



המרכז לדימות מוחי בילדים  
Educational Neuroimaging Center



המעבדה לעיבוד גיאומטרי של תמונות  
Geometric Image Processing Laboratory

# ZoomEmotion

## Winter 2020-2021



Efrat Israel  
Akiva Zonenfeld

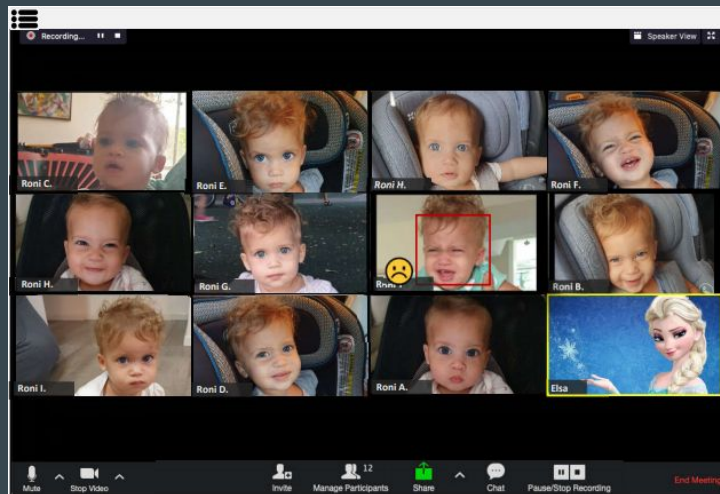
# Workflow



# Problem Definition

Create a desktop application that :

1. Analyzes children's emotions during virtual classes
2. Indicates the children that express negative emotions in real time
3. Creates a log with the collected data



# Initial Solution Definition

1. Get the video stream of each meeting participant from the Zoom meeting.
2. Capture image from the video stream.
3. Detect the faces of the children in the image.
4. Detect the emotion of each child.
5. Display the emotion gesture back on the video.



# Theoretical Knowledge Acquirement

- Stanford 231n course (Convolutional Neural Networks for Visual Recognition)
- RonNet Model - Convolutional Neural Network
- OpenCV
- Torch
- PyAutoGui
- (Tesseract-OCR)



# Research of Zoom Application

1. Zoom API
2. Zoom Client SDKs
3. Zoom Fully Customized SDKs

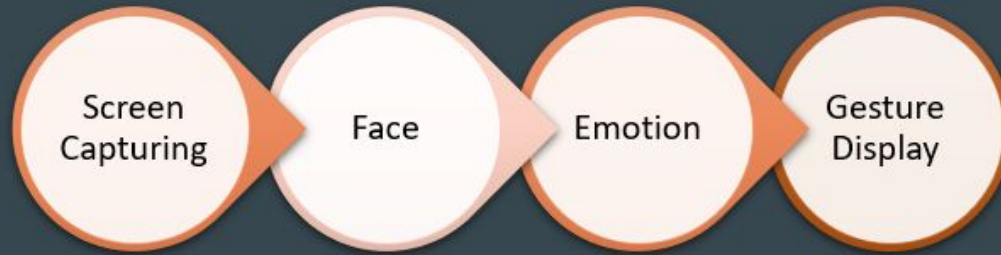


	Client SDKs	Fully Customized SDKs
<b>Image Processing</b>	Analysis of full screen	Analysis of each video stream separately
<b>Annotation Display on Screen</b>	Not supported	Supported - Display for all users
<b>Annotation Display on Video</b>	Not supported	Not supported
<b>UI Customization</b>	Not supported	Supported
<b>Addition Difficulties</b>	Non	Need to be purchased and Zoom selling is unresponsive

A New Solution is Required

# Pivot - Solution Redefinition

1. Capture of the video conference screen (Main change).
2. Detect the faces of the children in the image.
3. Detect the emotion of each child.
4. Display the emotion back on the video.



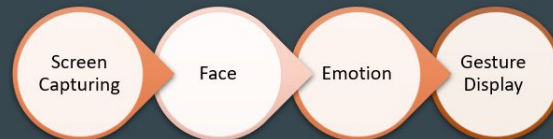
# Pivot - Solution Redefinition

## Advantages

1. Ease of Implementation
2. Ease of Use
3. Annotating on video is possible
4. Can be used with any video platform

## Disadvantages

1. Requires data matching of face, emotion and name







# First Version of the Application

Our mid-meeting milestone was a demo application that recognizes two children's faces, recognizes their emotions and displays relevant annotations on the video screen.

# Text Recognition and Logs Creation

We focused on building data structure for log creation.

Therefore, we added the text recognition feature which is used to identify the children face and emotion repeatedly.

In that way we can track each children emotion during the video session, and create logs accordingly.



# Parameters Interface

We added interface that enables the teacher to define different parameters regarding the emotions that are detected:

- Emotions Threshold
- Maximal faces to draw
- Gesture on screen time.





# Final Version of the Application

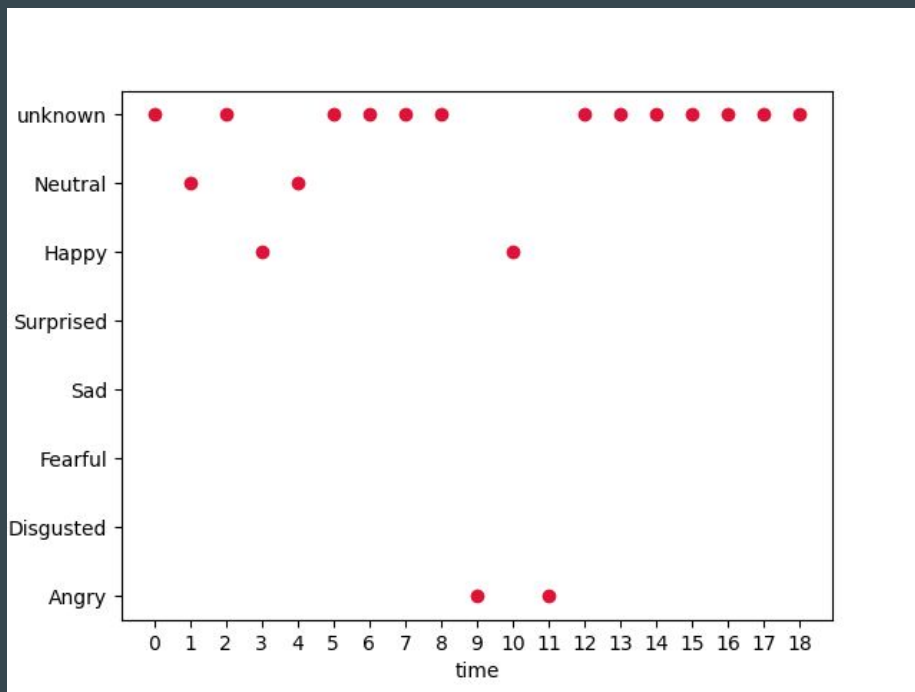
- Extension of the application capabilities to handle up to sixteen faces
- Addition of logs and graphs that show each student's emotion during the class
- Addition of graph that summarizes all emotions that were detected in the session, as a function of time

# Algorithm Description



1. Capture the Video Screen
2. Analyze the Captured Image:
  - 2.1. Find Video Frames in the Image
  - 2.2. For Each Frame:
    - 2.2.1. Detect Face In Frame (Using CV Face Detection)
    - 2.2.2. Recognize Name From Image And Match with the Face
    - 2.2.3. Detect Emotion Of Face (Using RonNet)
    - 2.2.4. Update Log with Relevant Data

# Demo Results



Child's Emotion Graph

	A	B	C
1		אלונה	איתי
2	5.568709135		
3	106.9580948		Neutral
4	114.1968858		
5	137.1638882	Fearful	Happy
6	147.9017015		Neutral

[Detailed log.csv](#)

	A	B	C
1	Name	Total	Angry
2	עדיסל בן משה	0.169486239	0
3	יובל	0.04237155974	0.02118577987
4	איתי	0.1271146792	0.04237155974

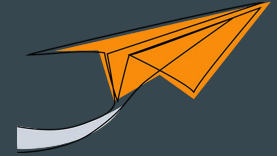
[Summary log.csv](#)



# Conclusions

- Our application assists teachers during virtual classes in real time, and helps researchers with analyzing offline
- We made pivots in order to find a solution that hits the mark. We dived into the world of image processing and data matching in order to get a complete application
- As our application is based on CNN, which is not 100% accurate, sometimes the emotion that is displayed on screen is not matching with the emotion that is recognized by human

# Future Work



1. Improve response time of the application
2. Extend emotions and behaviours that are detected
3. Consider implementing the initial solution since the last update in Zoom SDKs provides access to the raw video for all developers
4. Make adaptations to other video conference platforms, for example Teams and Hangouts
5. Extend application for analyzing adults emotions during video sessions





Thank You For Your Help  
And Attention!

...

Efrat Israel  
Akiva Zonenfeld