

# Target Scenario (Chapter 1)

- (Prof. Jae W. Lee) Clarify target scenarios. Edge device has a wide spectrum, which type of edge devices are you targeting?

# Thesis Statement & Storyline (Chapters 1,2)

- (Prof. Sung-Ju Lee) Add dissertation & contribution statement
- (Prof. Sung-Ju Lee) Refine the end-to-end storyline. Make the problems more tightly coupled.

# Technical Details (Chapters 4,5,6)

- **EagleEye**

- (Dr. Matthai Philipose) Clearly explain user study results
  - e.g., how crowded are the scenes? how much variation on the latency?
- (Dr. Matthai Philipose) Why additional model for face super-resolution? Why not integrate face upsampling process in the face recognition model (e.g., in the encoder)?
- (Dr. Matthai Philipose) Why not apply multi-exit model instead of variation-adaptive face recognition?
- (Prof Jae W. Lee) How is the face pose detected?
- (Dr. Matthai Philipose) What is “probes”?
- (Dr. Matthai Philipose) How to fine-tune ICN when there are multiple targets? Do we need 30 probes per each target?

# Technical Details (Chapters 4,5,6)

- **Pendulum**

- (Prof. Jae W. Lee) How to quantify “backbones” as numbers?
- (Dr. Matthai Philopose) Is the multi-user scheduling’s goal to make each user stay on the pareto-optimal curve?

- **Heimdall**

- (Dr. Matthai Philopose) How is GPU scheduling done in existing mobile DL frameworks?
  - e.g., at which granularity (e.g., kernel)? FIFO for multi-DNN concurrent inference scenarios?
- (Prof. Jae W. Lee) Is Heimdall an app/framework level scheduling solution? Is it simply making an API for priority scheduling?

# Discussion (Chapter 7)

- (Prof. Byung-Gon Chun, Prof. Sung-Ju Lee) Discuss generality across diverse task, workload, and hardware, including,
  - i) How to scale EagleEye for identifying 1,000 targets?
  - ii) Will GPU preemption still be important when NPUs are more prevalent?
- (Dr. Matthai Philipose) How practical would the proposed techniques be? Wouldn't production systems be provisioned with sufficient resource?
- (Dr. Matthai Philipose) Would next-generation network & compute technologies solve the resource bottleneck problem?

# Future Work (Chapter 7)

- (Dr. Matthai Philipose) Think of ambitious future research directions
  - e.g., where would 10-100x performance gain come from?
- (Prof. Sung-Ju Lee) Prepare a video demo