2024 Second International Conference on Inventive Computing and Informatics (ICICI) | 979-8-3503-7329-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICICI62254.2024.00054

Malpractice Detection in Examination Hall using Deep Learning

Aruna S K
Department of Computer Science and Engineering, School of Engineering and Technology, Bangalore, CHRIST University
India
sksaruna@yahoo.co.in

Madhumitha A
Department of Computer Science and Engineering, School of Engineering and Technology,Bangalore, CHRIST University
India
madhu.mm.mitha2@gmail.com

Selvanayaki Kolandapalayam Shanmugam
Department of Mathematics and Computer Science, Ashland University
Ohia, USA
skolanda@ashland.edu

Senthil Kumar Thangavel
Department of Computer Science and Engineering, Amrita School of Computing, Coimbatore Amrita Vishwa Vidyapeetham
India
t_senthilkumar@cb.amrita.edu

Maiga Chang
School of Computing Information and Systems Athabasca University
Alberta, Canada
maigac@athabascau.ca

***Abstract*— Various institutions administer tests at designated examination locations, chosen third-party and approved centers, and have established standards for installing CCTV cameras and conducting frisking under the supervision of designated personnel. Some institutions are using online proctoring, which enables students to take exams from any location. In all of the aforementioned scenarios, human monitoring is conducted, and maintaining a high level of vigilance may be challenging due to administrative oversight or intentional allowance of malpractice for personal gain. The malpractice detection may be attributed to acts like as plagiarism, unauthorized sharing of papers, and non-verbal communication. The study is conducted by capturing the dataset in the classroom of Christ University. The proposed approach is based on the YOLO framework. The movies are processed in real time to identify hand rotation, paper extraction, and classify the motion. The accuracy for the Head_right class is significantly higher than that of the Head_left class. The system is implemented using the programming language Python and has the potential for future expansion to provide real-time monitoring.**

Keywords*— *Malpractice detection, Feature extraction, deep learning, YOLO, Human monitoring.

I. Introduction

Surveillance systems, which have been around for ages, are an essential part of securing your home or business from any type of suspicious act that may occur. They have now taken precedence and become an integral part of life and infrastructure. The presence of surveillance can serve as a deterrent to would-be wrong-doing people and can protect them discreetly. " The digitalization and pandemic situation force institutions, governments, and corporations to deliver courses and trainings online; examination and evaluation are the two essential phases of completing a course successfully. Candidates are selected for jobs or higher education based on their score or cadence.

While the system of surveillance was highly demanding and emphasized over private and public sector companies, it has also started reshaping the educational institution and addressing its continuous and complex trends in wrong practices on campus, particularly in the exam centers, whether online or offline, as a way to support invigilators and their supervising tasks.

Detecting malpractices in the examination hall is vital to preventing wrong morals and ensuring the conduct of all the students appearing for the examination. With the advancement of technology and devices, there are multiple ways in which a student would attempt to engage in malpractice. However, the fundamentals of using sign languages, expression, object exchange, and much more remain in practice.

People use examinations to learn about themselves. Assessing a student necessitates comprehension of the concepts they are presenting. This also influences the student's final grade. The score largely reflects the student's abilities. Exams are also a way of seeing the continuous progress of the student in aptitude, logic, and reasoning. Assessments can be a continuous process through quizzes, exams, and vivas. The assignment is one way to collect some kind of performance data from students. Vivias allow students to actively evaluate their assignments to assess their effort.

Written exams can reflect what they really understand, what misconceptions they have, what they don’t understand, what kind of concepts they really know in the class, and their mastery levels in the particular concepts or skills. The score indicates their performance, skill level, knowledge level, and understanding level. The ranking of a student has to be considered with the effort of the student and the sincerity of the student.

979-8-3503-7329-5/24/$31.00 ©2024 IEEE
DOI 10.1109/ICICI62254.2024.00054

The proposed system identifies if any student engages in malpractice by interacting with other students through hand gestures, such as showing up numbers or other gestures; if any student attempts to pass answer scripts without the invigilator's knowledge; and if any student attempts to copy from another student.

Though online classes, trainings, and exams are the most efficient and viable ways to achieve academic goals, due to increased malpractice by fraudsters and cybercriminals, the evaluation criteria have become challenging. Failure to maintain the trust and integrity of the online examination and evaluation methods erodes the candidate's confidence, and legitimate candidates may lose their opportunities.

To overcome the challenge, this research proposes implementing an online vision-based malpractice detection system using machine learning, deep learning, and artificial intelligence techniques.

The rest of the study has been organized as follows: Section 2 reviews the related work in the areas of video processing, the YOLO algorithm, and motion analysis. Section 3 describes the proposed methodology, as well as the preprocessing. Section 4 discusses the experimental results in detail. Section 5 depicts the conclusion and future work.

II. Related Works

Senthil Kumar et al. proposed a system to monitor student misconduct in the examination hall and alert authorities [1]. This study describes three phases: feature extraction, classification, and cascading. The features are extracted using Haar features and classified using the AdaBoost classifier, and cascading is used to improve processing and performance. The system was successful in accurately detecting suspicious activities with less time. But the results of detections in low-luminated environments are the drawbacks of this algorithm.

[2] conducted a detailed survey of low-level techniques and high-level processing methods. It explains the real-time video analysis system that detects if an activity is suspicious or not. The implementation entails pre-processing with the Gabor filter and background and foreground feature extraction. The system applies a combination of detection techniques to identify and analyze human faces. The system detects and classifies suspicious activity with a red box and a green box when it records no such activity. The study attempts to compare the performance analysis with respect to precision and accuracy, which is satisfactory. The future prospects of this methodology are to improve detections in low-illumination and better feature extraction.

YOLO Architecture, its training, comparisons, performance, and results have been discussed in [3]. This study uses a single neural network with a unified architecture and the YOLO algorithm. The algorithm predicts using bounding boxes and their confidence scores. Optimizations are used to improve the predictions. When compared to other algorithms, YOLO has comparatively performed well, especially in real-time. Several combinations of algorithms, along with YOLO, are experimented with to test the mAP (mean average precision), which can significantly improve performance.

Detection and classification by using the YOLO Framework for cars and people were explained by [4]. In this study, a detailed presentation of a system that can assist drivers by detecting objects has been described. The implementation process includes data collection, framework training, and classification results. The author compares the detection results with varying grid cells. The system undergoes tests to evaluate its precision and accuracy. The results indicate that fewer layers and more grid cells detect objects more accurately. This is due to the fact that the detection of smaller objects increases with the number of grid cells. Hence, this work has been considered feasible for real-time processing.

[5] proposed a modified version of YOLO to count people in real-time. This study states that the YOLO-PC (YOLO-People Counting) approach is the best and fastest object detection algorithm. The computational requirements and time for detections are comparatively less when compared to other algorithms. The YOLO-PC uses a 9x9 grid cell approach to detect objects, with different threshold values. Real-time analysis counts the individuals detected within the boundary constraints based on their previous values. The YOLO-PC results outperform those of YOLO in terms of detection and confidence values. This model's applications extend to smart cities and entry points.

Unlike other algorithms, YOLO is considered to be fast, robust, and a generalized approach to object detection. Skin pixels could also be extracted, and the connected regions can be discriminated against to distinguish between humans [7]. Deep learning is also used in the medical field. It can have a variety of applications, including diagnosis and neural related problems. This aids in the early treatment of diseases, thereby improving patient treatment [8].

III. Proposed Methodology

A. Dataset Collection

The scope of the proposed research work is to detect and predict only the following malpractices: head rotation, paper passing, and showing hand gestures. The images and videos collected as a dataset must be captured with a resolution greater than 500 x 500. The examination hall must be in good lighting conditions. To prevent any disturbances, the camera must remain fixed and stationary. For the hardware, it is recommended to use a system with high graphical processing power for faster training and detection.

An examination hall with a group of students has been captured. Images and videos of students exhibiting malpractice and no malpractice were collected within the CHRIST (Deemed to be University) environment. For better performance, images and videos with a different batch of students were collected, and students were arranged in different combinations for variation. The original resolution of the images is 4496 x 3000. The recorded videos' resolution is 3840 x 2160. The frame rate of the video is 30 frames per second. The average duration of the video is 1 to 3 minutes.

B. Data Preprocessing

All the malpractice images are classified under paper passing, head rotation, and gesture recognition. Training images were labeled into four different classes, namely head-right, head-left, paper-pass, and gesture, for the identification of malpractice. The process of labeling involved using a tool to draw bounding boxes around the images and manually classifying them into classes. The tool outputs the labeled images in the form of an XML file with the coordinates of the bounding box and their classes. Fig. 1 shows the steps involved in preprocessing.

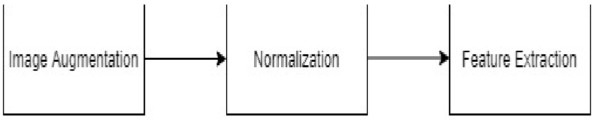

Fig. 1. Steps for preprocessing.

The labeled data was divided into training data and test data. In the proposed work, 449 images are used as training data. For the test data, 228 images are segregated for validation and evaluation.

The work's limited dataset necessitates the training of the model on a large dataset for any machine learning algorithm to function effectively. To expand the dataset, image augmentation has been used. The algorithm generates multiple images by rotating, zooming in and out, and flipping a single image. This leads to variations in algorithm training.

The resizing of all the images to a uniform size for the model to maintain uniformity for processing has been set to 352x448. For feature extraction, the YOLO algorithm learns from the training data. A certain pattern guides the extraction of the trainable parameters and features.

C. YOLO Architecture

This YOLO architecture is built upon a single convolutional layer of a convolutional neural network. This can be used for custom object detection since it is a generalized object detection algorithm. For the proposed automated system, the algorithm is trained with the custom labels or classes mentioned in Section Data Preprocessing to detect malpractices.

To make the process faster, the algorithm divides the images into smaller m x m grid cells. Instead of processing each pixel, it processes grids. Among the three versions of the YOLO algorithm, the latest improvised version, YOLOv3, is used for the work to produce better performance as it generates ten times more boxes to predict when compared to YOLOv2.

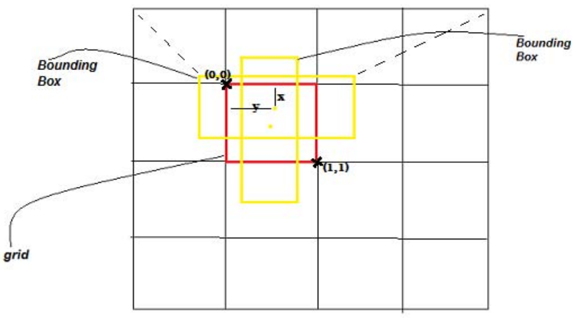

Fig. 2. Grid and Bounding Box using YOLO Algorithm.

- Objects in an image are of different sizes. Hence, detecting each object of varied sizes is important. Though the larger objects are easier to identify, it is important to detect the smaller objects as well. For this, YOLOv3 divides an image into three scales. For every scale, it is divided into three different grid cells to detect the objects. The first image is divided into 13x13 grid cells to detect larger objects. Next, we divide the image into 26x26 grid cells to identify medium-sized objects. Further, for the detection of smaller objects, the image is divided into 52x52 grid cells. Fig. 2 depicts the process of grid and bounding boxes using the YOLO algorithm.
- For each grid cell, there are anchor boxes that are formed. The anchor boxes help detect two things. One, if there is an object present in the grid cell or not. Two, if there is an object, what object is it, and what is the probability that it is a specific object. This gives a prediction about a possible object to be detected based on its probability score. The classes are then mapped to make a prediction with the help of a bounding box, class, and confidence score.

In YOLOv3, a total of nine anchor boxes are generated. The anchor boxes for the model can be generated with the help of the K-means clustering method. When anchor boxes are generated for a cell, each of them calculates their possible object predictions. These result in overlap and non-overlap scores. The Intersection Over Unit (IOU) provides the overlap value. This overlap value helps in detecting the object based on the most possible prediction by all boxes. If the threshold is set to 50%, then the probability of the object class having more than this threshold is classified. In other words, if the IOU of an object predicted by the boxes exceeds the threshold, it classifies it as that detected class

An anchor box calculates the following parameters:

- (x,y,h,w): The (x,y) coordinates of the box with its height and width.
- Cp: The confidence score, which is the probability of the predicted object class.

Hence, the shape of the network will be:

[mxmx(B*(5+C))] , where mxm is the number of divided grid cells, B is the number of bounding boxes, 5 for the five anchor box parameters (x,y,h, w), and Cp. And C for the number of labels or classes defined.

For a 13x13 grid cell, where B = 3 and C = 4 for four labeled classes,

[13 x 13 x (3*(5+4))]

Similarly, for a 26x26 grid cell,

[26 x 26 x (3*(5+4))]

Similarly, for a 52x52 grid cell,

[52 x 52 x (3*(5+4))]

This predicts the defined classes with a probability score to denote which malpractice has been exhibited. A non-max suppression is used to make predictions, which have multiple classes. It is used to improvise the results of an overlapped prediction.

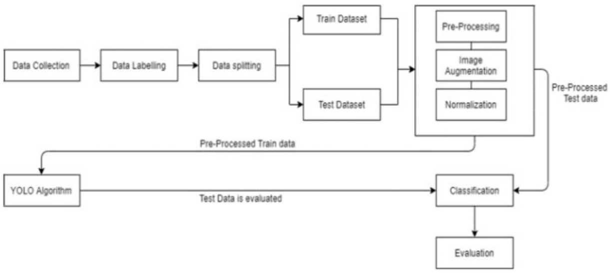

Fig. 3. Flow chart of the proposed system for examination malpractice detection and classification.

Fig. 3 elaborates on the process of the proposed system, which detects and classifies malpractices in the examination hall.

IV. Results and Discussions

A. Training

The YOLO architecture will be fed with the labeled training dataset. The XML files generated after the data labeling phase are given as input in the training phase. The Yolo architecture receives the XML files as input. After the completion of the training process, a weight file is generated,, which is saved as a.h5 file. This weight file includes all the trained parameters and the features that were extracted during the training process.

During training, instead of passing one image at a time, a group of images are passed. Training and testing data are generated in batches to feed into the model. In the proposed work, the batch size is taken as 6.

The entire life cycle of the training set is 15 times longer for smaller data sets. The number of epochs is set to 50. The threshold for prediction is set to 0.5, i.e., the algorithm detects malpractices with a confidence score greater than 50%. The number of initial epochs for matching with anchors is set to 3. Test images and their annotations are saved separately.

B. Validation

The trained YOLO algorithm is tested with test input files. The test input files are images and videos. The test image given as input will pass through the algorithm and generate a bounding box with the classified class and confidence score of the class. The test images were able to predict malpractices and classify them into defined classes.

If the student is attempting to exhibit malpractice by passing paper, the system classifies the malpractice under the defined class paper_pass and draws a bounding box. Similarly, a student who attempts to exhibit malpractice by rotating their head classifies them as head-right or head-left. The suspicious activity by a student showing gestures is detected by drawing a bounding box. Along with the bounding box, the system also displays the confidence score (that is, how confident the system is that the classification is true).

Fig. 4. Detection of Head Rotation and Gesture Malpractices.

Fig. 4 shows that the students communicate with hand gestures. The confidence scores of students exhibiting hand gestures are 94.03% and 88.98%, respectively. Head rotations to the left and right are detected with a confidence score of 83.02% and 97.15%, respectively.

Fig. 5. Detection of Head Rotation and Paper Exchange Malpractices.

A confidence score of 99.56% detects the exchange of paper between students. Head rotations to the left and right are detected with a confidence score of 83.02% and 98.52%, respectively. The same is illustrated in the Fig. 5.

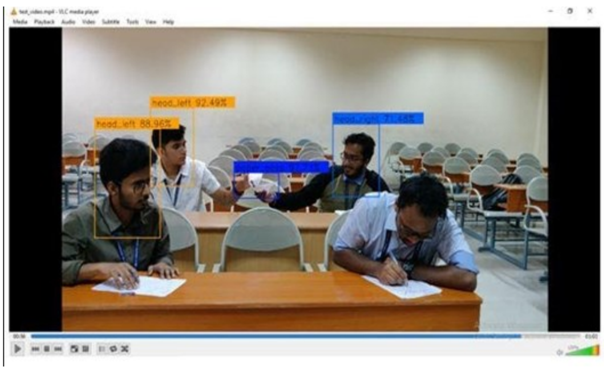

Fig. 6. Video analysis of malpractice detection and classification.

Fig. 6 illustrates that the paper exchanged from one student to another in the video is detected with a confidence score of 92.3%. Head rotations to the left and right are detected with a confidence score of 92.49% and 71.48%, respectively. Head rotation to the left exhibited by a student in the front is noted to be 88.96%.

The model converts the test video, frame by frame, into a series of images. These individual images are processed. The model performs detections by forming a bounding box along with the confidence score for each frame. The model then converts these into a video format. The output video that is generated has all malpractices detected with their probability score.

The threshold set for the algorithm is 0.5, which detects all the malpractices that have a confidence score a confidence score above 50%. The threshold can be varied accordingly, but it is ideally set at 50%. This means that all IOU greater than 0.5 in a particular grid are identified and classified.

The performance metric considered for the work is mAP (mean average precision). The average precision calculated for four labels is listed in Table 1.

TABLE I. AVERAGE PRECISION FOR EACH CLASS LABEL

Labels	**Average Precision**
Gesture	0.7086
Head_left	0.7514
Head_right	0.9670
Paper_pass	0.4404

The mean average precision is the sum of the average precision of all the defined class labels in the model multiplied by the number of labels in the model (n).

$$mAP = \frac{\sum AP}{n}$$

Hence, the mAP of the model is 0.7168. The average precision of the paper_pass label is comparatively less, as the instances of paper_pass used to train were less. But the mean average precision of the model makes it evident that the model has performed better.The model's performance is presented below:

TABLE II. MODEL PERFORMANCE

Execution time: 29secs per epoch			
Train best performance			
Loss	Accuracy	Precision	Recall
0.1364	0.9590	0.9629	0.9571
Validation best performance			
0.0568	0.9838	0.9838	0.9832
Test best performance			
0.0484	0.9862	0.9861	0.9846

The performance of the model are depicted in the Fig.7.

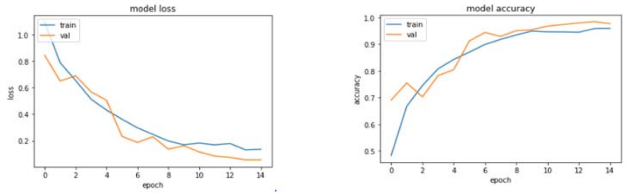

Fig. 7. Performance graphs.

When compared to the experimental approaches that were performed, YOLO outperformed other methods and techniques in terms of efficiency and performance. YOLO is a faster algorithm that was implemented with limited hardware configuration. The time taken to process during the training phase is comparatively shorter. The algorithm identifies and classifies a detected object per grid cell. Hence, it is more accurate in terms of its classification process. The architecture is more generalized, which provides scope for implementing different object classes. The confidence score, along with the prediction box and class, contribute to the work's accuracy.

V. CONCLUSION

The YOLO Algorithm was developed to identify instances of misconduct by students in examination rooms. It achieves this by accurately delineating the boundaries of the misconduct and providing a confidence score. The method exhibits enhanced efficiency by achieving quicker processing speeds and using reduced computer resources and time. It accomplishes this by accurately recognizing three distinct malpractices through the utilization of four classes. The system can be expanded to offer instantaneous resolutions, deterring dishonesty and guaranteeing equitable behavior. This automated system is a preliminary model, designed to operate within a restricted setting in order to conserve resources. Further research can expand to identify and classify other forms of misconduct, as well as perform object identification on electronic devices such as smartphones, smartwatches, and earbuds. With the growing shift towards online assessments, the algorithm may be expanded to identify instances of cheating during online exams and improve its accuracy.

Funding Statement: The authors received no specific funding for this study.

Conflicts of Interest: The authors declare that they have no conflicts of interest to report regarding the present study.

REFERENCES

[1] T S. Kumar and G. Narmatha, “Video analysis for malpractice detection in classroom examination,” in Advances in Intelligent Systems and Computing, vol. 397, pp. 135–146, 2016.

[2] D. Gowsikhaa, S. Abirami and R. Baskaran, “Automated human behavior analysis from surveillance videos: a survey,” Artificial Intelligence Revision., vol. 42, no. 4, pp. 747–765, Dec. 2014.

[3] J. Redmon, S. Divvala, R. Girshick and A. Farhadi, “You only look once: unified, real-time object detection.” Accessed: May 31, 2021.

[4] M. H. Putra, Z. M. Yussof, K. C. Lim, and S. I. Salim, “Convolutional neural network for person and car detection using yolo framework,” Journal Telecommunication Electron Computer Engineering, vol. 10, pp. 67–71, 2018.

[5] P. Ren, W. Fang, and S. Djahel, “A novel YOLO-based real-time people counting approach,” vol.1, pp. 675-689, . 2017.

[6] A. Khan, A. Sohail and U. Zahoora," A survey of the recent architectures of deep convolutional neural networks", *Artificial Intelligence Review,* pp.5455–5516, 2020 .

[7] S.K Patnaik, M S. Sidhu, Y. Gehlot, B. Sharma and P. Muthu, “ Automated skin disease identification using deep learning algorithm”, *Biomed Pharmacol J* ournal, vol.11, no.3, pp.876-889, 2018.

[8] A AA. Valliani, D. Ranti and EK. Oermann, “Deep learning and neurology: a systematic review”, *Neurol Thermal*, vol. 8, pp. 351–365, 2019.

[9] S. S. Chua, J. B. Bondad, Z. R. Lumapas and J. D. L. Garcia, "Online examination system with cheating prevention using question bank randomization and tab locking," *2019 4th International Conference on Information Technology (InCIT)*, 2019.

[10] S. V, R. A, B. Gope and S. S, "Detection of anomalous behaviour in online exam towards automated proctoring," *2021 International Conference on System, Computation, Automation and Networking (ICSCAN),* pp. 1-5, 2021.

[11] N. Soman, M. N. R. Devi and G. Srinivasa, "Detection of anomalous behavior in an examination hall towards automated proctoring," *2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, pp.1 -6, 2017.

[12] N. S. Malami, S. Thomas, U. U. Shua, Y. G. Mustapha, F. I. Onah and B. A. Jega, "Proposed iris movement tracking system to improve the security of online examinations in nigeria*," 2021 1st International Conference on Multidisciplinary Engineering and Applied Science (ICMEAS),* pp. 1-5, 2021.

[13] A. S. Kulkarni, N. E, M. Swetha and S. M. Kusuma, "Automated system for detection of suspicious activity in examination hall," *2021 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT),* pp.01-05, 2021.

[14] Anjali, P.C., Senthil Kumar .T, Lagisetty, R.K. , "Object Detection Using Mask R-CNN on a Custom Dataset of Tumbling Satellite. In: Reddy, V.S., Prasad, V.K., Wang, J., Reddy, K.T.V. (eds) Soft Computing and Signal Processing. ICSCSP 2022. Smart Innovation, Systems and Technologies, vol 313. Springer, Singapore. https://doi.org/10.1007/978-981-19-8669-7_34,2023

[15] Balamanikandan, J., Senthil Kumar.T, Maiga Chang ,"Eliminating Environmental Context for Fall Detection Based on Movement Traces. In: Reddy, V.S., Prasad, V.K., Wang, J., Reddy, K.T.V. (eds) Soft Computing and Signal Processing. ICSCSP 2022. Smart Innovation, Systems and Technologies, vol 313. Springer, Singapore. https://doi.org/10.1007/978-981-19-8669-7_31,2023

[16] 0.K.A. Dhanya, Sulakshan Vajipayajula, Kartik Srinivasan, Anjali Tibrewal, T. Senthil Kumar, T. Gireesh Kumar, Detection of Network Attacks using Machine Learning and Deep Learning Models, Procedia Computer Science,Volume 218,2023,Pages 57-66,ISSN 1877-0509, https://doi.org/10.1016/j.procs.2022.12.401.

[17] K. Suresh Kumar Patro,Vinod KumarYadav, Vidya S. Bharti , Arun Sharma,Arpita Sharma1 and T. Senthilkumar,IoT and ML approach for ornamental fish behaviour analysis,Scientific Reports 13, 21415 (2023). https://doi.org/10.1038/s41598-023-48057-w (Impact Factor : 4.997)

[18] P. Sridhar, Senthil Kumar Thangavel, Latha Parameswaran and Venkata Ramana Murthy Oruganti, "Fire Sensor and Surveillance Camera-Based GTCNN for Fire Detection System," in IEEE Sensors Journal, vol. 23, no. 7, pp. 7626-7633, 2023.