ATLANTIS PRESS

Atlantis Highlights in Computer Sciences, volume 4

Proceedings of the 3rd International Conference on Integrated Intelligent Computing Communication & Security (ICIIC 2021)

Automatic Invigilation Using Computer Vision

Manit Malhotra[1,*] Indu Chhabra[2]

[1,2] Department of Computer Science & Applications, Panjab University, Chandigarh, India
[*]Corresponding author. Email: manitmalhotra@rediffmail.com

ABSTRACT

Educational institutions determine students' strengths and weaknesses through exams. Students find numerous ways to cheat in physical exams like exchanging their sheets, using hidden notes, getting good grades, fulfilling their parents' expectations, and whatnot. Due to the physical limitations of human supervisors, typical invigilation methods cannot conduct successful exams while maintaining their integrity. An automated method based on computer vision to detect anomalous activities during exams is proposed in this study. This study centers around invigilating students' suspicious behaviour during physical exams through closed-circuit television (CCTV) cameras. The proposed method uses You Only Look Once (YOLOv3) with residual networks as the backbone architecture to inspect cheating in exams. The obtained results show the credibility and efficiency of the proposed method. The experimental results are promising and demonstrate the invigilation of the students in the examination. In this work, achieve 88.03% accuracy for the detection of cheating in the classroom environment

Keywords: *Cheating Detection, Deep Learning, Object Detection, Smart Invigilation, YOLOv3.*

1. INTRODUCTION

Exams are the most significant part of an education process. Assessment or exam is the method of seeking and interpreting the grades used by students and teachers to evaluate where the students are lying in their learning, in which areas of study they need more effort and how they get there. As more assessments or exams are taken, there are more chances to monitor and avoid the inappropriate means of cheating and preserve the integrity of the examination. Extensive developments in information and communication technologies (ICT) have impacted human life in every domain, particularly education.

Academic dishonesty and cheating have constantly been worrying factors for the educational institutes in the education system worldwide. It ruins the individuality of the students. The fact is that the students adopt different cheating approaches during the examination. There are psychological and social reasons why students opt to cheat in exams, such as parent's pressure, feeling of incompetency, want better grades, time constraints, fear of failure, and take a risk with less fear of being caught [1]. The research of Dr. Donald McCabe's and the International Centre for Academic Integrity survey and analysis states that about 68% of undergrads admit to using unfair means or cheating during exams [2]. There are various means of cheating that students adopt in exams. For example, in traditional physical exams, students use cheat sheets, writing on hands or arms, communicate to fellow students, hidden cell phones, etc.

In a traditional examination, human invigilators must be present for invigilation in the examination hall to monitor the students during exams. The more the number of students, the more invigilators are required in this system. More labor, energy, time, effort, and cost are the requisites of this system, making this traditional system burdensome. A proper invigilating system is needed to prevent cheating in examinations as it directly impacts the student's morality.

To ensure the principles and integrity of exams and to prevent cheating, a system based on computer vision is proposed in this paper. It will detect cheating by the detection of head and neck movements through a surveillance camera. It is more precise and error-free as compared to human labour. This system is better and more effective than the traditional invigilation system as it does not require as much labour, energy, effort, and time as needed in the conventional system.

In this paper, proposed a system that detects and recognizes the cheating done by the students in the classroom during exams. For the detection of the cheating implemented the Yolov3 [3] with ShuffleNet [4] as backbone architecture. In Yolov3, DarkNet-53 was used as the backbone. ShuffleNet was instigated as the backbone architecture in the proposed model. The architecture of yolov3 was changed and named it as modified yolo. The details regarding architecture were discussed in the methodology section 3. The rest of the paper is organized as follows. Literature work is discussed in section 2 of the article. In section 4, go through our results and explain what we've learned from our experiments. Finally, in section 5, come to a close with a brief conclusion and suggestions for future work.

2. RELATED WORK

Cheating during exams has become a significant challenge for academic institutions. On the one hand, where exams gauge a student's strengths and weaknesses, it plays an essential role in learning for educational institutions. In physical exams, students use unfair means to pass the examination. Students find their ways to cheat mainly due to the supervisor's negligence. Typical invigilation methods are human-dependent, time-consuming, require energy, and are not very successful in preventing students from cheating. Educational institutions need an automated and authentic way to prevent and detect different types of cheating during physical and online exams.

In the past few years, cheating and suspicious activities of students during online and physical exams have been detected multiple times by using machine learning and deep learning. Chang Liu *et al.* [5] presented a model using Spatial-temporal features to detect students' abnormal behaviour. Their proposed model successfully identified in movements like turning around, raising hands, etc. with 93.3% accuracy.

Plagiarism, another form of cheating or showing some other person's work as your own, is found by Li [6]. And designed a model using the RAE algorithm with the LSTM network and scored 79.4% accuracy in his work. Fang *et al.* [7] developed an intelligent monitoring system, applying supervised learning image recognition with an adaptive threshold to the streaming of an examination hall that divides the frames to the predetermined area range of the candidate's seat and sends a reminder to the supervisor if the predetermined limit is crossed. Haar-like features are commonly known for detecting a human face in object recognition. In [2], Adil *et al.* detects different suspicious activities in the exam. On the basis of a certain threshold, the model recognizes activities like hands in contact, sneaking the fellow student's paper. The project takes video input from a camera (CCTV) and converts it into frames. For pre-processing Gaussian Filter is used to remove background, foreground, and noise. The detection of suspicious activities is done using the Haar-like feature algorithm developed by Viola and Jones and identifies head with 70%, the hand with 72%, and face with 84% accuracy.

Working of VGG16 [8], MobileNet [9], and Inception V4 [10] models to detect cheating during examination have been compared by Kuin in [11]. For this work annotated frames from a video were passed to these models separately and showed that VGG16 and Inception V4 worked better on the dataset, they scored 96.8% and 96% accuracy than MobileNet that achieved 48.8% accuracy. Kulkarni [12], using computer vision techniques, developed a proficient examination invigilating model to detect the students' movements, poses, and expressions. The author suggested using Inception V3 for detection works better and produces a 10% less error rate instead of separately using segmentation, classification, and recognition algorithms. Nishchal *et al.* in [13] detected the cheating of students using the Openpose. The poses of the students were extracted by fetching the multiple joints of the body which increased the computation of the model and hence achieved only 63% of accuracy. As a result, system was not suitable for real time. The summary of research work shown in table 1.

Table 1. Summary of research work

Paper	Methodology	Accuracy
Automated invigilation system for detection of suspicious activities during examination [2]	Feature and AdaBoost	For face 84%, For hand 72%, For head 70%
Multi-index Examination Cheating Detection Method Based on Neural Network [6]	LSTM	79.4%
Realization of Intelligent Invigilation System Based on Adaptive Threshold[7]	EM algorithm and adaptive threshold	-
Real Time Automated Invigilator in Classroom Monitoring using Computer Vision [12]	Inception V3	Error rate less than 10% as compared to standard computer vision algorithms

Automated Cheating Detection in Exams using Posture and Emotion Analysis [13]	Open Pose	63%

Researcher around the world had done good job but there are some deficiencies that can be improved in term of accuracy, which the proposed study tries to improve. The following are the key contributions made to this research:

- A novel system has been developed that can identify and recognize students cheating in the examination. The yolov3 architecture was changed by replacing the parameters and backbone architecture.
- Generated the local dataset of invigilation of students in the examination.

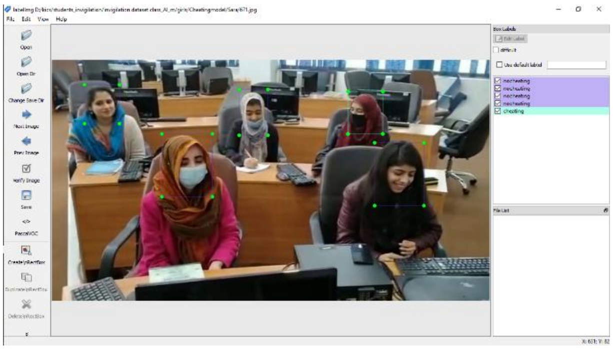
3. METHODOLOGY

3.1. Data preparation

For every deep learning, problem data is the heart and soul of the system. The dataset generated in the local environment for current work. A dataset of physical exams held in six classes was recorded/made through a camera of resolution 640 x 480 with a 25 fps of frame rate. The videos contained both normal and suspicious activities of the students. Thirty thousand frames were extracted from the recorded videos. For frame extraction used the python script with OpenCV library and extracted 5000 plus frames. The extracted frames were pre-processed by removing blurred frames from the folder manually. The dataset had 5693 labelled images after pre-processing. The sample frame cleaned dataset shown in Figure 1. Data was split into training validation and testing parts.

Figure 1 Sample Frame of dataset

The complete data was annotated with the LabelImg tool as shown in Figure 2 as supervised learning requires labelled dataset. This annotated dataset was then passed to the model for training.



labelImg D:/kics/students_invigilation/invigilation dataset class_AI_m/girls/Cheatingmodel/Sara/671.jpg

□ difficult

□ Use default label

☑ nocheating
☑ nocheating
☑ nocheating
☑ nocheating
☑ cheating

X: 631; Y: 82

Figure 2 Annotated Frame from dataset

3.2. Proposed model

Detect cheating in a classroom during exams is the aim of the problem. For this, modified yolov3 has been used. It will detect cheating by the head and neck movements of students in the examination. The proposed model considers cheating if students are looking around instead of doing their paper by detecting their neck movements. It considers students as no cheating who do not look around but do their paper as shown in Figure 3.

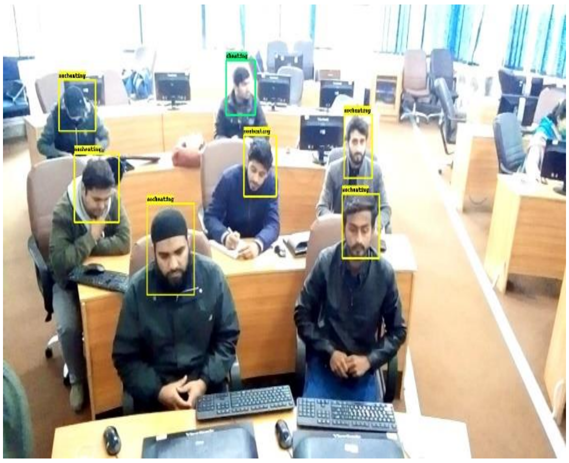


Figure 3 Automatic Cheating Detection

Yolov3 is an algorithm that directly predicts the location and class of objects, and it is a uni staged end-to-end algorithm for object detection. Compared to other object detection algorithms such as single shot detector (SSD) [14] and Faster RCNN [15], Yolov3 is very fast because of its single staged mechanism. A combination of a 1x1 convolutional layer, followed by 3x3 convolutional layers with residual networks [16], is employed as the backbone architecture of Yolov3 since the feature extraction yolov3 has many convolutional layers and parameters, which slows down the forward propagation.

To build a well-structured yolov3 object detector, a combination of ShuffleNet structure with yolov3 has been introduced to attain improved trade-off performance and efficiency [17]. The computational speed of convolution increases by ShuffleNet as the computational complexity reduces by using depth separable convolutional. There is an ability to map more channel features with less memory and computational power in our suggested object detector. ShuffleNet consists of two units, as shown in Figure 4. One unit is for the down sampling of the input, and the other one is for the sustainability of semantic information. To generate the feature map that is half in size, convolution is applied on the input in the first unit, whereas it increases the channels twice. The second unit sustains the information by split and splice of input channels without any change in output features. Both units of architecture manage and exchange the feature channels. The duplication and loss of information result from the splitting of channels reduced by this arrangement and the exchange of feature depth. The additional layer of the suggested detector is altered according to the original yolov3 [18-23].

There is considerable depletion in computational complexity by reducing the number of convolutional layers, whereas multiscale prediction is preserved. The backbone of the proposed object detector rested on the ShuffleNet. While in yolov3 the backbone architecture had a large number of training parameters as many convolution layers were used as compared to ShuffleNet architecture. Each unit contains multiple of convolution layers [24-28]. The red layer represents the 1x1 size of convolution layers, grey represent the 3x3 size of convolution layers and then yellow layer concatenate the these. Replacing the yolov3 backbone architecture with ShuffleNet made the calculation faster as a smaller number of layers in ShuffleNet. The back architecture of proposed model is shown in Figure 4. The ShuffleNet used with the Yolov3 for the detection of the cheating and no cheating. In yolov3 53 layered backbone were used which add up the additional parameter on network. By adding the ShuffleNet the trainable parameters were reduced as a smaller number of parameters utilized. The modification of the yolov3 was made by reducing the number of layers and named our architecture modified yolo [29].

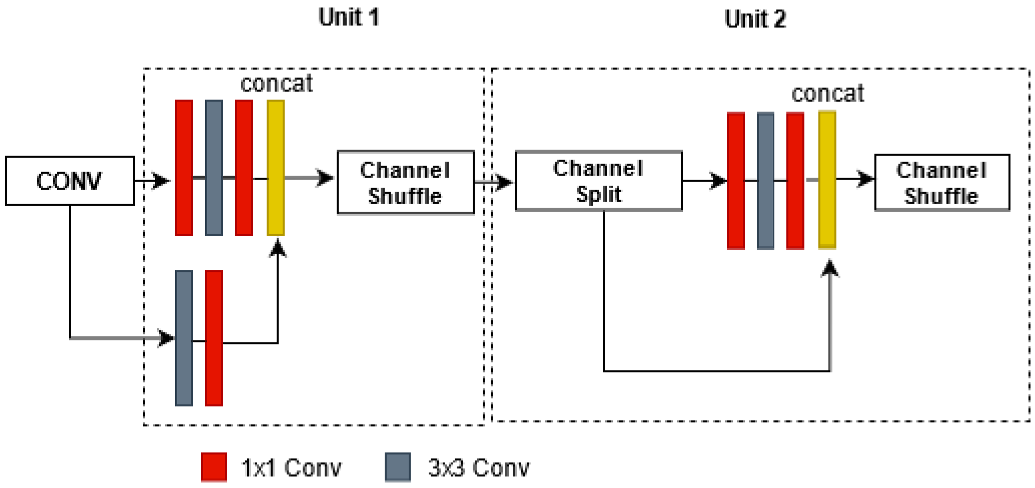


Figure 4 ShuffleNet network convolution block

Some asymmetrical performance was also observed in original yolov3 apart from efficiency. The localization of objects by bounding box position in yolov3 is somewhat inconsistent. Parameters of the model constantly update by backpropagation throughout the training of CNN by increasing the accuracy and minimizing the loss [30-31]. The loss function of original yolov3 comprises three parts: the bounding box prediction loss L bb, the confidence loss L conf, and the class prediction loss Lc. The distance between the detected and ground truth bounding boxes is measured as a mean squared error (MSE) to anticipate the loss of the bounding box. Yolov3 doesn't consider the intersection of union (IoU) that is significant, but, in our problem, it is concerned as for risk assessment, the depth of detected ROI has to be estimated. There may be different IoU of two bounding boxes if there is the same L2norm distance. It is required to incorporate IoU loss into the loss function. The formula of IoU for the calculations is given in equation 1.

$$IoU = |B^p \cap B^g| \div |B^p \cup B^g| \qquad (1)$$

Where B^p and B^g denoted as predicted bounding box and ground truth bounding box, respectively, IoU can remain the same while ground and predicted boxes are in distinct overlapping states. The generalized GIoU has been used for our detector as an optimized loss function to reduce the issue of standard IoU in equation 2.

$$GIoU = IoU - |C \backslash (B^p \cap B^g)| \div |C| \qquad (2)$$

Here C is the box containing both B^p and B^g with minute size, and IoU is the standard intersection over the union. The relation of both bounding boxes when they do not coincide can now be represented in relative terms by GIoU. Boxes are having the same IoU, but different overlapping states can be seen in Figure 5. The right state has a smaller GIoU value as compared to the left state. GIoU refers to the disarrangement between both bounding boxes. The loss function is given in equations 3 and 4.

$$Loss = L\ conf + L\ c + Lbb \tag{3}$$
$$L\ bb = 1 - GIoU \tag{4}$$

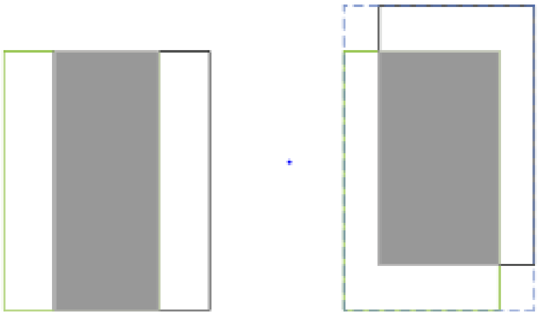


Figure 5 Two overlapping cases having the same IoU

4. EXPERIMENTS AND RESULTS

4.1. Results and discussion

The proposed model was implemented using python with Tensor Flow on core i7 with Nvidia 1080 ti system, and trained with a 0.0002 learning rate, batch size 1, 100000 epochs, and momenta of 0.091. The training began with a loss of 4.02323, which gradually reduced to 0.0023. The model was trained on 75 percent of a total dataset with 15 percent data for validation and 10 percent for testing. It achieved 88.03% accuracy on our dataset. The same dataset was trained on YOLOv3 with transfer learning and scored 82.06% accuracy. The proposed model has done well as compared to typical YOLOv3 both in terms of time and accuracy. Table 2 shows the detection results of the prepared dataset trained on two deep learning models. The average precision (AP) of both models indicates that the proposed model outperformed YOLOv3 with 88.03% accuracy in detecting cheating. The proposed network of YOLOv3 with ShuffleNet does well as its less complexity by using depth separable convolutional compared to the original YOLOv3.

Table 2. Average Precision of Models

Model	**Dataset**	**AP**
YOLOv3	Own Dataset	82.06%
Proposed	Own Dataset	88.03%

Average precision (AP), precision, and recall gauged the model performance in detecting cheating during exams. AP shows the model to object sensitivity and also the global performance measurement of the model. To calculate the AP using equation 5.

$$AP = \int_0^1 P(r)dr \tag{5}$$

4.2. Comparative analysis

For detecting the student's cheating in **[17]** used the Hidden Markov Model for temporal analysis of the head poses. The overall pose accuracy of their system was 79.8%. While the proposed model achieved the 88.03% for detecting the abnormal poses of students in the examination. In Figure 6, the graphs show the accuracy of the proposed model and yolov3. Both the models were trained using our generated data with the same training parameters. Proposed model outperformed the tradition yolov3 during training process at mostly epochs it has more accuracy. The proposed model has more accuracy as compared to the other one. Yolov3 has more trainable parameters as compared to the proposed one as a result it took more time to training.

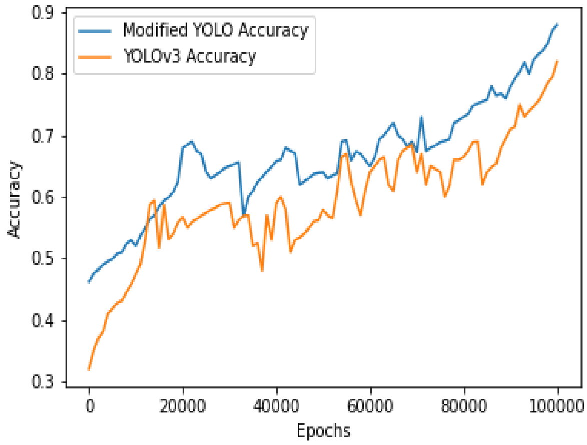


Figure 6 Accuracy Comparison Graph

4.3. Qualitative Analysis on dataset

Three annotators obtained the ground truth data manually by labeling the poses of the students as "cheating" and "no cheating". The extracted frames of five videos were initially considered normal and were closely observed in the examination context for detecting abnormal behavior. If any annotator labeled a frame as cheating, the final ground truth will be constructed considering it as cheating; justifying the videos must cover abnormal activities. Among five videos, a total of 1800 plus frames were extracted containing varying pose distributions from 6 videos. The pose distribution as "cheating" and "no cheating" in frames show variation among videos in terms of percentage as well, ranging from 72% to 88 % in case of modified yolo. Table 3 shows the detection of the cheaating and no cheating in the different videos using proposed approach and yolov3.

Table 3. Video Data Statistics on proposed model

Video	Frames	Modified Yolo (%)	Yolov3 (%)
1	330	78.33	75.29
2	230	72.42	69.36
3	246	77.91	78.88
4	311	88.03	79.71
5	389	86.54	82.06
6	321	87.12	80.14

5. CONCLUSION

This research proposed a novel model for students' invigilation in the examination using the deep learning and computer vision approach. In this work, implemented the YOLOv3 with ShuffleNet for automatic invigilation. The detection of the cheating was done on the base of the neck and head movement of the students. The data set was generated in a local environment for the experiments. The comparison was also made with the proposed model results with existing literature and as shown in the experiment section. The results show that proposed model achieved more accuracy as compared to other one. In future work, the system will also able to detect the other methods of cheating like exchanging sheet, wisping and gesture detection.

REFERENCES

[1] Ghizlane, Moukhliss, Belhadaoui Hicham, and Filali Hilali Reda. "A New Model of Automatic and Continuous Online Exam Monitoring." In 2019 International Conference on Systems of Collaboration Big Data, Internet of Things & Security (SysCoBIoTS), pp. 1-5. IEEE, 2019.

[2] Adil, Md, Rajbala Simon, and Sunil Kumar Khatri. "Automated invigilation system for detection of suspicious activities during examination." In 2019 Amity International Conference on Artificial Intelligence (AICAI), pp. 361-366. IEEE, 2019.

[3] Redmon, Joseph, and Ali Farhadi. "Yolov3: An incremental improvement." arXiv preprint arXiv:1804.02767 (2018).

[4] Zhang, Xiangyu, Xinyu Zhou, Mengxiao Lin, and Jian Sun. "Shufflenet: An extremely efficient convolutional neural network for mobile devices." In Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 6848-6856. 2018.

[5] Liu, Chang, Hao Zhou, Huan-Chen Xu, Bao-Yu Hou, and Lan Yao. "Abnormal Behavior Recognition in an Examination Based on Pose Spatio-Temporal Features." In 2020 IEEE 5th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA), pp. 380-386. IEEE, 2020

[6] Li, Zhizhuang, Zhengzhou Zhu, and Teng Yang. "A Multi-index Examination Cheating Detection Method Based on Neural Network." In 2019 IEEE 31st International Conference on Tools with Artificial Intelligence (ICTAI), pp. 575-581. IEEE, 2019.

[7] Fang, Yunjie, Jingcheng Ye, and Haoyu Wang. "Realization of Intelligent Invigilation System Based on Adaptive Threshold." In 2020 5th International Conference on Computer and Communication Systems (ICCCS), pp. 201-205. IEEE, 2020.

[8] Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." arXiv preprint arXiv:1409.1556 (2014).

[9] Howard, Andrew G., Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. "Mobilenets: Efficient convolutional neural networks for mobile vision applications." arXiv preprint arXiv:1704.04861 (2017).

[10] Szegedy, Christian, Sergey Ioffe, Vincent Vanhoucke, and Alex Alemi. "Inception-v4, inception-resnet and the impact of residual connections on learning (2016)." arXiv preprint arXiv:1602.07261 (2016).

[11] Kuin, Aiman. "Fraud detection in video recordings of exams using Convolutional Neural Networks." (2018).

[12] Kulkarni, Rutuja. "Real Time Automated Invigilator in Classroom Monitoring Using Computer Vision." In 2nd International Conference on Advances in Science & Technology (ICAST). 2019.

[13] Nishchal, J., Sanjana Reddy, and Priya N. Navya. "Automated Cheating Detection in Exams using Posture and Emotion Analysis." In 2020 IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT), pp. 1-6. IEEE, 2020.

[14] Liu, Wei, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, and Alexander C. Berg. "Ssd: Single shot multibox

detector.” In European conference on computer vision, pp. 21-37. Springer, Cham, 2016.

[15] Ren, Shaoqing, Kaiming He, Ross Girshick, and Jian Sun. “Faster r-cnn: Towards real-time object detection with region proposal networks.” arXiv preprint arXiv:1506.01497 (2015).

[16] He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. “Deep residual learning for image recognition. CoRR abs/1512.03385 (2015).” (2015): 646-661.

[17] Cote, Melissa, Frédéric Jean, Alexandra Branzan Albu, and David Capson. “Video summarization for remote invigilation of online exams.” In 2016 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1-9. IEEE, 2016.

[18] K. Yu, L. Tan, L. Lin, X. Cheng, Z. Yi and T. Sato, "Deep-Learning-Empowered Breast Cancer Auxiliary Diagnosis for 5GB Remote E-Health," IEEE Wireless Communications, vol. 28, no. 3, pp. 54-61, June 2021, doi: 10.1109/MWC.001.2000374.

[19] Rajendran, Ganesh B., Uma M. Kumarasamy, Chiara Zarro, Parameshachari B. Divakarachari, and Silvia L. Ullo. "Land-use and land-cover classification using a human group-based particle swarm optimization algorithm with an LSTM Classifier on hybrid pre-processing remote-sensing images." Remote Sensing 12, no. 24 (2020): 4135.

[20] Subramani, Prabu, Ganesh Babu Rajendran, Jewel Sengupta, Rocío Pérez de Prado, and Parameshachari Bidare Divakarachari. "A block bi-diagonalization-based pre-coding for indoor multiple-input-multiple-output-visible light communication system." Energies 13, no. 13 (2020): 3466.

[21] Rajendrakumar, Shiny, and V. K. Parvati. "Automation of irrigation system through embedded computing technology." In Proceedings of the 3rd International Conference on Cryptography, Security and Privacy, pp. 289-293. 2019.

[22] K. Yu, L. Tan, S. Mumtaz, S. Al-Rubaye, A. Al-Dulaimi, A. K. Bashir, F. A. Khan, “Securing Critical Infrastructures: Deep Learning-based Threat Detection in the IIoT”, IEEE Communications Magazine, 2021.

[23] Seyhan, Kübra, Tu N. Nguyen, Sedat Akleylek, Korhan Cengiz, and SK Hafızul Islam. "Bi-GISIS KE: Modified key exchange protocol with reusable keys for IoT security." Journal of Information Security and Applications 58 (2021): 102788.

[24] Nguyen, Tu N., Bing-Hong Liu, Nam P. Nguyen, and Jung-Te Chou. "Cyber security of smart grid: attacks and defenses." In ICC 2020-2020 IEEE International Conference on Communications (ICC), pp. 1-6. IEEE, 2020.

[25] Pham, Dung V., Giang L. Nguyen, Tu N. Nguyen, Canh V. Pham, and Anh V. Nguyen. "Multi-topic misinformation blocking with budget constraint on online social networks." IEEE Access 8 (2020): 78879-78889.

[26] Z. Guo, Y. Shen, A. K. Bashir, M. Imran, N. Kumar, D. Zhang and K. Yu, “Robust Spammer Detection Using Collaborative Neural Network in Internet of Thing Applications”, IEEE Internet of Things Journal, vol. 8, no. 12, pp. 9549-9558, 15 June15, 2021, doi: 10.1109/JIOT.2020.3003802.

[27] L. Tan, H. Xiao, K. Yu, M. Aloqaily, Y. Jararweh, “A Blockchain-empowered Crowdsourcing System for 5G-enabled Smart Cities”, Computer Standards & Interfaces, https://doi.org/10.1016/j.csi.2021.103517

[28] C. Feng et al., "Efficient and Secure Data Sharing for 5G Flying Drones: A Blockchain-Enabled Approach," IEEE Network, vol. 35, no. 1, pp. 130-137, January/February 2021, doi: 10.1109/MNET.011.2000223.

[29] N. Shi, L. Tan, W. Li, X. Qi, K. Yu, “A Blockchain-Empowered AAA Scheme in the Large-Scale HetNet”, Digital Communications and Networks, https://doi.org/10.1016/j.dcan.2020.10.002.

[30] Arun, M., E. Baraneetharan, A. Kanchana, and S. Prabu. "Detection and monitoring of the asymptotic COVID-19 patients using IoT devices and sensors." International Journal of Pervasive Computing and Communications (2020).

[31] Kumar, M. Keerthi, B. D. Parameshachari, S. Prabu, and Silvia liberata Ullo. "Comparative Analysis to Identify Efficient Technique for Interfacing BCI System." In IOP Conference Series: Materials Science and Engineering, vol. 925, no. 1, p. 012062. IOP Publishing, 2020.