[1]**ADAPTIVE ONLINE LEARNING PLATFORM TO ENHANCE PRIMARY EDUCATION**

**Attention monitoring system for primary students’ behavior recognition.**

Project ID: TMP-2023-24-91

**Project Proposal Report**

Samarakkody G.K.G.D.A – IT20619558

Supervisor – Mr. Samadhi Rathnayake.

Co-Supervisor – Dr. Dharshana Kasthurirathna.

BSc (Hons) in Information Technology Specializing in Data Science

Department of Information Technology

Sri Lanka Institute of Information Technology (SLIIT)

August 2023

**Declaration of the candidate & supervisor**

We declare that this is our own work, and this proposal does not incorporate without acknowledgement any material previously submitted for a degree or diploma in any other university or Institute of higher learning and to the best of our knowledge and belief it does not contain any material previously published or written by another person except where the acknowledgement is made in the text.

|  |  |  |
| --- | --- | --- |
| Name | Student Id | Signature |
| Samarakkody G.K.G.D.A | IT20619558 | **A close-up of a black background** |

The supervisor/s should certify the proposal report with the following declaration.

The above candidates are carrying out research for the undergraduate

Dissertation under my supervision.

Supervisor: Mr. Samadhi Rathnayake.

Signature: Date: 2023-08-25

Co-supervisor: Dr. Dharshana Kasthurirathna.

Signature: A signature on a white surface

Description automatically generated Date: 2023-08-25

**Abstract**

Online primary education has been increasingly popular in recent years. However, staying focused and involved is one of the most difficult aspects of online primary education. This is when attention-monitoring systems come into play. Attention-monitoring systems can be used to track students' engagement in online learning and identify pupils who are in danger of disengagement. We propose an attention-monitoring system using machine learning and deep learning approaches in this research proposal. A Generative Adversarial Network (GAN) will be used in our system to build a big dataset of student attention data. On this dataset, we will then train a model with a few-shot learning approach. This algorithm will be able to detect kids who are at risk of getting disengaged and will provide interventions to teachers to help these pupils stay focused and engaged. We feel that our suggested method has the potential to increase the quality of online primary education by assisting students in remaining focused and engaged. We intend to test our technique in a pilot study with a small group of primary kids. We will collect data on student participation and utilize it to evaluate our system's performance. If our approach is successful, we intend to make it available to other instructors and schools. We believe our method has the potential to significantly improve the quality of online primary education.

**Table of Contents**

[1. Introduction 5](#_Toc143898902)

[1.1 Background 5](#_Toc143898903)

[1.2 Literature survey 6](#_Toc143898904)

[1.3 Research Gap 9](#_Toc143898905)

[2. Research Problem 11](#_Toc143898906)

[3. Objectivities 12](#_Toc143898907)

[3.1 Main objective 12](#_Toc143898908)

[3.2 Sub objectives 12](#_Toc143898909)

[4. Methodology 13](#_Toc143898910)

[4.1 System Diagram 14](#_Toc143898911)

[5. Technology Selection 15](#_Toc143898912)

[6. Ability of commercialization 16](#_Toc143898913)

[7. Work breakdown plan. 17](#_Toc143898914)

[8. References 18](#_Toc143898915)

# 1. Introduction

## 1.1 Background

Online primary education has gained popularity due to its flexibility, convenience, and accessibility. However, staying focused and engaged in online learning can be challenging due to factors such as students' lack of familiarity with online learning, distractions, and motivation. Attention-monitoring systems can help track students' engagement and identify those at risk of disengagement. These systems use computer vision techniques to track students' eye movements, facial expressions, and head movements, generating real-time attention measurements. The proposed work aims to develop an attention-monitoring system using machine learning and deep learning techniques. The system will use a Generative Adversarial Network (GAN) to generate a large dataset of student attention data, and a few-shot learning algorithm to train a model on the dataset. The model will identify students at risk of disengagement and provide interventions to help them stay focused and engaged. The proposed system has the potential to improve the quality of online primary education by helping students stay focused and engaged. A pilot study with a small group of primary students will evaluate the system's performance, and if successful, it will be made available to other teachers and schools, potentially significantly impacting the quality of online primary education.

## 1.2 Literature survey

The attentiveness of primary students is one of the most important factors in online Learning. This research component will recognize the behavior of the student through an automated attention monitoring system. There are several research works carried out regarding the students’ attention monitoring.

According to “Machine learning applied to students' attentiveness detection using emotional and non-emotional” research by M. Elbawab, R. Henriques. They propose a machine-learning model that estimates students' attentiveness during e-learning classes only using a webcam. Combining physical and emotional facial features extracted from videos of students. Those videos are processed using drowsiness and head pose detectors as well as converting the videos into images representing one frame per second. Then the emotion detector is applied to the images and manual annotation is performed to detect the attentiveness level of the students. The drowsiness detector used OpenCV a computer vision library, to analyze the students' webcam footage and extract drowsiness features. The emotions detector employed a VGGnet, a deep learning model, to classify the students' emotions based on their facial expressions. The study utilizes a total of eleven variables. Including eye aspect ratio (EAR), yawn aspect ratio (YAR), head pose, and emotional states, to build a feature set characterizing a student's physical and emotional state based on their face. To estimate individual students’ attention levels, machine learning algorithms such as decision trees, random forests, support vector machines (SVM), and extreme gradient boosting (XGBoost) are used. Using those algorithms, the model achieved an average accuracy of 80.52% and an AUROC OVR of 92.12%. The study concludes that a machine learning model based on a combination of physical and emotional measures can accurately estimate students' attentiveness during e-learning classes using only a webcam. However, there are some limitations. The ground truth data of human annotation may not be completely reliable and is dependent on the observer. The training data set is limited to only seven student videos, which may limit the model's generalizability.

The research was done by Rahul RK, Shanthakumar S, Vykunth P, Sairamnath K is about real-time attention span tracking in online education. The paper discusses the recent coronavirus outbreak's increased demand for online education, as well as the need for an automated mechanism to monitor student activities during online classes. It emphasizes the problem of students losing concentration and engaging in activities such as starting a lecture and walking away or using proxies for online tests, which affects their knowledge level and results in low-skilled laborers. The paper proposes a mechanism that uses camera feed and microphone input to monitor students' real-time attention levels during online classes. As proposed methods, facial recognition is used to validate student attendance. Then calculate the attention span score blink rate detection, Emotion classification, eye gaze tracking, background noise, and body posture. The Viola-Jones algorithm is used to detect facial landmarks by scanning images to identify features of human faces and extracting 68 key points from the detected face image using OpenCV's dlib library. Eye-gaze tracking is a module that analyses eye movement to track the direction of a person's gaze. It is used as a parameter in calculating students' attention scores in online classes. The blink rate detection module computes the user's blink rate by dividing each eye in half and calculating the Eye Aspect Ratio (EAR) using Euclidean distances. It determines whether the eyes are open or closed and notifies the user if the eyes are closed for more than a certain amount of time. Emotion classification is categorized after determining the student's emotional state using facial landmarks and features extracted from the detected face image. For the body posture estimation, they have used TensorFlow pose-estimator (PoseNet), based on Mobilenet SSD to estimate the posture of the student. PoseNet uses heat maps to estimate pose differences, identifying 17 key points and assigning pixel similarity scores to predict student restlessness or focus during online lectures. The background noise is detected using PyAudio and calculates average noise levels every 5 seconds, indicating student concentration and noise levels. Overall Attention level detection calculates students' attention scores in online classes using non-verbal features, providing real-time feedback for students and organizations, and analyzing performance and teaching standards. The system evaluated 15 undergraduate students, comparing predicted attention scores with observed scores. Live graphs show predicted attention levels and real-time scores. The attention-tracking model is performed with minimal data, achieving an overall accuracy of 84.6233%.

This research article is about a real-time attention-monitoring system for classrooms using deep deep-learning approach for students' behavior recognition by Z Trabelsi, F Alnajjar, M M A Parambil, M Gochoo, and L Ali. This paper introduces an intelligent real-time vision-based classroom system that monitors students' emotions, attendance, and attention levels, even when wearing face masks. In the proposed system, the YOLOv5 deep learning models are used for object identification and behavior recognition. Different versions of the YOLOv5 model (v5m, v5n, v5l, v5s, and v5x) are tested and contrasted using various evaluation metrics like precision, recall, mAP, and F1 score. Deep learning models are trained with YOLOv5 pre-learned weights to improve performance in behavior recognition and attentiveness assessment. Multiple versions of the YOLOv5 model are assessed, and the findings reveal good performance with an average accuracy of 76%. The proposed model allows teachers to visualize students' behavior and emotional states, allowing them to handle teaching sessions more effectively. The technology creates live reports for instructors, offering real-time feedback on student activities, attention, and emotions. The proposed strategy improves instructor effectiveness while also improving students' academic experience. The proposed model uses student action datasets from web and classroom sources, utilizing search engines and databases.

X. Ling, J. Yang, J. Liang, H. Zhu, and Hui Sun, propose a deep learning-based method for the analysis of students' attention in offline class. Initially, the classroom is divided into different periods and categorized them into 4 states such as lecturing, interaction, practice, and transcription. then deep learning methods are used to detect faces and extract head-pose parameters, while automatic speech recognition models determine class states. Use Retinaface for face detection, which produces good results and is fast, and use Vision Transformer as the head pose regression model to estimate the head pose of students. Used sound type clarification model and a speech recognition model to identify different states in the class. A multi-model analysis is performed by merging the class state information with head pose parameters to determine whether each student's specific head pose is engaged in learning. This proposed method accurately calculates the learning attention of each student by combining the class-state sequence and individual head-pose parameters. The paper emphasizes the method's potential for use in high school and university classrooms and laboratories.

Experimental work in the paper is divided into two parts such as model training and model inference, which are completed on various software and hardware platforms. The video and audio development platforms were Pytorch 1.7.1 and PaddlePaddle, both using the Python programming language. The portable inference system made use of an NVIDIA Xavier AgX 32 g high-performance edge terminal, an Ubuntu 18.04 system, minicanda3, Pytorch 1.7.1, and PaddlePaddle.Video and audio signals were collected using a 1080p high-resolution industrial camera and a high-sensitivity pickup head. However, the paper does not explicitly mention numerical results and statistical analysis conducted in the paper.

## 1.3 Research Gap

Most of the Existing research on attention-monitoring systems has focused on adult students in traditional classroom environments. It’s challenging when developing these attention monitoring systems for primary students. Most adult students are familiar with online quizzes and assessments. But when compared to primary students they haven’t been very much engaged with activities like online quizzes and assessments, it's more important to recognize the behavior of the students to check whether they are paying attention to those quizzes and assessments or not.

While we are mainly focusing on online e-learning some of the research is focused on offline classroom environments. In the physical learning environment, they have divided the time durations into different states like lecturing, interaction, practice, and transcription. Then the learning attention of each student is detected using video and audio information which are taken from IoT devices based on their head pose parameter and class state sequence. This proposed work provides insights for teachers to improve their teaching methods and has potential applications in schools and smart education programs.

One of the significant drawbacks of online e-learning is when compared to the traditional classroom the teacher cannot monitor the student attentiveness accurately. Our proposed work will provide a solution for that as well. One of the research projects that we reviewed on estimating students' attentiveness during e-learning classes using a machine learning model based on physical and emotional measures using only a webcam. However, the study collected videos from seven students for the development and analysis of the machine-learning model. There is a limitation in their proposed work which is the size of the training data set is limited to seven students’ video data. If the model is trained on a dataset. The data of that dataset is collected from a specific population the model may not be able to generalize well for other population data. Most of the research purposes for collecting more relevant data can be considered as a drawback. In our proposed work we are mainly focusing on training a machine-learning model for a small amount of data using Few hot learning algorithms without affecting the generalizability of the model.

One of the research projects that I reviewed provides a mechanism to use a camera feed and microphone input to monitor the real-time attention level of the students during online classes address issues like distractions and improve the quality of online learning. They use various image-processing techniques and machine-learning algorithms to calculate attention scores based on non-verbal features. In their proposed work in performance evaluation of the research project, they have identified as future work, they are going to improve the system performance by training the model using more data. In our research work, we are going to specifically focus on improving the performance of the attention monitoring system even though there is a lack of training data.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Features | Research  [1] | Research  [2] | Research  [3] | Research  [4] | Proposed  work |
| Behavior recognition through attention monitoring | **yes** | **yes** | **yes** | **yes** | **yes** |
| Attention monitoring for primary students (age 7 -10) | **no** | **no** | **no** | **no** | **yes** |
| Train model for large amounts of data | **no** | **yes** | **no** | **no** | **yes** |
| System performance is high for a small amount of trained data. | **no** | **no** | **-** | **no** | **yes** |
| Online attention monitoring | **yes** | **no** | **no** | **yes** | **yes** |

Table 1 – Research Gap

# 2. Research Problem

The COVID-19 pandemic forced most schools worldwide to move to online learning. When compared to the adult learners for primary students this was a challenge because they are not very familiar with online learning as adult learners. One of the biggest challenges in online primary education is staying focused and engaged in online learning. This is where attention-monitoring systems come from. In our proposed work our attention monitoring system is mainly focusing on recognizing the behavior of the primary student through attention monitoring.

Through online e-learning platforms, primary students will be given online quizzes, and assessments to enhance their personalized skills. Therefore, the teachers/ parents have a responsibility to monitor students' engagement through online learning by checking the attentiveness of the students. In online learning, most teachers have difficulties checking the students' engagement with the activities they are doing for the students.

This attention-monitoring system can be effective for teachers in recognizing the students' behavior. As data science students we are going to propose an attention-monitoring system based on machine learning and deep learning techniques. Through the literature review, we have identified that several research work projects have gone through this topic. However, in most of the research we have identified that they are having difficulties with finding a large data set to train their machine-learning model. Collecting a large data set is not an easy task and it is time-consuming. While we are training our model for a large data set that is generated using a Generative Adversarial Network (GAN) we are going build a model using Few shot learning algorithms which helps to train the model on small data sets and improve the performance of the proposed system.

# 3. Objectivities

## 3.1 Main objective

* This research project aims to develop an adaptive online learning platform for primary students to enhance their primary education.
* The main objective is to develop an attention-monitoring system for primary students to recognize their behavior.

## 3.2 Sub objectives

* Monitoring the attention of the students while they are engaged with the online quizzes and the assignments.
* Identify the attentiveness of the students after doing the image processing according to the measurements such as eye gaze tracking, head pose detection, and drowsiness.
* Improve the performance of the model by generating new data using existing data. While training the model for the new data set. Analyze the performance of the model using few shot learning algorithms for a small amount of existing data.
* Monitor the student's attention and provide the attention level prediction for the teacher whether that student is engaged with the work or not.

# 4. Methodology

1. Data collection.

* While the student is engaged with the online quiz or the assessments using a webcam, images are taken from each student and create image dataset.
* And store them in the database.

1. Data Preprocessing.

* Before developing the model, data preprocessing must be done. There can be several preprocessing techniques to apply.
* Image resizing should be done to get images into common size. That could be important for training the model and ensuring all the images are of the same size and feature extraction can be used to extract the most important features from the images.

1. Data annotation.

* Since there is a small amount of existing data, manual annotation can be more efficient and accurate.
* While annotating the image data non-verbal features like eye gaze (looking left, looking right, or looking forward), drowsiness, and head pose (looking up looking down, looking left, and looking right).
* The images are annotated as attention-giving, not giving attention or neutral.

1. Training and fine-tuning the model.

* To implement the model, we specifically use Few shot learning algorithms.
* As we have identified, having a small amount of training data affects the generalizability of the model.
* To check the performance of the model with the large dataset we use a generative adversarial network (GAN) to generate new data using existing data.
* Few shot models can learn from a small amount of label data and perform well. Therefore, we use Few shot learning algorithms like MAML (meta-learning algorithm) and prototypical networks.
* We should choose the model with high accuracy and fine-tune the model using hyperparameters.

1. Testing and evaluating the model.

* Once the model is trained. We should test the model using validation set data.
* To evaluate the performance of the model, Metrics like F1 score, precision, and recall are used. Cross-validation strategies can also be used.

1. Deploying the model.

* After deploying the model, the images are taken from students using a webcam while they are engaged with the web platform.
* Those images can feed into the model and the teacher should be able to recognize the attentiveness of the students whether that student is paying attention, not paying attention or natural.

## 4.1 System Diagram

**A diagram of a data processing process

Description automatically generated**

Figure 1 : High level system architecture

# 5. Technology Selection

For the development of the proposed work, the following technologies are required. And those technologies are expected to be used.

1. Python – which is an open-source and well-documented language. It is a general-purpose language and versatile language for model building.
2. REST API - enabling the backend server and web application's communication.
3. Google Colab – Jupyter Notebook environment that is hosted in google cloud platform.
4. Visual studio code.

**Frameworks**

* OpenCV: OpenCV is a well-known computer vision library for extracting features from images**.**
* TensorFlow is a popular machine-learning library for training and deploying machine-learning models**.**
* PyTorch: PyTorch is a popular machine learning library for training and deploying machine learning models**.**
* Scikit-learn is a well-known machine learning library for training and evaluating machine learning models.

**Libraries**

* NumPy is a well-known scientific computing package that may be used to manipulate arrays.
* Pandas is a popular data analysis package that can read and write data from databases.
* Matplotlib: Matplotlib is a popular data-plotting library that may be used to visualize experiment findings.

# 6. Ability of commercialization

* We believe that our proposed system has the potential to be a successful project.
* The system can be marketed directly to schools and teachers, allowing them to select the best solution for their specific needs.
* Partnerships with other educational technology businesses can help you reach a larger audience and provide a more comprehensive product.
* Licencing allows companies to create their own products and services based on the system, whereas franchising allows other companies to utilize the system's name and brand to offer their own products and services.

# 7. Work breakdown plan.

Test the model using validation data and evaluate the performance of the model and deploy the model

Train the model and hyperparameter tuning to improve accuracy of the model

Analyze the data and process the data before training the model and identifying suitable algorithms

Data collection while detecting the attentiveness of the students

develop an attention-monitoring system for primary students to recognize their behavior.

# 8. References

|  |  |
| --- | --- |
| [1] | T. Zhang, A. El Ali , A. Hanjalic and P. Cesar , "Few-shot learning for fine-grained emotion recognition using physiological signals," *IEEE Transactions on Multimedia,* 2022. |
| [2] | Z. Trabelsi , F. Alnajjar , M. Parambil, M. Gochoo and L. Ali , "Real-Time Attention Monitoring System for Classroom: A Deep Learning Approach for Student’s Behavior Recognition," *Big Data and Cognitive Computing,* 2023. |
| [3] | X. Ling , J. Yang , J. Liang , H. Zhu and H. Sun , "A Deep-Learning Based Method for Analysis of Students’ Attention in Offline Class.," *Electronics,* 2022. |
| [4] | M. Elbawab and R. Henriques , "Machine Learning applied to student attentiveness detection: Using emotional and non-emotional measures," *Education and Information Technologies,* pp. 1-21, 2023. |
| [5] | N. Kavyashree , L. K. Shailaja , J. Anitha and P. Sindhu , "Real-time Attention Span Tracking in Online Education," 2022. |