$$mAP = \frac{1}{n} \sum_{i=0}^n AP_i \quad (5)$$

Birleşim Üzerinden Kesişme (IoU): IoU, öngörülen sınırlayıcı kutunun ve gerçek durum sınırlayıcı kutunun kesişim alanı ve bağlantı alanının oranı ile verilir (6).

$$IoU = \frac{|A \cap B|}{|A \cup B|} \quad (6)$$

Modelin eğitimi 5000 epoch süresince çalıştırılmıştır. Eğitim sonunda elde edilen bulgulara göre modelin ortalama kaybı, hataların zamanla azalması, ortalama kesinlik ortalaması ve ortalama kesinliğin artması modelin başarısı için esastır. Eğitim başarı ortalaması %82 test başarı ortalaması %78 düzeyinde olmuştur. Epoch sayısının artmasıyla ortalama kayıp ve hata miktarlarının azaldığı ve bu düşüşlerin modelin testte iyi performans göstermesini sağladığı gözlemlenmiştir.

Test sonuçları incelendiğinde hassasiyet olarak göz ve ağız tespitinin diğer uzuvlara göre başarılı olduğu, f1 skoru olarak tüm uzuvlarda %70 ve üzeri bir performansın sergilendiği görülmektedir. Göz ve ağızın daha iyi sonuçlar vermesi kamera açısıyla doğrudan ilişkilidir. Bunlardan en zor olanının aslında kulak tespiti olduğu söylenebilir. Kulak diğer uzuvlar ile kolayca karıştırılabilmekte ve modelin performansını düşürebilmektedir.

Tablo 1. Yüz uzuv tespitlerinde performans değerleri

<i>Model</i>	<i>Precision</i>	<i>Recall</i>	<i>F1-score</i>
Göz Tespiti	0.85	0.88	0.81
Burun Tespiti	0.74	0.75	0.70
Ağız Tespiti	0.82	0.88	0.81
Kulak Tespiti	0.71	0.75	0.74

Conclusions

İş sağlığı ve güvenliği konusunda işletmelerde uygulanan en önemli önlemlerden biri çalışanların kurallara riayet edip etmemesidir. İşletmelerde yetkin kişiler tarafından periyodik olarak denetlenmesi çalışanların iş güvenliği kurallarına uyması açısından yeterli değildir. Sürekli ve otonom bir şekilde bu sürecin yürütülmesi dalgınlık ile oluşabilecek kazaların da önüne geçecektir. Bu çalışma, bu tür bir denetim sürecinin gerçek zamanlı yürütülmesini önermiştir. Bu çalışma, işletmedeki çalışanların koruyucu kullanıp kullanmadığını kontrol ederek gerçek zamanlı bir sistemde YOLOv4 algoritmasını kullanarak %70 ve üzeri genel bir elde etmiştir. Düşük maliyetli ve hızlı bir tespit öneri sistemi olduğu için birçok farklı sektörde rahatlıkla kullanılıp başarılı sonuçlar elde edilebilir.

Teşekkür

Bu çalışma Sakarya Üniversitesi Bilimsel Araştırma Projeleri Komisyonu (BAPK) tarafından 2022-7-24-40 proje numarası ile desteklenmiştir.

Referanslar

- [1] H. Ceylan, "Türkiye'de İnşaat Sektöründe Meydana Gelen İş Kazalarının Analizi," *Uluslararası Muhendis. Arastirma ve Gelistirme Derg.*, vol. 6, no. 1, pp. 1–6, Jan. 2014, doi: 10.29137/umagd.346068.
- [2] S. Karahan and Y. S. Akgul, "Eye detection by using deep learning," in *2016 24th Signal Processing and Communication Application Conference (SIU)*, May 2016, pp. 2145–2148, doi: 10.1109/SIU.2016.7496197.
- [3] M. F. ADAK, "Identification of Plant Species by Deep Learning and Providing as A Mobile Application," *Sak. Univ. J. Comput. Inf. Sci.*, vol. 3, no. 3, pp. 231–237, Dec. 2020, doi: 10.35377/saucis.03.03.773465.
- [4] C. Cao et al., "An Improved Faster R-CNN for Small Object Detection," *IEEE Access*, vol. 7, pp. 106838–106846, 2019, doi: 10.1109/ACCESS.2019.2932731.
- [5] F. Sultana, A. Sufian, and P. Dutta, "A Review of Object Detection Models Based on Convolutional Neural Network," 2020, pp. 1–16.
- [6] M. YU, Y. LIN, and X. WANG, "An efficient hybrid eye detection method," *TURKISH J. Electr. Eng. Comput. Sci.*, vol. 24, pp. 1586–1603, 2016, doi: 10.3906/elk-1312-150.

- [7] Y. Jia *et al.*, “Caffe: Convolutional Architecture for Fast Feature Embedding,” Jun. 2014, [Online]. Available: <http://arxiv.org/abs/1408.5093>.
- [8] V. Jain and E. Learned-Miller, “FDDB: A Benchmark for Face Detection in Unconstrained Settings.”
- [9] B.-C. Chen, C.-S. Chen, and W. H. Hsu, “Cross-Age Reference Coding for Age-Invariant Face Recognition and Retrieval,” 2014, pp. 768–783.
- [10] E. Civik and U. Yuzgec, “Deep Learning Based Continuous Real-Time Driver Fatigue Detection for Embedded System,” in *2020 28th Signal Processing and Communications Applications Conference (SIU)*, Oct. 2020, pp. 1–4, doi: 10.1109/SIU49456.2020.9302035.
- [11] H. Aung, A. V. Bobkov, and N. L. Tun, “Face Detection in Real Time Live Video Using Yolo Algorithm Based on Vgg16 Convolutional Neural Network,” in *2021 International Conference on Industrial Engineering, Applications and Manufacturing (ICIEAM)*, May 2021, pp. 697–702, doi: 10.1109/ICIEAM51226.2021.9446291.
- [12] D. Garg, P. Goel, S. Pandya, A. Ganatra, and K. Kotecha, “A Deep Learning Approach for Face Detection using YOLO,” in *2018 IEEE Punecon*, Nov. 2018, pp. 1–4, doi: 10.1109/PUNECON.2018.8745376.
- [13] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” Apr. 2020, [Online]. Available: <http://arxiv.org/abs/2004.10934>.
- [14] A. M. Roy, R. Bose, and J. Bhaduri, “A fast accurate fine-grain object detection model based on YOLOv4 deep neural network,” Oct. 2021, [Online]. Available: <http://arxiv.org/abs/2111.00298>.
- [15] X. He, J. Wang, C. Chen, and X. Yang, “Detection of the Floating Objects on the Water Surface Based on Improved YOLOv5,” in *2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA)*, Dec. 2021, pp. 772–777, doi: 10.1109/ICIBA52610.2021.9688111.
- [16] J. G. Shanahan and L. Dai, “Introduction to Computer Vision and Real Time Deep Learning-based Object Detection,” in *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, Aug. 2020, pp. 3523–3524, doi: 10.1145/3394486.3406713.
- [17] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia, “Path Aggregation Network for Instance Segmentation,” Mar. 2018, [Online]. Available: <http://arxiv.org/abs/1803.01534>.