MTH2245
Supervised by: Dr. Samah El-Shafiey

Deep Learning in Students' Disengagement Detection

Authors
Mostafa Ali
Shahd Ahmed
Mohamed Hisham
Nouran Hani
Ahmed Al-Deeb
Meram Mahmoud

Abstract

Students' disengagement is one of the main challenges faces **online learning** especially after its rapid growth since Covid-19. This research proposes a system that detects students' **disengagement** in real-time. This system depends on using **deep learning** to analyze facial expressions and recognize signs of **disengagement** like yawning or **drowsiness**. Two models, **VGG16** transfer learning and **Facial Landmarks** neural network, were compared. The **VGG16** transfer learning model performed better achieving **93.64%** total accuracy and is used in our real-time **disengagement** detection system.

Problem definition

Disengagement, which is the lack of participation in **online classes**, can lead to a poor understanding of the material and even cause students to drop out of the course. Teachers face difficulties in tracking and monitoring their students in virtual settings. Therefore, automatic assessment of **disengagement** during online sessions is necessary in solving the **disengagement** problem.

Literature review

- We are concerned about the automatic detection of **student disengagement** using **Deep Learning** models.
- Out of 272 search results only 37 study followed our inclusion and exclusion criteria and was examined.
- Most approaches of **disengagement detection** are face dependent and divided according to features into:

1. Part based: eyes and mouth. **2. Appearance based:** the entire face.

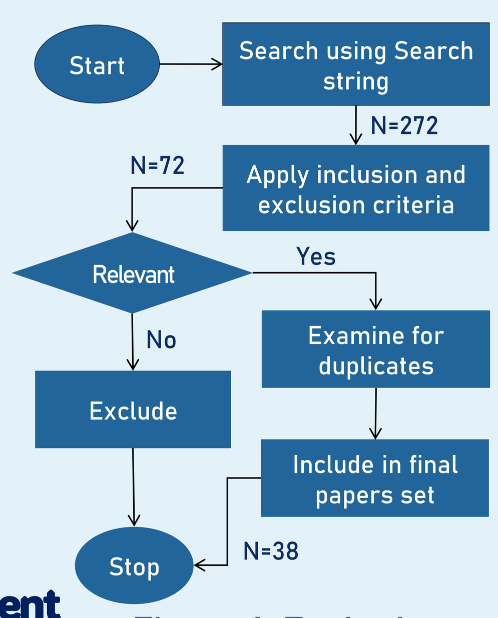


Figure 1: Exclusion Criteria

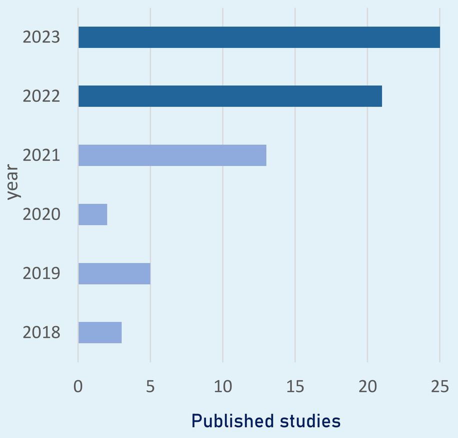


Figure 2: Number of Studies Published Last 5 Years

Data description

The dataset is organized into three primary categories: open-source datasets, images from various open-source databases, and a distinctive dataset developed by the authors.

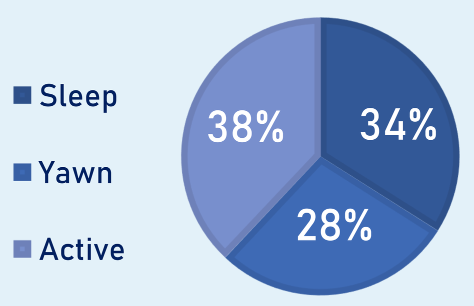


Figure 3: Categories

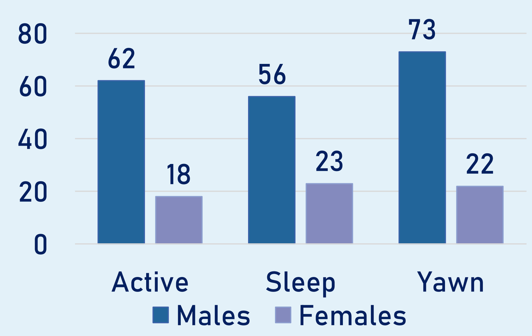


Figure 4: CUFE Data

Math modeling

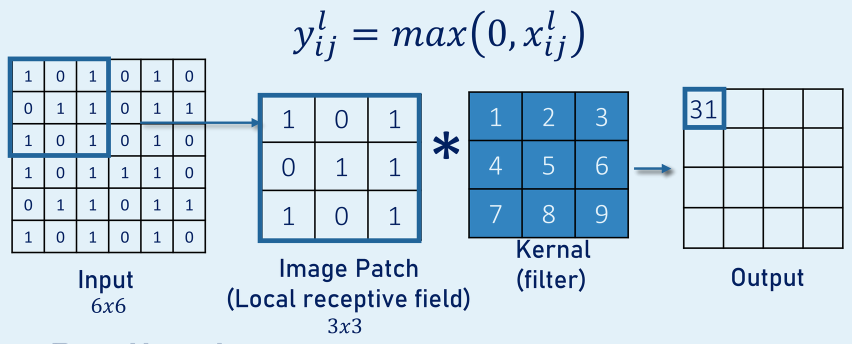
VGG16 is a pre-trained **CNN** model that imitates how the visual cortex of the brain processes and recognizes images. It consists of 3 types of layers (convolution – pooling – fully connected) layers.

Forward propagation:

Convolution layers:

$$x_{ij}^{l} = \sum_{a=0}^{m-1}\sum_{b=0}^{m-1} w_{ab}\, y_{(i+a)(j+b)}^{l-1} + b^{l}$$

$$y_{ij}^{l} = max\left(0, x_{ij}^{l}\right)$$



Pooling layers:

$$x_{ijc}^{l} = Avg\ or\ max\left(x_{i:i+f,j:j+f,c}^{l-1}\right)$$

FC layers:

$$z^{(i)[l]} = W^{[l]} a^{(i)[l-1]} + b^{[l]}$$

$$a^{(i)[l]} = g\left(z^{(i)[l]}\right)$$

Cost function:

$$J = -\sum_{i=1}^{c} y_i . log\left(\frac{e^{z_i}}{\sum_{i=1}^{n} e^{z_i}}\right)$$

Back propagation:

Fully Connected layers:

$$\frac{\partial J}{\partial w_{ab}^{l}} = \frac{\partial J}{\partial y_i} . \frac{\partial y_i}{\partial w_{ab}^{l}}$$

$$\frac{\partial J}{\partial b^{l}} = \frac{\partial J}{\partial y_i} . \frac{\partial y_i}{\partial b^{l}}$$

Convolution layers:

$$\frac{\partial J}{\partial w_{ab}^{l}} = \sum_{i=0}^{N-f}\sum_{j=0}^{N-f} \frac{\partial J}{\partial x_{ij}^{l}} \frac{\partial x_{ij}^{l}}{\partial w_{ab}^{l}}$$

$$\frac{\partial J}{\partial b^{l}} = \sum_{i=1}^{m} \frac{\partial J}{\partial x_{ab}^{l}} \frac{\partial x_{ab}^{l}}{\partial b^{l}}$$

Updating w, b using Adam optimizer

$$w_{ab}^{l} = w_{ab}^{l} - \alpha \frac{v'_{dw}}{\sqrt{s'_{dw} + £}}$$

$$b^{l} = b^{l} - \alpha \frac{v'_{db}}{\sqrt{s'_{db} + £}}$$

Experimental work

Data preprocessing:

- Preprocessing steps were applied to the data to make it appropriate for the models.
- The first steps were common to both models, but each model required additional preprocessing to work effectively, as illustrated.
- Two different scenarios were used to achieve the highest accuracy.

Scenarios:

- **VGG16** was employed. By applying fine tunning, the top layers of the model were removed, and new layers were added to get the target characteristics.
- **Landmark** detection was performed using Dlib and an NN (Neural Network) model.

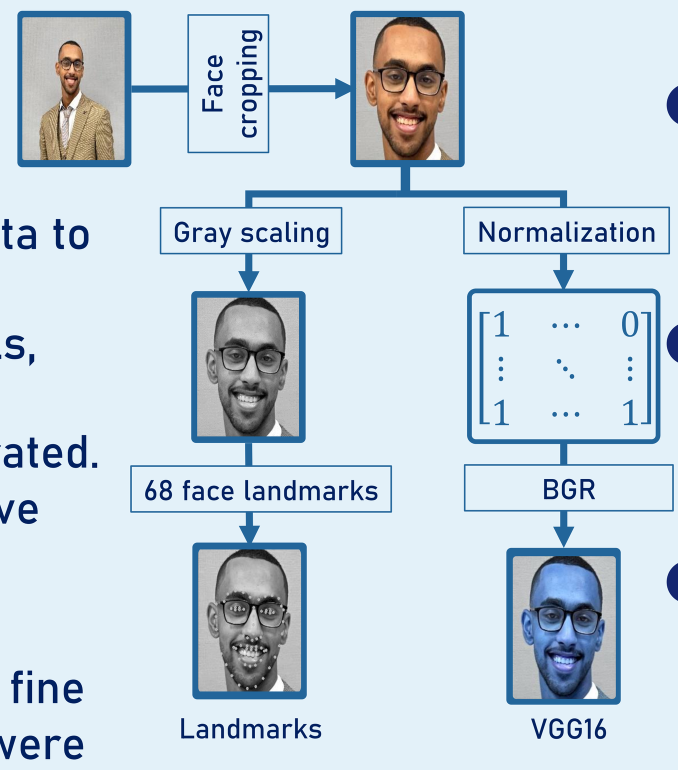


Figure 5: Data preprocessing

Layer (type)	Output shape
dense (Densel)	512
dense_1 (Densel)	512
dense_2(Densel)	512
dense_3(Densel)	512
dropout (Dropout)	512
dense_4(Densel)	3

Table: 1 NN Summary

Layer (type)	Output shape
Vgg16 (Functional)	**512**
Flatten (Flatten)	**512**
dense (Densel)	**512**
dense_1 (Densel)	**512**
dropout (Dropout)	**512**
dense_2 (Dense)	**3**

Table 2: Fine Tunning

Results

- The **VGG16 transfer learning** model achieved better results and less overfitting.
- It has achieved good performance at detecting the 3 levels of **engagement.**
- The **VGG16** is preferred to be used in the proposed system for **disengagement detection .**

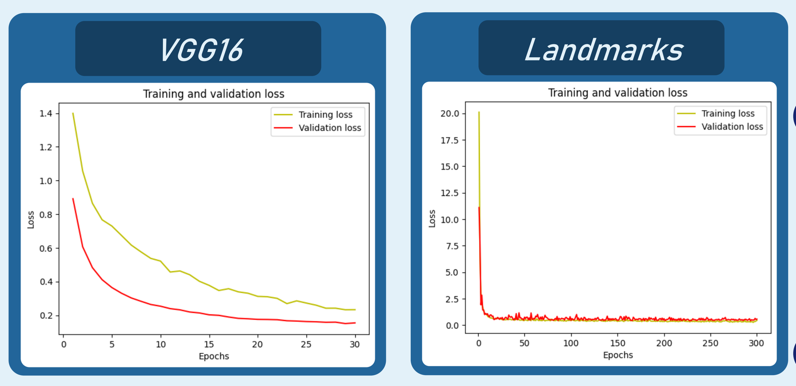


Figure 7: Training and Validation loss

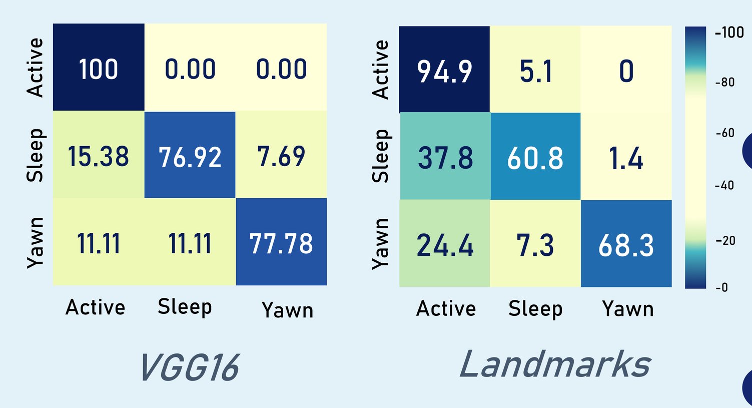


Figure 8: Confusion Matrix Heatmap and Classes Accuracy

POC	VGG16	Landmarks
Accuracy	93.83%	80.6%
F1 Score	84.21%	76.52%
Average accuracy	89.58%	84.19%
Micro precision	84.38%	76.29%
Macro precision	85.11%	83.15%

Table 3: Models Compression

Conclusion

- We compared two models: **VGG16** transfer learning and **Facial Landmarks** neural network.
- We trained them on diverse datasets. Additionally, we gathered images of students at Cairo University.
- Based on results, the **VGG16** transfer learning model performed better and is used in our real-time **disengagement detection** system.

Future work

- Enhancing **DL** models performance by:
  - Acquiring more students' data.
  - Training on various **disengagement** behaviors.
- Implementing a web app solution which is accepted in INJAZ company program and middle east competition.

References

P. Buono, B. De Carolis, F. D'Errico, N. Macchiarulo, and G. Palestra, "Assessing student engagement from facial behavior in on-line learning," Multimedia Tools and Applications, vol. 82, no. 9, pp. 12859–12877, Oct. 2022, doi: 10.1007/s11042-022-14048-8.

K. Simonyan, "Very deep convolutional networks for Large-Scale image recognition," arXiv.org, Sep. 04, 2014. https://arxiv.org/abs/1409.1556

Scan the QR to get the full list of all references