

Deep Vision for First Person Action Understanding

PhD Candidate:

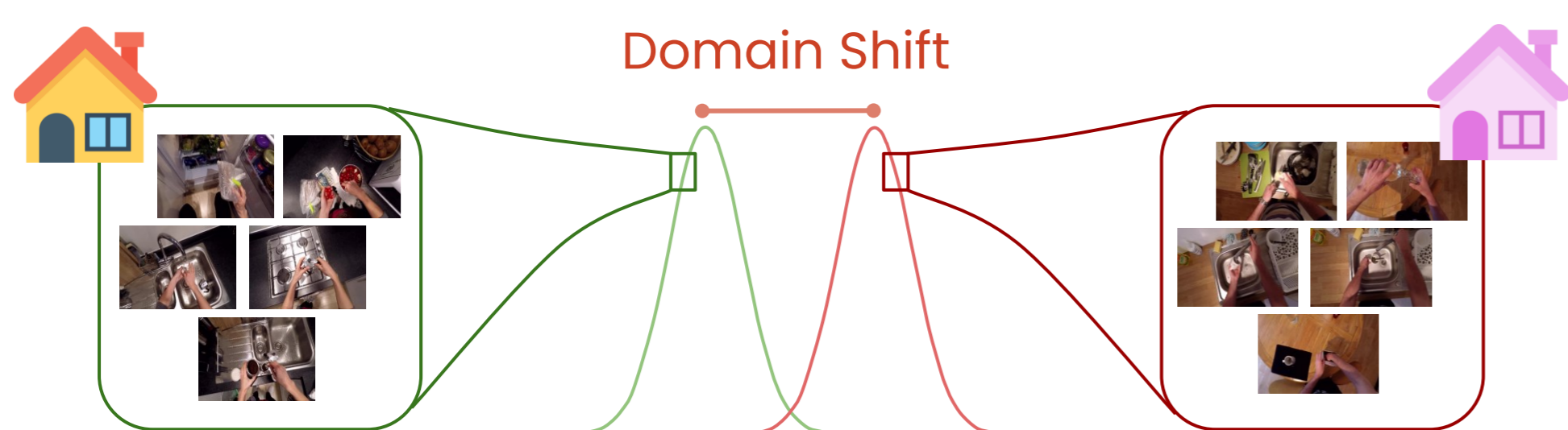
Mirco Planamente

1. Egocentric Vision

Egocentric vision is a sub-field of computer vision that analyzing images and videos captured by a wearable camera, which is typically worn on the head and approximates the visual field of the camera wearer. It allows us to study human behavior in a much more direct manner, giving the opportunity to learn how humans perceive the world and interact with the environment.

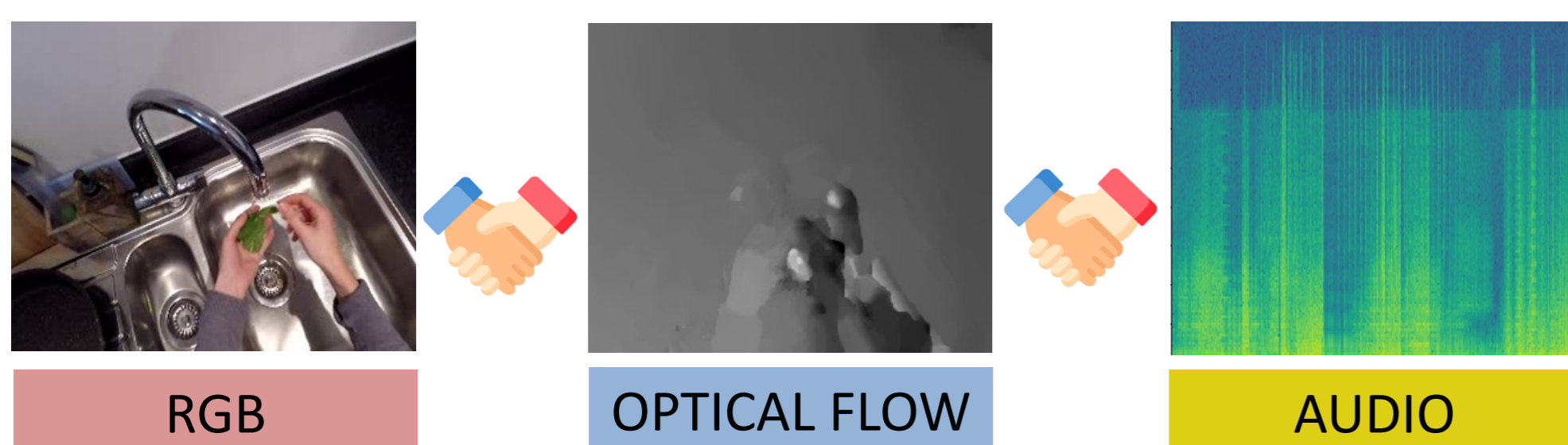


2. Main problems



3. General Idea

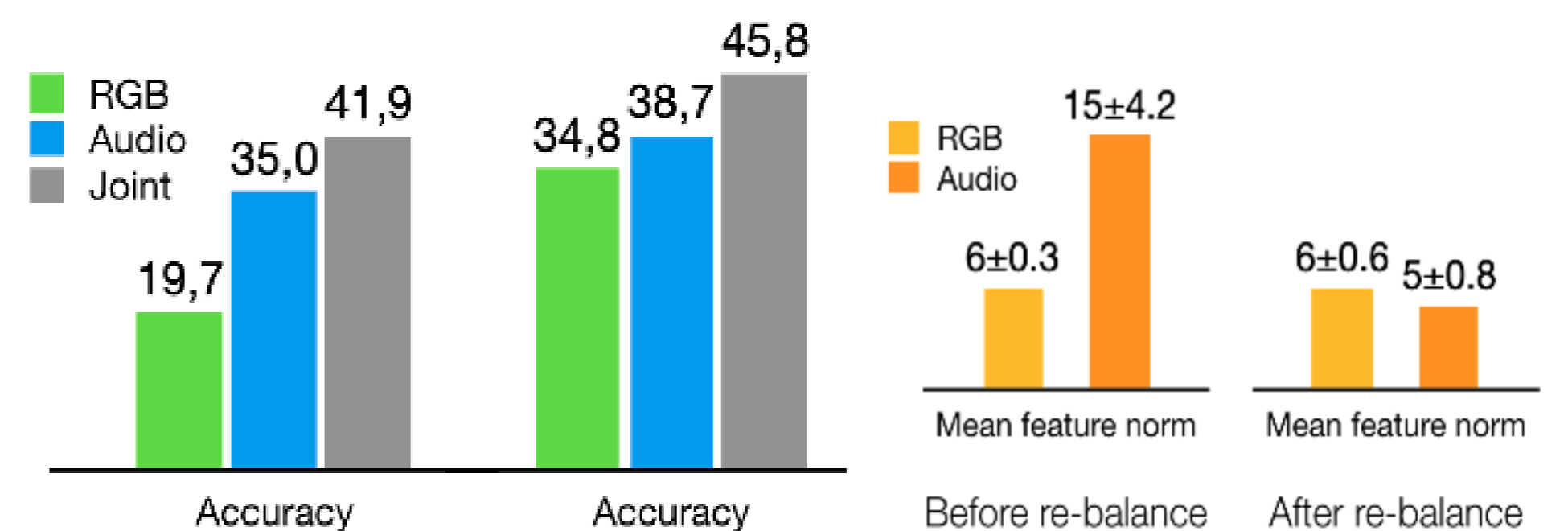
Multimodal Learning



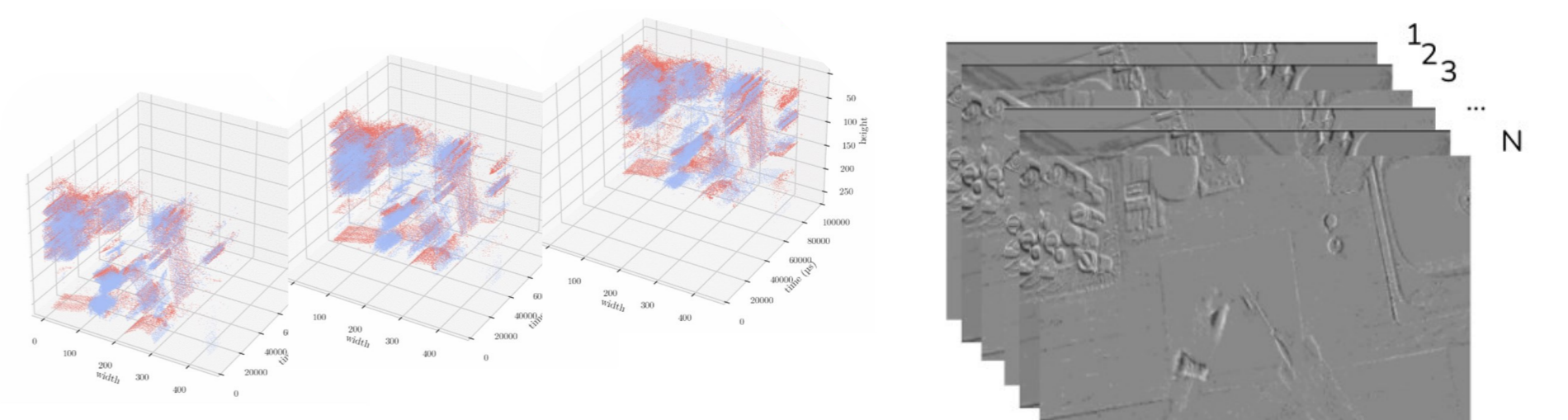
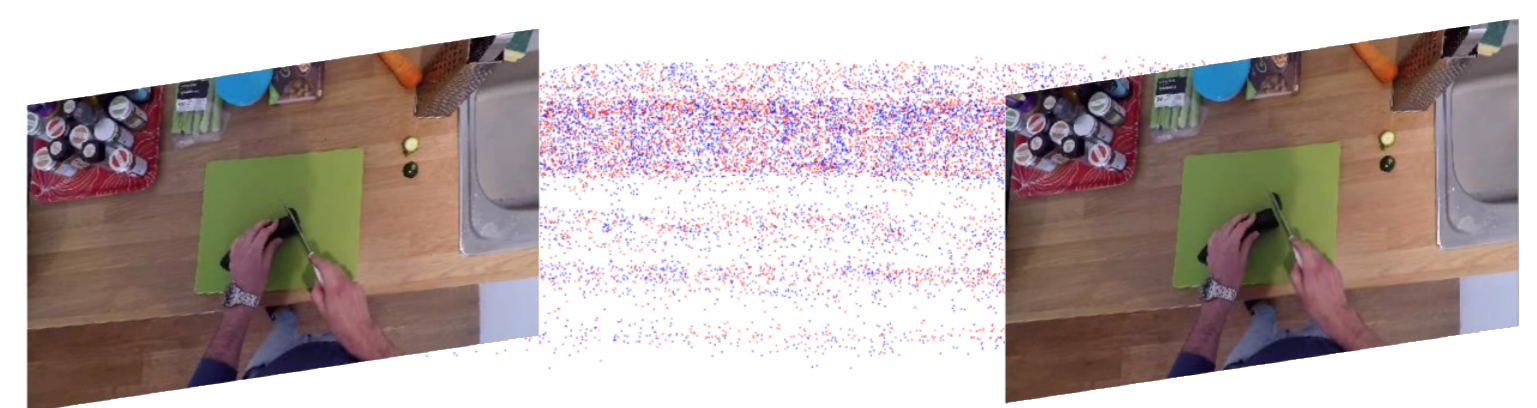
Unsupervised Domain Adaptation (UDA). Align the data distributions of the **source** and **target** domains.

4. Methods

Relative Norm Alignment (RNA). Re-balance the two modality contributions during training.



Event camera is novel bio-inspired sensors, which asynchronously capture pixel-level intensity changes in the form of "events".



5. References

- M Planamente et al. "Domain Generalization through Audio-Visual Relative Norm Alignment in First Person Action Recognition" WACV22
- Chiara Plizzari*, Mirco Planamente* et al. "E2 (GO) MOTION: Motion Augmented Event Stream for Egocentric Action Recognition" CVPR 2022