SAKARYA UNIVERSITY JOURNAL OF COMPUTER AND INFORMATION SCIENCES
http://saucis.sakarya.edu.tr/
e-ISSN: 2636-8129
Publisher: Sakarya University
Vol. 7, No. 2, 2024, 253-263
DOI: 10.35377/saucis...1461268

RESEARCH ARTICLE

Cigarette Detection in Images Based on YOLOv8

Yerniyaz Bakhytov[1], Cemil Öz[2]

[1] Sakarya University, Graduate School of Natural and Applied Sciences, Computer and Information Engineering, Sakarya, Türkiye
[2] Sakarya University, Faculty of Computer and Informatics, Department of Computer Engineering, Sakarya, Türkiye

Corresponding author:
Yerniyaz Bakhytov, Sakarya University, Graduate School of Natural and Applied Sciences, Computer and Information Engineering, Sakarya, Türkiye
yerniyaz.bakhytov@ogr.sakarya.edu.tr

Article History:
Received: 29.03.2024
Accepted: 22.07.2024
Published Online: 26.08.2024

ABSTRACT
This study investigates methods to develop and test the automatic detection of cigarettes in images using modern deep learning models such as YOLOv5 and YOLOv8. The study's primary aim is to improve the accuracy and reliability of recognizing **objects** associated with smoking, which could significantly enhance the monitoring of public places, media content analysis, and support for anti-smoking campaigns. Tobacco use poses a serious threat to public health, causing numerous diseases and resulting in millions of deaths annually. Advanced technologies such as computer vision and artificial intelligence offer new opportunities for more effective monitoring and analysis, which can help mitigate the negative effects of tobacco use. The training results are presented, with the YOLOv8 model achieving an accuracy of 87.4% and the YOLOv5 model slightly outperforming it with an accuracy of 89.6%. In conclusion, the article thoroughly explores the use of the YOLOv8 model in images for cigarette identification. It contributes to the existing body of knowledge by presenting a comparative analysis of the performance of the YOLOv8 and YOLOv5 models, thereby providing valuable insights for future research.

Keywords: Deeplearning, Object detection, YOLOv8, Smoking detection

1. Introduction

Cigarette smoking, a major global health concern, is responsible for numerous deaths and illnesses worldwide, including lung cancer, cardiovascular disease, and chronic obstructive pulmonary disease. As per the World Health Organization's data, in 2023, the global smoking population was around 1.3 billion, with over eight million fatalities attributed to smoking-related diseases [1]. Therefore, it is important to develop effective methods to control and prevent smoking, and to raise awareness about the dangers of smoking.

One of the ways to control smoking is to identify cigarettes by image, which can be used to monitor people smoking in public places, detect violations of smoking laws, and also to analyze the behavior and habits of smokers. Identifying cigarettes from images can also be useful for research in psychology, sociology, and medicine related to smoking.

Cigarette detection from an image is a computer vision task that involves localizing and classifying cigarettes in an image. This task is a subtask of object detection, which consists of finding bounding boxes and class labels for all objects in an image. Object detection is one of the most rapidly growing and challenging areas of computer vision, which has many applications in various fields such as security, medicine, robotics and entertainment.

There are many object detection methods that can be divided into two main categories: region-based and single-pass based. Region-based methods first generate candidate regions that may contain objects and then classify them using convolutional neural networks (CNNs). Examples of such methods are R-CNN and Faster R-CNN. Single pass based methods perform object detection in a single pass of the network, using anchor boxes or center points to predict bounding boxes and class labels. Examples of such methods are YOLO, SSD and RetinaNet [2].

In this work, YOLOv5 and YOLOv8 were selected for the cigarette detection task due to their high performance, accuracy, flexibility, and community support. These models provide the optimal balance between processing speed and detection accuracy, making them ideal for use in monitoring and video surveillance systems where fast and accurate object detection is required. Methods such as Faster R-CNN provide high accuracy but require more processing time due to the two-step process. YOLO solves these problems in a one-step process, speeding up processing time. YOLOv5 and YOLOv8 generally outperform SSDs in accuracy due to their optimized architectures [3]. EfficientDet is resource efficient, but YOLOv8 offers better performance thanks to the latest architectural improvements.

YOLO is one of the most widely used and popular object detection methods today. YOLO differs from other object detection algorithms in its speed and accuracy, and it has become famous for these properties. This method was proposed by Joseph Redmon and Ali Farhadi in 2015. He then released YOLOv2 [4] in 2016, developing the model architecture by adding batch normalization, helper frames, and cluster dimensions. YOLOv3 [5] was released in 2018. Compared to its predecessor, this feature further improves model performance by using a more efficient backbone network, multiple reference systems, and spatial pyramid aggregation. YOLOv4 [6] was released in 2020, introducing innovations such as tiled data augmentation, a state-of-the-art anchor-free detection head, and an improved loss function. In the realm of computer vision, YOLOv5 [7] stood out by boosting model performance and incorporating cutting-edge functionalities such as hyperparameter optimization, integrated experiment tracking, and automatic export to widely used formats. Meituan introduced YOLOv6 [8] in 2022, and it now powers a significant portion of the company's autonomous delivery robots. YOLOv7 [9] introduced new functionalities, such as pose estimation using the COCO key points dataset. The latest release from Ultralytics, YOLOv8, represents a significant advancement in the YOLO series, integrating cutting-edge features and enhancements to elevate its performance, versatility, and efficiency. Supporting a wide spectrum of artificial vision tasks including detection, segmentation, pose estimation, tracking, and classification, YOLOv8 empowers users with a multifaceted solution adaptable to various use cases and industries [10].

This article presents work on cigarette detection from images using the YOLOv8 and YOLOv5 algorithms. The structure of the two algorithms was also analyzed and the similarities and differences between them were analyzed. We trained the models on separate datasets collected from different sources. The results of the two models were then analyzed and compared.

This article explores cigarette detection from images using the YOLOv8 and YOLOv5 algorithms, highlighting their structural similarities and differences. Utilizing state-of-the-art models ensures high speed and accuracy, crucial for public health monitoring and automated surveillance. By training on diverse datasets, we aim for robust performance in real-world scenarios. Comparing the results of the two models provides insights into their respective strengths and weaknesses, contributing to the optimization of object detection tasks and advancing the field of computer vision.

2. Literature review

Cigarette smoking is one of the leading causes of death and morbidity in the world, so it is important to develop effective methods to detect and prevent smoking in public places. In recent years, a lot of research has emerged on the use of deep convolutional neural networks and other computer vision techniques to solve this problem. In this review, we will look at several such studies and compare their approaches, methods, problems and results.

One such system is Eye-Smoker. Based on computer vision and neural networks, the Eye-Smoker system is innovative in detecting smoking in images. Eye-Smoker uses YOLOv3 architecture to detect people smoking in images, focusing on the nose region. The system analyzes signatures such as hand movements, smoke, and posture to detect smoking accurately. Training occurs on various data, including nose images, video, and real detection from a webcam. The system's advantages are its high accuracy and ease of use in public places. However, the system has a disadvantage: if the nose is not visible in the image, then the system cannot detect smoking [11].

Another work on this topic uses deep convolutional neural networks to detect people smoking in public places. This approach emphasizes using a small amount of training data, which is a key feature. The system can detect smoking in images by focusing on the area of the face and hands where signs of smoking often appear [12].

The following work on this topic is the YOLO-Cigarette system. The main architecture of this model is based on YOLOv5, an improved version of YOLO, which is known for its high speed of object detection. One of the critical features of YOLO-Cigarette is to solve the problem of low detection accuracy of small objects such as cigarettes. To achieve this, the model includes a new FSPP (Fine-Grained Spatial Pyramid Pooling Module) module, which improves the accuracy of detecting small targets. In addition, using the MSAM (Multi-Spatial Attention Mechanism) mechanism improves the model's ability to focus on essential parts of the image, which also helps improve detection accuracy. The model parameters are quantized to reduce computational complexity and speed up the detection process. This model demonstrates superiority over the original YOLOv5 model in the detection accuracy of people smoking and achieves high performance in actual outdoor conditions [13].

The following work describes a smoking detection model based on a convolutional neural network called SmokingNet. This model automatically detects smoking in video content through images. Unlike traditional methods based on cigarette smoke detection algorithms, SmokingNet can detect smoking images using only human smoking gesture information and cigarette image characteristics without the need to detect cigarette smoke. It shows high accuracy and superior performance for real-time monitoring [14].

The following paper presents an improved YOLOv5-based algorithm to detect smoking behavior in public places like hospitals, schools, and stations to enhance health and safety. The algorithm improves detection by replacing the C3 module with the CoT module and integrating the CBAM attention mechanism before the SPPF structure, improving feature extraction. The improved algorithm significantly enhanced detection performance through experimental comparisons, raising the accuracy from 61.3% to 62.9%, demonstrating its effectiveness and real-time detection capabilities [15].

The following article proposes an improved YOLOv5 algorithm for smoking detection in public places, addressing low detection accuracy for small and medium targets in complex scenes. The enhancements include using Mosaic-9 for better data enhancement, introducing the Convolutional Block Attention Module (CBAM) to improve focus on target locations, replacing traditional upsampling with transpose convolution for better semantic recognition, and incorporating the SPD-Conv module to enhance recognition of low-resolution and small target images. Experimental results show a 3% improvement in detection accuracy for smoking targets [16].

The subsequent work delves into the details of an advanced algorithm for detecting smoking by substation personnel. The algorithm, based on GhostNetV2-YOLOv5, enhances the original YOLOv5 by improving detection accuracy and speed. The specific improvements include a 2.58% increase in total mean Average Precision (mAP) and a 1.61-fold increase in prediction speed. These enhancements are crucial in preventing equipment damage and ensuring safety [17].

These works all use deep learning to detect smoking, but they use different approaches and techniques. Eye-Smoker and YOLO-Cigarette use variations of the YOLO model for smoking detection, while the other uses its deep learning model. All these works demonstrate high accuracy in their results, highlighting the effectiveness of using deep learning for smoking detection.

3. Methodology

The methodology of our approach is based on using YOLOv8 and YOLOv5 to detect cigarettes in images. However, detecting cigarettes is challenging because cigarettes may be small, partially hidden, or have different shapes and colors. We have developed our algorithm for solving these problems.

3.1. Dataset

The first step in our methodology is to collect a database of images. This is an essential step because the quality and variety of images in the database directly affect the training efficiency of the model. In our case, we collected an image database containing images of cigarettes in various contexts and lighting conditions. We used the Roboflow Universe [18] web application to collect the dataset and label objects in the images.

The dataset we compiled is a testament to our meticulous approach, containing 1200 images of people smoking, all collected for the specific task of detecting cigarettes in photographs. The dataset is divided into three subsets: a training set (70%) with 840 images, a validation set (15%) with 180 images, and a test set (15%) also with 180 images. Each image in the dataset is meticulously annotated, indicating the presence of a cigarette in the photo in the form of bounding boxes around the cigarette. This detailed annotation process ensures the dataset is ready for robust model training. The dataset is rich in diversity, encompassing a variety of scenes, lighting, and poses of people smoking. It also includes background variations such as streets, cafes, and houses, making it a comprehensive representation of various cigarette detection scenarios.

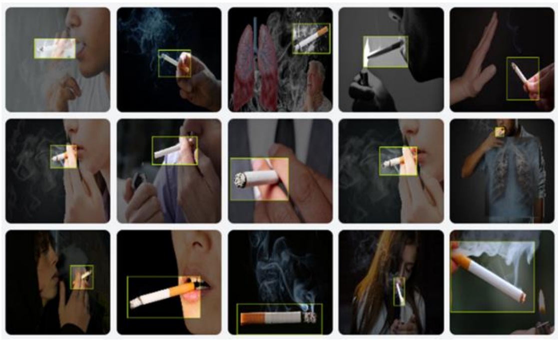


Figure 1. Smoking Cigarettes Dataset

3.2. Detection Models

In this work, using YOLOv5 and YOLOv8 models to detect cigarettes in images represents an essential aspect of computer vision research. The YOLO architecture, known for its high accuracy and speed, is the basis for these models in object recognition. Convolutional neural networks form its basis. The architectures of the YOLOv8 and YOLOv5 models consist of the backbone, neck, and head, as shown in Figure 2 and Figure 3.

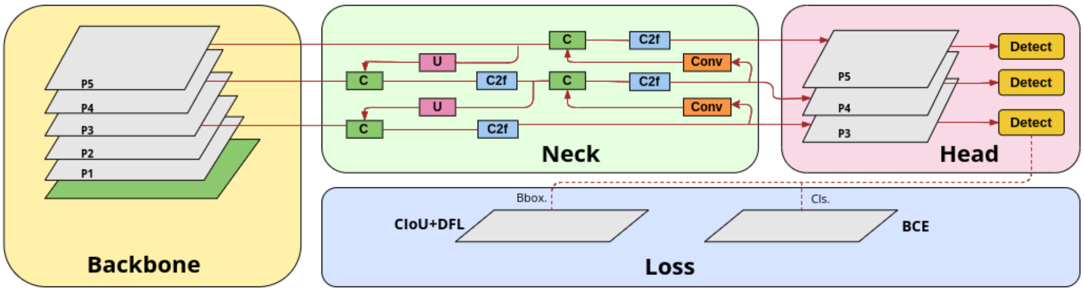


Figure 2. The Structure of the YOLOv8 Algorithm

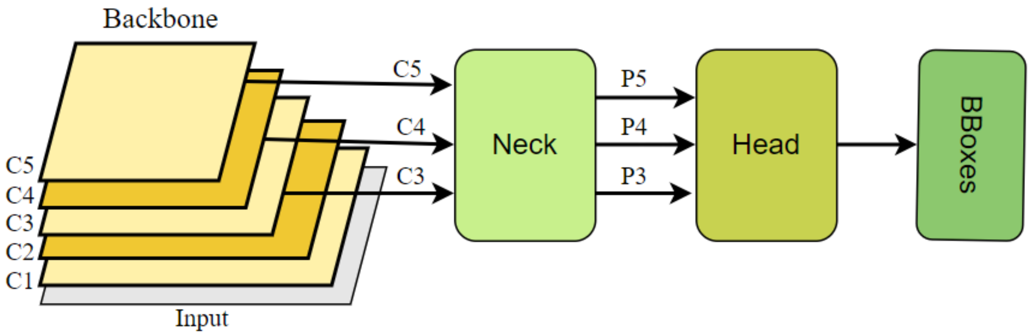
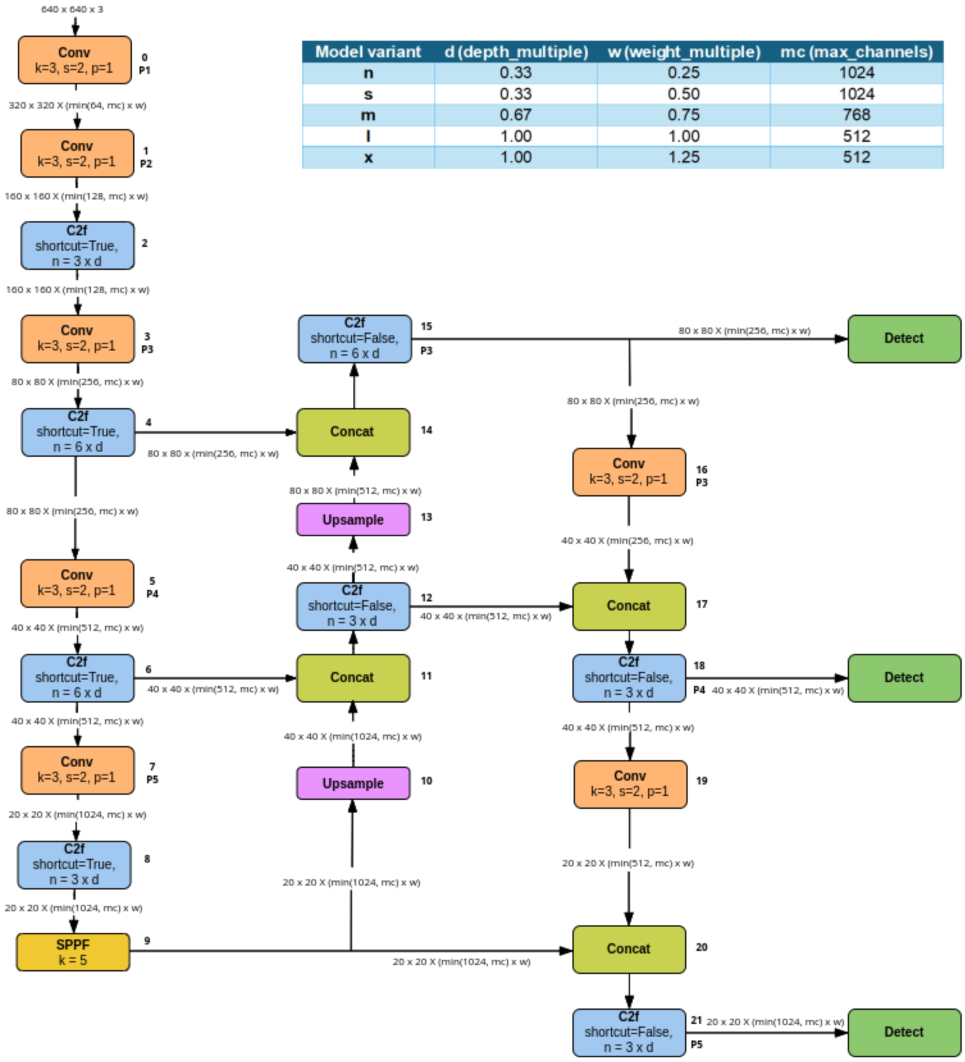


Figure 3. The Structure of the YOLOv5 Algorithm

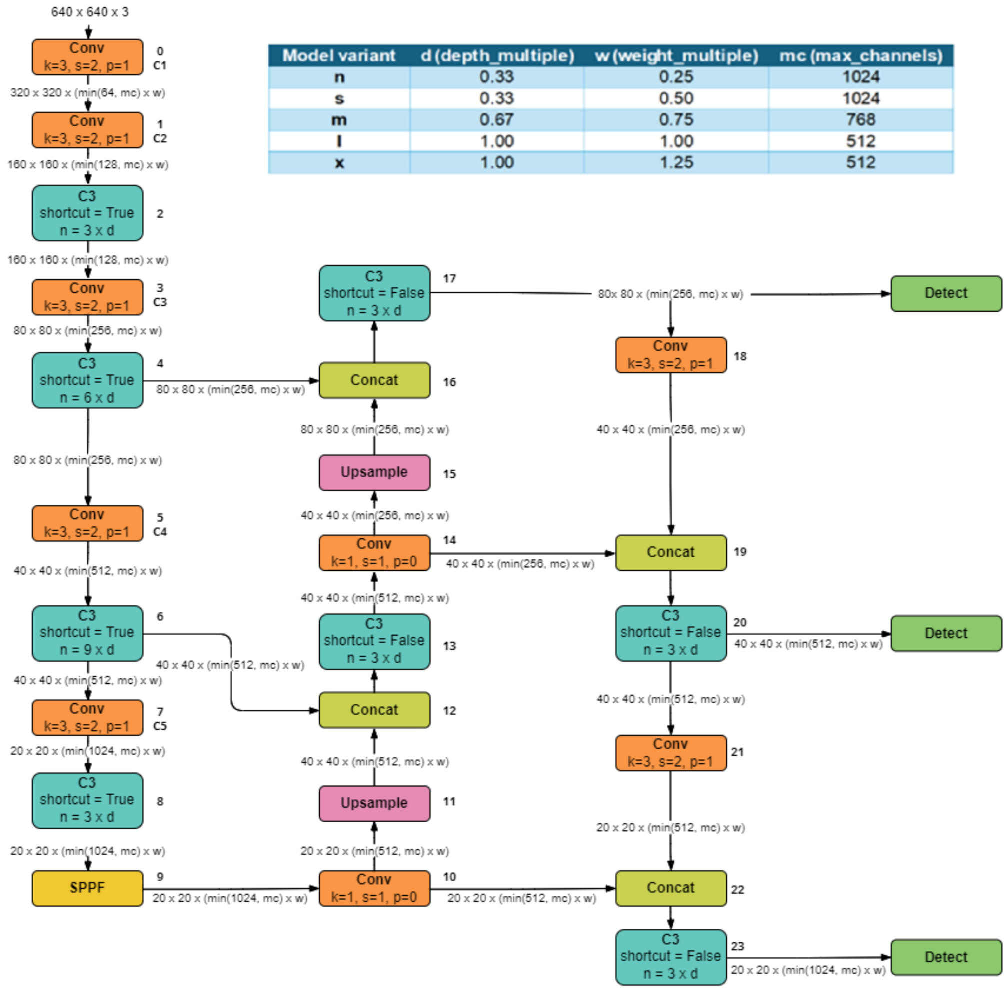
3.3 Backbone

The two models use the Cross Stage Partial (CSP) [19] architecture, which divides the feature map into two components. Convolution operations are applied to the first part, and the second part is fused with the outcomes from the preceding stage. As a result, the CSP architecture enhances CNN training effectiveness while reducing computational overhead. Unlike YOLOv5, in YOLOv8, the first convolutional kernel was increased from (1x1) to (3x3), and the primary building block C3 was replaced with C2f [20]. In Figure 4 and Figure 5, we use k, s, and p to represent kernel, stride, and padding, respectively. n is a depth parameter that determines the number of BottleNeck stacks by adding additional layers. This varies across the model scales: nano, small, medium, large and extra-large.



Model variant	d (depth_multiple)	w (weight_multiple)	mc (max_channels)
n	0.33	0.25	1024
s	0.33	0.50	1024
m	0.67	0.75	768
l	1.00	1.00	512
x	1.00	1.25	512

Figure 4. The Layout of the YOLOv8 Algorithm’s Architecture



Model variant	d (depth_multiple)	w (weight_multiple)	mc (max_channels)
n	0.33	0.25	1024
s	0.33	0.50	1024
m	0.67	0.75	768
l	1.00	1.00	512
x	1.00	1.25	512

Figure 5. The Layout of the YOLOv5 Algorithm's Architecture

The difference between YOLOv5 and YOLOv8 is that one uses C3, and the second uses C2f. Block C3 significantly reduces the number of model parameters without loss of accuracy. Thus, it reduces computational complexity and information loss. Block C3 is shown in Figure 6. The C2f module, which is a combination of the C3 module and the ELAN concept from YOLOv7, is introduced in YOLOv8. This enables the model to gather more intricate details about the gradient flow [21]. The C2f module, as depicted in Figure 6, comprises 2 ConvModules and n DarknetBottleNecks, which are interconnected via Split and Concat operations [22].

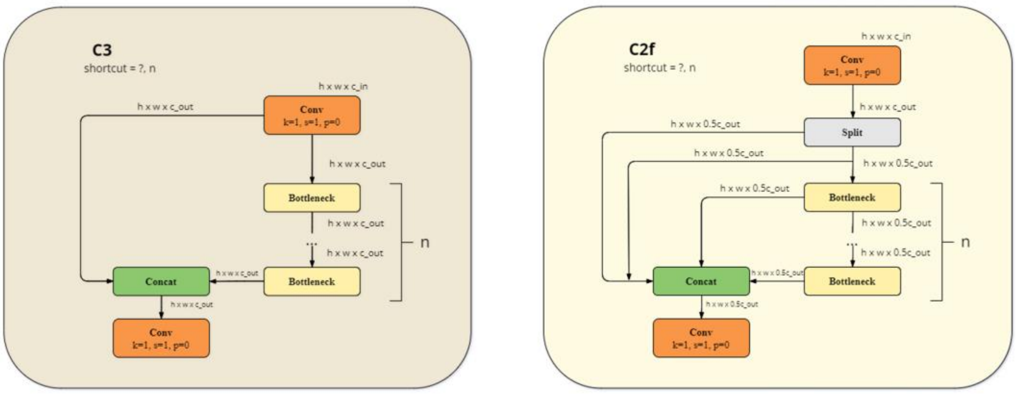


Figure 6. The Structure of the C3 and C2f Blocks

3.4. Neck

The Backbone links to the Neck at three distinct depths to combine features from various network layers. These fused features are then transmitted to the Head. The Neck incorporates path aggregation network (PAN) [23] and feature pyramid network (FPN) [24] structures to mitigate information loss caused by multiple convolutions. The Feature Pyramid Network (FPN) performs upsampling from top to bottom, enhancing the feature information in the lower-level feature map. Meanwhile, the Path Aggregation Network (PAN) downsamples from bottom to top, capturing additional information from the top-level feature map. These two feature outputs are combined to ensure accurate predictions for images of different sizes.

3.5. Head

The Head consists of three detection modules, which have been intentionally separated into distinct classification and regression tasks. This decoupling technique was initially proposed in YOLOX and YOLOv6 [25] specifically for anchor-free detection. In contrast to the YOLOv5 model's coupled Head, YOLOv8 adopts a decoupled head architecture. Specifically, we separate the classification and detection heads to enhance performance.

3.6. Evaluation Metrics

The performance of the model was assessed using metrics such as precision (P), recall (R), Intersection over Union (IOU) and mean average precision (mAP).

In the context of target detection, Intersection over Union (IOU) quantifies the agreement between the predicted and actual detection frames [26]. The IoU, represented by Equation 1, involves the comparison of the target box (Bgt) and the prediction box (B).

$$IoU = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \tag{1}$$

Precision assesses the model's ability to predict positive instances correctly. It calculates the proportion of correct positive predictions out of all positive predictions made [27]. A higher precision score implies fewer false positives, indicating that the model is more effective at correctly identifying true positives. The formula for calculating precision is given in Equation 2.

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

In machine learning, recall assesses the model's ability to capture all relevant positive examples. It is computed as the ratio of true positives to the sum of true positives and false negatives [27]. A higher recall score implies fewer false negatives, indicating that the model is more effective at identifying all positive instances. Equation 3 provides the formula for recall calculation.

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

Mean Average Precision (mAP) is a metric that evaluates the model's effectiveness in terms of object detection accuracy. It represents the average of precision scores at different levels of recall and is calculated using Equation 4 [28].

$$mAP = \frac{1}{N}\sum(precision\ at\ each\ recall\ level) \tag{4}$$

The Mean Average Precision (mAP) is represented as mAP (50), which signifies the average precision value when the IoU threshold is set at 0.5. Additionally, mAP (50-95) denotes a range of IoU thresholds from 0.5 to 0.95, incrementing in steps of 0.05.

4. Results

Our training setup comprised Ultralytics YOLOv8.0.194 for object detection, Python 3.8.18 for scripting, PyTorch 2.1.2 with CUDA 11.8 for deep learning, and an NVIDIA GeForce RTX 3060 Laptop GPU for accelerated computations. The specific training parameters included 50 epochs, stochastic gradient descent (SGD) optimization, an initial learning rate of 0.01, and a momentum of 0.937.

In this work, the model was trained using different sizes of YOLOv5 and YOLOv8. The Table 1 shows the results of our trained models. Comparing the results, YOLOv5 models have higher accuracy than YOLOv8. More precisely, the highest result was 89.6%. And for YOLOv8, the highest accuracy is 87.4%. That is, the YOLOv5 model performed 3% better than YOLOv8.

Table 1. Results of our Trained Models

Model	Parameters	mAP(50)	mAP(50-95)
YOLOv8n	3011043	0.822	0.390
YOLOv8s	11135987	0.877	0.415
YOLOv8m	25856899	0.845	0.416
YOLOv8l	43630611	0.867	0.416
YOLOv8x	68153571	0.874	0.429
YOLOv5n	2508659	0.834	0.406
YOLOv5s	9122579	0.846	0.396
YOLOv5m	25065715	0.896	0.434
YOLOv5l	53164115	0.889	0.417
YOLOv5x	97200371	0.881	0.412

Figure 7 and Figure 8 show the mAP0.5 performance of the YOLOv8 and YOLOv5 models trained in the experimental environment of this work.

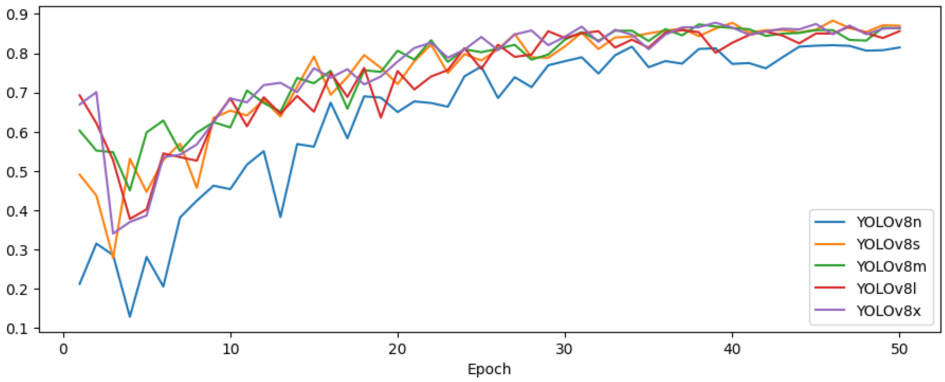


Figure 7. mAP with IoU=0.50 with the Training YOLOv8 Over 50 Epochs

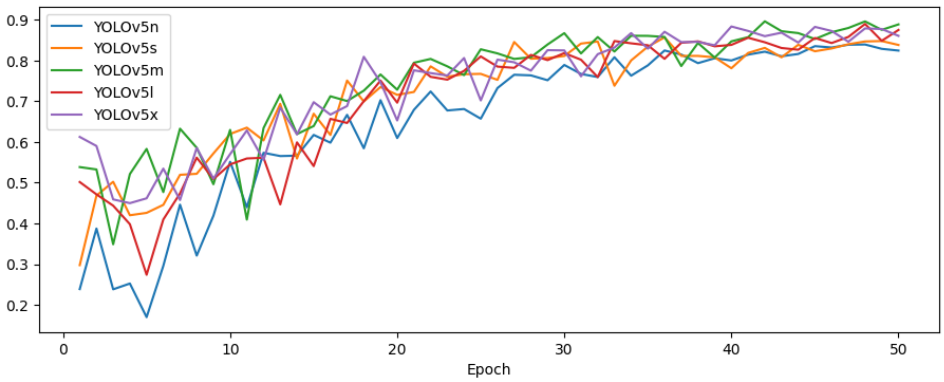
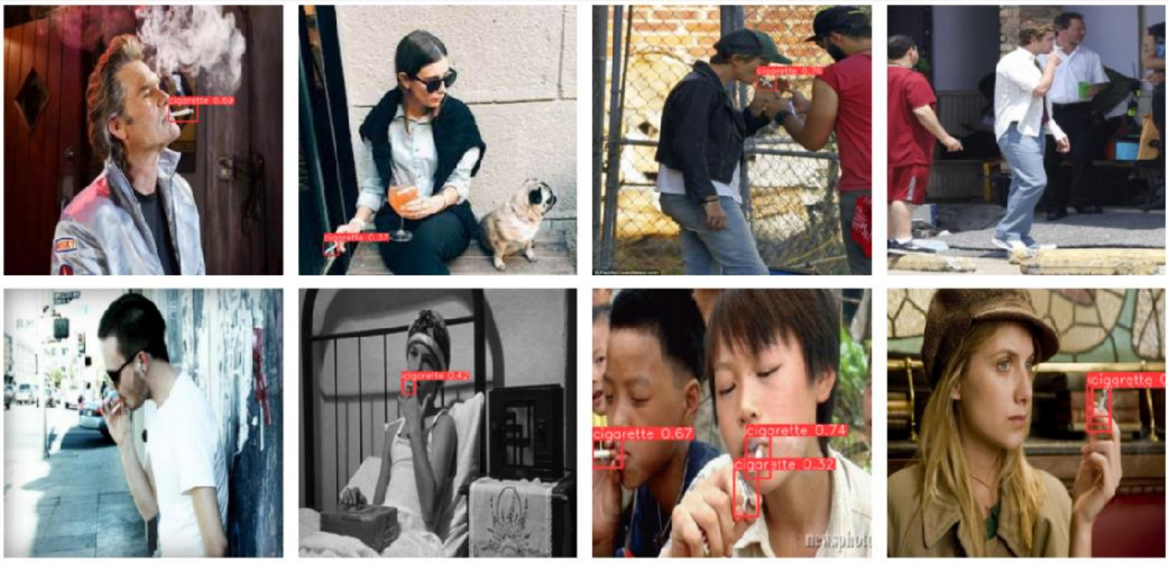


Figure 8. mAP with IoU=0.50 with the Training YOLOv5 Over 50 Epochs

Visualization of the detection effect of YOLOv8 and YOLOv5 models, several images are selected from the test dataset for detection comparison. Figure 6 in the results section shows cigarette detection using the YOLOv8 and YOLOv5 algorithms on multiple images. The apparent superiority of YOLOv5 over YOLOv8 in cigarette detection can be attributed to several factors, such as model architecture, hyperparameter tuning, object detection algorithm, and implementation details. Overall, the superior performance of YOLOv5 in Figure 9 suggests that it exhibits better accuracy and reliability in detecting cigarettes within the images than YOLOv8.



YOLOv8

YOLOv5

Figure 9. The Result of the Proposed Model

5. Conclusion

This article provides a significant contribution to the field of object detection, specifically in the identification of cigarettes in images. The study leverages the YOLOv8 algorithm, a powerful tool for object detection, and applies it to a critical public health issue: tobacco use. The comprehensive literature review offers a robust understanding of the current state of research, while the detailed methodology section ensures the study's replicability.

First, our model is trained using the public dataset of similar jobs [29], achieving 78% accuracy. Then, we created our dataset to improve the accuracy of the model. First, the YOLOv8 model is trained using the public dataset [29] and achieved 78% accuracy. Second, the dataset is created to improve the accuracy of the model. Then, both models, YOLO8 and YOLO5, were trained using the dataset to perform a comparative analysis. The YOLOv8 model achieved an accuracy of 87.4%, while the YOLOv5 model reached 89.6% accuracy.

This study provides a comparative analysis of the YOLOv8 and YOLOv5 models, offering valuable insights for future research on cigarette detection. These findings could guide the development of more accurate and efficient models for cigarette detection in images, thereby aiding in the broader fight against tobacco use.

References

[1] World Health Organization, "Tobacco" 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/tobacco.

[2] E. Arkin, N. Yadikar, X. Xu, A. Aysa and K. Ubul. "A survey: object detection methods from CNN to transformer," Proc. - Multimedia Tools and Applications, 2023.

[3] J. Redmon and A. Farhadi. "YOLO9000: Better, Faster, Stronger" Proc. - arXiv preprint arXiv: 1612.08242, 2016.

[4] J. Redmon, S. Divvala, R. Girshick and A. Farhadi. "You Only Look Once: Unified, Real-Time Object Detection," Proc. - arXiv preprint arXiv: 1506.02640, 2016.

[5] J. Redmon and A. Farhadi. "YOLOv3: An Incremental Improvement," Proc. - arXiv preprint arXiv: 1804.02767, 2018.

[6] A. Bochkovskiy, C. Wang and M. Liao. "YOLOv4: Optimal Speed and Accuracy of Object Detection," Proc. - arXiv preprint arXiv: 2004.10934, 2020.

[7] D. Jayakumar and S. Peddakrishna. "Performance Evaluation of YOLOv5-based Custom Object Detection Model for Campus-Specific Scenario," Proc. - International Journal of Experimental Research and Review, 2024.

[8] C. Li, L. Li, H. Jiang, K. Weng, Y. Geng, K. Ke, Q. Li, M. Cheng, W. Nie, Y. Li, B. Zhang, Y. Liang, L. Zhou, X. Xu, X. Chu, X. Wei and Wei, X. "YOLOv6: A Single-Stage Object Detection Framework for Industrial Applications," Proc. - arXiv preprint arXiv: 2209.02976, 2022.

[9] C. Wang, A. Bochkovskiy and M. Liao. "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," Proc. - arXiv preprint arXiv: 2207.02696, 2022.

[10] Ultralytics YOLOv8 Docs, "Introducing Ultralytics YOLOv8" 2023. [Online]. Available: https://docs.ultralytics.com/#where-to-start.

[11] J. R. Macalisang, N. E. Merencilla, D. Ligayo. "Eye-Smoker: A Machine Vision-Based Nose Inference System of Cigarette Smoking Detection using Convolutional Neural Network," Proc. - 7th International Conference on Engineering Technologies and Applied Sciences (ICETAS), 2020.

[12] A. Khan, S. Khan, B. Hassan, and Z. Zheng. "CNN-Based Smoker Classification and Detection in Smart City Application,". Sensors 2022.

[13] C. Santiago, M. Reyes, L. Tria. "Deep Convolutional Neural Network for Detection of Cigarette Smokers in Public Places: A Low Sample Size Training Data Approach," Proc. - 2022 International Conference on Decision Aid Sciences and Applications (DASA), 2022.

[14] D. Zhang, C. Jiao, S. Wang, "Smoking Image Detection Based on Convolutional Neural Networks," Proc. - 2018 IEEE 4th International Conference on Computer and Communications (ICCC), 2018.

[15] C. Wang, T. Zheng, F. Sun and H. Lia "A Smoking Detection Algorithm Based on Improved YOLOV5," Proc. - 2023 IEEE 3rd International Conference on Power, Electronics and Computer Applications (ICPECA), 2023.

[16] Q. Ding, X. Dong, W. Guo, W. Zheng and Y. Pan, "Smoking Detection Algorithm Based On Improved YOLOv5," Proc. - 2023 IEEE 7th Information Technology and Mechatronics Engineering Conference (ITOEC), 2023.

[17] J. Peng, C. Wang, Y. Li, H. Chen, "Substation Personnel Smoking Detection Based On GhostNetV2-YOLOv5," Proc. - 2023 6th International Symposium on Autonomous Systems (ISAS), 2023.

[18] F. Ciaglia, F. S. Zuppichini, P. Guerrie, M. McQuade, and J. Solawetz. "Roboflow 100: A Rich, Multi-Domain Object Detection Benchmark," Proc. - arXiv preprint arXiv: 2211.13523, 2022.

[19] C. Wang, H. Mark Liao, Y. Wu, P. Chen, J. Hsieh,and I. Yeh. "CSPNet: A New Backbone that can Enhance Learning Capability of CNN," Proc. - CVPR 2020 open access, 2020.

[20] T. Huang, M. Cheng, Y. Yang, X. Lv, J. Xu. "Tiny Object Detection based on YOLOv5," Proc. - 5th International Conference on Image and Graphics Processing, 2022.

[21] K. Jiang, T. Xie, R. Yan, X. Wen, D. Li, H. Jiang, N. Jiang, L. Feng, X. Duan, and J. Wang. "An Attention Mechanism-Improved YOLOv7 Object Detection Algorithm for Hemp Duck Count Estimation,". Proc. - Internet and Computers for Agriculture, 2022.

[22] R. Ju, W. Cai, "Fracture Detection in Pediatric Wrist Trauma X-ray Images Using YOLOv8 Algorithm," Proc. - arXiv preprint arXiv:2304.05071v5, 2023.

[23] S. Liu, L. Qi, H. Qin, J. Shi, and J. Jia. "Path Aggregation Network for Instance Segmentation," Proc. - arXiv preprint arXiv:1612.03144, 2017.

[24] T. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie. "Feature Pyramid Networks for Object Detection," Proc. - arXiv preprint arXiv:1803.01534, 2018.

[25] J. Terven, D. Cordova-Esparza, and J. Romero-Gonzalez. "A Comprehensive Review of YOLO Architectures in Computer Vision: From YOLOv1 to YOLOv8 and YOLO-NAS," Proc. - Machine Learning and Knowledge Extraction, 2023.

[26] M. Everingham, L. Gool, C. Williams, J. Winn and A. Zisserman. "The PASCAL Visual Object Classes (VOC) Challenge," Proc. - International Journal of Computer Vision, 2009.

[27] J. Davis and M. Goadrich. "The relationship between Precision-Recall and ROC curves," Proc. - 23rd international conference on Machine learning, 2006.

[28] R. Padilla, S. Netto and A. Eduardo. "A Survey on Performance Metrics for Object-Detection Algorithms," Proc. -

2020 International Conference on Systems, Signals and Image Processing (IWSSIP), 2020.
[29] Dataset. "Cigarette Computer Vision Project", 2022. [Online]. Available: https://universe.roboflow.com/smoke-tamtu/cigarette-6ubdv.

Author(s) Contributions
Yerniyaz Bakhytov: Data generation, Software, Methodology, Writing – Original Draft.
Cemil Öz: Methodology, Writing, review, and editing.

Conflict of Interest Notice
Authors declare no conflict of interest regarding the publication of this article.

Ethical Approval
It is declared that during the preparation process of this study, scientific and ethical principles were followed, and all the studies benefited from are stated in the bibliography.

Availability of Data and Material
Not applicable.

Plagiarism Statement
This article has been scanned by iThenticate ™.