IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 31 December 2023, accepted 25 January 2024, date of publication 30 January 2024, date of current version 14 February 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3359689

RESEARCH ARTICLE

Effectiveness of Pre-Trained CNN Networks for Detecting Abnormal Activities in Online Exams

MUHAMMAD RAMZAN[1,3], ADNAN ABID[2,3], (Senior Member, IEEE), MUHAMMAD BILAL[4], KHALID M. AAMIR[5], SUFYAN A. MEMON[6], AND TAE-SUN CHUNG[7]

[1]Department of Software Engineering, Faculty of Computing and Information Technology, University of Sargodha, Sargodha 40100, Pakistan
[2]Department of Data Science, Faculty of Computing and Information Technology, University of the Punjab, Lahore 54000, Pakistan
[3]Department of Computer Science, University of Management and Technology, Lahore 54770, Pakistan
[4]Department of Computing and Information Systems, School of Engineering and Technology, Sunway University, Petaling Jaya, Selangor 47500, Malaysia
[5]Department of Information Technology, Faculty of Computing and Information Technology, University of Sargodha, Sargodha 40100, Pakistan
[6]Department of Defense System Engineering, Sejong University, Seoul 05006, South Korea
[7]Department of Artificial Intelligence, Ajou University, Suwon-si 16499, South Korea

Corresponding authors: Muhammad Bilal (muhammadb@sunway.edu.my) and Tae-Sun Chung (tschung@ajou.ac.kr)

This work was supported in part by the Institute of Information and Communications Technology Planning and Evaluation (IITP) under the Artificial Intelligence Convergence Innovation Human Resources Development under Grant IITP-2023-RS-2023-00255968, and in part by the Information Technology Research Center (ITRC) Support Program funded by the Korean Government [Ministry of Science and ICT (MSIT)] under Grant IITP-2021-0-02051.

ABSTRACT Online exams are growing increasingly popular in organizations and educational institutes because they are more flexible and cost-effective than conventional paper-based exams. When face-to-face exams are not possible, such as during floods, unexpected situations, or pandemics like COVID-19, this exam mod has become even more popular and important. However, online exams may have difficulties, such as the need for a reliable internet connection and the possibility of cheating. Because there is no human supervisor present to monitor the exam, so cheating is a major concern. The environment employed for the online exams ensures that every student finalizes the evaluation process without using any type of cheating. This study investigates the detection and recognition of unusual behavior in an academic setting, such as online exams, to prevent students from cheating or engaging in unethical behavior. After consulting with experts and reviewing the online exam held in Covid-19 and other online exams, selected the four most common cheating activities found in the online exam. The study extracts key frames using motion-based frame extraction techniques before employing advanced deep learning techniques with various convolutional neural network configurations. This study presents several deep learning-based models that analyze the video exam to classify four categories of cheating. This method extracts key frames from a video sequence/stream based on human motion. This research developed a real dataset of cheating behaviours and conducted comprehensive experiments with pre-trained and suggested deep-learning models. When evaluated using standard performance criteria, the YOLOv5 model outperforms other pre-trained and fine-tuned approaches for detecting unusual activity.

INDEX TERMS Online exams, abnormal activities, cheating, computer vision, deep learning, CNN.

I. INTRODUCTION

In this decade, information and communication technologies have advanced quickly and directly impacted people's lives, particularly in education. The most popular way to evaluate student performance is through exams. Exams, on the other hand, can be classified into two categories: conventional and online. A traditional exam is described as a collection of questions given out in class. Every student's exam is generated using fixed questions, and students have a set amount of time to begin and complete it. On the other hand, an online exam is a sort of evaluation that is given by the internet or intranet.

The associate editor coordinating the review of this manuscript and approving it for publication was Francisco J. Garcia-Penalvo.

Using a computer or another internet-capable device, such as a smartphone or tablet, students or candidates take the exam. It is a complex task that requires the integration of various technologies and methodologies to monitor and evaluate student behavior to detect abnormal activities during online exams. A comprehensive approach incorporates several techniques, including eye tracking, facial recognition, mouse and click pattern monitoring. Students reply to exam questions using a computer or other device, and the exam questions are given electronically. Online programs have become a viable educational option. For example, during the COVID-19 epidemic, this platform was progressively recognized in colleges and higher education institutions like universities and even adopted in elementary schools. Maintaining a safe and fair testing environment is one of the numerous issues associated with online exams. These exams are remote, so it is challenging to stop cheating or getting illegal assistance. Ensuring the integrity of exam questions, keeping an eye on and authenticating students' identities, and preventing technological malpractice like the use of unauthorized equipment or external resources all become challenging tasks. However, because of the disconnected structure of online education, there are problems with the likelihood of academic dishonesty, especially when students take exams from remote areas without the disciplinary procedures generally used at exam halls [1], [2]. Online education is rapidly expanding in student enrollment. Coursera is an e-learning system that provides diverse courses, certificates, and degree programs from world-renowned colleges and institutes. Where 19 million users were registered in 2022. However, existing research suggests that online cheating is common, involving academic dishonesty on the part of both instructors and students [3], [4] Although online education offers excellent learning opportunities for people unable to attend traditional, high-quality schools due to scheduling conflicts or other physical limitations. Still, if academic dishonesty problems are not addressed, its legitimacy may be at risk. Cheating and unusual activities such as impression, whispering, and hand connection are common in offline and online examinations worldwide, compromising the dignity and morality of fair examination administration.

The study [5] aimed to investigate academic cheating practices and the perceived impact of online learning on academic performance among Pakistani students enrolled in high schools, colleges, and universities during COVID-19. According to the findings, 30% of students acknowledged cheating on at least one online exam, while 60% of students admitted to cheating frequently. According to a survey and study by Dr. Donald McCabe [6] International Center for Educational Honesty, which has researched academic dishonesty patterns for over a decade, about 68 % of undergraduate students confess to cheating on tests or written work. Data has been gathered [7] for all students who were involved in cheating activities at Virtual University Pakistan over 6 years, including 13 semesters, using a computer-based examination system. The number of cases and modes of cheating used by students involved in unfair means were discussed in this study.

Although online courses grew in popularity during the COVID-19 epidemic, they should not be regarded as a standard means of education without careful consideration in the case of a pandemic revival. For instance, e-cheating has been documented, and the disengaged atmosphere of online classes has led to serious problems. As a result, the reliability of online courses could be questioned.

This study presents a deep learning-based framework for detecting abnormal behavior during exams, such as cheating and misconduct. The proposed method extracts keyframes from a video sequence or stream based on human motion to analyze students' unusual online behavior. Deep learning models, pre-trained models, including YOLOv5, and a suggested CNN are utilized for the detection of cheating activities. As a result, the following are the primary research contributions:

- The dataset has been collected from the Universities for online examinations during COVID-19. The datasets include where students are involved during the following activities: a) By using an external device, b) Head movement, c) Multiple people, d) Talking to others, and e) normal exams.
- Expert annotators have labeled the data as part of the data set processing. The dataset is available on request for educational purposes.
- Selecting only the most important frames from the video sequence using a motion-based keyframe extraction approach.
- Different Pre-trained models including YOLOV5 and Convolutional Neural Network (CNN) models are utilized for online examination to improve the accuracy and efficiency of cheating detection in online environments

The rest of the paper is organized as follows: Related work is addressed in section II. Section III describes the proposed method for identifying online cheating activity. Section IV analyses the results, and Section V comes to a conclusion and discusses future studies.

II. BACKGROUND

This section divides the existing literature work into three parts. First started with cheating activities in physical exams, then moved on to online exam cheating, and finally in those studies where both online and physical exams were taken using machine learning, deep learning, and hybrid models.

Ramzan et al. [8] identified and recognized anomalous activities of students in exam rooms, to assist invigilators in observing students from cheating or employing unfair techniques. Their research compared the effectiveness of several CNN networks that have been used to look for unusual physical exam activity. It extracts keyframes using motion-based frame extraction approaches before applying the deep learning-based technique in various settings. According to the performance methods, this model improved the previously described strategies for abnormal activity

recognition. To distinguish each form of anomalous behavior, previous research in this field relied exclusively on handcrafted features, hard-coded algorithms, and computer vision-based. Senthilkumar et al. [9] developed a mechanism for recognizing abnormal behavior in the lecture hall. The framework comprises three elements that collaborate to control student behavior throughout the analysis exam. The facial area of the student is detected initially, followed by the student's hand contact and signal. The authors [10] identified the face area using Haar features and proposed a system for monitoring student behavior during tests. They also identified and warn of cheating practices such as hand touch and hand messaging. A goal was to detect suspicious behavior during offline academic assessments. Three components have been used in this study for identifying facial malpractices and testing using a PCA-based face recognition technique such as students speaking with one another, and recognizing illegal objects or equipment.

The proposed system would immediately alert the students if it notice any dubious behavior. This work [4] suggested a mechanism for analyzing and recognizing student activity from videos captured by video surveillance during the test. Automatic video surveillance is an excellent method for monitoring students and identifying unusual conduct. Students could exchange papers and look at one another's answer sheets to get copies of answers from them or copies of answers from secret sources including handwritten answers on hands, cell phones, summarize papers, and textbooks.

This paper [11] provided a method for eliminating a proctor's physical presence during the exam by using a multimodal framework. Using a camera and an active window capture, recorded video. The test-face takers were identified and examined to predict their actions. His feature points are determined to identify his head position. A cell phone, a notebook, or the presence of other people from the various things that have been detected. A rule-based inference system that has ascertained whether exam fraud occurred was produced by the integration of these models.

The suggested [12] detection of laughter, eye gaze tracking to establish the applicant's direction of glance, eyes blinking/close duration, and head activity/head position detection are all included in the testing monitoring approach. Artificial intelligence models have been used in the task to categorize applicant activities. Exams that were remotely proctored cut down on logistical work, sped up evaluations, and made it simpler to connect with test-takers who were far away.

This study [13] proposed an intelligent system that automatically identifies cheating in online exams. The students involved in cheating activities were identified based on examining their eye-gaze and head pose.

In this method [2], four data sets are extracted from an exam video as part of the method the authors suggested for identifying cheating behaviors. These data sets are then put into a trained classification system. By converting each video into a multivariate time series that displays the time-varying event data retrieved from each video frame, they may have used the cheating activities detection method as a multivariate time-series classification problem.

This method [14] used an e-cheating intelligence source to detect online cheating techniques, consisting of two primary components: the internet protocol analyzer and the behavior analyzer. The intelligence agent identified the actions of the students, recognized them, and prevented any undesirable activity. It has been used in a course assessment to assign randomized multiple-choice questions and can be connected with online learning programs to track students' behavior. The method has been evaluated by testing it on various data sets. This research [15] described an automated system for proctoring online exams using multimedia analytics. One microphone, a webcam, and a wearable camera are part of the system hardware used to record the audio and visual environment of the testing site. One of the six essential elements that continuously estimate the primary behavioral indications is user identification, followed by text recognition, voice tracking, operating window detection, observation, and mobile phone recognition. They were able to develop higher-level characteristics to define whether the test-taker is cheating at any time during the exam by combining the continuous estimating components and using a temporal sliding window in this work.

This study [16] aimed to assess the types, frequency, and variability of electronic cheating among technical university students, their views regarding different examination practices, and instructor attitudes against cheating as reported by students. Traditional monitoring methods focus on tester identification and lack efficient detection of anomalous tester actions. This research provided a solution to the challenge of identifying abnormal examiner conduct in online tests: use a camera to capture the examinee's head position and mouth condition as well as identify the examiner's improper behavior during the test [17].

Fayyoumi et al. [18] validated students for online tests using facial recognition, and the main goal was to provide a solution for online test systems. More importantly, throughout the test, the system constantly verifies that the student who began the exam is the same one who finished it, preventing cheating by peering at other PCs or reading from another paper.

This study [19] examined two basic models including physical and online exams. According to the claim, cheating online should be recognized. Due to a growth in the quantity of exam cheating information queries made in Spain during the period immediately preceding the pandemic, the authors [20], employed the newly developed research technique of search engine data analysis. The findings suggested that the use of Internet data analytics as an approach for academic integrity research should be expanded.

The study [21] developed an online assessment framework based on the unique contextual characteristics of South African institutions. Universities can implement the initial set

of suggestions for implementing online examinations due to the recommended approach. The study discussed [22], the classification methods for cheating online exams. To have a thorough understanding of cheating avoidance, prevention, and detection, educators and researchers working in the field of online learning may find the study to be a useful resource.

This approach [23] was based on online examinations, which are recorded by distance learning platforms. To calculate the student's risk of cheating, tailored IP geolocation and additional data. In 22 courses with around 3600 students, where the partial or final online tests were not invigilated, the strategy. Examples of the identified cheating were shown together with the discovered cheating risk ratings. The technique could be used to choose students for knowledge re-validation or to compare instances of student cheating across institutions, nations, age groups, and courses. The likelihood of students cheating was examined over four academic terms including the two when the institution was closed due to COVID-19. In this paper [24], a deep feature extraction approach for suspicious activity identification was proposed and 63 layers-based CNN-centered deep architecture was proposed for feature acquisition for this reason. A feature selection method optimizes the collected deep features. The objective of this study [25] was to automatically identify and classify candidates in real-time videos based on their behavior in the examination environment. The results were obtained utilizing their own datasets and ensemble learning approaches based on deep learning. The prohibited items and suspicious head motions have been observed using a modified Faster RCNN algorithm. To do this, the region of interest (ROI) of detected objects, as well as the intersection over union (IOU) between the posture points for the hands and the face, have been identified. the face ROI's features were recorded after the final convolution layer.

Using Machine Learning (ML) techniques, [26] presented a novel method for detecting potential exam cheating situations. It collects information about general behavior, student attendance, and academic performance. The dataset has been used in research on student behavior and performance to create models for forecasting academic success, identifying at-risk students, and detecting problematic behavior. The proposed model employed a long short-term memory (LSTM) strategy with a dropout layer, thick layers, and an optimizer named Adam, outperforming all previous three-reference efforts by 90%. Accuracy is said to improve as more complex, optimized architecture and hyperparameters are used. To find students who were cheating online, this study employed a data-driven methodology [27]. The K-means clustering method is used to first identify student clusters based on three different types of information. The results point to the existence of two distinct clusters, one of which possesses characteristics that are strongly associated with online cheating and the other of which is made up of morally upright student behavior. The results of the clustering analysis are supported by the personality data.

Based on the literature review we concluded that limited work has been done in the domain of abnormal activities detection of students in the online exam. One significant problem is a shortage of labeled datasets containing various cheating behaviors, which limits the model's ability to generalize effectively. The existing system that has been used for the detection the cheating activities mainly depends on the technology, tools, and platform used. It is challenging to keep an eye on software or hardware due to restricted system settings. Or it is difficult to identify clever cheating. Collaboration during an online exam might be challenging, even with available communication tools. Sometimes female students may have privacy concerns while using monitoring tools and webcam recording for cheating detection. It's critical and takes careful consideration to strike a balance between protecting student privacy and maintaining exam integrity. It can be difficult to obtain full and reliable data on online cheating actions. Individuals may be unwilling to declare their involvement due to the restricted nature of cheating behaviors.

III. PROPOSED METHODOLOGY

This section has described the proposed technique for identifying cheating behaviors during an online exam using deep learning-based models. Fig. 1 shows the basic steps of the technique and a generic framework of the proposed system. We utilized deep learning-based fine-tuned CNN and pre-trained models for the detection of online exam cheating activities. The basic steps of the proposed methodology are discussed in the following sections.

A. DATASETS

The classification model's performance is influenced by the quality of data used in the learning process. This study created a video dataset for detecting unusual behavior in online exams. There is no standardized dataset for academic examination invigilation to detect cheating during an online exam. The dataset consists of student exam cheating behaviors. Gathering relevant videos for the classes is the first step in creating a video dataset. Following the preprocessing of the videos, the next step is to separate them into separate classes so that they can be implemented. To obtain the real video datasets, we requested videos of the online exam during COVID-19 where students were engaging in cheating activities from the university. The Student Online Cheating Activity (S_OCA) dataset categorizes activities as normal or abnormal. The abnormal cheating activity was detected through four classes: external devices, head movement, multiple persons, and talking to others.

- Data Preparation

In academia, two types of examinations are considered: physical and online. As a result, the analysis of students' unusual activity during the Examination may be physical or online, but only online exam activities are taken in this research work. During an online exam, abnormal behavior can include

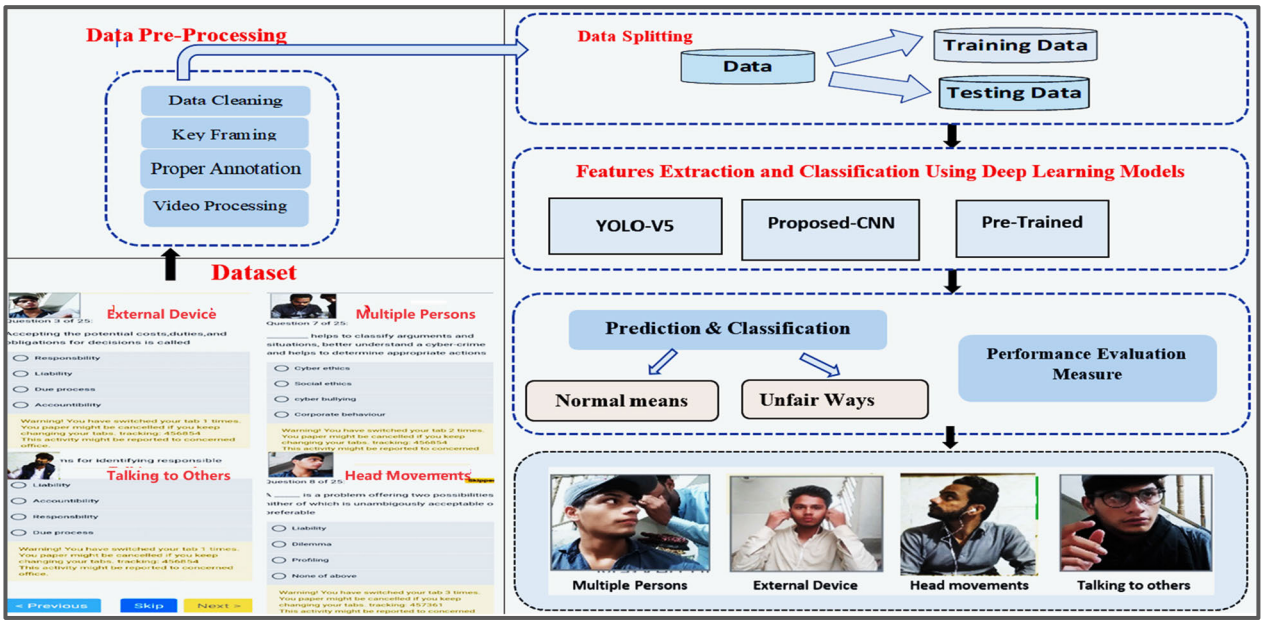


FIGURE 1. The basic flow diagram for cheating activities detection in online exams using deep learning.

irregular answer patterns, multiple logins, suspicious access patterns, copying and pasting from outside sources accessing unauthorized websites, Multiple videos of online Examinations from some universities collaborating with outside help, and using messaging, and audio chat. Academic integrity policies are common at educational institutions and outline what is and is not permitted during tests. To maintain regularity and fairness, atypical actions may be chosen following these principles and instructions. After consulting with experts and reviewing the real-time online exam held in COVID-19 and other online exams, cheating activities were selected. Sample images of online and physical are shown in Fig. 2.

Fig. 3 depicts the four most common types of cheating activities discovered in online exams. The scope has been defined from the perspective of the front camera and video camera. We also did not consider software-based cheating; instead, we only considered student involvement. The following steps were involved in preparing quality content in the dataset of cheating activities for the online exam.

- Exam type selection
- Collection of Videos
- Preprocessing
- Divide into classes

These exam videos are collected on a special request for research purposes from the University. The validation of the dataset is the next step in the preparation process. The total number of classes, number of videos, and number of images used in the prepared dataset are discussed in Table 1.

The videos are preprocessed and ready for training the classification model in this task. The brightness, hue, and saturation of the video are then adjusted. Quality, lightness,

TABLE 1. Statistics of the S_OCA dataset.

Type of Activities	CLASS LABELS	No of videos	No. of Keyframes	Training (70 %)	Testing (30 %)
Use of the external device	0	15	440	308	132
Head movement	1	12	480	336	144
Multiple person	2	13	434	304	130
Talking to other	3	12	373	261	112

variation, saturation, shade, and quantity are defined for this dataset by a few quality metrics. The data set contains 52 videos divided into four categories of unusual activities with the help of an expert from academia in the university. Parameters for the dataset preparation have been discussed in Table 2.

This dataset will be accessible for use in research. When publicly available, information about datasets' descriptions, sizes, numbers of classes, labeling procedures, various scenarios in which data is captured, terms for using the data, relevant resources for data access, ethical considerations for data uses, limitations, challenges, and citations for the data are provided. Ensure that the dataset is representative and diverse by correcting for imbalances in class distributions using techniques such as oversampling or undersampling. The type and format of the data, as well as ensuring consistency, cleanliness, and accurate labeling, are all factors to consider when preparing datasets for classifying online exam cheating activities. The data's size and diversity, ensure that it is large enough to encompass the various degrees of cheating behaviors, diverse environments, candidates, and so on. Aim for a well-balanced, representative dataset that

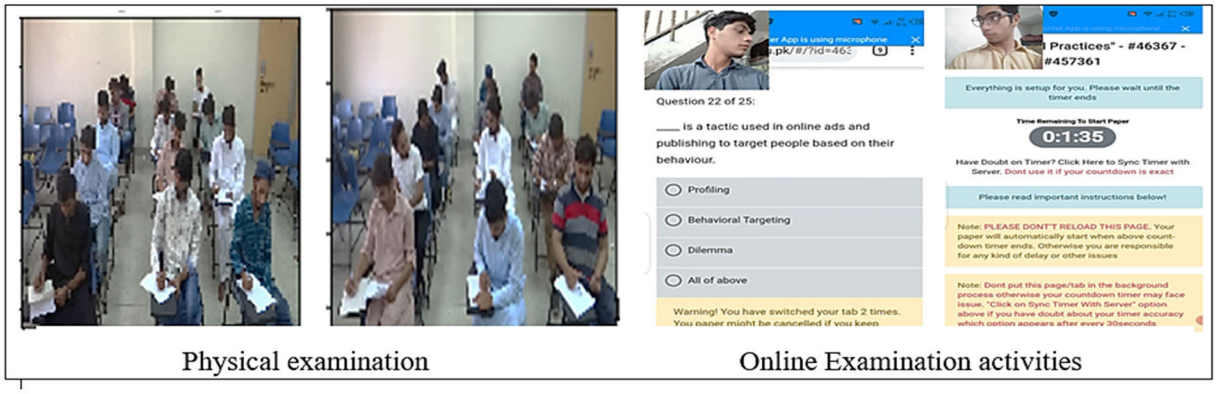


FIGURE 2. Sample images from two types of examination physical and online.

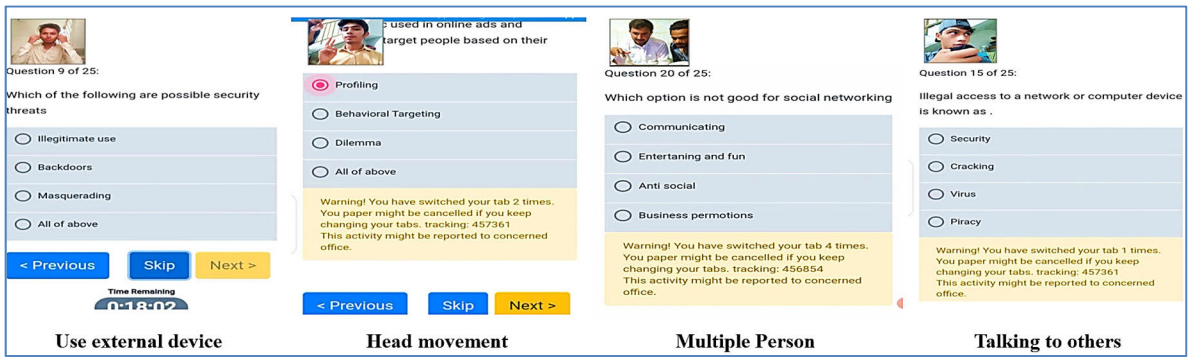


FIGURE 3. Sample frames from S_OCA dataset online examination of four classes.

reflects the target population and scenario. The data has been obtained from the University and also with the candidates' explicit consent and awareness while respecting their privacy and confidentiality. Additionally, during data collection and usage, ensure compliance with applicable laws and regulations regarding data protection and security.

B. VIDEO PRE-PROCESSING

Long-duration videos are taken and then cut into frames. The Gaussian filter is used in preprocessing to remove noise from video frames, and video frames are given to histogram equalization. For the online exam, we resize the datasets (224,224,3).

1) KEY FRAME EXTRACTION

In the Keyframing redundant frames have been removed and extracted only those frames containing an unusual series of actions in a video sequence. Each video in this dataset is composed of a 30-frame-per-second sequence of frames. Due to the similarity of the frames, a video frame's information is highly duplicated, needing only a few frames having relevant information. Histogram, color histogram, frame difference, correlation, entropy difference, and other keyframe extraction techniques are available. First, sample all the videos with a three-frame skip factor for crucial frame extraction. The skipping factor helps get rid of unnecessary frames. A motion-based key frame extraction technique was used [23]. This technique uses the absolute difference in pixels between two successive frames. All video frames are retrieved and utilized for training purposes. The model's complexity and computing cost grow because of the repeated duplicate frames. The number of frames separated from the video depends upon the threshold value. The keyframe extraction approach is utilized in this study to remove duplicate consecutive frames sssmany frames [28], [29].

The formula T= (mean of absolute difference + standard deviation of absolute difference) is used to get the threshold value.

$$(absdiff)_f = absdiff(cf_{i+1}, Pf_i) \tag{1}$$

In the equation above, *Pfi* stands for the prior and present frames. The average difference of the matrix achieved from the equation is then calculated (1).

$$Avg_{diff} = Avg((absdiff)_f) \tag{2}$$

If the current frame is either chosen as a keyframe or skipped if the Avgdiff value exceeds a predefined threshold (T).

$$KF_i = if \begin{cases} Avgdiff > T \ key\ frames \\ Avgdiff < T \ no\ key\ frames \end{cases} \tag{3}$$

The current frame is used to update the previous frame, and the entire process is repeated. For frame-level classification, our keyframe extraction techniques retrieve 1727 from

TABLE 2. Parameters for the S_OCA dataset.

S.No.	PARAMETERS FOR DATA PREPARATION	Description
1	Annotators/ Examiner	Three examiners have analyzed the cheating activities.
2	Experience of Examiner	The examiner who analyzed the activities had more than 10 years of experience in university for conducting the exam.
3	Qualification of Examiner	Two examiners have Ph.D. degrees, and the 3rd has MS degree.
4	Job description of Examiner	Identity cheating activities, Handling misconduct of students during exams, Exam conducting Results preparation
5	Data sources	University Of Sargodha, CS & IT department.
6	Collection methods	The dataset has been taken from the real online exams conducted during COVID-19, which also includes cheating activities.
7	Data preprocessing	Different pre-processing steps, including keyframes and resizing, have been used.
8	Data quality control measures	For the quality of the dataset, data quality control measures have been considered: Data Cleaning, Data Profiling, Data Standardization, Data Validation, Data Accuracy, and Data Security Measures.
9.	Scope of the camera	scope the front camera and video camera perspective, We did not focus on software end base cheating. We only considered student involvement.

11500 frames; for video-level classification, 52 video clips are obtained. The number of frames extracted is depends on the threshold value. However, if we increase the threshold value, the number of extracted frames is also reduced. In this research study threshold value is set up as 340000, which gives optimal key frame extraction.

2) PRE-TRAINED AND FINE-TUNED DEEP LEARNING MODELS FOR CLASSIFICATION

The suggested CNN with the optimized configuration and pre-trained models are utilized to compare results for detecting and classifying online exams. The following deep learning models have been used.

- Pre-trained YOLOv5 [30]
- Pre-trained models: Incetion_Resnet_v2 [31]
- Pre-trained models: DenseNet121 [32]
- Pre-trained Inception-V3
- Fine-Tuned CNN: Fine-Tuned CNN models with the specific arranged of layers are utilized.

C. PRE-TRAINED YOLOV5 FOR ONLINE EXAM

The YOLOV5 deep learning-based model is widely used in object identification and classification. This model is more accurate and faster, with a lighter and more portable model. The primary detection features are the global receptive field, anchor frame matching, grid division, and multi-semantic fusion. Unlike traditional object recognition algorithms, the YOLO model uses CNN to predict the bounding box directly and the probabilistic likelihood of visual objects, significantly improving detection accuracy. The YOLO model predicts a set of class probabilities and bounding boxes. In order to create anchor boxes in each input image grid, it first separates the complete image into several grids of various sizes. These anchor boxes are generated according to a preset scale and size. Each anchor box predicts the abjectness score, box width, box height, box center offset, and class scores simultaneously, as opposed to a two-stage detector. The main architectural block of the YOLO family includes:

- YOLOv5 Backbone: Backbone is utilized for pre-training purposes. CSP Darknet serves as the backbone for extracting features from images composed of cross-stage partial networks.
- YOLOv5 Neck: It builds a feature pyramids network with PANet for feature aggregation and sends it to the Head for prediction.

YOLOv5 Head: It has layers that provide predictions for object detection based on the anchor boxes. Aside from that, YOLOv5 makes use of the following training methods.

- Activation and optimization: YOLOv5 employs leaky ReLU, sigmoid activation, SGD, and ADAM as optimizers.
- Loss function: it employs binary cross-entropy in conjunction with logit loss.

As a result, YOLO is a one-stage object detector that quickly finds objects from start to finish. YOLO manifests itself in a variety of ways. The main distinction between the architectures is the feature extraction and the convolutional kernel. There are four YOLOV5 architectures available: 1) YOLOV5s, 2) YOLOV5m, 3) YOLOV5l, and 4) YOLOV5x. These architectures differ in terms of feature extraction and convolutional kernel. The YOLOv5 model is also used for unusual activity during the online exam. The challenges for image recognition increase because the student's environment differs during this online exam. We utilized transfer learning to YOLOV5 to evaluate students' unusual behavior.

YOLOv5 has been shown to significantly improve the processing time of deeper networks. This feature become increasingly important as the technology advances to larger datasets and real-time detection. YOLOv5, a large dataset for object identification, segmentation, and labeling, was trained using the Common Objects in Context (COCO) dataset. As a result, YOLOv5 could be used to detect such unusual behavior during the exam. The steps below are essential for preparing the dataset for training.

- label text file creation
- Data Splitting into the training & validation set
- Data.Yaml file creation

D. PRE-TRAINED INCEPTION_ResNet_v2 ARCHITECTURE

The CNN model Inception_ResNet_v2 has 164 layers and requires a 299 by 299 input image. This model was already trained using millions of images from the ImageNet dataset.

TABLE 3. Description of CNN model.

S.No.	NAME OF LAYERS	PARAMETERS
1	Conv2D	Filters used 32,kernel_size is 3, the padding value is same", the activation used is "relu",input_shape is (224,224,3))
2	MaxPooling2D	pool_size is 2
3	Conv2D	Filters is 32, kernel_size is 3, padding used is "same", and activation "relu" is used
4	MaxPooling2D	the pool_size is 2
5	Conv2D	(the filters is 32, the kernel_size is 2, the padding used is "same"and activation is "relu"
6	MaxPooling2D	The pool_size is 2
7	Conv2D	the filters used is 32, kernel_size is 2, the padding is"same" ,and the activation="relu"
8	MaxPooling2D	The pool_size is 2
9	Conv2D	the filters used 32, kernel_size is 2, padding is "same", And the activation is "relu"
10	MaxPooling2D	The pool_size is 2)
11	Conv2D	the filters= size is 32, the kernel_size is 2, padding is "same", and the activation is "relu"
12	MaxPooling2D	The pool_size is 2
13	Dropout	Value is 0.5
14	Conv2D	The filters is 32,the kernel_size is 2,the padding is "same", and the activation is "relu"
15	MaxPooling2D	The pool_size is 2
16	Flatten()	
17	Dense	50, and activation is "relu"
18	Dense	4 and activation is "softmax" the activity_regularizer=l1(0.001)

The model can classify 1000 different types of objects. Other advantages of IRNV2 include the transformation of inception modules into residual inception blocks, the addition of extra inception modules, and the addition of a brand-new type of inception module (inception A) immediately after the stem module.

E. PRE-TRAINED DenseNet121 ARCHITECTURE

DenseNet121 is the third pre-trained model for analyzing cheating activities during online exams. Each layer in DenseNet, a convolutional neural network, is connected to all layers below it. The first layer is linked to the second, third, fourth, and so on levels, while the second layer is linked to the third, fourth, fifth, and so on levels.

F. PRE-TRAINED INCEPTION-V3 ARCHITECTURE

Google introduced Inception-V3, a convolutional neural network (CNN) architecture, in 2015. The Inception-V3 architecture's convolutional layers, pooling layers, and fully connected layers are all assembled sequentially. Parallel convolutions with variable filter sizes are followed by pooling operations and feature map concatenation in a typical Inception module. Because it can collect data at many scales and resolutions, the network can more accurately reflect the world. To boost performance and avoid overfitting, Inception-V3 additionally incorporates batch normalization, dropout, and weight decay.

G. CNN MODEL

CNN model with the combination of the different layers is also utilized on the online exam. In this proposed method, we used 2-hidden layers of CNN, each layer comprising (Convolutional, pooling, and non-linear Leaky ReLU), and fully connected layers comprising (flatten, dense, dropout, dense with SoftMax activation function). Table 3 shows the structure and layers details of CNN model.

IV. RESULT AND DISCUSSION

The results for the pre-trained and suggested CNN model for the online exam are given.

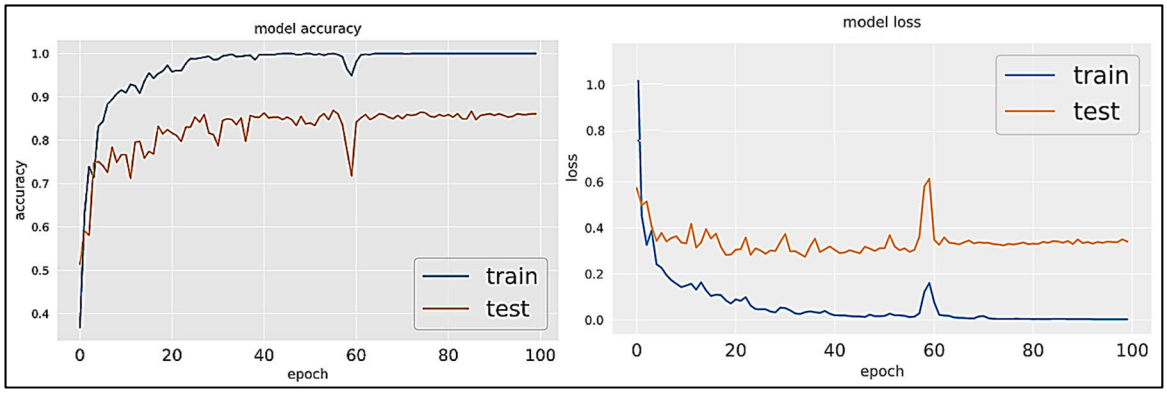


FIGURE 4. Model accuracy and model loss using DenseNet121.

A. PRE-TRAINED MODELS RESULTS

The following pre-trained models, DenseNet121, Inception_ResNet_v2, Inception-V3, and YOLOv5, are utilized to analyze online exam datasets. The model accuracy, loss, precision, recall, and F-measure are shown below.

1) RESULTS OBTAINED USING DenseNet121

The 1st pre-trained model was used to analyze the performance on online exams. The model accuracy and model Loss are shown in Figure 4, Table 4 shows the Classification reports, confusion matrix, and ROC curve, which are shown in Figure 5.

TABLE 4. Classification report using DenseNet121.

	:ECISION	Recall	F1-score	Support
Talking to others	0.93	0.79	0.85	127
External	0.84	0.88	0.86	149
Head-movement	0.84	0.86	0.85	118
Multiple persons	0.86	0.91	0.88	124
Accuracy			**0.86**	**518**

2) RESULTS OBTAINED USING INCEPTION_ResNet_v2

The 2nd pre-trained model was used for the analysis of the performance on online exams. Figure 6 displays the model

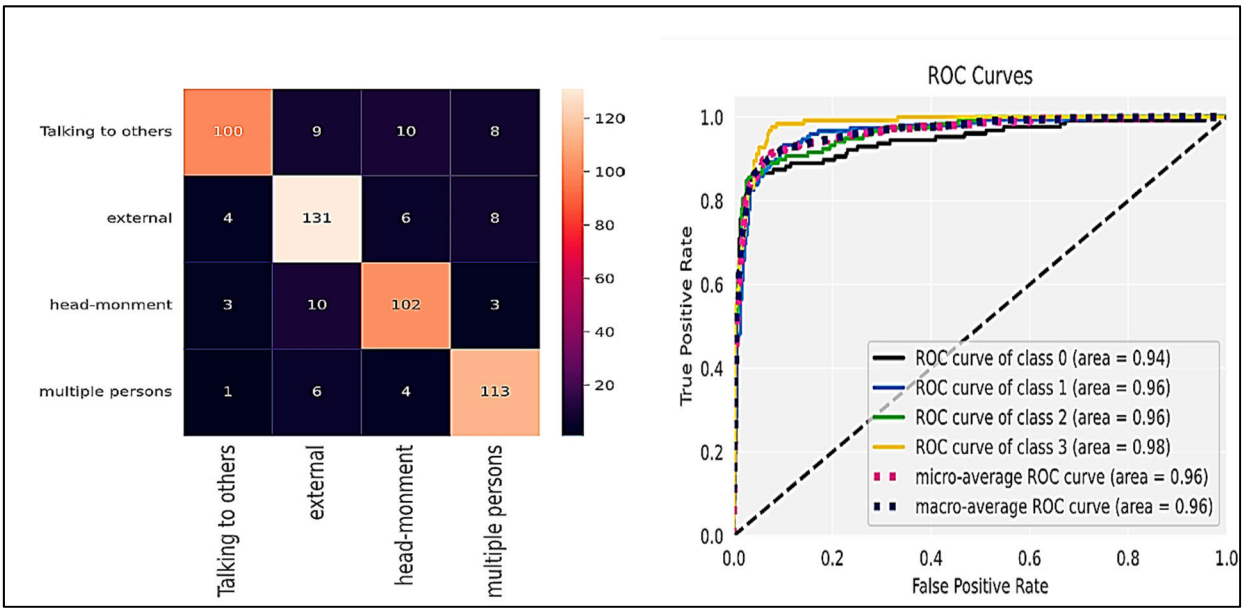


FIGURE 5. Confusion matrix and ROC Curve using DenseNet121.

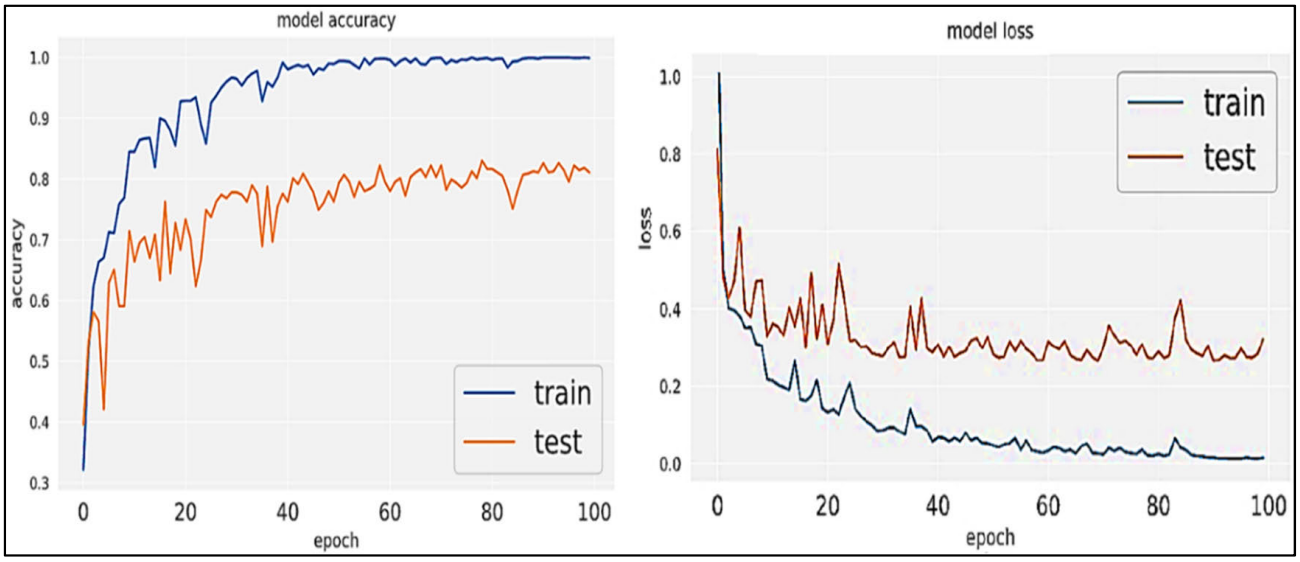


FIGURE 6. Model accuracy and model loss for Inception_Resnet_v2.

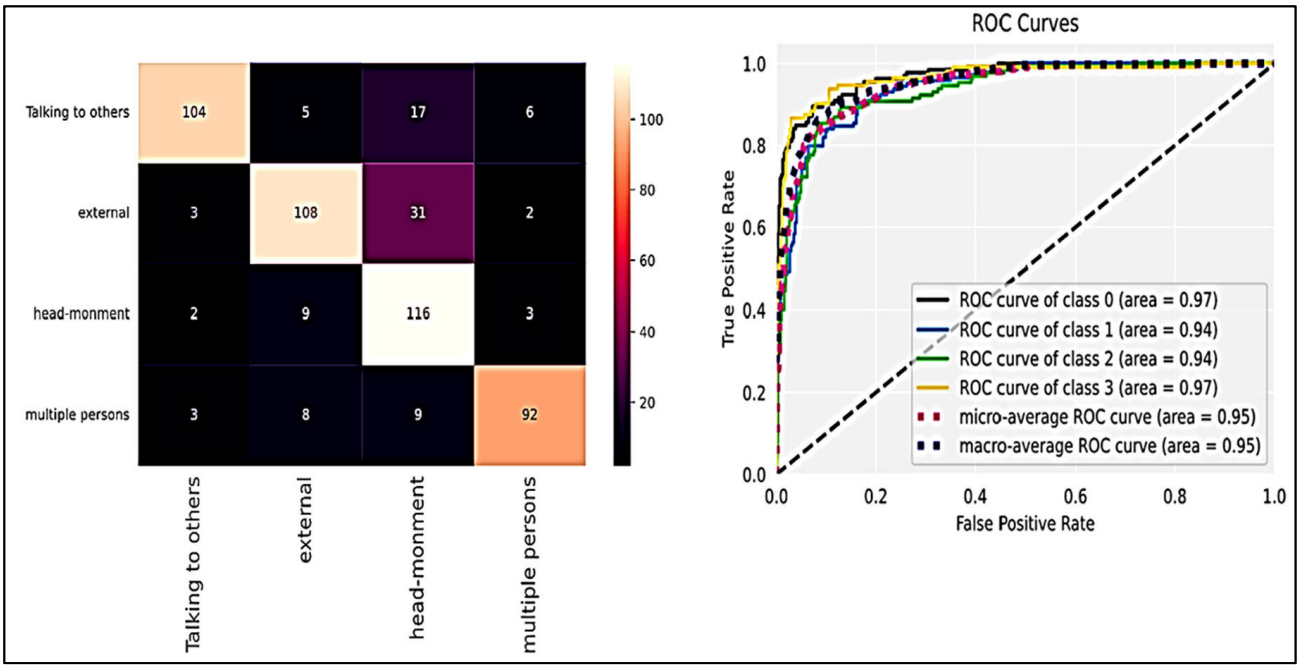


FIGURE 7. Confusion matrix and ROC curve for Inception_Resnet_v2.

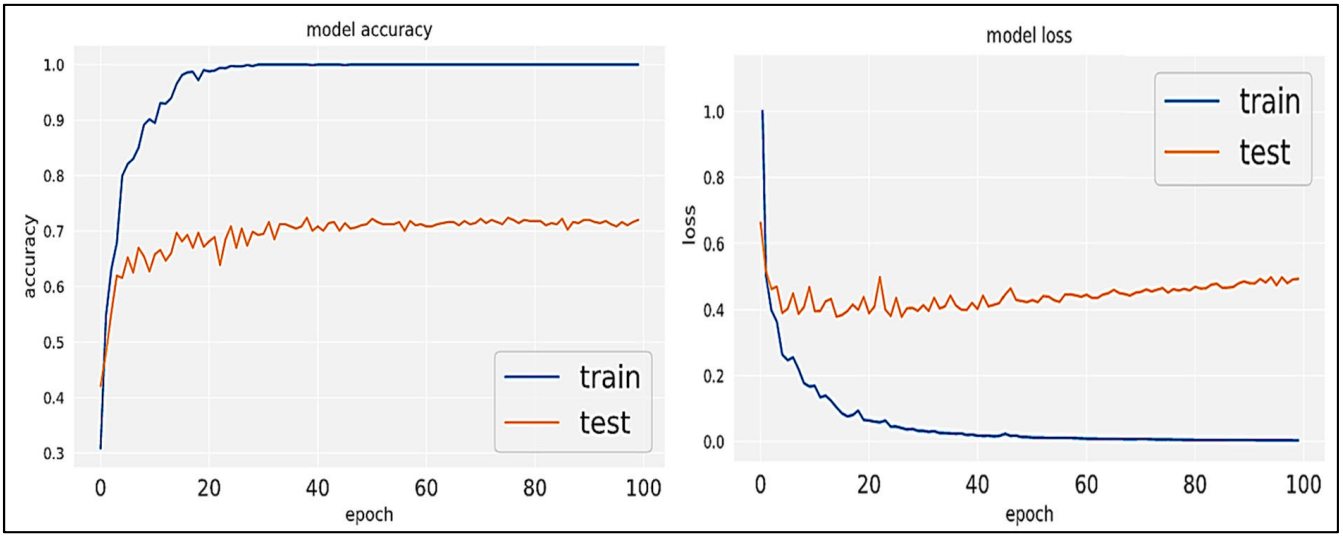


FIGURE 8. Model accuracy and model loss for Inception-V3.

TABLE 5. Classification report for inception-ResNet-v2.

	'RECISION	Recall	F1-score	Support
Talking to others	0.93	0.79	0.85	132
External	0.83	0.75	0.79	144
Head-movement	0.67	0.89	0.77	130
Multiple persons	0.89	0.82	0.86	112
Accuracy			**0.81**	**518**

loss and accuracy. The classification reports are presented in Table 5; the confusion matrix and ROC curve are shown in Figure 7 after that.

3) RESULTS USING INCEPTION-V3

The third pre-trained model was used to analyze the performance on online exams. Figure 8 displays the model loss and accuracy.

The classification reports are presented in Table 6. Also, the confusion matrix and ROC curve are shown in Figure 9 after that.

B. FINE-TRAINED CNN MODEL

Fine-tuned CNN model is used to analyze the performance of online exams. Figure 10 shows model accuracy and

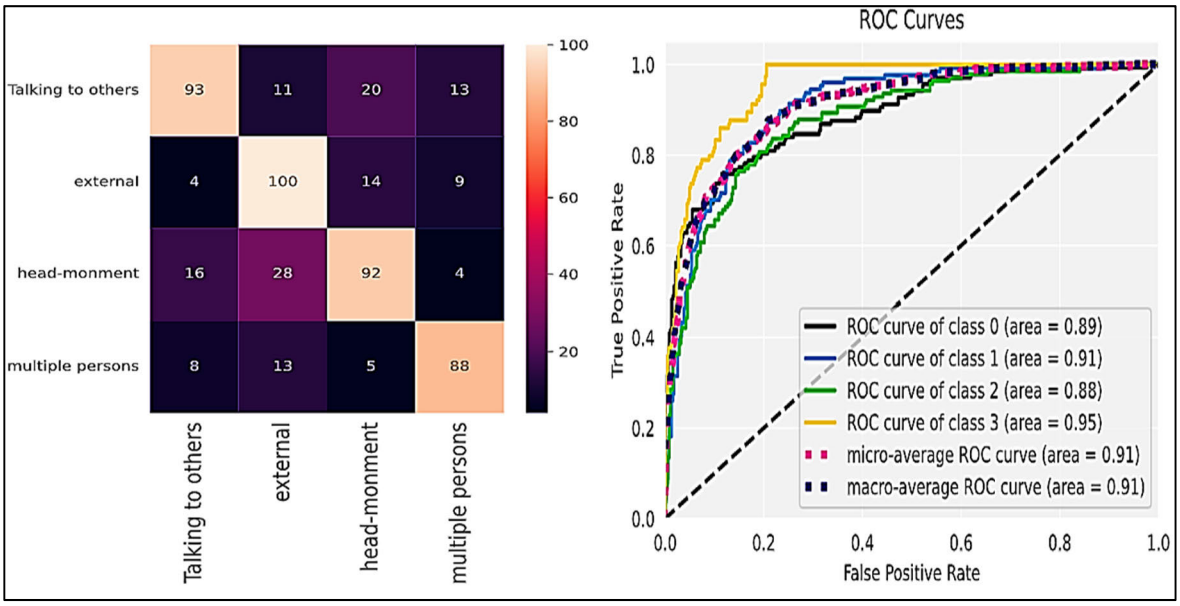


FIGURE 9. Confusion matrix and ROC curve for Inception-V3.

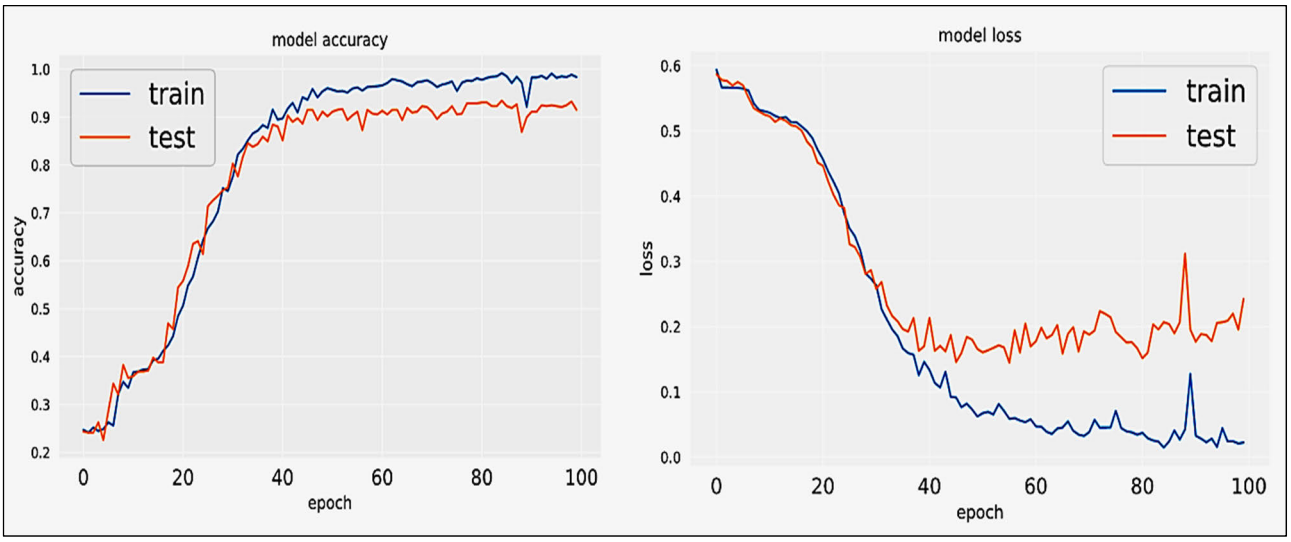


FIGURE 10. Model accuracy and model loss by using CNN.

TABLE 6. Classification report for Inception-V3.

	'RECISION	Recall	F1-score	Support
Talking to others	0.77	0.68	0.72	137
External	0.66	0.79	0.72	127
Head-movement	0.70	0.66	0.68	140
Multiple persons	0.77	0.77	0.77	114
Accuracy			**0.72**	**518**

TABLE 7. Classification report using CNN model.

	'RECISION	Recall	F1-score	Support
Talking to others	0.98	0.81	0.89	140
External	0.92	0.97	0.95	136
Head-movement	0.91	0.92	0.92	125
Multiple persons	0.85	0.97	0.90	17
Accuracy			0.92	518

loss, while Figure 11 illustrates confusion matrix and ROC curve for CNN. The classification reports are shown in Table 7.

C. YOLOv5 MODEL RESULTS ANALYSIS

The YOLOv5 model was used to analyze the performance of online exams. The following figures are shown in the model

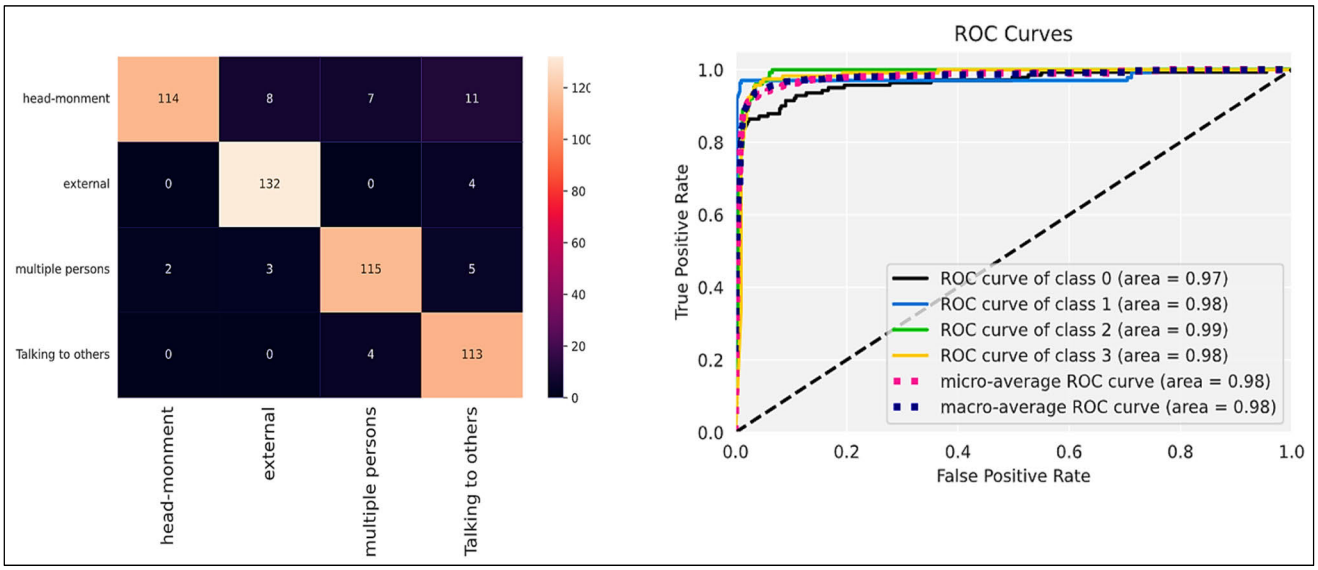


FIGURE 11. Confusion matrix and ROC curve for CNN.

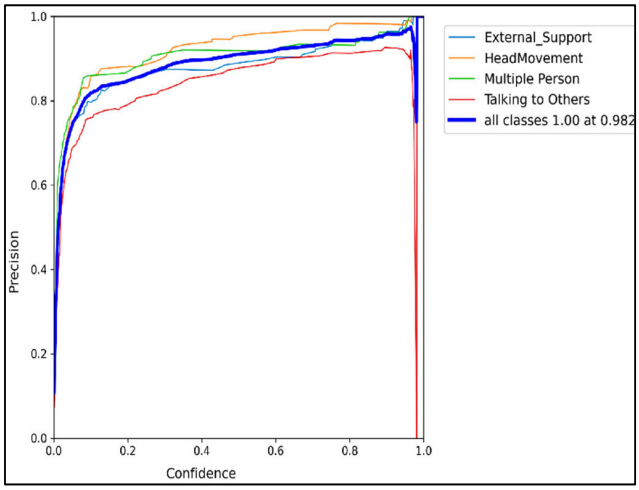


FIGURE 12. Precision graph for online exam datasets using YOLOV5 model.

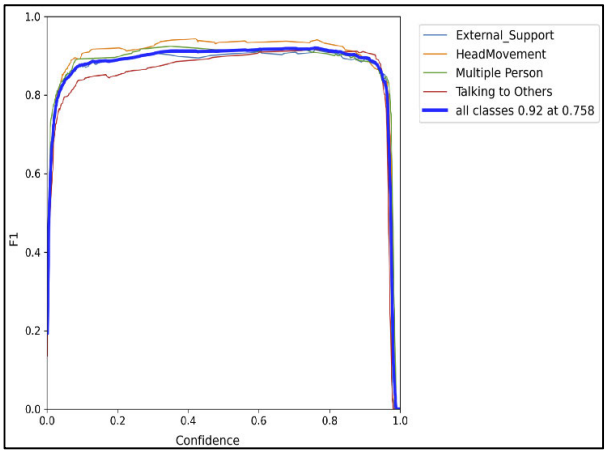


FIGURE 13. F1 graph for online exam datasets using the YOLOV5 model.

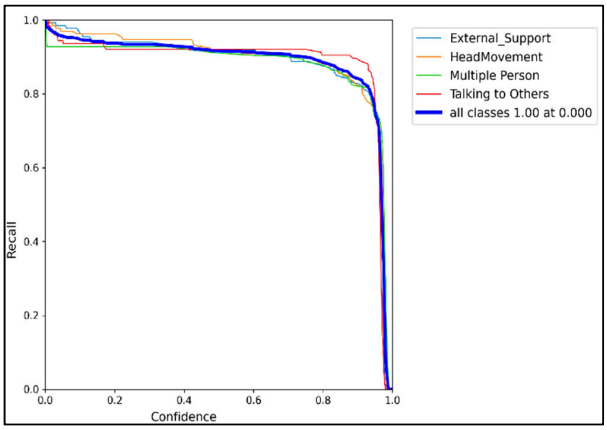


FIGURE 14. Recall graph for online exam datasets using YOLOV5 model.

accuracy, model loss, classification reports, confusion matrix, and ROC curves.

a) Training Process: The period span is 400 epochs. In the following section, the training loss curves, the validation loss function curves, the number of epochs increases, and the loss curve regularly become stable, indicating that the model training result improves.

b) Model Measurement: YOLOv5's pre-training weight training was used in this training to obtain the F1, Precision, Recall, and evaluation indexes for precision Recall. The precision graph, F1 graphs, and Recall graph are shown in the following Figures 12 to 14.

Labelling anomalous behaviors using bounding boxes during online tests with YOLOv5 entails using the YOLOv5 object detection model to automatically recognise and label certain abnormal or suspicious actions in exam recordings or screen captures.

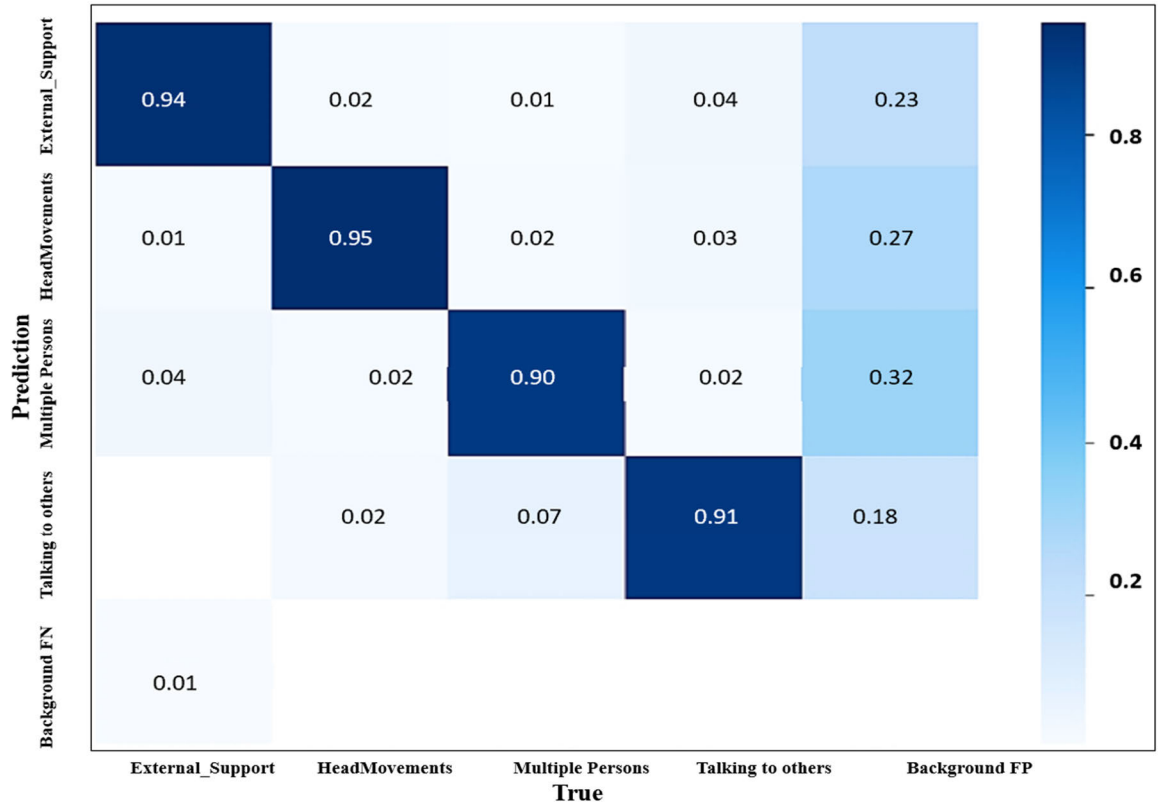


FIGURE 15. Confusion matrix for online exam dataset using YOLOV5.

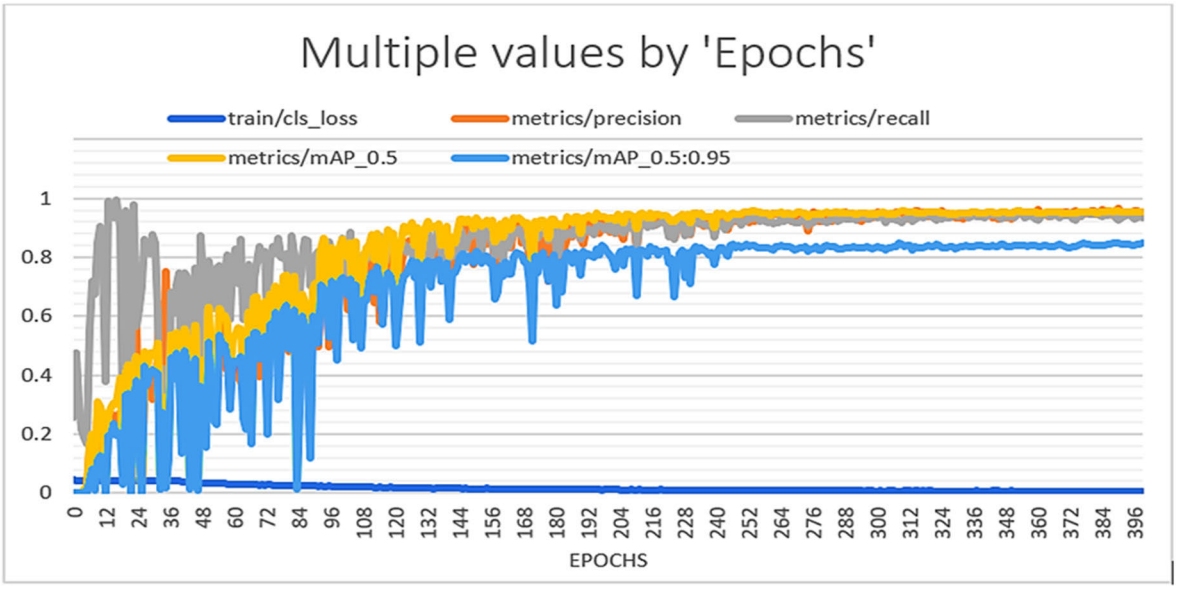


FIGURE 16. Visualization of performance measure using multiple values of Epochs.

Table 8 contains the assessment data for every index following model training. The trash classification model's mAP (IoU[0.5]), mAP (IoU [0.5: 0.95], Recall, and Precision are shown to be 95.40 %, 84.65 %, 93.16 %, and 95.54 %, respectively.

In Figure 15, the confusion matrix is shown. The confusion matrix's columns indicate actual values, while its rows represent predicted values. Ideally, all items would be appropriately classified, which means that the diagonal should include all categories in its row and column, which

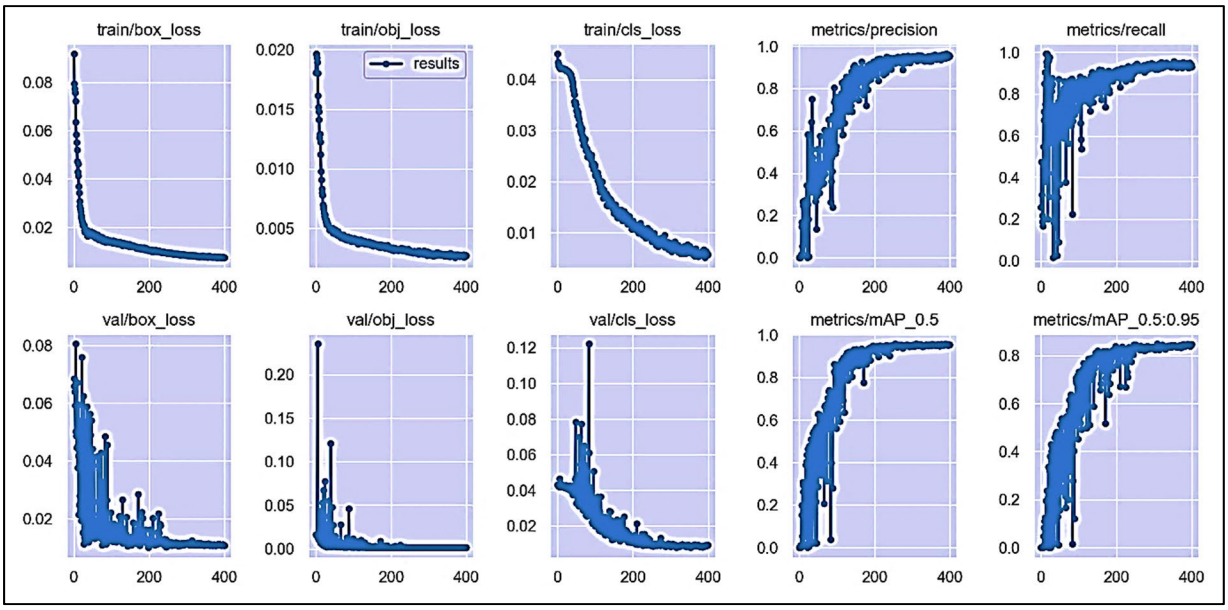


FIGURE 17. Visualization of mAP, recall, precision, classification loss, abject loss, and box loss by using different. epochs

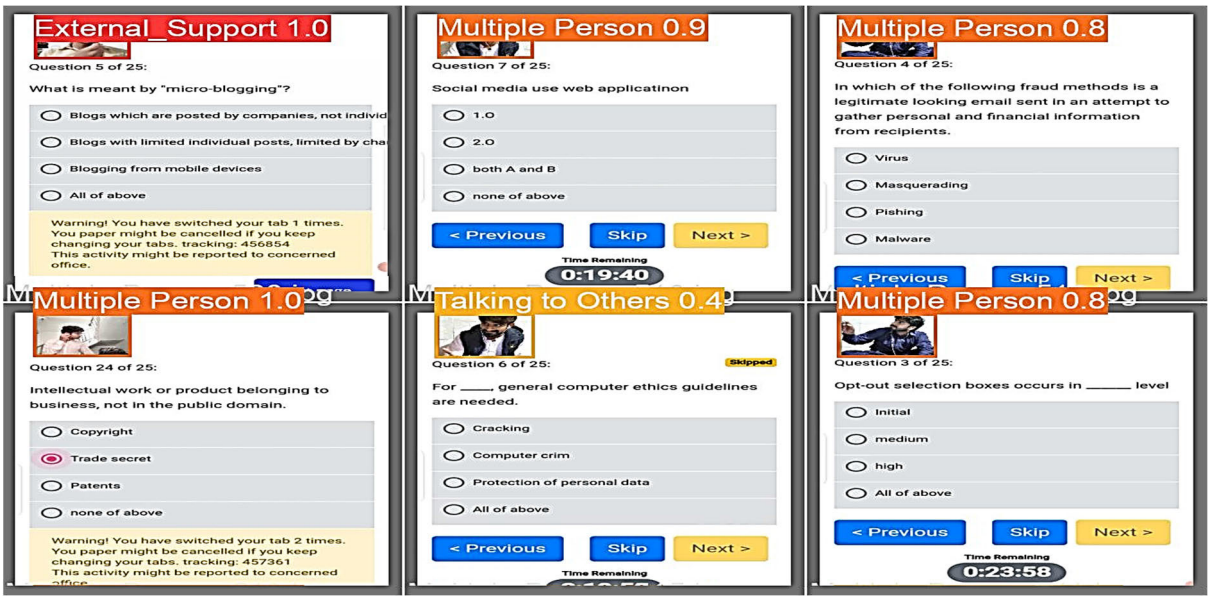


FIGURE 18. Detection results on online exams datasets with best train weights using YOLOV5.

TABLE 8. Performance measures value by using Yolo-V5.

S. No.	Performance measures	Results
1	Precision	0.9554
2	Recall	0.9316
3	mAP_0.5	0.9540
4	mAP_0.5:0.95	0.8465

has been roughly true for all classes except thrusters thus far.

The bottom row depicts false negatives (FN), which are things in the frame but not detected by the algorithm. The rightmost column represents false positives (FP), which occur when the algorithm detects an object when there is none.

The visualization of different performance measure values by using the multiple values of Epochs is shown in Figure 16.

Figure 17 depicts three types of loss: box loss, objectness loss, and classification loss. The box loss measures how well the algorithm can detect an object's center and how well the predicted bounding box covers an object. Objectness is a measure of the probability that an object exists in a suggested zone of interest. If the objectivity is high, this indicates that an item is likely to be present in the image window.

TABLE 9. Performance comparison of Inception-V3, Inception_ReseNetV2, DenseNet121, YOLOv5 and CNN model.

CNN Models	Accuracy	Macro Average			Weighted Average			Support
		Precision	Recall	F1-Score	Precision	Recall	F1-Score	
Inception-V3	0.72	0.73	0.72	0.72	0.72	0.72	0.72	518
Inception_ReseNetV2	0.81	0.83	0.81	0.82	0.83	0.81	0.81	518
DenseNet121	0.86	0.86	0.86	0.86	0.86	0.86	0.86	518
Suggested CNN	0.92	0.92	0.92	0.91	0.92	0.92	0.91	518
YOLOv5		**0.9554**	**0.9316**	**mAP_0.5= 0.9540**		**mAP_0.5:0.95 = 0.8465**		**518**

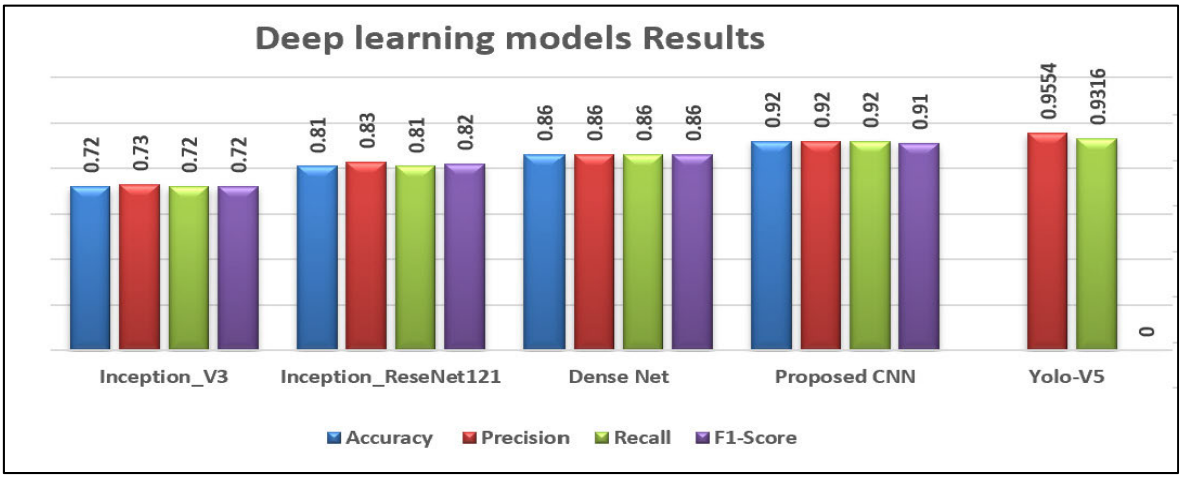


FIGURE 19. Deep learning models results for the detection of cheating activities in online exam.

Classification loss measures how well the algorithm predicts the correct class of an object.

Precision, recall and mean average precision improved quickly before plateauing after about 400 epochs. The validation data's box, objectiveness, and classification losses declined rapidly until about 400. To choose the best weights, we used early stopping.

Detection Results: The test pictures of unusual activities during online examination are shown in Figure 18. Table 9 shows the Performance Comparison of Inception-V3, Inception_ReseNetV2, DenseNet121, YOLOv5 and CNN Mode Compared to other pre-trained and CNN models, the results obtained with YOLOv5 are better. The suggested CNN models were tested using various configurations of layers and parameters over more than 100 epochs, but the results did not improve. Table 9 shows a comparison of the results of these models. The deep learning models results for detecting cheating in online exams are shown in Figure 19.

YOLOv5 outperforms its rivals. The performance metrics employed by YOLOv5 and other deep learning models differ only marginally. A specific feature, such as Accuracy, is missing from the Yolov5 model. Instead, we used the confidence level of each detected object in the frame. The YOLOv5 model is approximately twice as fast as the others.

V. CONCLUSION

In this study pre-trained and fine-tuned deep learning-based models has used for the cheating activity detection model in online exams. Pre-trained models Inception-V3, Inception _Resnet_v2, DenseNet121, YOLOv5, and Fine-tuned CNN are used for performance analysis. Datasets of cheating activities in online exams are obtained from the university on a special request for research purposes only. The datasets contain four types of unusual online examination activities. The suggested CNN models were evaluated over more than 100 epochs with various layer and parameter combinations, but the results of YOLOv5 remain superior. Yolov5 and other deep-learning models use the same performance metrics. The detection results of deep learning models Yolov-V5 performed better on our produced dataset. The Yolov5 model does not contain a specific feature, such as Accuracy. Instead, we used the confidence level of each recognized object in the frame, and mAP 0.5 and mAP 0.5:0.95 are used in YOLOv5. The speed of the YOLOv5 model is roughly twice that of the others.

REFERENCES

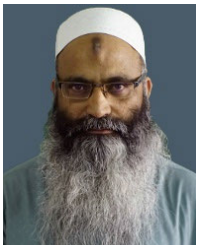
[1] A. Gupta and A. Bhat, "Bluetooth camera based online examination system with deep learning," in *Proc. 6th Int. Conf. Intell. Comput. Control Syst. (ICICCS)*, May 2022, pp. 1477–1480, doi: 10.1109/ICICCS53718.2022.9788147.

[2] M. M. Masud, K. Hayawi, S. S. Mathew, T. Michael, and M. E. Barachi, "Smart online exam proctoring assist for cheating detection," *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13087. Cham, Switzerland: Springer, 2022, pp. 118–132, doi: 10.1007/978-3-030-95405-5_9.
[3] R. M. Al_airaji, I. A. Aljazaery, H. T. S. Alrikabi, and A. H. M. Alaidi, "Automated cheating detection based on video surveillance in the examination classes," *Int. J. Interact. Mobile Technol. (iJIM)*, vol. 16, no. 08, pp. 124–137, Apr. 2022, doi: 10.3991/ijim.v16i08.30157.
[4] R. M. Alairaji, I. A. Aljazaery, and H. S. Alrikabi, "Abnormal behavior detection of students in the examination Hall from surveillance videos," in *Advanced Computational Paradigms and Hybrid Intelligent Computing*. Singapore: Springer, 2022, pp. 113–125, doi: 10.1007/978-981-16-4369-9_12.
[5] A. A. Malik, M. Hassan, M. Rizwan, I. Mushtaque, T. A. Lak, and M. Hussain, "Impact of academic cheating and perceived online learning effectiveness on academic performance during the COVID-19 pandemic among Pakistani students," *Frontiers Psychol.*, vol. 14, Mar. 2023, Art. no. 1124095.
[6] D. L. McCabe, "Cheating among college and university students: A north American perspective," *Int. J. Educ. Integrity*, vol. 1, no. 1, Nov. 2005, doi: 10.21913/ijei.v1i1.14.
[7] S. A. Butt, "Analysis of unfair means cases in computer-based examination systems," *Pacific Sci. Rev. B, Humanities Social Sci.*, vol. 2, no. 2, pp. 75–79, Jul. 2016.
[8] M. Ramzan and A. Abid, *Automatic Unusual Activities Recognition Using Deep Learning in Academia*. Accessed: Jun. 20, 2022. [Online]. Available: https://www.academia.edu/download/74918847/pdf.pdf
[9] T. S. Kumar and G. Narmatha, "Video analysis for malpractice detection in classroom examination," in *Proc. Int. Conf. Soft Comput. Syst.*, in Advances in Intelligent Systems and Computing, vol. 397, 2016, pp. 135–146, doi: 10.1007/978-81-322-2671-0_13.
[10] Z. Li, Z. Zhu, and T. Yang, "A multi-index examination cheating detection method based on neural network," in *Proc. IEEE 31st Int. Conf. Tools Artif. Intell. (ICTAI)*, Nov. 2019, pp. 575–581.
[11] N. Malhotra, R. Suri, P. Verma, and R. Kumar, "Smart artificial intelligence based online proctoring system," in *Proc. IEEE Delhi Sect. Conf. (DELCON)*, Feb. 2022, pp. 1–5, doi: 10.1109/DELCON54057.2022.9753313.
[12] D. Komosny and S. U. Rehman, "A method for cheating indication in unproctored on-line exams," *Sensors*, vol. 22, no. 2, p. 654, Jan. 2022, doi: 10.3390/s22020654.
[13] A. Singh and S. Das, "A cheating detection system in online examinations based on the analysis of eye-gaze and head-pose," in *Proc. Int. Conf. Emerg. Trends Artif. Intell. Smart Syst.*, Jun. 2022, doi: 10.4108/EAI.16-4-2022.2318165.
[14] L. C. Ow Tiong and H. J. Lee, "E-cheating prevention measures: Detection of cheating at online examinations using deep learning approach—A case study," 2021, *arXiv:2101.09841*.
[15] G. Kasliwal, "Cheating detection in online examinations," Master's Projects, San José State Univ., San Jose, CA, USA, Tech. Rep., 2015, doi: 10.31979/etd.y292-cddh.
[16] D. Dobrovska, "Technical student electronic cheating on examination," in *Proc. Int. Conf. Interact. Collaborative Learn.*, in Advances in Intelligent Systems and Computing, vol. 544, 2017, pp. 525–531, doi: 10.1007/978-3-319-50337-0_49.
[17] S. Hu, X. Jia, and Y. Fu, "Research on abnormal behavior detection of online examination based on image information," in *Proc. 10th Int. Conf. Intell. Hum.-Mach. Syst. Cybern. (IHMSC)*, vol. 2, Aug. 2018, pp. 88–91, doi: 10.1109/IHMSC.2018.10127.
[18] A. Fayyoumi and A. Zarrad, "Novel solution based on face recognition to address identity theft and cheating in online examination systems," *Adv. Internet Things*, vol. 4, no. 2, pp. 5–12, 2014, doi: 10.4236/AIT.2014.42002.
[19] E. Bilen and A. Matros, "Online cheating amid COVID-19," *J. Econ. Behav. Org.*, vol. 182, pp. 196–211, Feb. 2021, doi: 10.1016/j.jebo.2020.12.004.
[20] R. Comas-Forgas, T. Lancaster, A. Calvo-Sastre, and J. Sureda-Negre, "Exam cheating and academic integrity breaches during the COVID-19 pandemic: An analysis of internet search activity in Spain," *Heliyon*, vol. 7, no. 10, Oct. 2021, Art. no. e08233, doi: 10.1016/j.heliyon.2021.e08233.
[21] T. Ngqondi, P. B. Maoneke, and H. Mauwa, "A secure online exams conceptual framework for south African universities," *Social Sci. Humanities Open*, vol. 3, no. 1, Jan. 2021, Art. no. 100132, doi: 10.1016/j.ssaho.2021.100132.
[22] F. Noorbehbahani, A. Mohammadi, and M. Aminazadeh, "A systematic review of research on cheating in online exams from 2010 to 2021," *Educ. Inf. Technol.*, vol. 27, no. 6, pp. 8413–8460, Mar. 2022, doi: 10.1007/s10639-022-10927-7.
[23] S. Gopane and R. Kotecha, "Enhancing monitoring in online exams using artificial intelligence," in *Proc. Int. Conf. Data Sci. Appl.*, in Lecture Notes in Networks and Systems, vol. 287, 2022, pp. 183–193, doi: 10.1007/978-981-16-5348-3_14.
[24] M. D. Genemo, "Suspicious activity recognition for monitoring cheating in exams," *Proc. Indian Nat. Sci. Acad.*, vol. 88, no. 1, pp. 1–10, Mar. 2022, doi: 10.1007/S43538-022-00069-2.
[25] A. R. Khan, T. Saba, M. Z. Khan, S. M. Fati, and M. U. G. Khan, "Classification of human's activities from gesture recognition in live videos using deep learning," *Concurrency Comput., Pract. Exper.*, vol. 34, no. 10, p. e6825, May 2022.
[26] W. Alsabhan, "Student cheating detection in higher education by implementing machine learning and LSTM techniques," *Sensors*, vol. 23, no. 8, p. 4149, Apr. 2023.
[27] M. Garg and A. Goel, "Detection of internet cheating in online assessments using cluster analysis," in *Proc. Int. Conf. Data Manage., Analytics Innov.* Singapore, Springer Nature, 2023, pp. 77–90.
[28] M. Huang, H. Shu, and J. Jiang, "An algorithm of key-frame extraction based on adaptive threshold detection of multi-features," in *Proc. Int. Symp. Test Meas.*, vol. 1, 2009, pp. 149–152, doi: 10.1109/ICTM.2009.5412976.
[29] Z. Wang and Y. Zhu, "Video key frame monitoring algorithm and virtual reality display based on motion vector," *IEEE Access*, vol. 8, pp. 159027–159038, 2020. [Online]. Available: https://ieeexplore.ieee.org/abstract/document/9178317/
[30] G. Jocher, K. Nishimura, T. Mineeva, and R. Vilariño, "YOLOV5," Code Repository, Tech. Rep., 2020. [Online]. Available: https://github.com/ultralytics/yolov5
[31] C. Szegedy, S. Ioffe, V. Vanhoucke, and A. Alemi, "Inception-v4, inception-ResNet and the impact of residual connections on learning," in *Proc. AAAI Conf. Artif. Intell.*, 2017, vol. 31, no. 1, pp. 1–14.
[32] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, "Densely connected convolutional networks," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 4700–4708.

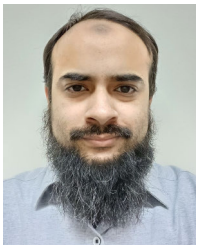
MUHAMMAD RAMZAN received the Ph.D. degree in CS from the University of Management and Technology, Lahore, Pakistan. He is currently a Lecturer with the CS and IT Department, University of Sargodha. Previously, he was a Lecturer with the Virtual University of Pakistan. He was also with the Higher Education Department as a Database/Network Administrator. Before this he was a Lecturer with Minhaj University Lahore (Sharia College). He has authored more than 40 research articles published in reputed peer-reviewed journals. His research interests include medical imaging, deep learning, machine learning, video surveillance, activity recognition, and computer vision.

ADNAN ABID (Senior Member, IEEE) received the Ph.D. degree from Politecnico di Milano, Italy, in 2012. He is currently a Professor with the Department of Data Science, Faculty of Computing and Information Technology, University of Punjab, Lahore, Pakistan. He is a member of the ACM. He has been associated with editorial boards of well-reputed journals in the area of computer science. His research interests include adaptive and self-organizing, systems computational, linguistics computer education, data mining and machine learning, data science, databases, and software engineering.

MUHAMMAD BILAL received the Ph.D. degree in computer science from Taylor's University, Malaysia, in 2021. He is currently a Senior Lecturer with the School of Engineering and Technology, Sunway University. Previously, he was an Assistant Professor with the University of Southampton Malaysia. His earlier roles include an Assistant Professor and an incharge with the FAST National University of Computer and Emerging Sciences, a Tutor with Taylor's University, a Lecturer with the University of Sargodha, and a Software Developer with Sofizar. He has published several research papers in journals and conferences. His research interests include data mining, social computing, machine learning, information processing, social media data analytics, and software engineering.



KHALID M. AAMIR received the M.S. degree in systems engineering from the Pakistan Institute of Engineering and Applied Sciences (PIEAS), Quaid-e-Azam University, Islamabad, Pakistan, in 1999, and the Ph.D. degree in computer engineering from the Lahore University of Management Sciences (LUMS), Lahore, Pakistan, 2008. Currently, he is an Assistant Professor with the Department of Computer Science and IT, University of Sargodha, Sargodha, Pakistan. His research interests include machine learning, bioinformatics, and medical signal processing.



SUFYAN A. MEMON received the Ph.D. degree in electronic systems engineering from Hanyang University, Republic of Korea, in 2016.

He has been an Assistant Professor with the Department of Defense Systems Engineering, Sejong University, Seoul, Republic of Korea, since March 2021. His research interests include tracking, estimation, guidance, navigation, and control.

TAE-SUN CHUNG received the B.S. degree in computer science from KAIST, in February 1995, and the M.S. and Ph.D. degree in computer science from Seoul National University, in February 1997 and August 2002, respectively. He is currently a Professor with the Department of Software, Ajou University. His current research interests include flash memory storages, query processing in spatial databases, machine learnings, and general database systems.