ToMCAT Offline Viz

Caleb Shibu Jones, Harshita Narnoli, Rupal Jain

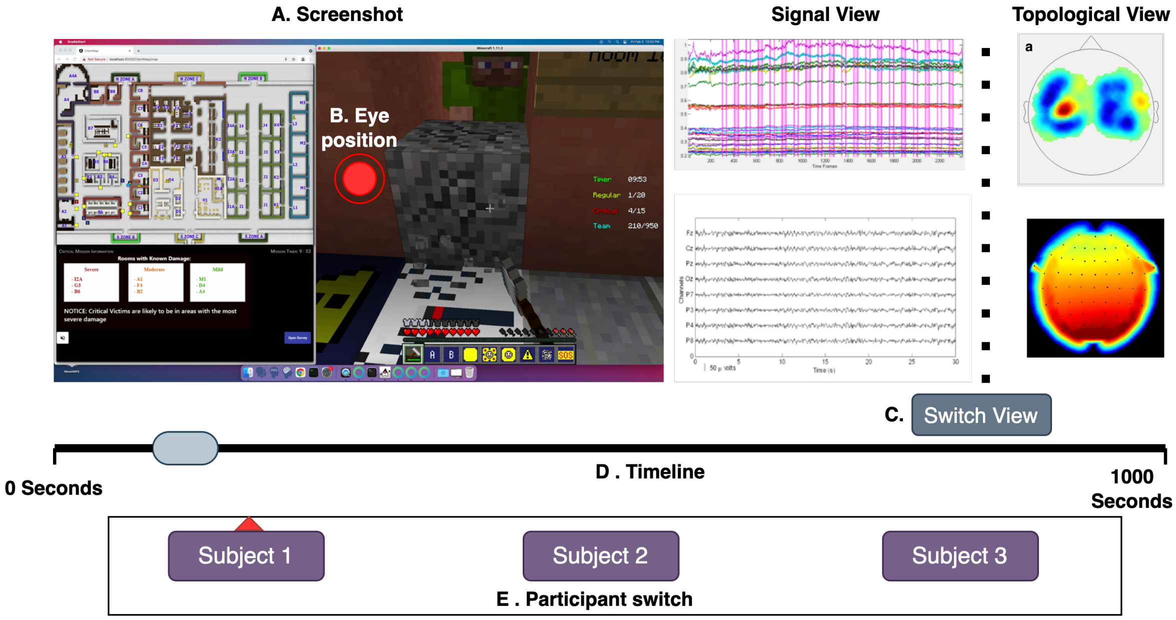

Fig. 1: Tomcat Offline Viz tool

Abstract— ToMCAT (*Theory of Mind-based Cognitive Architecture for Teams*) offline viz tool is a collaborative project to develop an all-in-one multimodal (fNIRS, Eye tracker, EEG) visualization tool. To achieve this goal, the project accepts experimental data in the form of CSV files and screenshots (PNG) and visualizes this data to the users in a way that is easy to understand and analyze. The tool utilizes multiple modes of data, including functional near-infrared spectroscopy (fNIRS), eye tracking, and electroencephalography (EEG), to provide a complete picture of the cognitive and perceptual processes, such as attention, working memory, and emotion. By using this tool, researchers can gain valuable insights into how team members interact with each other through varying tasks and environments. The tool presents the data in a visually intuitive manner, by superimposing screenshots with eye-tracking data on the left and fNIRS/EEG data on the right. The ToMCAT Offline Viz tool is a valuable resource for social and computer scientists who are working on the ToMCAT project. It helps the users to read and analyze multimodal data in a way that is both efficient and informative, and to use this data to design an AI agent that can facilitate team coordination and global team optimization.

1 Introduction

To enhance the collaborative capabilities of AI agents in team settings, The Theory of Mind-based Cognitive Architecture for Teams (ToMCAT) [5] project proposes an integrated approach that encompasses all critical abilities. These abilities include accurately deducing the internal states of other agents, collaborative problem-solving, and socially aware communication. The ultimate goal of the project is to design an AI agent that concentrates on a select set of essential abilities required for efficient collaboration and seamlessly integrates them into a unified framework.

We are developing an innovative visualization tool, called "*ToMCAT Offline Viz*," which combines multiple modalities to assist researchers in the analysis of experimental data. This solves the challenge of understanding the complex cognitive and neural mechanisms underlying teamwork. The tool enables researchers to superimpose eye-tracking data onto screenshots to gain insights into the visual attention of team members. It also displays functional Near-Infrared Spectroscopy (fNIRS) [4] and Electroencephalography (EEG) [3] data on the right-hand side of the screen, providing information on neural activity.

By analyzing two types of signals, namely fNIRS and EEG patterns with eye-tracking data, researchers can determine how visual stimuli affect the brain. This information is particularly valuable to social scientists who study the variation of signal waves as the subject's gaze shifts across different areas of a screenshot.

Both signal and topological views of the fNIRS and EEG data are available, and users can switch between them by clicking a button based on their preferred visualization. A slider is utilized to compare the continuous changes in each modality working in sync to facilitate the analysis of variations and further enhance visualization. Referencing Fig. 2, the slider is mapped with the timestamp of each task and the timestamp of each physiological data. Additionally, the tool enables changing the subject to facilitate comparing the similarity or variation across subjects.

Our current objective is to address the issue of a lack of comprehensive and user-friendly visualization tools for researchers studying

- *Harshita Narnoli is a Ph.D. Graduate student at the University of Arizona. E-mail:harshitanarnoli@arizona.edu.*
- *Rupal Jain is a MS Graduate student at the University of Arizona. E-mail:jainrupal@arizona.edu.*
- *Caleb Jones Shibu is a MS Graduate student at the University of Arizona. E-mail:calebshibu@arizona.edu.*

cognitive and neural mechanisms. This problem holds significance as it directly impacts the ability of social and computer scientists working on the ToMCAT project to gain a better understanding of how individuals interact within teams. The development of effective visualization tools enables researchers to gain deeper insights into the cognitive and neural mechanisms involved in effective coordination within teams.

In summary, our team is in the process of developing a user-friendly and innovative visualization tool that can analyze multiple modalities, such as eye-tracking, fNIRS, and EEG, to provide a comprehensive understanding of the cognitive and neural mechanisms that underlie teamwork. Enabling researchers to gain insights by providing a compact approach to visualizing complex data, which includes where team members direct their visual attention and how effective their coordination is by showing the change of trend in physiological data. Aiding the development of effective AI agents for team coordination and global team optimization.

2 Background

The ToMCAT [5] project is comprised of a set of local agents (one for each human teammate) equipped with cameras and microphones to capture facial expressions and speech, as well as virtual sensors that record the local environment, the actions performed by human teammates, and chat exchanges between them.

The multimodal data includes:

a. Functional Near-Infrared Spectroscopy (fNIRS) is a technology that can identify which parts of the brain are being used during an activity. When a specific part of the brain is activated, the blood flow to that region increases. By detecting changes in the concentration of the blood in the brain, fNIRS can determine which parts of the brain are most active [4].

b. Electroencephalography (EEG) is a technique used to record the spontaneous electrical activity of the brain in the form of an electrogram [3].

c. Eye tracking involves the detection of the human pupil and the subsequent recording of eye movements and fixations while viewing images or any other content on screen [2].

2.1 Related Work

Visualization Related: **Pupil Player** is a crucial tool for researchers who use Pupil Capture to collect eye-tracking data. It provides a user-friendly interface for visualizing recorded data and analyzing it. Additionally, Pupil Player enables researchers to export their data into various formats, making it easier to analyze and share with other researchers. Visualization plugins include Vis Circle, Vis Cross, Vis Polyline, Vis Light Points, and Vis Eye Video Overlay. Analysis Plugins include Surface Tracker, Fixation Detector, Blink Detector, Head Pose Tracking, IMU Timeline [2].

The **nirsLAB package** is an adaptable software analysis environment designed to facilitate the examination of time-varying near-infrared measurements of tissue. It is particularly well-suited for analyzing data acquired using NIRx systems [6].

BrainVision Analyzer is used by Scientists to process a variety of neurophysiological data. For ToMCAT project, EEG data is being collected. Analyzer is easy to use and offers a variety of powerful features for processing and visualizing EEG data [1].

Based on our extensive literature survey, we did not come across any existing visualization that integrates fNIRS and EEG data with eye-tracking data and compares them.

Domain Specific Related: Referring to Fig. 2, each experiment usually comprise of 3 participants and they follow a fixed instructed paradigm containing eight different tasks done in sequence:

A. Rest state: The participants are requested to sit back and relax for 300 seconds.

B. Finger-tapping individual task: The participants tap the spacebar on the keyboard for 10 seconds.

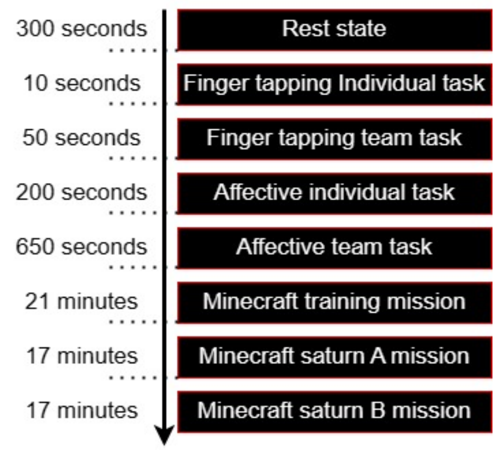

Fig. 2: ToMCAT Experimental Paradigm

C. Finger-tapping team task: The participants tap the spacebar on the keyboard for 50 seconds in sync with each other based on visual feedback.

D. Affective individual task: The participants see an image for a few seconds on screen and they rate it based on its arousal and valence score. This is done for a total of 15 images. This lasts for roughly 3 minutes.

E. Affective individual task: The participants see an image for a few seconds on screen and they rate it based on its arousal and valence score after discussing it with each other. This is done for a total of 15 images. This lasts for roughly 11 minutes.

F. Minecraft training mission: This is for a mission tailored to train participants based on their role as medic, transporter, and mechanic.

G. Minecraft Saturn A mission & H. Minecraft saturn B mission : These are two missions with different environments where participants with their roles as medic, transporter, and mechanic have to team up and collaborate to rescue critical victims placed in their environments.

All the task explained above have their own csv files that contains the timestamps at which they started and ended.

3 Preliminary Results

3.1 Data Preprocessing

The Lang Lab uses lab recorder software to collect physiological data and stores it in XDF file which requires parsing and conversion to a readable format such as CSV file. For each experiment, 3 XDF files are generated, one for each participant, containing EEG, fNIRS and Eye-tracking data in the form of a dictionary, which has to be decoupled into EEG.csv, fNIRS.csv and gaze.csv. To process this data, a script was written that performs the following tasks:

1. Reads all XDF files.
2. Iterates thought all Physio streams streams individually.
3. Reads start and stop time stamps from XDF file. Uses these timestamps to create a date-timestamp distribution.
4. Creates a data frame with all the physio stream data with the date-timestamp distribution.
5. Creates a label distraction from start to stop timestamp Baseline task and Minecraft metadata. Then, syncs all this information with data frame.
6. Removes data before start and stop time stamps from XDF file.
7. Exports the data frame as CSV file and pickle file.

Overall, this script enables efficient processing of physiological data for the Lang Lab's experiments, allowing for easier analysis and interpretation.

The fNIRS data contains a lot of physiological noise such as heartbeat, breathing, and other motion artifacts, which needs to be filtered out. To remove these physiological noises a script was developed that implements a Butter-bandpass filter. The low-pass filter was set to 0.01 Hz and the high-pass filter was set to 0.2 Hz. This filtering process effectively removes unwanted noise from the data, resulting in a cleaner and more accurate representation of the fNIRS signals.

3.2 Data Visualization

Referencing Section 1, so far we have made progress in visualizing screenshots and fNIRS data along with the timeline for each subject.

Eye Tracker The tool would have the following sections:

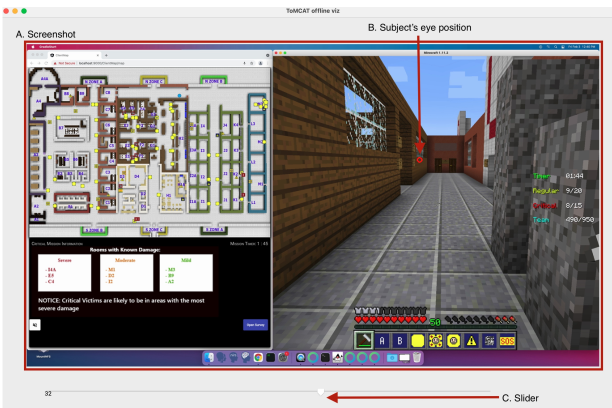

Fig. 3: Screenshots overlaying with eye tracking.

Fig. 3 shows a slider that is numerically mapped to the sequence of screenshots and we have superimposed the eye tracking data as well onto the screenshot.

- Figure 3 A. Screenshot: The red bounding box shows the screenshot of the resolution (1280 X 720) of an iMAC display.
- Figure 3 B. Selected Eye Position: The red dot on the screenshot shows what the subject is looking at, at that particular instance of time.
- Figure 3 C. Slider: Currently, the slider controls the sequence of screenshots along with the superimposed position of the subject's eye.

Fig. 4 shows a slider that is numerically mapped to the sequence of fNIRS data over time. This plot is being shown over a window of 100 samples for each channel. The y-axis displays the fNIRS channel name and the x-axis shows the sample count. The plot includes HbO (Oxygenated Hemoglobin) and the HbR (De-oxygenated Hemoglobin) data, represented by red and blue datalines respectively.

3.3 Multiple Views Relation

The study of how visual stimuli impacts the brain is a fascinating and complex field of research. In recent years, advancements in technology have enabled researchers to analyze two types of signals, fNIRS and EEG patterns, in conjunction with eye-tracking data to better understand this phenomenon. By combining these methods, researchers can determine the specific effects of visual stimuli on different regions of the brain, such as changes in blood flow and electrical activity.

Social scientists have a particular interest in this type of research, as it allows them to study how the brain responds to visual stimuli in a

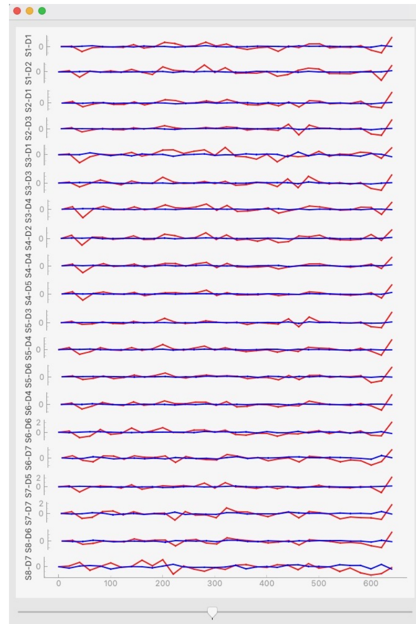

Fig. 4: fNIRS

variety of contexts. For example, they can analyze the variation of signal waves as the subject's gaze shifts across different areas of a screenshot, providing valuable insights into how people process visual information. This information can be used to inform a wide range of fields, from advertising and marketing to psychology and education. Ultimately, by understanding how visual stimuli affects the brain, researchers can gain a deeper understanding of human behavior and cognition, opening up new avenues for exploration and discovery.

3.4 Code Execution

The sample data included in the repository contains dummy values, and cannot be used to reproduce the results shown in the screenshot. To run the code and generate results, follow the steps below:

- Create a virtual environment and run the command

```
pip3 install -r requirements.txt
```

- In order to view the screenshots along with the eye-tracking data, run the command

```
python3 screenshot_eye_track_slider.py
```

- In order to view the fNIRS data, run the command

```
python3 fnirs_slider.py
```

4 Research Plan

Referencing Section 1, our goal is to develop an innovative visualization tool to read and analyze complex data, including eye-tracking, fNIRS, and EEG, to provide a comprehensive understanding of teamwork. By developing ToMCAT offline viz, we plan to provide an innovative and user-friendly approach thereby supporting the research team in analyzing experimental data.

Fig. 1 gives a brief overview of ToMCAT offline viz tool.
The tool would have the following sections:

- **(Label A & B)** Fixation of Eye position superimposed over the screenshots with the diameter of circle changing based on pupil dilation on the screen.
- **(Label C)** A button that would switch between topological and signal view for EEG and fNIRS data.

- **(Label D)** A slider to scrub through the experiment timeline.
- **(Label E)** Three buttons to switch between participants.

4.1 Data

We have retrieved the data from LangLab arranged by Caleb who is an active member of the lab working under the supervision of Dr. Adarsh Pyarelal and Dr. Kobus Barnard. We have taken permission from the lab regarding the usage of the data.

The data consists:

Data type	**Data Description**
Experimental Images	PNG image format (Resolution: 1280X720, File size $\sim$ 600 kB) Sampled at 10 Hz
fNIRS	CSV Format (File size approx 53 MB) Sampled at 10.2 Hz
Eye Tracker	CSV Format (File size approx 140 MB) Sampled at 200 Hz
EEG	CSV Format (File size approx 1.2 GB) Sampled at 500 Hz

4.2 Evaluation

LangLab has been collecting multimodal data for the ToMCAT project since the Fall of 2021, providing us with an abundance of data to work with. Our only constraint will be ensuring that the ToMCAT offline visualization tool functions smoothly with limited system resources.

To evaluate the project, we plan to take feedback from computer and social scientists regarding their experience using the tool after the complete development. Specifically, we will be asking for their opinions on the tool's user-friendliness and overall effectiveness.

After the completion of the project, the tool will be continued and maintained by Caleb, which will be potentially used by the ToMCAT social scientists.

4.3 Technology

Programming language: Python 3.9

Libraries: PyQt, CV2, Pandas, Matplotlib, MNE-Python PyQt5 is a cross-platform GUI toolkit that will help us to develop an interactive desktop application. CV2 is an OpenCV package for Python that serves for image loading, image compression, etc. Pandas will be used to load data from CSV files allowing the working of "relational" or "labeled" data.

Matplotlib would allow us to create static, animated, and interactive visualizations. MNE-Python might be used for a topological view of the brain for EEG and fNIRS signals.

4.4 Deviations from Proposal

The project has remained mostly in line with the original proposal. However, as we continue to make progress, we have decided to prioritize the completion of the signal view for both fNIRS and EEG data. Once this is completed, we plan to move on to developing the topological view for both types of data. We believe that focusing on the signal view first will allow us to better understand the underlying data and make informed decisions regarding the design and implementation of the topological view. Overall, these changes are expected to improve the quality and usefulness of the final product.

4.5 Timeline

Milestone 1: We have developed a skeleton visualization tool that enables us to scrub through the timeline and visualize screenshots like video playback.

Milestone 2: The eye positions have been overlaid onto the screenshots, allowing us to observe shifts in gaze location and corresponding changes in screenshots as we navigate through the timeline.

Milestone 3: We have added a different section that shows a signal view for fNIRS signals which is currently decoupled with the eye-tracking data.

Milestone 4: We need to create a separate section that displays a line plot of EEG data. Both fNIRS and EEG data will be then synced with the eye tracking data based on timestamps.

Milestone 5: Add a section to the right that shows the topological view for fNIRS signals along with a button to toggle between the topological view and signal view.

Milestone 6: Similarly, a section needs to be added on the right-hand side that presents the topological view of EEG signals, accompanied by a button to switch between the topological and signal views.

Milestone 7: Working on the project report by presenting comprehensive information regarding the project's scope, objectives, methodology, findings, and recommendations in a clear, concise, and structured manner.

Milestone 8: Working on completing the project report contents as well as preparing the project presentation.

Project Milestones ([*] Marked as complete)

Milestone	Description
March 17	Skeleton of the viz tool (Basic structure design) [*]
March 24	Eye positioning for gaze fixation [*]
March 29	Progress Update with the addition of fNIRS signals [*]
April 7	Adding EEG signal view using line plot
April 14	Topological view implementation to switch from signal view
April 21	Project Initial Report of viz tool with defined subsections
May 3	Final Report Completion

4.6 Team Member Roles

Since Caleb is a member of the ToMCAT project, he is involved in data pre-processing of the physiological data and syncing it with the timestamps of the task done during the experiment. Apart from this, he has completed the skeleton structure of the tool which includes the slider. Harshita is involved in translating the surface coordinates of the eye-tracking data to the resolution of the screenshot and superimposing it on the top of the screenshot. Rupal is involved in the development of the signal view of the fNIRS data and linked to the slider. The signal view of both fNIRS and EEG will now be linked to the eye-tracking data.

5 Impacts

The development of the ToMCAT Offline Viz tool has a significant impact on the study of cognitive and neural mechanisms involved in teamwork by providing a comprehensive approach to visualizing complex data. The tool enables researchers to gain insights into where team members direct their visual attention and how effective their coordination is, facilitating the development of effective AI agents for team coordination.

6 Challenges

All the team members have signed a Non-Disclosure Agreement with The University of Arizona eIRB, so actual data is not allowed to be shared outside ToMCAT team members. For gaining the access, Rupal and Harshita have completed all the related CITI and COI training. But we are still waiting for approval from eIRB for data access due to the internal structure change of superiors of ToMCAT team.

The progress of the project is facing obstacles due to a data-related problem, as only two team members are tasked with working on components using simulated data, while the testing responsibility rests solely on one team member.

References

[1] EEG-BrainVision. Brainvision, 1997.

[2] M. Kassner, W. Patera, and A. Bulling. Pupil: An open source platform for pervasive eye tracking and mobile gaze-based interaction. *UbiComp 2014 - Adjunct Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, 04 2014. doi: 10.1145/2638728.2641695

[3] J.-P. Lachaux, N. Axmacher, F. Mormann, E. Halgren, and N. E. Crone. High-frequency neural activity and human cognition: past, present and possible future of intracranial eeg research. *Progress in neurobiology*, 98(3):279–301, 2012.

[4] S. Tak and J. C. Ye. Statistical analysis of fnirs data: a comprehensive review. *Neuroimage*, 85:72–91, 2014.

[5] ToMCAT. Ml4ai, 2019.

[6] Y. Xu and H. Graber. nirslab: A computing environment for fnirs neuroimaging data analysis. 04 2014. doi: 10.1364/BIOMED.2014.BM3A.1