**Introduction**

Facial expression detection, recognition, and understanding happens naturally for most, but individuals with Autism Spectrum Disorder (ASD) can have trouble inferring emotions based on these expressions, which can cause depression, loneliness, and a disconnect with society.

However, studies have shown that through the use of Assistive Technology (AT) software, individuals can see an improvement in their emotion recognition abilities [1].

We propose AEGIS (Augmented-reality Expression Guided Interpretation System), a multimodal AT system deployable on a wide range of user devices, designed in order to assist in learning to better identify expressions and thus improve their societal experiences.

AEGIS leverages the use of TimeConvNets [2], our novel deep convolutional neural network, in order to achieve real-time facial expression classification in a seamless manner.

**What is AEGIS?**

AEGIS requires a device with a screen and video streaming camera, which can range from handheld devices such as mobile phones and tablets, to wearable technology such as smartglasses. It can also be easily integrated into video conference systems. To use AEGIS, an intuitive software app will be installed and run when the user requires it. The device should be positioned in such a way that the user can see the screen at all times, while the video streaming camera faces another person’s face.

Running in real-time, AEGIS passes the real-world frames captured from the camera into our novel TimeConvNet neural network architecture, where both spatial and temporal information is analyzed by sets of learned convolutional filters.

The resulting expression classification is produced and mapped to a set of preselected emoticon images. The selected emoticon is added to the original real-world image, hovering over the target individual’s forehead. Upon seeing this augmented frame, the user can better infer the target person’s emotional state. This entire process takes milliseconds to complete, allowing for a seamless experience for the user.

**Visualization Options**

The way the information is presented to the user is a key factor in how well they understand the given message. It must be unambiguous, and require very little conscious thought to understand. Visual based feedback was chosen due to studies showing that individuals diagnosed with ASD are highly visual-oriented and respond better compared to auditory based stimuli [1].

With the growing popularity of instant messaging and smartphones, emotions can be seen in almost every internet chat. They span across all genders and ethnicities, and are also starting to spread among all age ranges as well. They can immediately convey certain tones with a minimal delay, which is perfect for our needs.

Alternate visualization methods were also considered, but can be implemented in the future as custom options. These include color based augmentation, or text based augmentation where a word would float alongside a facial bounding box.