Federated Personalized Scanpath Prediction for Privacy-Preserving UI Optimization

Md. Akil Raihan Iftee
Department of Computer Science and Engineering
Khulna University of Engineering & Technology, Bangladesh
iftee1807002@stud.kuet.ac.bd

I. Problem Specification

- **Individual Variability in Scanpaths:**
  Eye Scanpaths Are Unique to Individuals, for instance, ***Novices*** *scan more;* ***experts*** *jump directly to relevant areas. Personalized scanpaths help* ***Prioritize*** *content in areas of predicted attention.*
- **Privacy Sensitivity of Gaze Data:**
  Gaze Data is Highly Sensitive and Privacy-Critical as *Eye movements can reveal User identity* ***(biometric patterns)****, Cognitive states (confusion,* ***interest****, fatigue), User intent (what they focus on).*
- **Privacy Risks in Traditional Training:**
  In traditional training, *multiple users' raw gaze data would have to be directly shared together to the trainer model*. This exposes different user's data to the trainer.
- **Legal & Ethical Constraints:**
  Regulations like ***GDPR*** *[1] and* ***HIPAA*** *[2] discourage or ban collection of personally identifiable data like gaze traces unless explicitly consented*, which is *hard to enforce at scale*.

II. Objective

Federated Learning (FL) allows training on-device with no need to export personal gaze or usage data, compliant with privacy regulations.

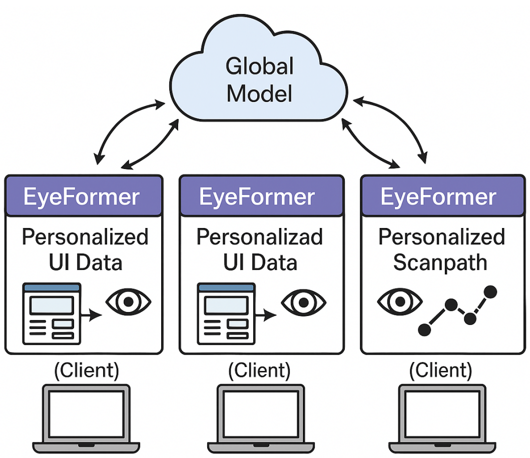

Fig. 1: The Framework of Proposed Gaze-Enhanced Multi-modal Interaction

III. Method

We use **EyeFormer** [3] as a baseline model to generate scanpaths (sequence of fixations with durations). Instead of centralized training, each device collects a few scanpath samples from its user. The model is fine-tuned locally (few-shot adaptation), and Federated Learning combines these updates into a global model that still preserves user individuality. Given a new UI, the model can produce **personalized scanpaths** conditioned on the **user's behavior**.

IV. System Overview

Privacy Personalization: Users' Raw eye-tracking and screen data never leave the device. The global model only aggregates learned parameters, **never sees raw sequence data**. This protects gaze privacy and captures personal viewing strategies. The global model can generate generic or user-specific predictions depending on whether user data is provided during inference.

Datasets/Models: EyeFormer's own training (various stimulus types, including UIs) provides a strong starting point. The **UEyes** [4] data could also seed pretraining. The project could compare a federated-trained EyeFormer vs. a centrally-trained version.

Visual Behavioral Aspects: It relies on UI pixel features and temporal gaze behavior. Applications include reordering or resizing UI elements based on predicted user scanpaths, thus **optimizing "visual flow"** for each person. The method bridges visual prediction with user-specific behavioral patterns.

References

[1] J. Kröger, O. Lutz, and F. Müller, "What does your gaze reveal about you? on the privacy implications of eye tracking," in *IFIP International Summer School on Privacy and Identity Management*, pp. 226–241, Springer, Cham, 2020.

[2] E. Larsen, J. Kolman, F. Masud, and F. Sasangohar, "Ethical considerations when using a mobile eye tracker in a patient-facing area: lessons from an intensive care unit observational protocol," *Ethics & Human Research*, vol. 42, no. 6, pp. 2–13, 2020.

[3] Y. Jiang, Z. Guo, H. R. Tavakoli, L. A. Leiva, and A. Oulasvirta, "Eyeformer: Predicting personalized scanpaths with transformer-guided reinforcement learning," in *Proceedings of the ACM Symposium on User Interface Software and Technology (UIST)*, 2024.

[4] Y. Jiang, L. A. Leiva, H. R. Tavakoli, P. R. B. Houssel, J. Kylmala, and A. Oulasvirta, "Ueyes: Understanding visual saliency across user interface types," in *Proceedings of the ACM CHI Conference on Human Factors in Computing Systems (CHI)*, 2023.