ISSN: 2093-5374 / E-ISSN: 2093-5382

Multi-Scale Attention-based Wireless Network Algorithm for Enhancing Language Learning Outcomes

Dr. Salima Rustamiy[1*], Nargis Kurbanazarova[2], Gulbakhor Atayeva[3], Dildora Khashimova[4], Mohinakhon Khamidova[5], Mohira K. Usmanova[6], Feruza Mamatova[7], and Dildora Avezova[8]

[1*]Doctor of Philology, Professor, Oriental University, Uzbekistan. salimarustamiy@gmail.com, https://orcid.org/0009-0003-3605-7340

[2]Associate Professor, Department of Uzbek Linguistics, Termez State University, Uzbekistan. qurbonazarovan@tersu.uz, https://orcid.org/0000-0003-1384-8302

[3]Associate Professor, Candidate of Philological Sciences, Department of English, Samarkand State University named after Shar of Rashidov, Uzbekistan. atayevagulbaxor2025@gmail.com, https://orcid.org/0009-0005-8457-2798

[4]DSc, Professor, Department of Uzbek Language and Literature, Tashkent State University of Law, Tashkent, Uzbekistan. d.xashimova@tsul.uz, https://orcid.org/0000-0002-8542-611X

[5]Kimyo International University in Tashkent, Uzbekistan. m.khamidova@kiut.uz https://orcid.org/0000-0002-7681-0996

[6]Doctor of Philosophy (PhD) in Philological Sciences, Dotsent, Karshi State University, Uzbekistan. usmanovamohira@gmail.com https://orcid.org/0009-0006-1846-7719

[7]PhD in Philological Sciences, Associate Professor, Department of English Linguistics, The National University of Uzbekistan named after Mirzo Ulugbek, Uzbekistan. feruzamakhammadovna@gmail.com, https://orcid.org/0000-0003-0977-9682

[8]PhD student, National University of Uzbekistan, Uzbekistan. avezova_d@nuu.uz, https://orcid.org/0009-0008-1744-7227

Received: October 21, 2024; Revised: December 09, 2024; Accepted: January 08, 2025; Published: March 31, 2025

Abstract

Identifying students' behaviors in language learning classrooms can serve as a criterion for evaluating the efficacy of instructional methods. This research introduces an algorithm for detecting language learning classroom behavior with an enhanced object detection framework (i.e., YOLOv5) through wireless networks. The feature pyramidal framework in the cervical network of the initial YOLOv5 system is integrated with a balanced bidirectional feature pyramidal network. Their subsequent processing involves feature fusion across several object sizes to extract fine-grained characteristics of distinct actions. A spatial and channel Multi-Scale Attention Method (MSAM) is incorporated between the neck and prediction networks to enhance the model's concentration on object data, improving the detection precision. The initial minimal suppression is enhanced by employing the distance-based intersecting ratio to augment the differentiation of occluded items. Several studies on the newly established database encompassed four behaviors: listening, gazing

Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), volume: 16, number: 1 (March), pp. 94-103. DOI: *10.58346/JOWUA.2025.I1.005*

*Corresponding author: Doctor of Philology, Professor, Oriental University, Uzbekistan.

down, lying down, and rising. The findings indicated that the algorithm introduced in this work can precisely identify diverse student actions, exhibiting greater accuracy than the YOLOv5 framework. Upon evaluating the impact of student behavior recognition across several scenarios, the enhanced system demonstrated a mean precision of 89.2% and a recall of 91.3%, surpassing the performance of the systems under comparison.

Keywords: Multi-Scale Attention, Wireless Network, Language Learning, Education.

1 Introduction

The ongoing advancement of Artificial Intelligence (AI) (Zhai et al., 2021) has rendered intelligent education a prominent subject in the past few decades. Children constitute the primary participants in classroom learning activities (Wen, 2021). Integrating AI in classroom instruction and applying Deep Learning (DL) to analyze student behaviors would enhance understanding of learner conditions in language learning and increase teaching effectiveness through wireless networks (Surendheran & Prashanth, 2015). If a pupil is captivated by intriguing substances, he will exhibit indicators of attentiveness, such as seated, listening, or participating in question-and-answer sessions with the teacher. When a pupil experiences boredom in class, he lowers his head, becomes preoccupied, and falls asleep at his desk. The observation of pupil conduct in the classroom is of considerable importance. Many pupils in classroom settings might experience considerable object obstruction, presenting a notable obstacle to behavior identification. Conventional techniques include manually recording students' conduct in language learning, leading to inadequate documentation and significant consumption of human resources (Ferdowsi & Moradi, 2014; Bisrat et al., 2021) The use of Flanders' interactive structure signified the inception of the contemporary statistical analysis school. Numerous optimization techniques have been developed; most depend on educators' assessments or post-class observations. Conventional teaching assessments lack automated analytical instruments for classroom conduct.

Detecting classroom activity is a complex problem due to multiple pupils in a shared space, resulting in unequal pixel representation between the front and rear rows and significant occlusion challenges through wireless networks (Balamurugan et al., 2017). Identifying the nuanced properties associated with these obstructions is a critical challenge to enhance the accuracy of classroom behavior identification. The pupils' conduct exhibits diminished variability due to their static placements during the session. The object categories consist of static postures or displays, such as listening, gazing down, and seated, which do not necessitate multiple representations of time-related data (Demirel et al., 2022. This study excludes the temporal features of activity to preserve recognition speed and precision.

The authors employed an object identification methodology and introduced a Convolutional Neural Network (CNN) (Kattenborn et al., 2021) method that integrates Region of Interest (ROI) (Hossain et al., 2023) pooling and localized hold training to categorize language learning classroom behavioral pictures. They converted the issue of school behavior detection into a detailed categorization challenge of behavioral photographs in language learning (Kurbanazarova et al., 2024). The inference rate of the two-stage detecting architecture is inadequate, rendering it unsuitable for practical applications (Abdullah, 2024). The present research addresses the problem by employing an object identification method to identify students in school and classify their behavior through wireless networks.

The research offers an enhanced approach to educational behavior detection by integrating both detecting speed and precision. To address authentic educational situations in language learning and to improve identification efficiency, the study concentrates on single-stage identification computations, asserting that their performance is primarily influenced by two variables: the acquisition of Multi-Scale Attention Method (MSAM) characteristics and non-maximum suppression post-processing techniques.

The extraction of multiple scales is essential for identifying the intricate characteristics of dense objects. At the same time, minimal suppression post-processing techniques are vital for addressing occlusion issues and enhancing detection rates.

2 Background

The smart classroom has progressed swiftly in recent years, resulting in substantial breakthroughs in DL (Sarker, 2021) techniques for identifying pupil actions in educational environments. This subsection examines some significant studies that have advanced this discipline. The research employed the Region-based Fully Convolutional Network (R-FCN) design (Vijaya Kumar & Mahammad Shafi, 2023) to identify yawning habits in students, including pruning techniques and an innovative mouth-fitting approach to improve detection rates and ensure accurate behavioral analysis in classroom settings through wireless networks. The research presented a model utilizing coordinate-based attention integrated with fused residual systems. This algorithm markedly enhances the precision of classroom conduct detection by leveraging channel and spatial data to emphasize relevant aspects in student photos.

Subsequent contributions developed an innovative network utilizing You Only Look Once (YOLO) (Gallagher & Oughton, 2025) as the foundation. Through the development of an independent student behavior database and the implementation of data augmentation methods, they attained enhanced detection results (Chakma, 2025). The study progressed the field by incorporating a channel attention strategy into their framework, hence improving the understanding of student behaviors in classroom settings. These studies demonstrate the continuous advancement of methods for identifying pupil actions while simultaneously revealing a deficiency in the implementation of end-to-end models in educational contexts through wireless networks. In response, the research presents an innovative target identification methodology tailored to analyzing classroom behavior among pupils.

Object detection methods, fundamental to computer visual perception, are typically classified into two primary categories: two-phase and one-phase sensors. The two-phase sensors, exemplified by Region CNN (R-CNN) (Yi et al., 2021), were the inaugural implementation of DL in identifying objects. R-CNN employs targeted searches to identify potential areas from a picture, which are classed using CNN. Fast and Faster R-CNN were developed to improve the speed and effectiveness of the initial R-CNN. Faster R-CNN introduced Region Proposal Networking, which markedly expedited the detection procedure by allowing the concurrent prediction of object boundaries and scores for every category.

One-phase detectors like YOLO streamline the detection procedure by omitting the proposal phase, directly forecasting item categorization and localizations from whole photos. This method provides expedited processing speeds but has historically exhibited reduced accuracy relative to two-phase techniques. Advancements improved this methodology in algorithms, which progressively enhanced precision and speed, impacting later versions of YOLO.

The emergence of the Transformer design established a new paradigm in identifying objects, utilizing self-attention techniques to augment detection skills in language learning. This evolution highlights ongoing creativity in the industry, aiming to reconcile the trade-offs between detecting speed and precision. The studies and evaluations reveal persistent deficiencies in classroom behavior monitoring. (1) Despite the high identification rate of hardware-based behavior detection techniques, their reliance on hardware for multi-modal data gathering renders them impractical for real classroom environments and unsuitable for widespread implementation. (2) Certain DL algorithms employed in educational settings through wireless networks cannot simultaneously identify multiple objects within a single image frame, rendering them unsuitable for environments with numerous students; alternatively, while they detect multiple individuals in the same frame, they often lack real-time processing capabilities. Adequate real-time student conduct monitoring systems remain absent. The absence of extensive educational environment databases hinders researchers' ability to train DL models, as there are currently no publicly available large-scale classroom scene databases. This research annotates datasets for behavioral recognition in an educational context and refines the behavior recognition model in response to challenges encountered in these settings. The enhanced method was evaluated and yielded favorable results.

3 MSAM-based Wireless Network Algorithm

The proposed MSAM architecture consists of many essential elements: a feature-extracting system, the architecture, a decoder, and an estimation engine. The process commences with the foundation necessary for data extraction networks. This backbone examines input photos and extracts intricate characteristics, which are then organized into four-dimensional map structures in language learning. The mappings of features are transmitted to the element of the design. A normalization-based MSAM method is applied to the characteristics. This attention process is precisely engineered to improve the encoding of geographic data inside the characteristic maps, which is essential for precisely identifying targets.

The enhanced features advance to the subsequent stage of the architecture, where four encoding devices employing a multi-scale flexible MSAM system are implemented. This arrangement has been tailored to sparsify characteristics and augment detection capabilities across various scales. This architectural feature is crucial for managing the complications of differing object dimensions within the photos through wireless networks. The enhanced features are transmitted to the decoding and the forecasting head in language learning. Like that, the decoder analyzes these features and uses a Feed-Forward Network (FFN) to generate the final grouping results and item box boundaries.

An MSAM system is incorporated to enhance the YOLOv5 model's focus on pupil areas in a picture, addressing the presence of walls, chairs, tables, doors, and other backdrop elements in a conventional classroom setting, particularly when encountering extensive untargeted areas. The MSAM process identifies a limited subset of essential details from a vast array and concentrates on this crucial data, disregarding most unimportant data. The Convolutional Block Attention Modules (CBAM) presented in this article is a mixed MSAM system that amalgamates spatial and channel-related MSAM, exclusively emphasizing channel-related MSAM. CBAM incorporates both channel-related and spatial MSAM mapping procedures, thereby preserving a more significant amount of feature data. Figure 1 illustrates the construction of CBAM.

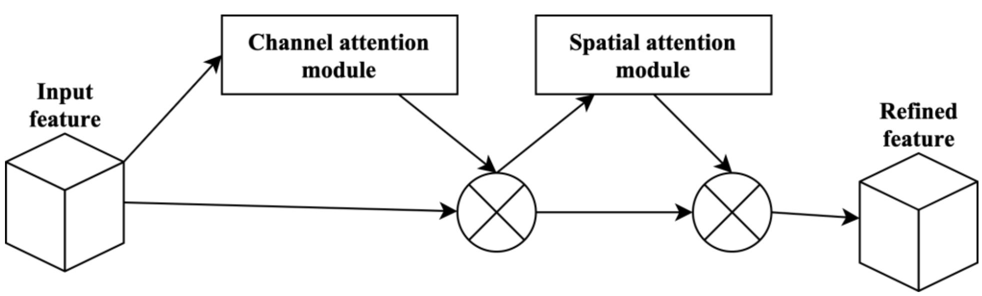


Figure 1: Attention Module

The graphic above illustrates the comprehensive flow architecture of the CBAM module, comprising two distinct submodules: the channelized MSAM component and the spatial MSAM component. The input characteristics initially traverse a channel MSAM component, which assigns related weights for

MSAM to every feature channel. Following the acquisition of the weighted outcomes, they proceed through a spatial MSAM component, which applies to the relevant MSAM-heavy objects to various places of the map of features, ultimately yielding the output product of the convolutional layers. Every characteristic in a channel corresponds to a distinct detector; hence, it is logical for channel MSAM to concentrate on the traits. The channel MSAM component condenses the list of features in the spatial domain to generate one-dimensional vectors and then performs operations.

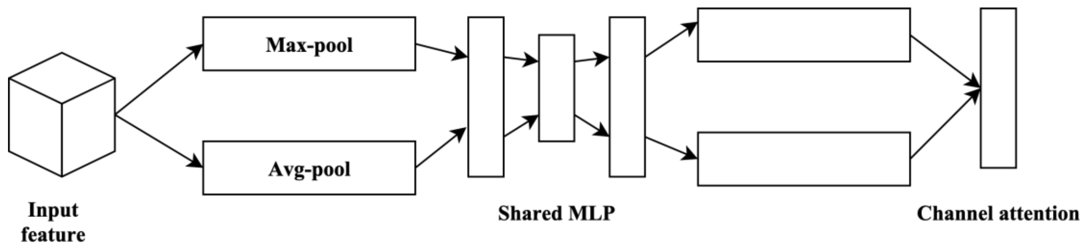


Figure 2: Channel Attention Module

The channel MSAM component depicted in Figure 2 initially executes global maximum pooling and mean pooling on the input characteristic map across the top and bottom parameters, respectively, to consolidate the spatial data of the map's features. The outcome is transmitted to the Multiple Layer Perceptron (MLP) via the standard, Fully Connected (FC) level. The last channel MSAM map of features $M_c(F)$ is produced using the sigmoid activating process. After being weighted with the input characteristic mapping, the result of this feature mapping serves as the input characteristic for the spatial MSAM component. The mathematical representation of the channel-related MSAM mapping $M_c(F)$ is:

$$M_c(F) = \sigma\big(MLP(avg - pool(F)) + MLP(max - pool(F))\big) \quad (1)$$

Where σ denotes the sigmoid activating operation, MLP signifies the relationship weight function across layers, and avg-pool and max-pool refer to the mean and maximum pooling activities.

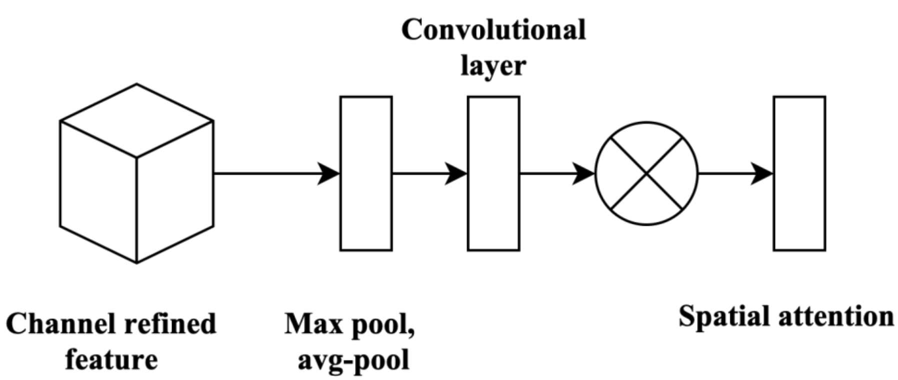


Figure 3: Spatial Attention Module

After the channel MSAM component, the research presents the spatial MSAM modules illustrated in Figure 3 to emphasize the relevant locations of items. Analogous to channel focus, the spatial MSAM method condenses streams and executes average and maximum pooling along the channel scale. The input characteristic layer computes the highest and median values for every characteristic point across the channel in the language learning class. The two outcomes are combined, and the channel count is modified using convolution with a single channel. The sigmoid operation produces the spatial focus map, represented mathematically as

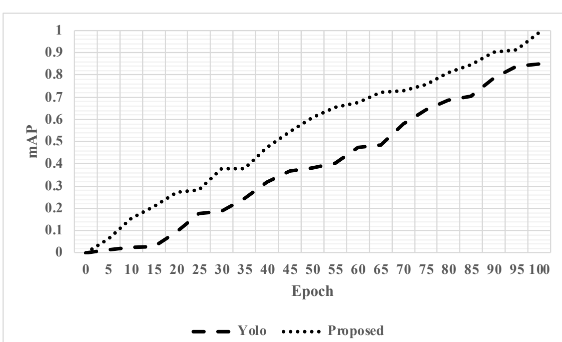
$$M_s(F) = \sigma\big(f^{6\times6}(avg - pool(F), max - pool(F))\big) \quad (2)$$

In the formula, 6 × 6 denotes the dimensions of the convolution kernel, indicating that a 6 x 6 convolution process is executed on the feature mapping. Empirical evidence suggests that a 6 x 6 convolutional kernel outperforms a 3 x 3 convolution kernel. This research incorporates the MSAM system across the neck and prediction networks. Following the incorporation of the MSAM system, the method prioritizes pupil targets within the learning environment, disregarding extraneous elements in language learning class. Utilizing the saliency of items enhances the acquisition of pupil target attributes and increases detection precision through wireless networks.

4 Results

The research has established a strong hardware and software environment for the tests. The hardware configuration comprises a server featuring 32-core Processors and a solitary Graphical Processing Unit (GPU), providing adequate computing capacity for demanding DL applications. The studies utilize PyTorch, Python, and CUDA to capitalize on recent improvements and assist DL systems. In the training stage, the research uses the AdamW optimization, a variation of the Adam optimization process incorporating decoupled weight decay. This decision is based on AdamW's shown efficacy in managing sparse slopes and mitigating overfitting in intricate models. The curriculum is set with a batch size of 2, a rate of learning of 0.0003, and a decay of weight coefficient of 0.0002. These parameters are carefully selected to balance training rate and convergence security, maximizing model variable adjustments in language learning classes.

The comparative outcomes of the average accuracy value curve between the enhanced network framework presented in this paper and the YOLOv5 version are as follows: In Figure 4(a), mean Absolute Precision (mAP) denotes the mean preciseness at an Intersection of Union (IoU) threshold of 0.5, Figure 4(b) indicates the mAP across IoU thresholds ranging from 0.5 to 0.9. The result illustrates a substantial enhancement in the median precision of this method relative to YOLOv5. Upon reaching the 70th iteration, the technique presented in this study achieves a mAP=0.5 of approximately 0.81, then increasing consistently to around 0.82, whereas the algorithm used in YOLOv5 completes 65 iterations. The mAP=0.5 increased to 0.82 during the round and ultimately stabilized at approximately 0.85.



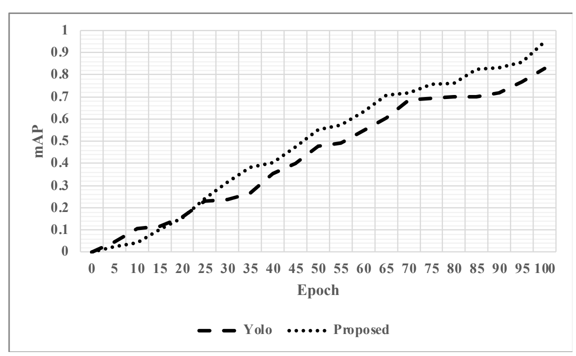


Figure 4: Map Analysis (a) 0.5 (b) 0.5 to 0.9

Figure 5 illustrates that the method exhibited strong performance in the educational setting of a behavior recognition database, with a mAP=0.5 of 88.5 and a median recall rate of 91.3. The detection precision for sitting down is elevated due to its more prominent picture area, achieving an accuracy rate of 94.1 in language learning class. The precision in identifying hearing and laying down actions is inferior to other actions due to the reduced area covered by the head.

In occluded settings with several student objects disrupting, the YOLOv5 model fails to eliminate redundant recognition frames, resulting in two detection sessions for the obscured objects. This study introduces DIoU to enhance minimal suppression by calculating the distance among the center locations in consecutive target detection structures, enabling better target predictions and eliminating redundant detecting frames through wireless networks.

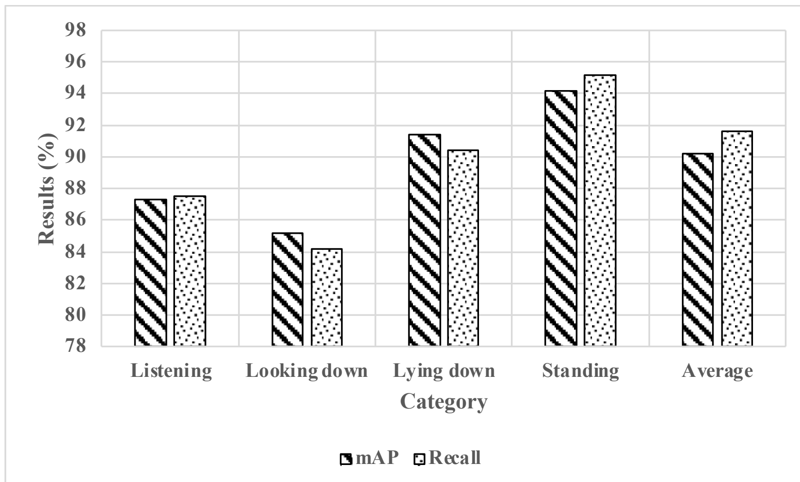


Figure 5: Precision and Recall Analysis

This research objectively assesses the algorithm's benefits by comparing it with conventional detection techniques. To guarantee equity, every algorithm in the study utilizes identical training variables and data, with the experimental findings in Figure 6. The mean precision of the enhanced method presented in this research surpasses that of the other three techniques in the language learning class. Due to the enhanced feature pyramid architecture and the incorporation of the MSAM system, the method's detection rate is marginally inferior to that of the YOLOv5 approach; however, it can still process 35 frames per second, successfully meeting the objective of real-time behavior in student identification.

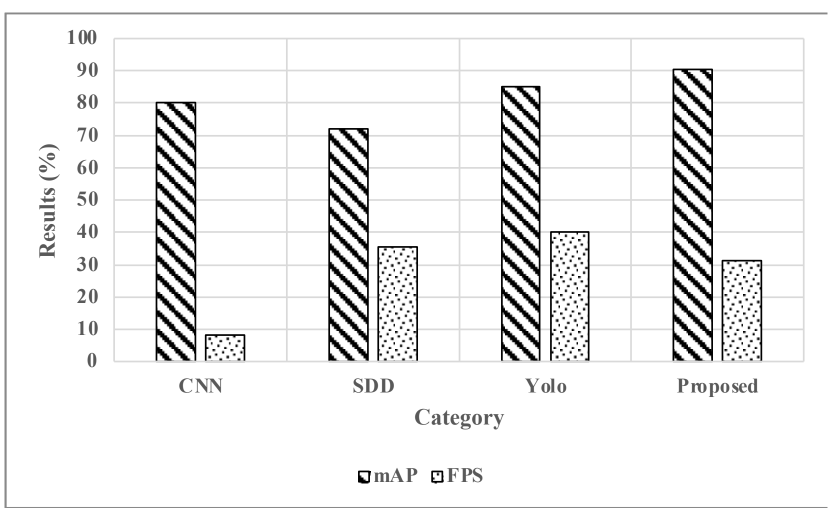


Figure 6: Detection Analysis

5 Conclusion

This study identifies students' classroom conduct utilizing the enhanced YOLOv5 method and reframes the issue of classroom behavior identification as a fine-grained classification challenge of behavioral photographs. The weighted bidirectional feature pyramid networks (BiFPN) facilitate an adequate representation of multi-scale object characteristics, extracting the nuanced aspects of object actions and diminishing the missed findings of small target students situated in the back part of the learning environment through wireless networks. Integrating the mechanism for MSAM into the YOLOv5 method enhances the prominence of the object area within complicated backgrounds, significantly augmenting the detector's precision. Additionally, the application of DIoU refines the original non-maximum repression, further enhancing the detection rate while addressing the issue of missed detections arising from classroom obstruction. The experimental findings indicate that the improved algorithm presented in this research can precisely identify various language learning classroom actions and demonstrates superior accuracy compared to the YOLOv5 method in crowded classroom environments and occluded scenarios, exhibiting enhanced accuracy and resilience. The mAP levels of the proposed approach surpass those of Faster R-CNN, SSD, and the original YOLOv5 method, signifying superior detection accuracy. The suggested approach attains an inference speed of 33 frames per second, satisfying the criteria for real-time video recognition in language learning classes. Despite the promising results, the study is constrained by the following factors. Initially, students' behaviors frequently manifest cohesively, and acquiring video footage from real schools proves challenging. It is difficult to account for the temporal context of behavioral occurrences to enhance accuracy through wireless networks. The research did not account for the temporal variables preceding and after each behavior model; instead, the study concentrated on the behavior categorization issue inside single-frame pictures. In contrast to sprinting and jumping, which require many frames to contain data, the actions exhibited by pupils in educational settings are frequently immobile and are assessed using single-frame photographs.

References

[1] Abdullah, D. (2024). Recent advancements in nanoengineering for biomedical applications: A comprehensive review. *Innovative Reviews in Engineering and Science, 1*(1), 1-5. https://doi.org/10.31838/INES/01.01.01

[2] Balamurugan, S., Manikandan, R., & Manigandan, R. (2017). Modular Multilevel Cascaded Inverter Based on Capacitor Voltage Balancing Control using Carrier Phase Shifted Sinusoidal Pulse Width Modulation Technique. *International Journal of Advances in Engineering and Emerging Technology*, *8*(4), 20-30.

[3] Bisrat, A., Minda, D., Assamnew, B., Abebe, B., & Abegaz, T. (2021). Implementation challenges and perception of care providers on Electronic Medical Records at St. Paul's and Ayder Hospitals, Ethiopia. *BMC medical informatics and decision making, 21*, 1-12.

[4] Chakma, K. S. (2025). Flexible and wearable electronics: Innovations, challenges, and future prospects. *Progress in Electronics and Communication Engineering, 2*(2), 41–46.

[5] Demirel, H. O., Ahmed, S., & Duffy, V. G. (2022). Digital human modeling: a review and reappraisal of origins, present, and expected future methods for representing humans computationally. *International Journal of Human–Computer Interaction, 38*(10), 897-937. https://doi.org/10.1080/10447318.2021.1976507

[6] Ferdowsi, S. K., & Moradi, A. (2014). Investigating of the effects of using conceptual plans in improving science lesson learning for male students of sixth grade in primary schools of Ferdows Town. *International Academic Journal of Innovative Research, 1*(1), 21–29.

[7] Gallagher, J. E., & Oughton, E. J. (2025). Surveying You Only Look Once (YOLO) Multispectral Object Detection Advancements, Applications and Challenges. *IEEE Access*. https://doi.org/10.1109/ACCESS.2025.3526458

[8] Hossain, M. S., Shahriar, G. M., Syeed, M. M., Uddin, M. F., Hasan, M., Shivam, S., & Advani, S. (2023). Region of interest (ROI) selection using vision transformer for automatic analysis using whole slide images. *Scientific Reports*, *13*(1), 11314.

[9] Kattenborn, T., Leitloff, J., Schiefer, F., & Hinz, S. (2021). Review on Convolutional Neural Networks (CNN) in vegetation remote sensing. *ISPRS journal of photogrammetry and remote sensing*, *173*, 24-49. https://doi.org/10.1016/j.isprsjprs.2020.12.010

[10] Kurbanazarova, N., Shavkidinova, D., Khaydarov, M., Mukhitdinova, N., Khudoymurodova, K., Toshniyozova, D., ... & Alimova, R. (2024). Development of Speech Recognition in Wireless Mobile Networks for An Intelligent Learning System in Language Education. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, *15*(3), 298-311. https://doi.org/10.58346/JOWUA.2024.I3.020

[11] Sarker, I. H. (2021). Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions. *SN computer science*, *2*(6), 1-20.

[12] Surendheran, A. R., & Prashanth, K. (2015). A survey of energy-efficient communication protocol for wireless sensor networks. *International Journal of Communication and Computer Technologies*, *3 (2), 50, 57*. https://doi.org/10.31838/IJCCTS/03.02.01

[13] Vijaya Kumar, D. T. T., & Mahammad Shafi, R. (2023). A fast feature selection technique for real-time face detection using hybrid optimized region based convolutional neural network. *Multimedia Tools and Applications*, *82*(9), 13719-13732.

[14] Wen, Y. (2021). Augmented reality enhanced cognitive engagement: Designing classroom-based collaborative learning activities for young language learners. *Educational Technology Research and Development*, *69*(2), 843-860.

[15] Yi, D., Su, J., & Chen, W. H. (2021). Probabilistic faster R-CNN with stochastic region proposing: Towards object detection and recognition in remote sensing imagery. *Neurocomputing*, *459*, 290-301. https://doi.org/10.1016/j.neucom.2021.06.072

[16] Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., ... & Li, Y. (2021). A Review of Artificial Intelligence (AI) in Education from 2010 to 2020. *Complexity*, *2021*(1), 8812542. https://doi.org/10.1155/2021/8812542

Authors Biography

Salima Rustamiy is a highly accomplished Doctor of Philology and Professor at Oriental University, Uzbekistan. With extensive experience in linguistic research and teaching, she has made significant contributions to the field of philology. Her research interests encompass language structure, literature analysis, and cross-cultural linguistics. She has published numerous papers in national and international journals.

Nargis Kurbanazarova is an esteemed Associate Professor specializing in Uzbek Linguistics at Termez State University. With a deep interest in the linguistic heritage of Uzbekistan, she has been actively involved in research and academia. She has authored several research articles, contributing to the preservation and advancement of Uzbek linguistic studies.

Gulbakhor Atayeva is an Associate Professor and Candidate of Philological Sciences at Samarkand State University named after Shar of Rashidov. Her academic work focuses on English linguistics and its application in multidisciplinary contexts. She is known for her research publications that bridge linguistic theory and practical usage.

Dildora Khashimova is a Doctor of Science (DSc) and a Professor at the Tashkent State University of Law. Her work primarily revolves around Uzbek language and literature, with a focus on language policy and literary analysis. She has contributed significantly to Uzbek linguistics through her research and publications.

Mohinakhon Khamidova is a dedicated researcher at Kimyo International University in Tashkent. Her academic interests include linguistics and intercultural communication. She has authored various research papers addressing contemporary linguistic challenges.

Mohira K. Usmanova is a Doctor of Philosophy in Philological Sciences and a Dotsent (Associate Professor) at Karshi State University. Her research encompasses philology, linguistics, and the evolution of language. She has presented her work at numerous conferences and published in reputable journals.

Feruza Mamatova is an accomplished Associate Professor at the National University of Uzbekistan named after Mirzo Ulugbek. She holds a PhD in Philological Sciences with expertise in English linguistics and language teaching methodologies. Her research focuses on linguistic diversity and pedagogy.

Dildora Avezova is a promising PhD student at the National University of Uzbekistan. Her research interests lie in linguistics and language education. Despite being in the early stages of her academic career, she has already contributed to research forums and published insightful papers.