Volume 1 Issue 2 2023
Pages: 83-92
ISSN: 2980-3152

CURRENT TRENDS IN COMPUTING

http://ctc.karabuk.edu.tr
RESEARCH ARTICLE

DEEP LEARNING-BASED HUMAN DETECTION FOR FALL INJURIES

BUSE SARICAYIR[1], ESMANUR ALICAN[1] AND CANER OZCAN[2*]

[1] *Computer Engineering Department, Karabuk University, Karabük, 78050, Turkiye*

[2] *Software Engineering Department, Karabuk University*

ABSTRACT. Falls are a significant public health problem, especially among the elderly and people with limited mobility. A fall may seem like a minor accident, but the injuries that can result from a fall and the underlying health problems that can cause falls have a significant impact on people's lives. Especially in elderly individuals, such accidents occur more frequently and lead to more severe consequences. Research shows that one-third of homebound older adults and more than half of hospitalized older adults are at risk for falls. Falls can result in impaired balance and gait, fear of falling, disability, and a decline in daily activities and quality of life. This fear adversely affects the daily lives of elderly individuals. Therefore, real-time fall detection systems contribute to preventing more severe injuries. Our proposed method uses state-of-the-art deep learning techniques to detect and localize people in video streams. Its goal is to ensure the rapid provision of assistance to the person who has fallen after the incident. In the development stage of the paper, YOLOv7 and YOLOv8 architectures have been utilized. Furthermore, we discuss the potential applications of this approach in real-world scenarios, such as fall detection systems for elderly care, surveillance, and automated emergency response. The main contributions of this work are a novel deep learning approach to human detection in the context of fall injuries, practical applications of the proposed approach, and its potential to improve safety and quality of life for at-risk populations, especially the elderly and those with limited mobility.

1. INTRODUCTION

Falls are unpredictable events that can have serious consequences, including injury and even accidental death. Following injuries from a fall, a person may be unable to call for help or, in the worst case scenario, lose consciousness. Fall injuries are a particularly significant problem for the elderly population. In Turkey, the elderly population is increasing compared to previous periods. If this trend continues, approximately one-third of the country's population is expected to be aged 65 and older in the coming years. As a result, the number of elderly care facilities is expected to increase [1].

Care centers for the elderly play a significant role in ensuring the well-being and safety of the elderly. One of the biggest concerns in such facilities is the occurrence of fall-related injuries that can affect

E-mail address: canerozcan@karabuk.edu.tr(C.Ozcan) esmnralicann@gmail.com(E.Alican), busesaricayir@gmail.com (B.Saricayir).
Key words and phrases. YOLOv7,YOLOv8, Elderly People, CNN, Fall Detection.

the elderly. While a fall may seem like a minor accident, fall injuries and the underlying health issues contributing to falls can significantly impact human lives. Falls are among the most common causes of injury, regardless of age. Such accidents are more prevalent, and their consequences are more severe, particularly among the elderly population. In 2017, the elderly population constituted 8.9% of the world's population. The top three countries with the highest proportion of elderly citizens were Monaco with 32.2%, Japan with 27.9%, and Germany with 22.1%. Turkey ranked 66th out of 167 countries [2].

Research indicates that one-third of the elderly living at home or in nursing homes, and more than half of those in hospitals, are at risk of falling [3]. Research shows that an increasing number of older people are afraid of falls and the complications that can arise as a result. This fear has a negative impact on the daily lives of older people [4]. Falls can lead to a loss of independence among the elderly. They might encounter difficulties in their daily activities and become reliant on assistance. Older people's daily routines and lifestyles may change. Falls can also have an emotional impact. Fear, anxiety, and insecurity can lead to social isolation among the elderly. Injuries resulting from falls can increase healthcare utilization and costs.

Therefore, our study aims to develop a system that employs deep learning using the dataset we have prepared, aiming to detect potential fall incidents within elderly care centers. The primary goal is to reduce the intervention time for fall injuries among the elderly and to provide caregivers with real-time information about accident victims through an application that facilitates patient access. The motivation behind this study is rooted in addressing a critical and pressing public health concern. Falls are a significant source of injuries, particularly among the elderly and individuals with limited mobility, and can lead to severe health complications, increased healthcare costs, and even loss of life. Recognizing the urgency and importance of this issue, the paper is motivated key factors such as advances in deep learning, elderly care and independent living, potential for automated emergency response and, real-world applications. This paper is organized as follows. Related work is provided in Section 2, followed by methods and data collection and preparation in Section 3. Experimental results and research discussion are presented in Section 4, and the conclusion is presented in Section 5.

2. Related Work

Detection of falls among the elderly in care centers is pivotal due to their significant contribution to injuries and fatalities. Falls are the primary cause of injury in older individuals and the leading cause of accidental death among those aged 75 or older [5]. These incidents result in physical harm and evoke psychological and emotional repercussions for the elderly and their families [6]. Introducing innovative fall detection systems offers the potential to mitigate these adversities. Falls can lead to reduced independence, hinder daily activities, and necessitate assistance, altering the elderly's routines and lifestyles. Emotional implications encompass fear, anxiety, and insecurity, fostering social isolation. Fall-related injuries escalate healthcare use and expenditures, impacting emergency visits, hospitalizations, rehabilitation, and caregiver responsibilities. Preventive interventions hold promise for preventing falls in older adults.

Numerous fall detection systems have been proposed, employing diverse techniques, including the noteworthy application of machine vision technology. These systems rely on digital image processing and high-resolution image acquisition for accuracy. The human activity recognition project is the first work to come to the fore in research on this subject. In this paper, 7 data sets and 7 different methods were tested, and since the results were not related to real life, it was decided to try different methods in future studies [7]. However, as this research shows, not all imaging systems offer high-quality images, and investing in higher-resolution devices is costly. Obstructions caused by objects or architectural elements lead to reduced event perception. As a result of the research that considered this event, another study that detects human falls based on body posture was examined. In experiments with the model developed in this study, a performance value of 95.1% was achieved by filtering invalid objects, and a sampling frequency of 5 fps (frames per second) was obtained [8]. Similar limitations apply to environmental-based systems, which require the installation of acoustic or vibration sensors throughout the interior. The results obtained from one of the sample studies on this subject showed that the fall identification accuracy reached 90.53% with only 20% labeled data using the proposed algorithm. It was observed that the accuracy increased from 90.53% to 98.74% with the increase in labeling rates from 20-80% [9].

In another paper on this topic, the proposed model was compared with 8 different models, and the highest accuracy value of 83.2% was obtained [10]. However, the application requirements and cost hinder the feasibility. One notable study focuses on a multi-class deep learning framework for fall event classification using Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) architectures [11]. In experiments involving eleven subjects, LSTM achieved a 92.1% accuracy in fall event detection, paralleled by CNN models. Predominant image-based fall detection systems use motion-derived or background subtraction techniques. Identifying abrupt aspect ratio changes is significant, warranting a synthesis of Hidden Markov Models [12], adaptive thresholding, and personalized user data for video-based fall detection. Velocity discrimination discriminates between motion types in video segments, albeit spatially constrained. A separate investigation employs Gaussian Mixture Models within the YCbCr color space for background subtraction in object detection [13]. This two-stage approach extracts three object attributes, with the first two integrals used for fall detection and the last one for fall event confirmation. A toggle Finite State Machine oversees individuals' activities. This study leverages contemporary advancements in object detection methodologies, specifically within the You Only Look Once (YOLO) [14] architecture. YOLOv7 [15] and YOLOv8 [16] architectures were emphasized. One of the key advantages of the YOLO algorithm is its inherent computational efficiency and real-time processing capabilities. Executing object detection in a single pass enables YOLO to achieve remarkable computational swiftness while maintaining precision. This makes YOLO suitable for scenarios requiring rapid responses, as seen in fall injury detection in elderly care facilities. The study explores operationalizing YOLOv7 and YOLOv8 architectures in deep learning-infused object detection systems. Through a comparative analysis, the primary goal is to identify the superior architecture for effective implementation of fall detection in elderly care facilities.

TABLE 1. Related works compare table.

References	Year	Method/Architecture	mAP
Zhang [8]	2020	Sensors	95.1
Liu[9]	2019	SVM, Sensors	98.74
Sim[10]	2015	Sensors	83.2
Cardenas[11]	2023	LSTM and CNN	92.1
Our work	2023	YOLOv8 and YOLOv7	v8: 78, v7:81

3. METHODS

3.1. Data Collection and Preparation.

In this study, we use the fallen_new_version dataset [17], which contains 3290 samples categorized into four different classes: "falling", "sitting", "squatting" and "standing", representing various human activities. The multi-class structure of dataset enhances classification efficacy, fostering algorithmic discernment. Sample allocation adheres to best practices: 2445 samples for training, 503 for validation, and 342 for testing, facilitating model refinement. This strategic allocation ensures empirical rigor, mitigating overfitting. Table 2 provides a summary of the dataset arrangement, outlining the training, testing and validation roles that enhance algorithmic performance calibration with a mean Average Precision (mAP) value.

TABLE 2. Number of samples on fallen_new_version dataset [17].

fallen_new_version dataset	
Training	2445
Validation	503
Test	342
Total:	3290

3.2. YOLOv7 and YOLOv8 Algorithms.

YOLOv7 [15] is an advanced real-time object detection algorithm compared to previous versions in terms of speed and accuracy. As shown in Figure 1, YOLOv7 achieves this by having fewer parameters than other YOLO versions, making it capable of operating efficiently with low computational power. In order to improve the detection performance and accuracy of YOLOv7's design, significant changes were made to create specific modules. The Extended Efficient Layer Aggregation Network (E-ELAN) is integrated into the CSPDarknet backbone of YOLOv7. The Efficient Layer Aggregation Network (ELAN) is designed to efficiently construct networks by following the longest short gradient path. E-ELAN is a modified version of this structure. YOLOv7 incorporates the pre-processing method from YOLOv5. For identifying sections (modules) within the overall model that require re-parameterization, YOLOv7 employs gradient flow propagation methods [19]. In general, YOLOv7 employs a more effective feature integration method, resulting in more accurate object detection performance. As a result, YOLOv7 requires less expensive computational hardware compared to other deep learning models. It can be trained much faster on small datasets without pre-trained weights.

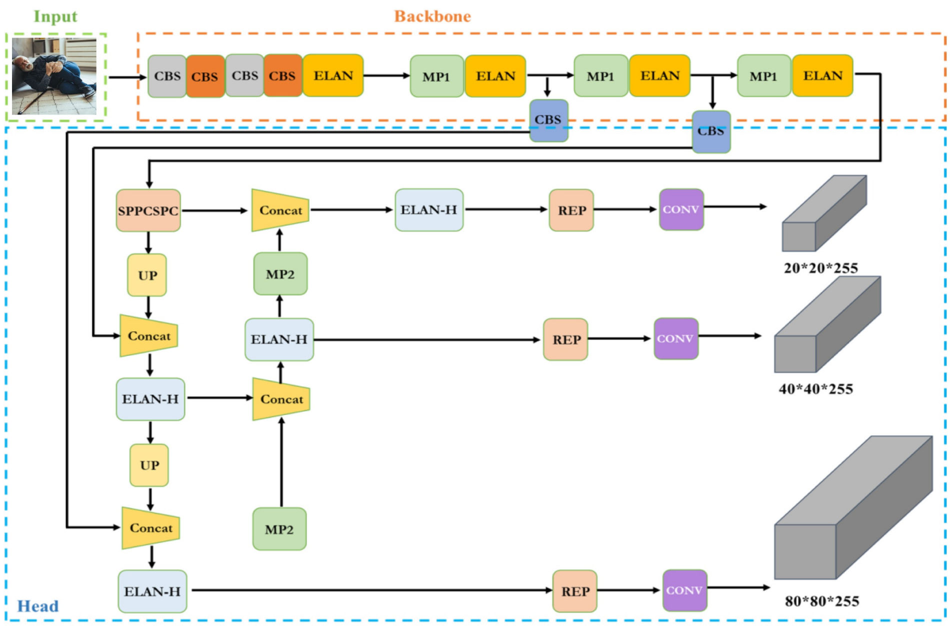

FIGURE 1. YOLOv7 architecture [18].

The YOLOv8 model architecture, given in Figure 2, introduces an anchor-free approach, which means it no longer relies on predicting an object's distance from predefined anchor boxes. Instead, it directly estimates the object's center. This innovation reduces the number of predicted bounding boxes, leading to a more efficient Non-Maximum Suppression process during post-processing, which is typically a complex step for filtering detections after inference [21]. In terms of model variations, there are five distinct architectures (YOLOv8n, YOLOv8s, YOLOv8m, YOLOv8l, and YOLOv8x), each designed for specific tasks like object identification, segmentation, and classification. Among these, YOLOv8x stands out for its high precision in detection but sacrifices speed compared to the others. On the opposite end of the spectrum, YOLOv8 Nano is the fastest and most compact model in terms of size [22]. This innovation shows great potential for enhancing object detection tasks and delving into the intricacies of the various YOLOv8 iterations to identify the most suitable model for specific applications.

4. EXERIMENTAL STUDIES

The primary objective of this scientific endeavor is to perform an exhaustive comparative analysis between two different architectural configurations of the YOLO framework, as shown in Table 3. This

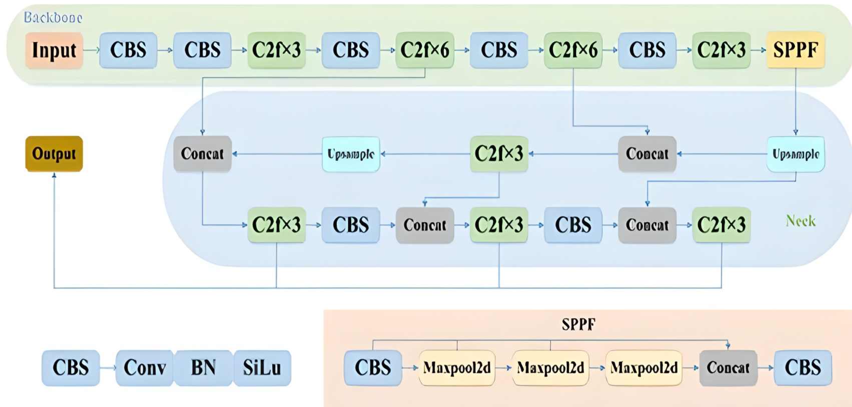

FIGURE 2. YOLOv8 architecture [20].

investigation is motivated by the overarching goal of determining which of these architectural variants demonstrates superiority in both precision and computational efficiency. Through a rigorous process of evaluation and experimentation, this study seeks to identify the architectural configuration that achieves an optimal balance, maximizing accuracy in object detection while simultaneously optimizing computational processing speed. In the pursuit of this research objective, significant emphasis is placed on the utilization of well-established metrics, such as mAP0.5:0.95 and mAP0.5, within the domain of computer vision. These metrics, widely adopted for the comparative analysis of diverse object detection models and algorithms, play an indispensable role in evaluating the performance of the YOLO architectures under investigation.

Both mAP0.5:0.95 and mAP0.5 serve distinct purposes in the assessment of object detection models. mAP0.5:0.95 offers a comprehensive evaluation across a spectrum of Intersection over Union (IoU) thresholds, ranging from 0.5 to 0.95, providing a nuanced appraisal of performance at varying levels of agreement with ground truth annotations. In contrast, mAP0.5 restricts its evaluation to a fixed IoU threshold of 0.5, yielding a more streamlined yet specific evaluation of detection accuracy under particular conditions of overlap with ground truth annotations. Both metrics hold valuable significance in the context of model evaluation, enabling researchers and practitioners to gauge model performance according to their specific requirements and the characteristics of the given task.

TABLE 3. Experimental results

Model	mAP0.5	mAP0.5:0.95
YOLOv7	81%	0.495
YOLOv8	78%	0.454

In most academic studies, rather than evaluating model performance using only a single metric, it is common to consider learning rate control and to include multiple metrics, such as precision (which emphasizes the model's ability to find positive instances and the accuracy of positive predictions, respectively) (1), recall (2), mAP (3), and IoU (4), are commonplace. These metrics not only offer a quantitative means of gauging model performance but also facilitate an equitable and extensive comparison across various tasks and domains [23]. Furthermore, they play an important role in model refinement by providing insight into the rate at which a model learns, allowing for necessary adjustments. Consequently, a holistic evaluation of all facets of model performance becomes imperative, particularly in the case of deep learning models characterized by intricate and multi-layered architectures. It's important to note that precision and recall are often used in conjunction, and there is typically a trade-off between them; as one metric improves, the other may deteriorate. Balancing precision and recall in classification tasks is critical to tailoring model performance to the specific goals and requirements of a given application.

$$Presicion = \frac{TP}{TP+FP} \tag{1}$$

$$mAP = \frac{1}{N}\sum_{i=1}^{N} APi \tag{2}$$

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

$$IoU = \frac{TP}{TP+FP+FN} \tag{4}$$

As shown in Figure 3, the graph illustrates the outputs stemming from trainings with the most elevated mAP values for both architectural configurations. An examination of this graphs reveals a notable distinction. While the mAP values within the mAP50 graph for YOLOv8 hover around 0.77, the corresponding metric rises to its zenith, reaching 0.81 within the mAP50 graph for YOLOv7. This pattern, however, does not hold true for the mAP50-95 graphs, where YOLOv7 reaches its highest mAP value at 0.495, slightly eclipsed by YOLOv8 at 0.454. Precision and recall graphs further underscore differences between the algorithms, with YOLOv8 showing superior values relative to YOLOv7.

As part of the experimental works, we also used the Fallsdata2 [24] dataset in the testing phase. We do not provide detailed information about this dataset, especially since it was not used in the training phase. The accuracy values obtained using the Fallsdata2 dataset are shown in Figure 4.

Accordingly, the YOLOv7 architecture exhibited higher accuracy on the test data. To compare the architectures, they went through different processes. First, they were compared in terms of accuracy factor. According to the tests performed, YOLOv7 showed higher accuracy compared to YOLOv8. Therefore, YOLOv7 is a better choice for this data set in terms of accuracy. Another aspect taken into consideration is speed. In this context, after training, YOLOv8 yielded a speed of 1.6ms, whereas YOLOv7 achieved a

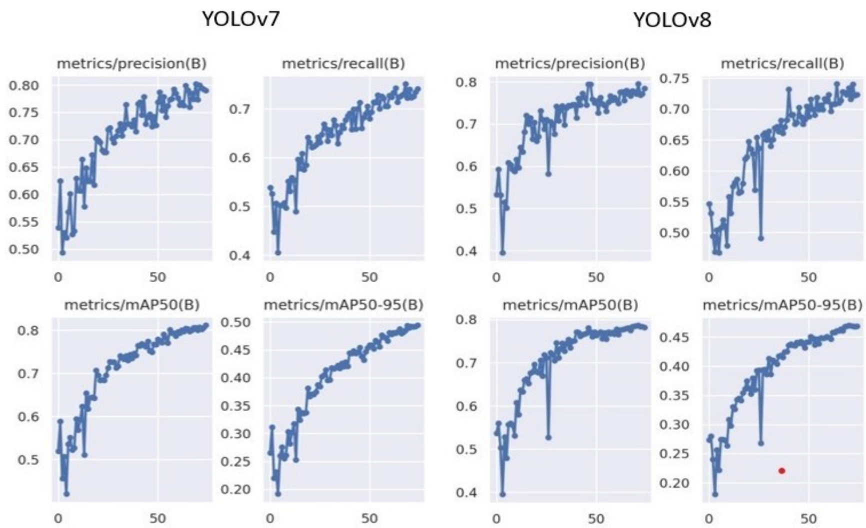

FIGURE 3. YOLOv7 and YOLOv8 metric results for train data in the fallen_new_version dataset.

FIGURE 4. YOLOv8 vs YOLOv7 test results[23].

training speed of 1.7ms. In this architectural comparison, YOLOv8 is the better choice in terms of speed, as it is faster than YOLOv7.

Since the results of comparing these two factors did not definitively determine the architectural choice, another factor was considered for comparison. Robustness was chosen as this factor, and included testing on challenging data. As a result of these performance comparisons, YOLOv7 showed better results than YOLOv8, leading to its selection as the architectural choice for the paper.

5. Conclusion

From the available data, it is clear that falls are one of the most common causes of injuries spanning all age groups. Despite their seemingly harmless nature, falls can lead to injuries with far-reaching consequences, which are exacerbated by underlying health conditions. This impact is particularly pronounced in the elderly, for whom such accidents occur at alarming rates and have significant consequences. The novelty of this paper is derived from its innovative deep learning model, rigorous comparative analysis, focus on practical applications, and its potential to significantly impact public health and the well-being of individuals at risk of fall injuries. Therefore, the proposed study makes an important contribution to the literature. The proposed model excels in terms of precision, recall, mAP, robustness, real-time responsiveness, adaptability, its data-driven approach, potential for automation, and its substantial potential to make a positive impact on public health. As a result, real-time fall detection systems play a critical role in reducing the severity of resulting injuries. The proposed work aims to be field tested in elderly care centers to quickly assist individuals following a fall incident. To achieve this goal, the paper utilizes the YOLOv7 and YOLOv8 architectures and strives to bring innovation to the field compared to previous studies. While this study presents a promising approach, it is important to be aware of some limitations when considering its practical application and to continuously address and mitigate them to improve the model's effectiveness and ethical use. The study's performance heavily relies on the quality and diversity of the dataset used for training and evaluation. If the dataset is biased or not fully representative of real-world scenarios, the model's performance might be limited in terms of generalizability. While the study aims to reduce false alarms, there is still a possibility of false positives where the system incorrectly identifies a non-fall event as a fall. This can be a limitation, particularly if the system triggers unnecessary alerts or interventions.

Data Statement

During the present investigation, no datasets were generated. Instead, for the purposes of training and testing, the study made use of two Roboflow-based datasets, namely 'fall_new_version[17]' and 'Falls-data2 [24]'.

Conflict of Interest

The authors have no affiliation with any organization with a direct or indirect financial interest in the subject matter discussed in the manuscript.

References

[1] Fevzi, K. A. Y. A.,Elderly Population and Nursing Homes in Turkey ,Academic Perspective International Peer-Reviewed Journal, (61), 423-440. , (2017).

[2] Tuik. https://data.tuik.gov.tr/Bulten/Index?p=Statistics-on-the-Elderly-2017. [Accessed 08-August-2023].

[3] Güner, S. G., Ural, N. , Falls in the Elderly: Situation Assessment within the Scope of Thesis Studies Conducted in Our Country. Izmir Katip Celebi University Journal of Health Sciences Faculty, 2(3), 9-15.,(2017).
[4] Şentürk, A. Y., Rates of Falls in the Elderly and Fall Prevention Measures. Anadolu Current Medical Journal, 2(2), 47-52,(2020).
[5] Ambrose, A. F., Geet, P., and Jeffrey M. H. ,Risk Factors For Falls Among Older Adults: A Review Of The Literature, Maturitas 75.1, 51-61, (2013).
[6] El-Bendary, N., Tan, Q., Pivot, F. C., Lam, A., Fall Detection and Prevention for the Elderly: A Review of Trends and Challenges, International Journal on Smart Sensing and Intelligent Systems, vol.6, no.3, 3913, pp.1230-1266. ,(2013). https://doi.org/10.21307/ijssis-2017-588
[7]] Zhang, S.; Wei, Z.; Nie, J.; Huang, L.; Wang, S.; Li, Z., A Review on Human Activity Recognition Using Vision-Based Method. J. Healthc. Eng. 2017, 3090343. (2017).
[8] Zhang, J.; Wu, C.; Wang, Y. ,Human fall detection based on body posture spatio-temporal evolution, Sensors 2020, 20, 946., (2020).
[9] Liu, C.; Jiang, Z.; Su, X.; Benzoni, S.; Maxwell, A. , Detection Of Human Fall Using Floor Vibration And Multi-Features Semi-Supervised SVM. Sensors, 19, 3720., (2019).
[10] Sim, J.M.; Lee, Y.; Kwon, O. , Acoustic Sensor Based Recognition of Human Activity in Everyday Life for Smart Home Services. Int. J. Distrib. Sens. Netw. 2015, 679123. , (2015).
[11] Cardenas, J.D.; Gutierrez, C.A.; Aguilar-Ponce, R., Deep Learning Multi-Class Approach for Human Fall Detection Based on Doppler Signatures, Int. J. Environ. Res. Public Health, 20, 1123., (2023).
[12] Eddy, Sean R., Hidden Markov Models, Current Opinion In Structural Biology 6.3: 361-365.(1996).
[13] Jiang, P., Ergu, D., Liu, F., Cai, Y., Ma, B. A Review of Yolo Algorithm Developments, Procedia Computer Science, 199(Complete), 1066–1073.,(2022). https://doi.org/10.1016/j.procs.2022.01.135
[14] Redmon, J., Divvala, S., Girshick, R., Farhadi, A. ,You Only Look Once: Unified, Real-Time Object Detection, arXiv [Cs.CV] (2016). Retrieved from http://arxiv.org/abs/1506.02640.
[15] Wang, Chien-Yao, Alexey Bochkovskiy, and Hong-Yuan Mark Liao., YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Vancouver, BC, Canada (2023).
[16] GitHub. https://github.com/ultralytics/ultralytics, [Accesssed 20-October-2023].
[17] ByoungWook. fallen_new_version Dataset. Roboflow Universe. Retrieved from https://universe.roboflow.com/byoungwook-b5cd6/fallen_new_version, [Accesssed 08-August-2023].
[18] Jiang, K., Xie, T., Yan, R., Wen, X., Li, D., Jiang, H., Wang, J. , An attention mechanism-improved YOLOv7 object detection algorithm for hemp duck count estimation. Agriculture, 12(10), 1659. (2022).
[19] Karadağ, B., Ali, A. R. I., Object Detection on Smart Mobile Devices Using YOLOv7 Model. Politeknik Dergisi, 1-1. (2023).
[20] Lou, H., Duan, X., Guo, J., Liu, H., Gu, J., Bi, L., Chen, H. DC-YOLOv8: Small-Size Object Detection Algorithm Based on Camera Sensor. Electronics, 12(10), 2323. (2023).
[21] Talaat, F. M., & ZainEldin, H. , An improved fire detection approach based on YOLO-v8 for smart cities. Neural Computing and Applications, 1-16. (2023).
[22] Reis, D., Kupec, J., Hong, J., & Daoudi, A. , Real-Time Flying Object Detection with YOLOv8. arXiv [Cs.CV]. Retrieved from http://arxiv.org/abs/2305.09972 , (2023).
[23] Mungoli, N. , Adaptive Ensemble Learning: Boosting Model Performance through Intelligent Feature Fusion in Deep Neural Networks. arXiv preprint arXiv:2304.02653. (2023).
[24] Roboflow, fallsdata2_dataset Dataset. Roboflow Universe. Retrieved from https://universe.roboflow.com/zsd/fallsdata2, [Accesssed 2-September-2023].