

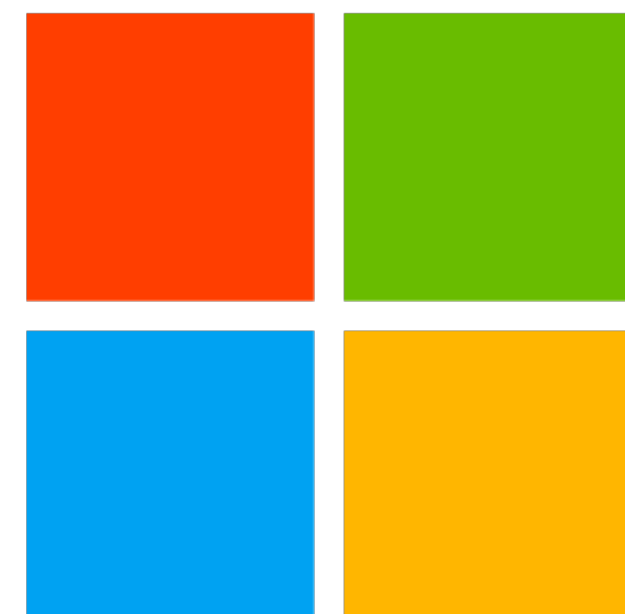
EMO: Real-Time Emotion Recognition from Single-Eye Images for Resource-Constrained Eyewear Devices

Hao Wu¹, Jinghao Feng¹, Xuejin Tian¹, Edward Sun¹, Yunxin Liu², Bo Dong³, Fengyuan Xu^{1*}, Sheng Zhong¹

¹Nanjing University



²Microsoft Research



³Kitware

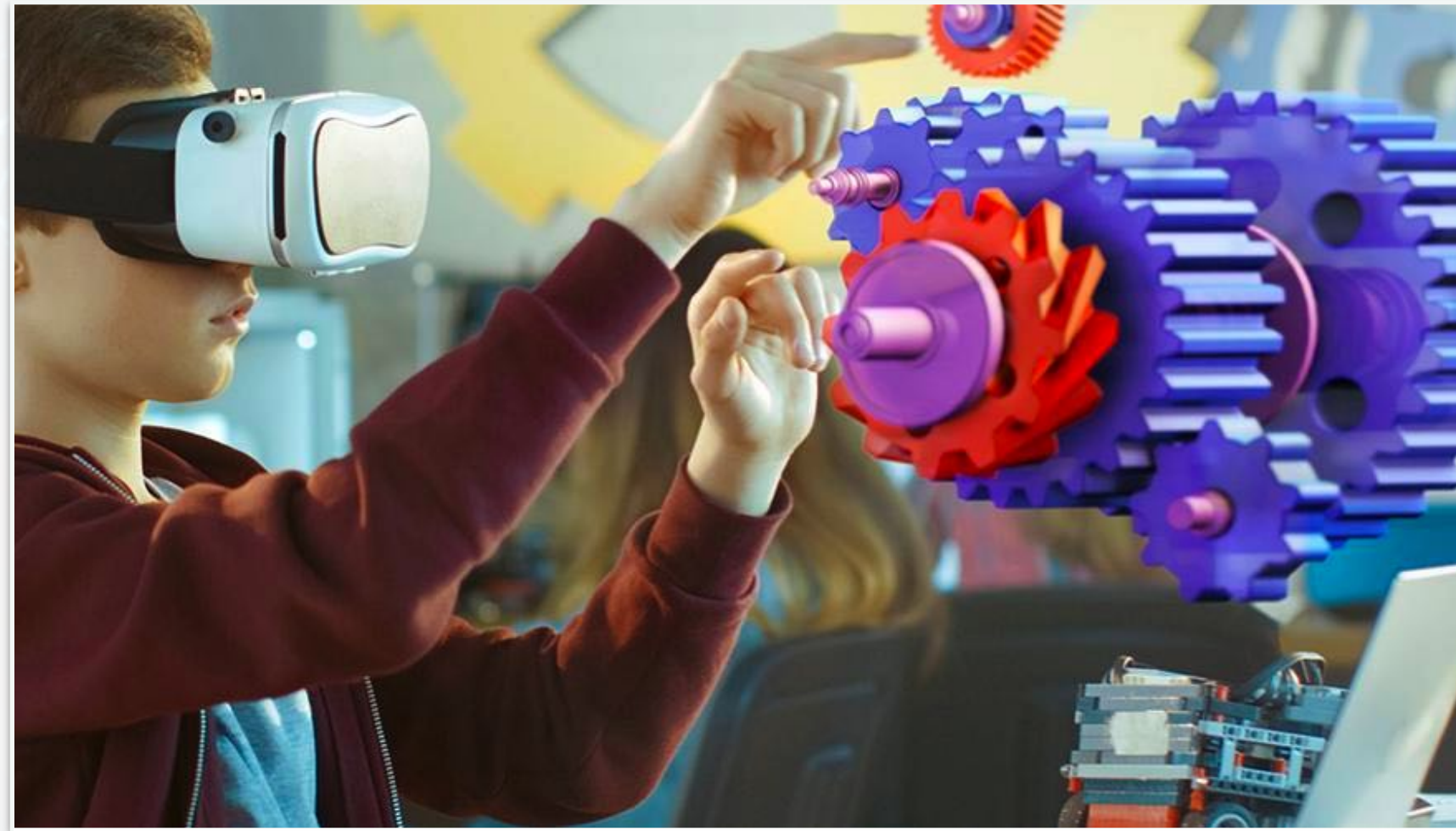




Eyewear: wearable devices for eyes, e.g., AR, VR, Eye-tracker.

Eyewear devices paint a magical future for HCI

Eyewear devices paint a magical future for HCI



Education



Commerce

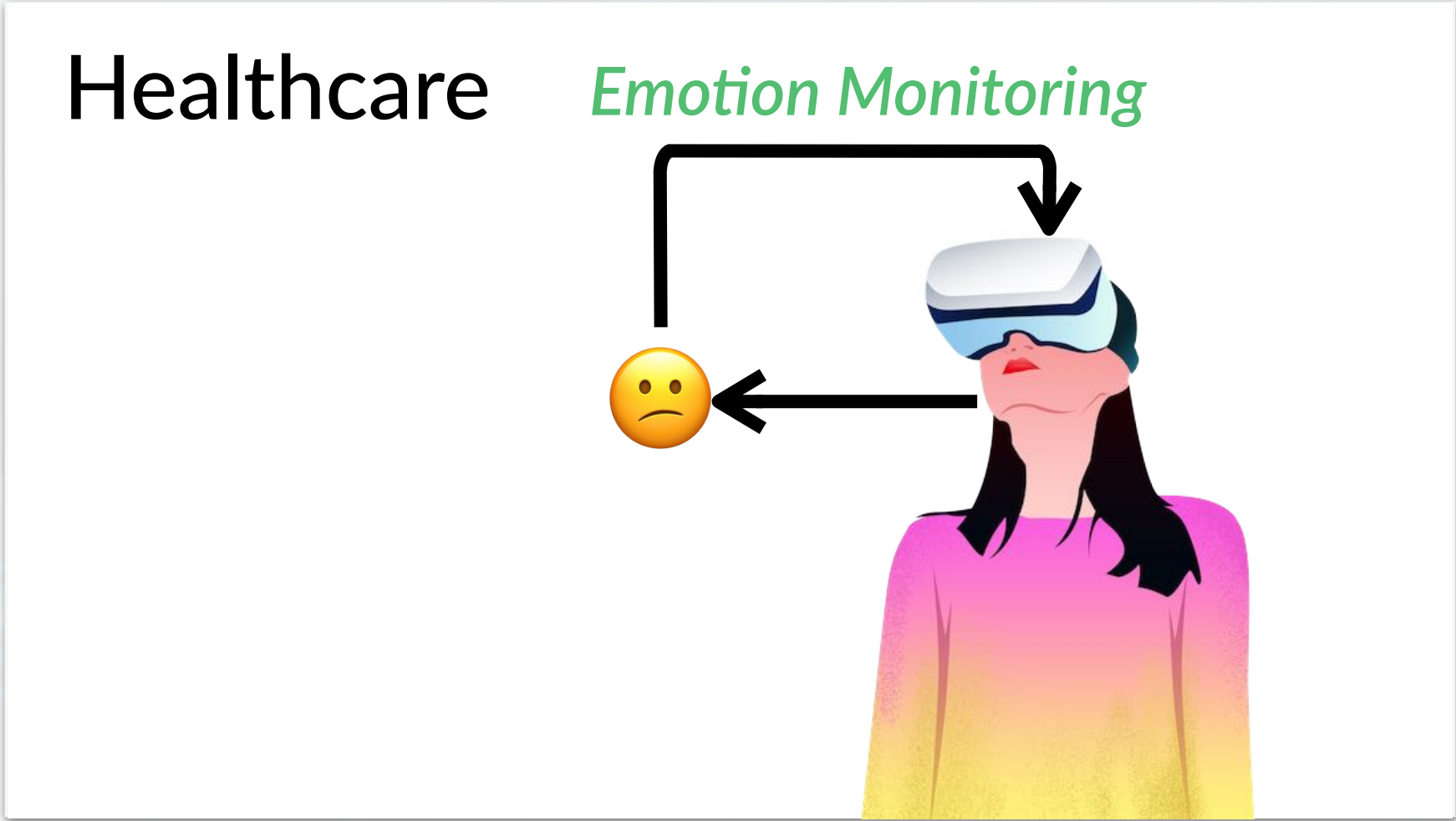
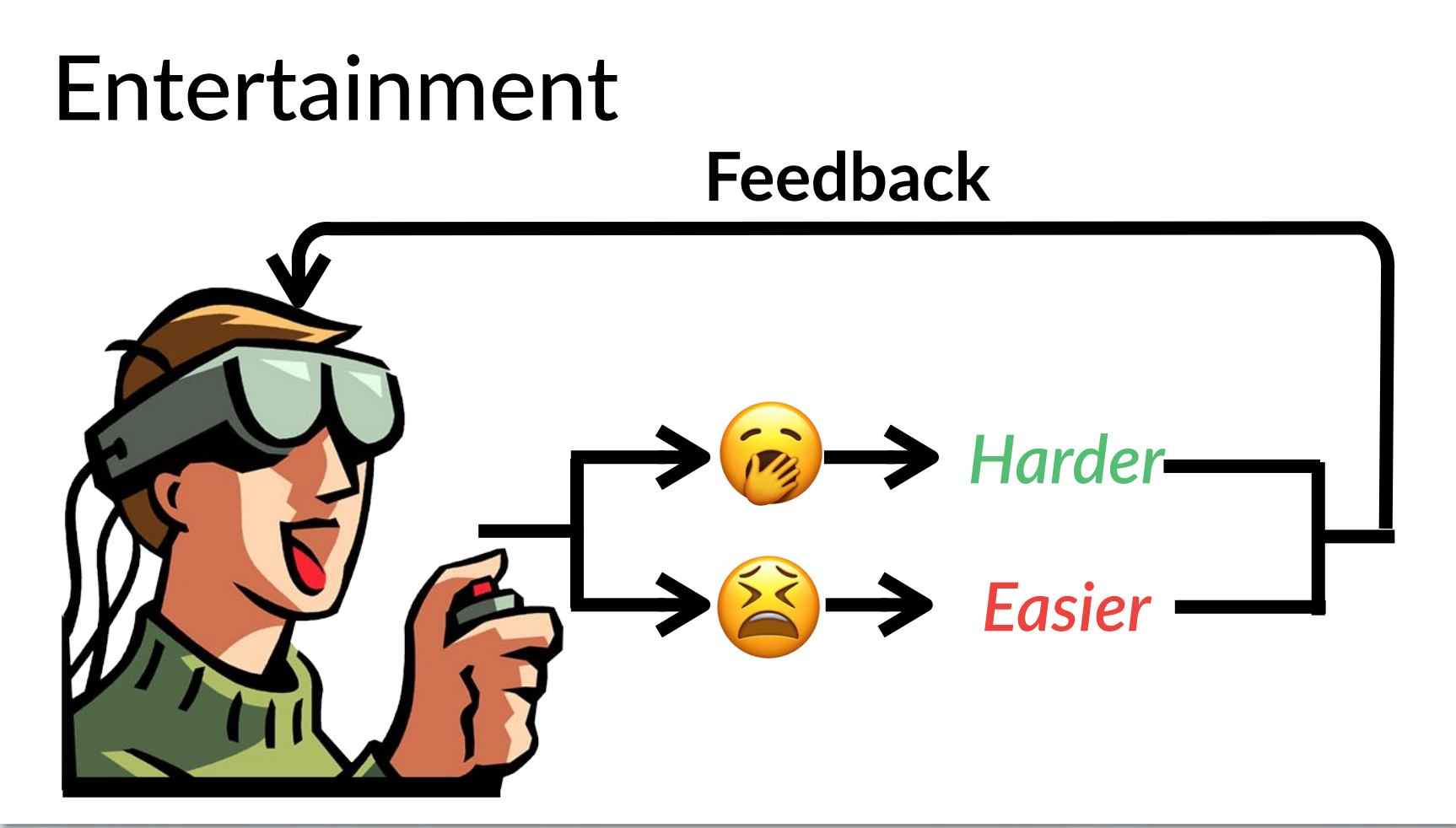
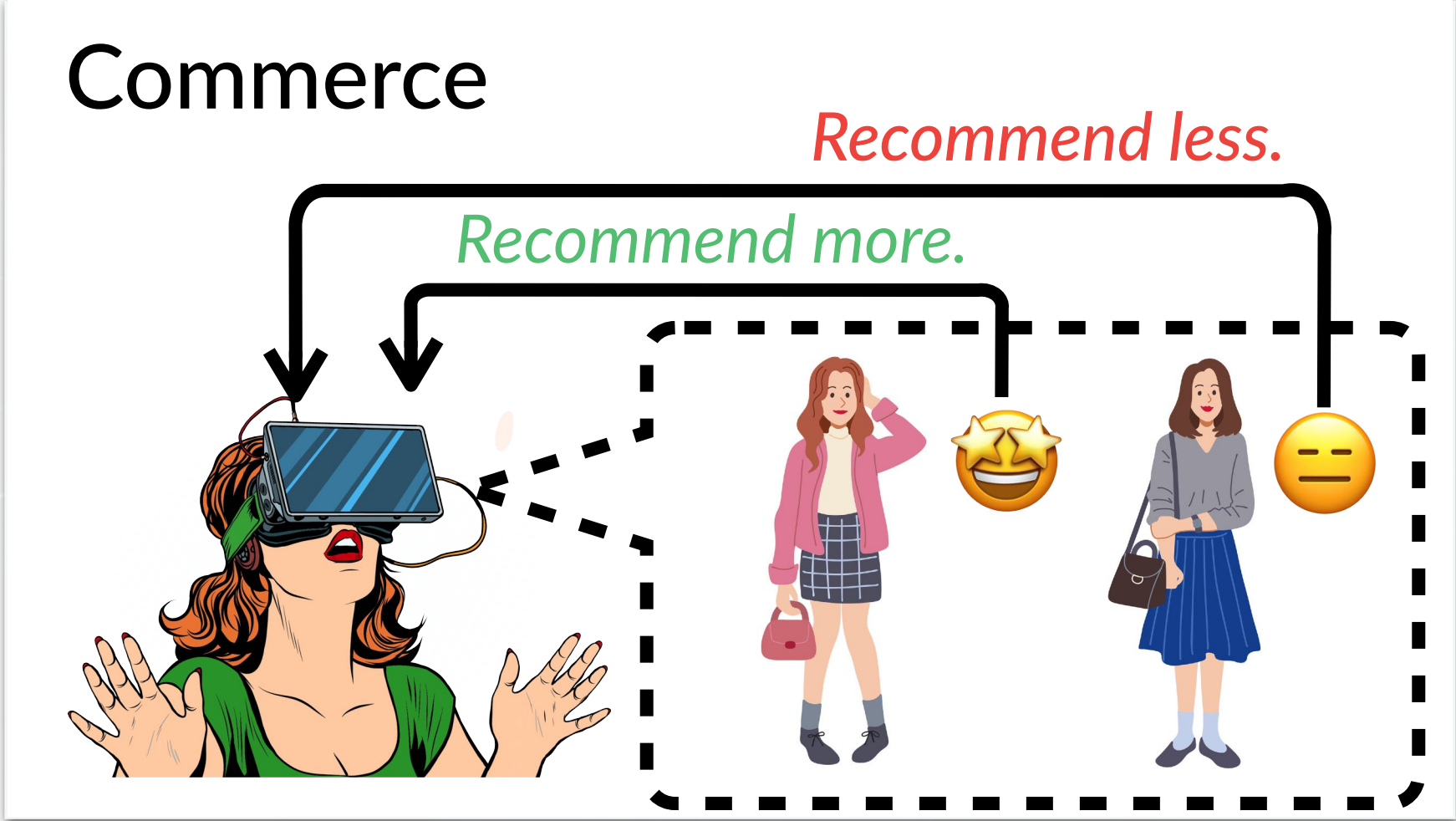


Entertainment



Healthcare

Emotion recognition is highly desirable for eyewear



EMO: Emotion Recognition on Eyewear



EMO: Emotion Recognition on Eyewear

Eyewear



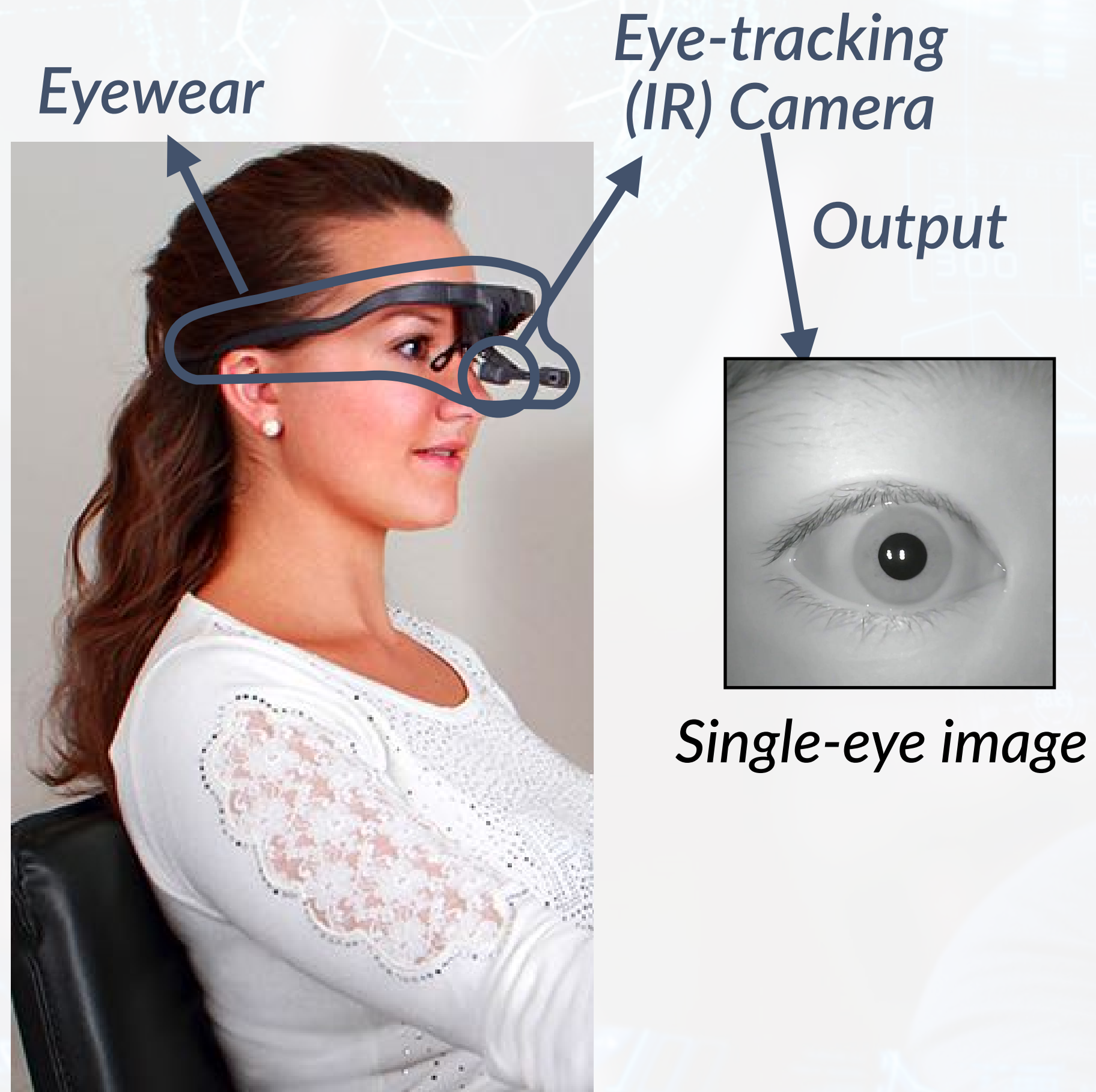
EMO: Emotion Recognition on Eyewear

Eyewear

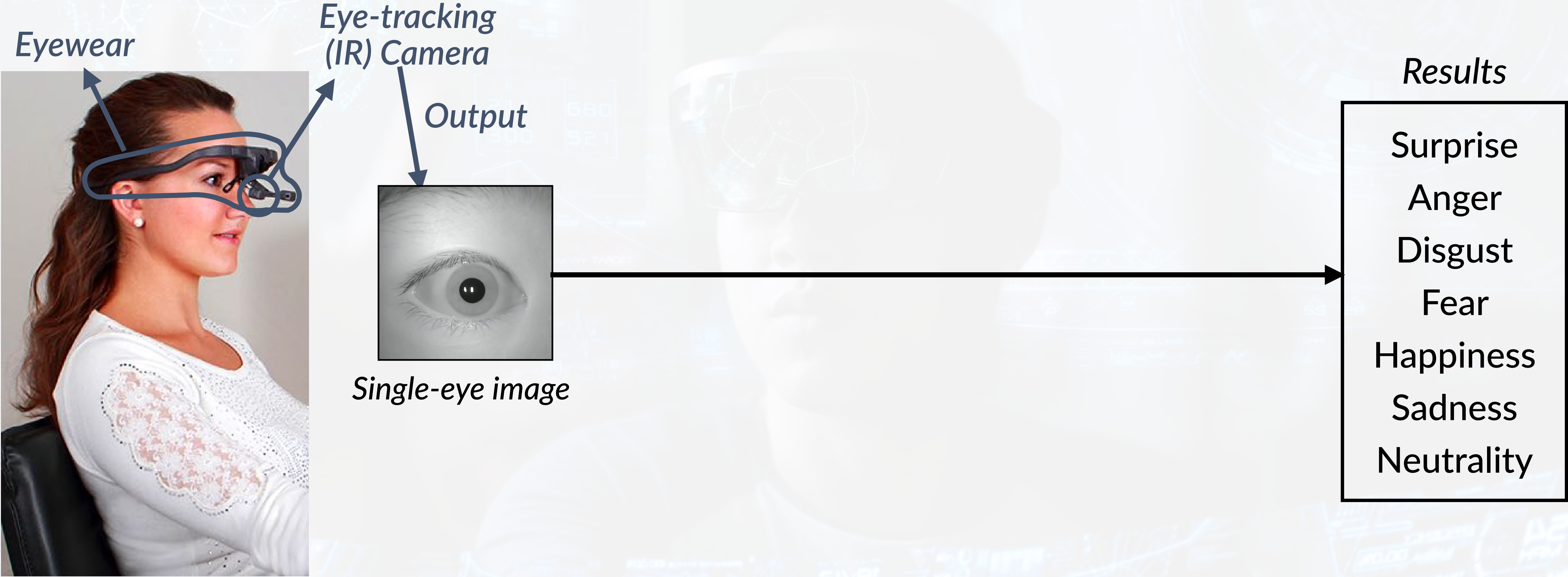
*Eye-tracking
(IR) Camera*



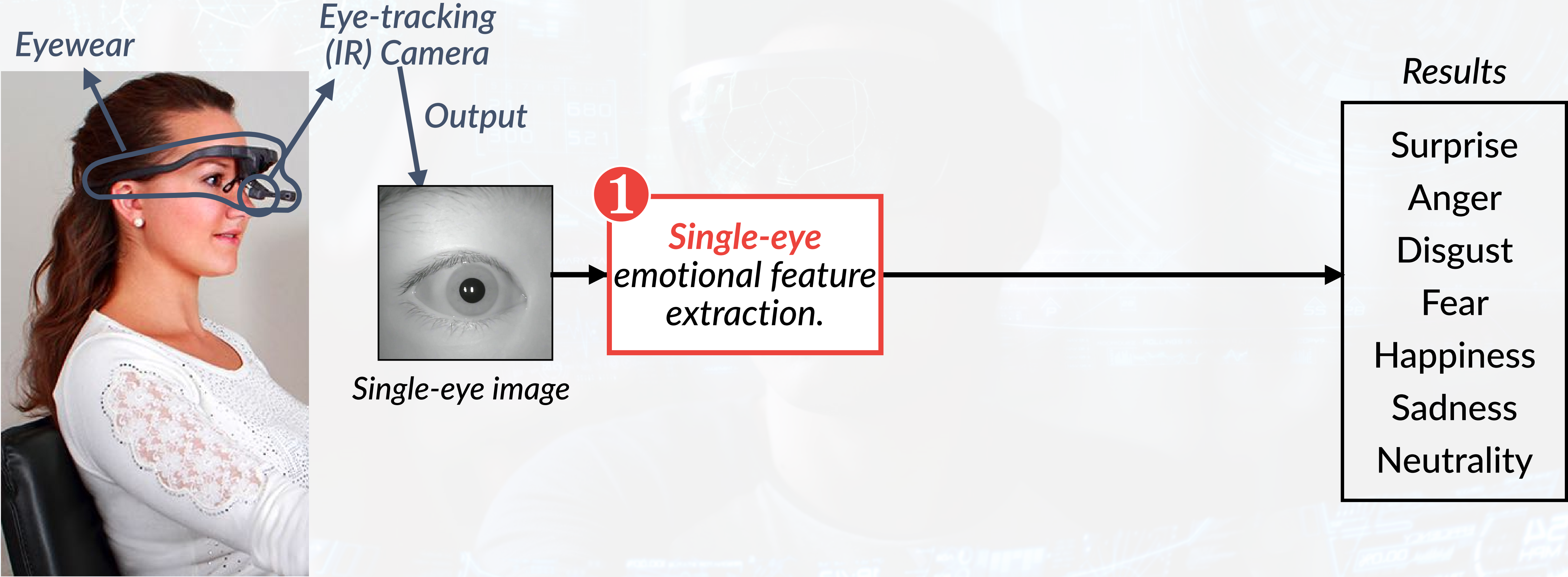
EMO: Emotion Recognition on Eyewear



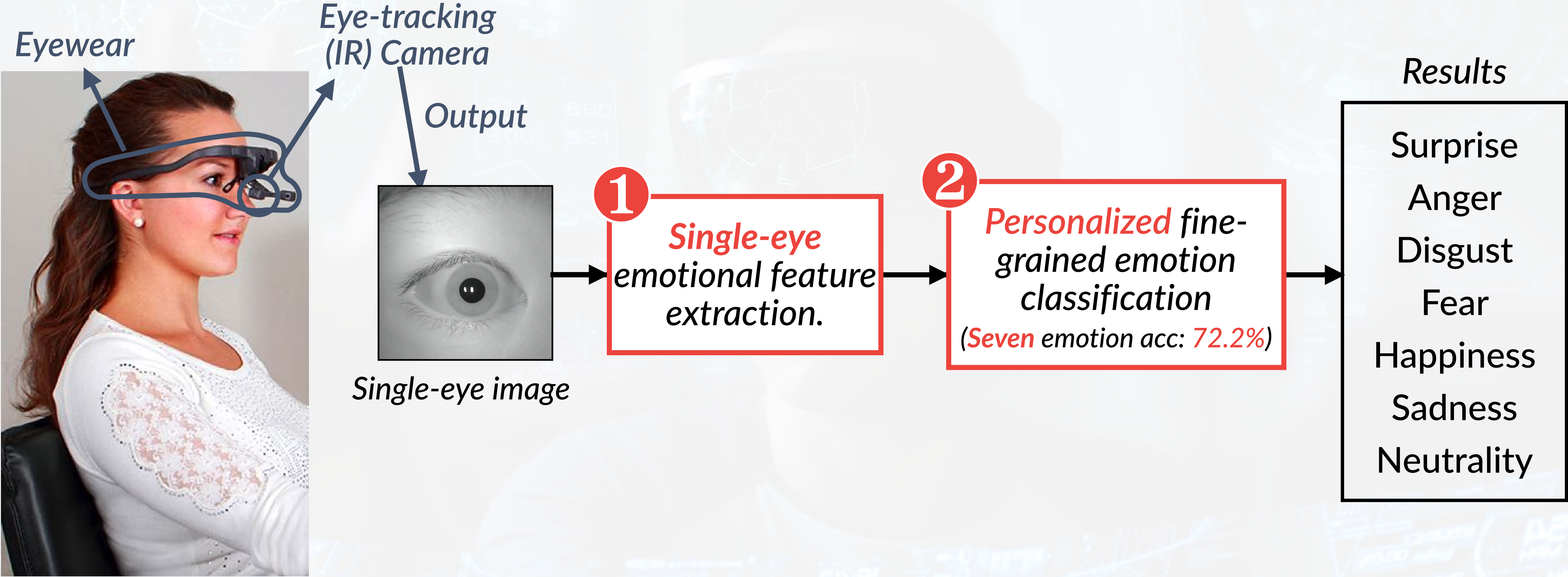
EMO: Emotion Recognition on Eyewear



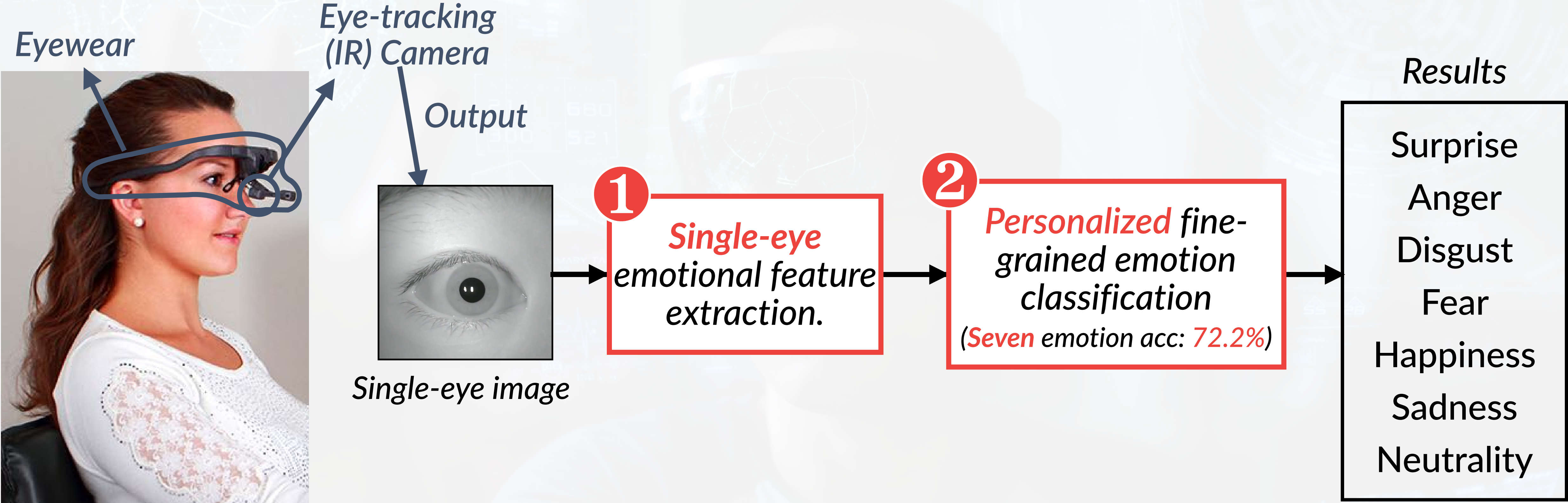
EMO: Emotion Recognition on Eyewear



EMO: Emotion Recognition on Eyewear



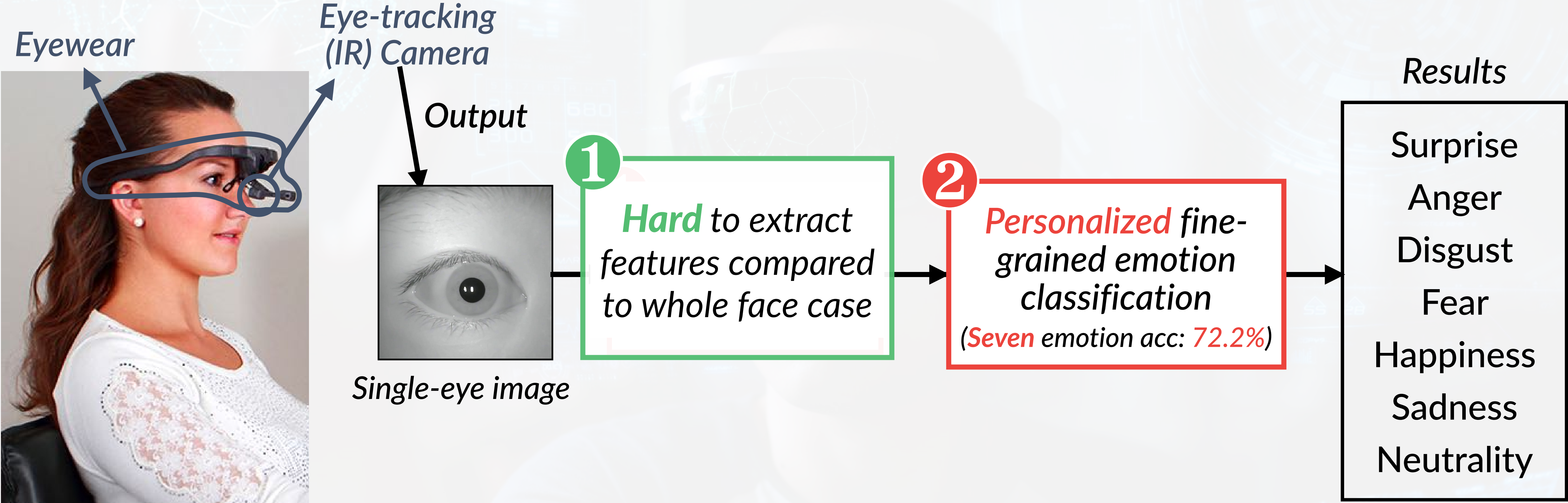
EMO: Emotion Recognition on Eyewear



3 *Real-time untethered recognition system.*

(CPU: *Open-Q820/12%*, Memory: *73MB*, Speed: *12.8fps*)

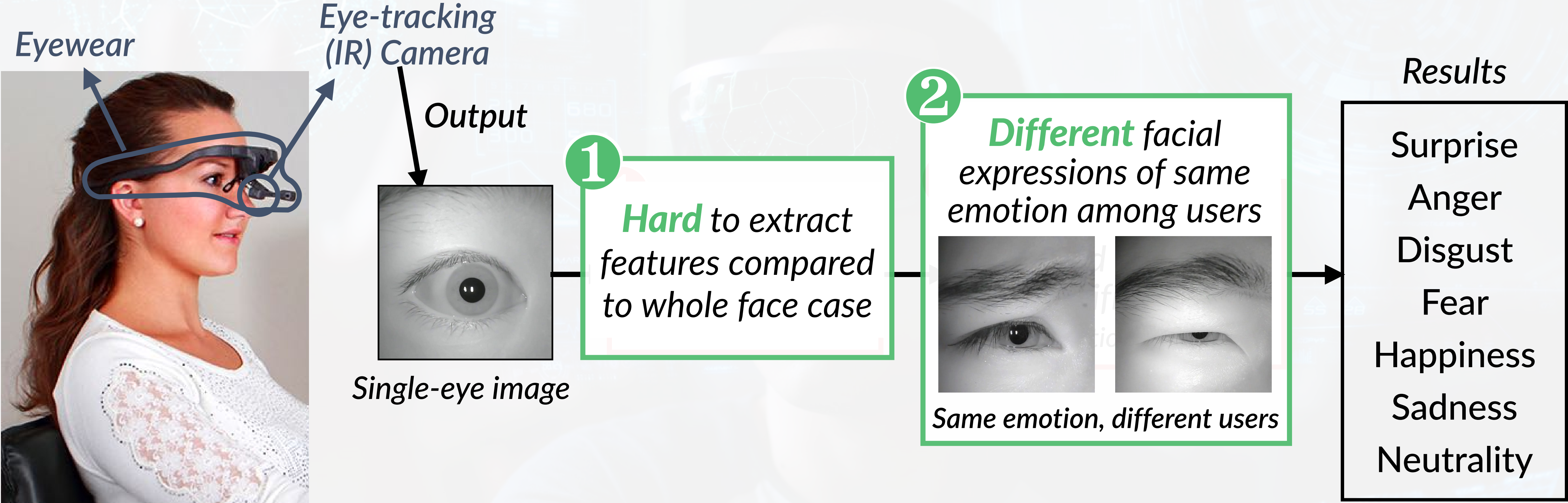
EMO: Emotion Recognition on Eyewear



3 *Real-time untethered recognition system.*

(CPU: *Open-Q820/12%*, Memory: *73MB*, Speed: *12.8fps*)

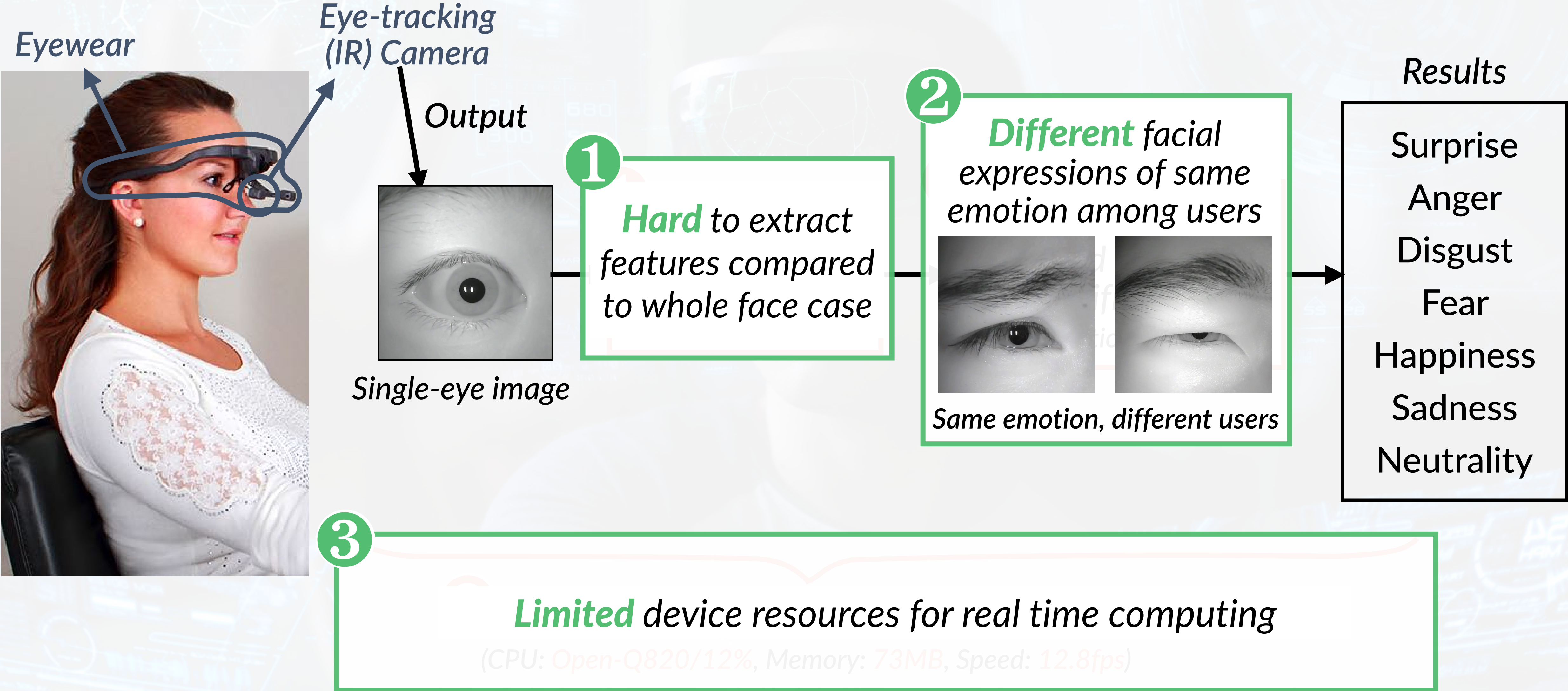
EMO: Emotion Recognition on Eyewear



3 *Real-time* untethered recognition system.

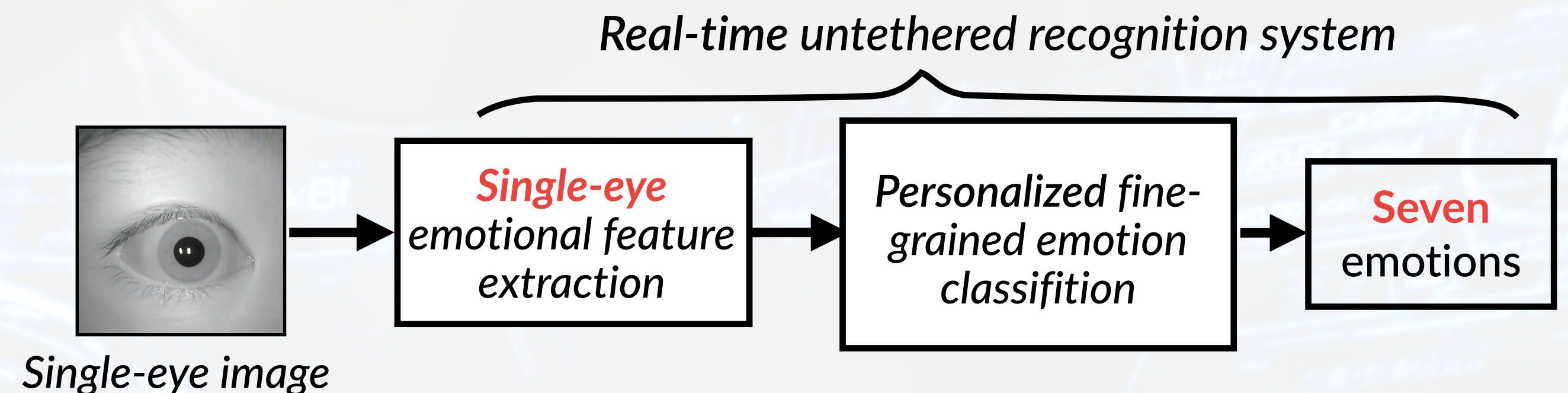
(CPU: *Open-Q820/12%*, Memory: *73MB*, Speed: *12.8fps*)

EMO: Emotion Recognition on Eyewear



Today's talk

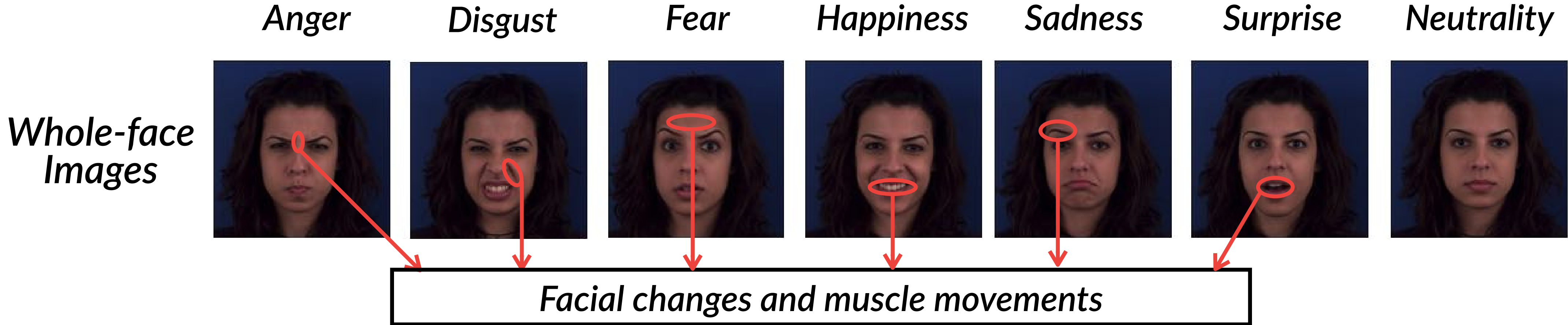
1. *Single-eye emotional feature extraction.*
2. *Personalized fine-grained emotion classification.*
3. *Real-time untethered recognition system.*
4. *The prototype and evaluation.*



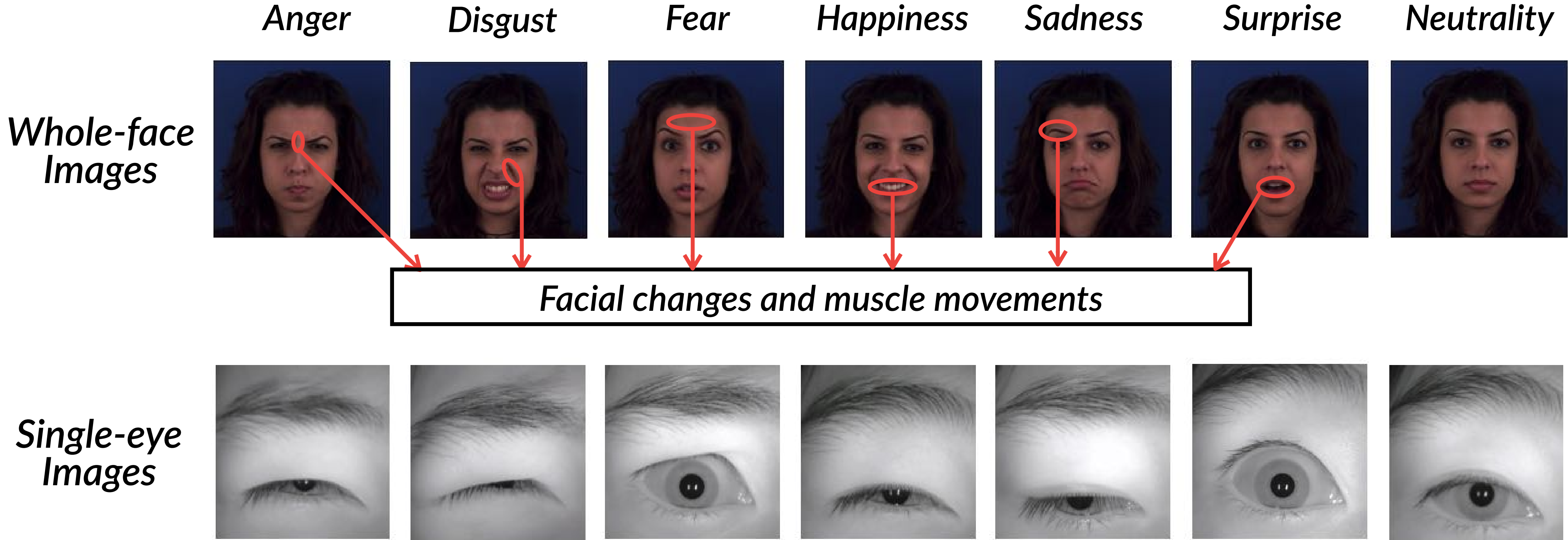
Hard to extract features compared to whole face case



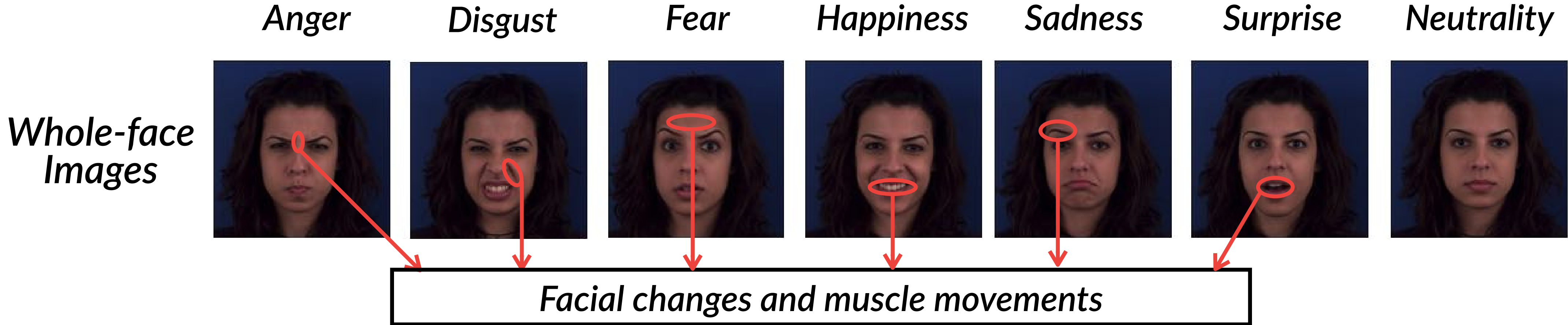
Hard to extract features compared to whole face case



Hard to extract features compared to whole face case

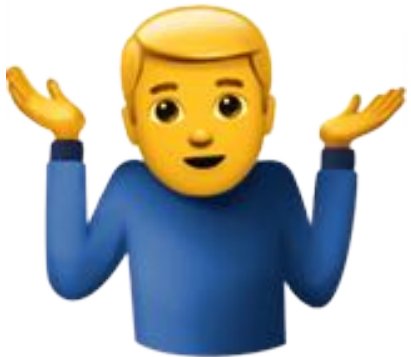
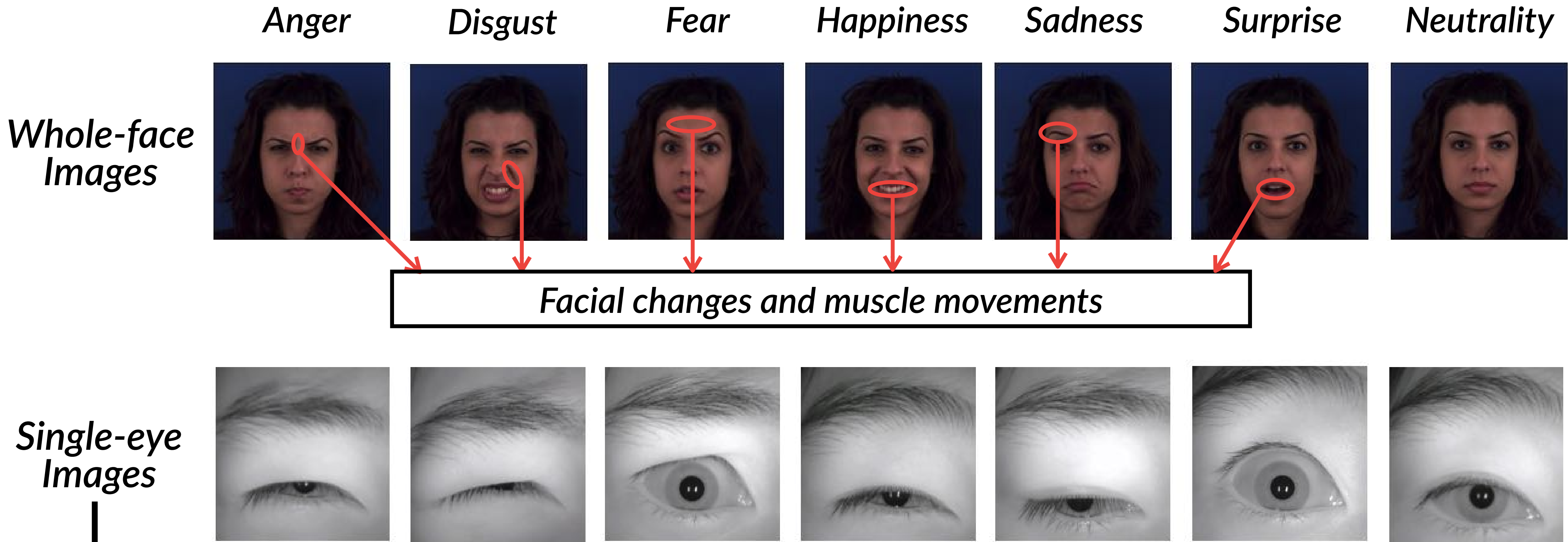


Hard to extract features compared to whole face case



Features are quite *less*

Hard to extract features compared to whole face case



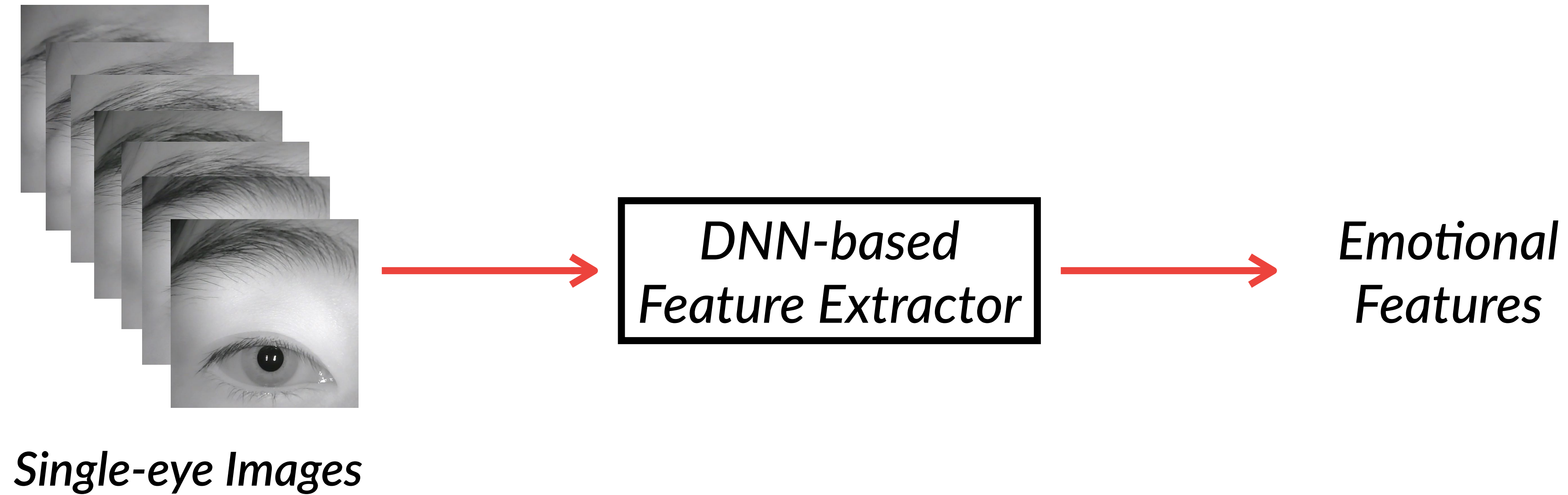
Features are quite **less**



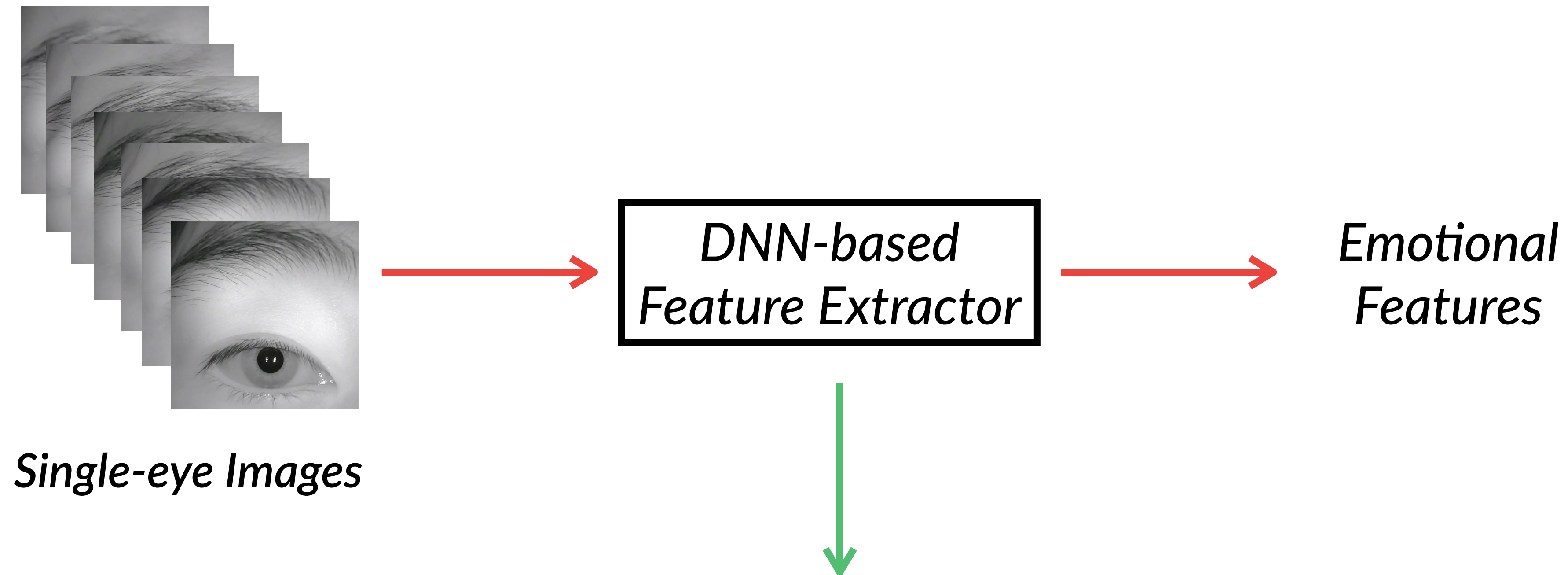
- 1. 65% emotional features
- 2. Close view of fixed area



Single-eye emotional feature extraction.



Single-eye emotional feature extraction.



Single-eye Images

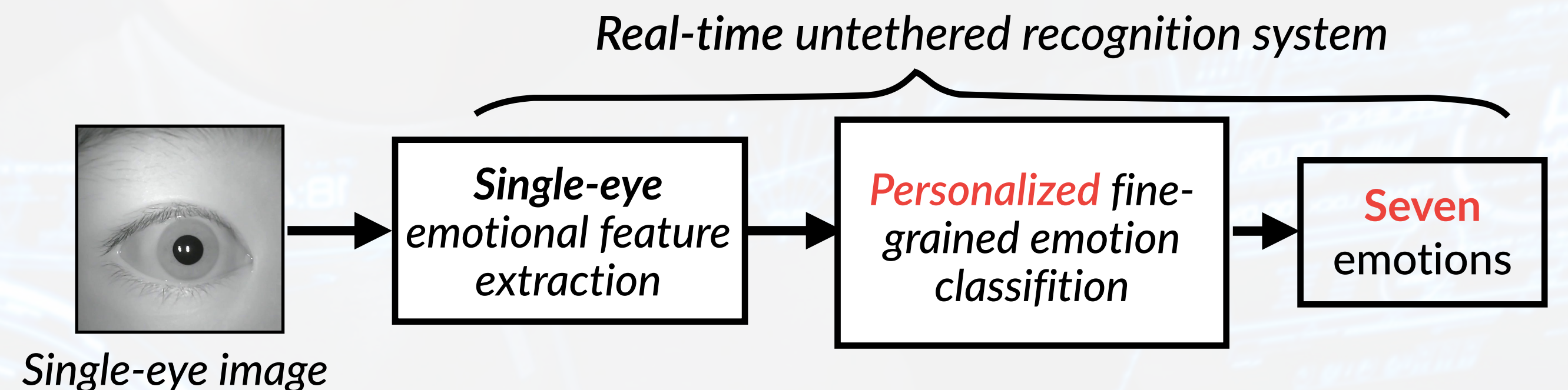
*DNN-based
Feature Extractor*

*Emotional
Features*

1. An *improved* ResNet. (Single-eye image Friendly)
2. Train on imperfect datasets. (*Two-phase* training)

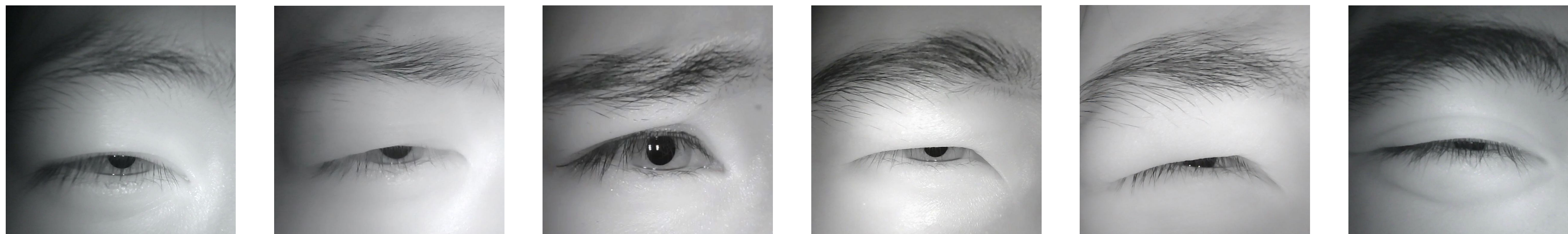
Today's talk

1. *Single-eye emotional feature extraction.*
2. *Personalized fine-grained emotion classification.*
3. *Real-time untethered recognition system.*
4. *The prototype and evaluation.*



Different facial expressions of same emotion among users

Different people express the same emotion



Different facial expressions of same emotion among users

Different people express the same emotion



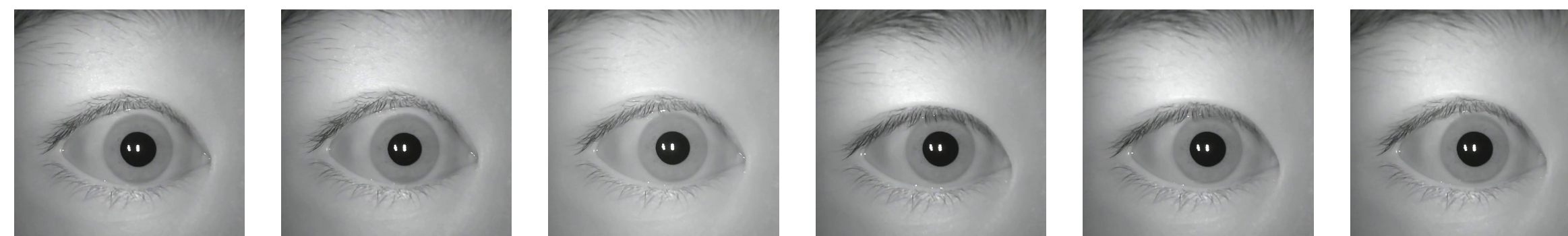
Quite different



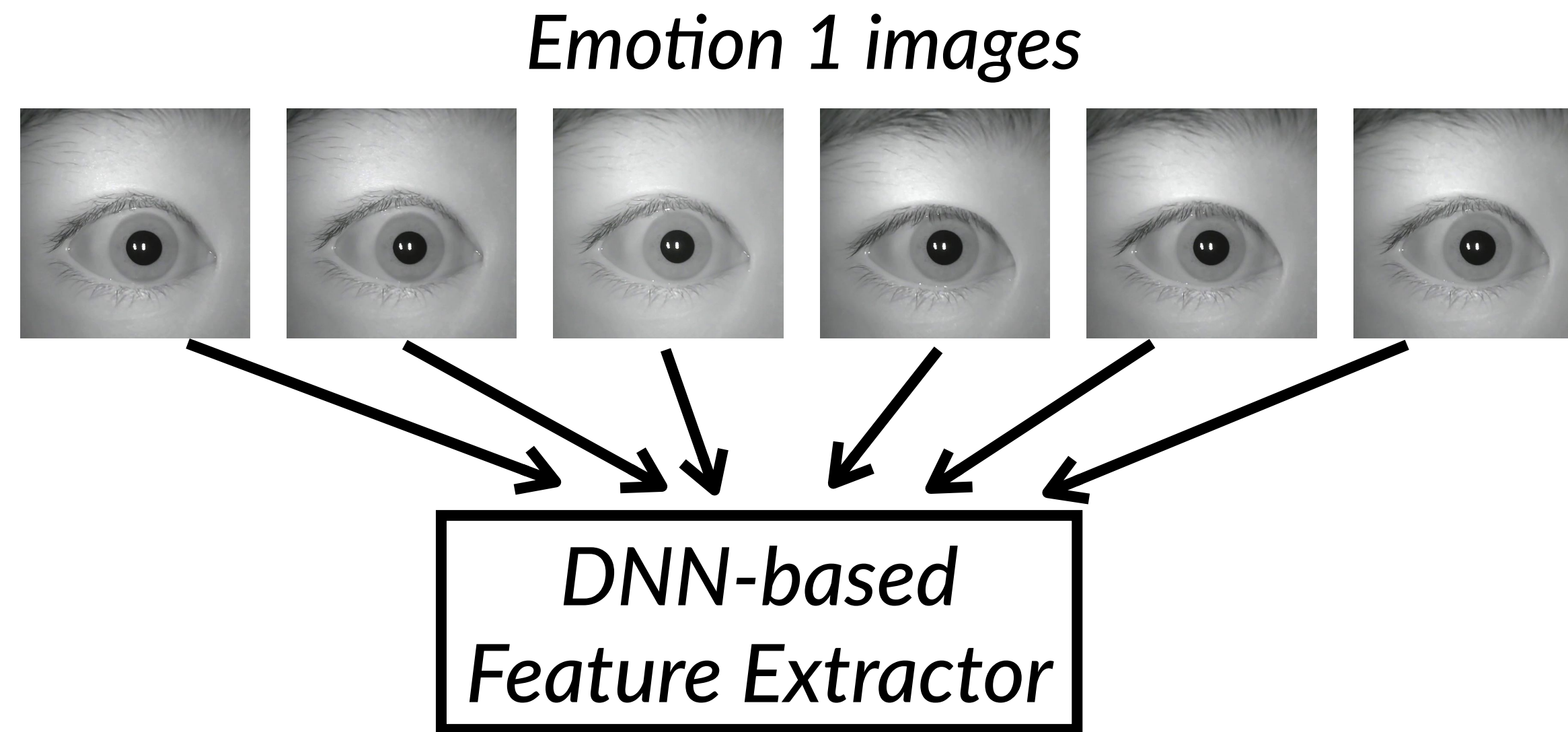
One-size-fits-all strategy is not suitable

Personalized fine-grained emotion classification

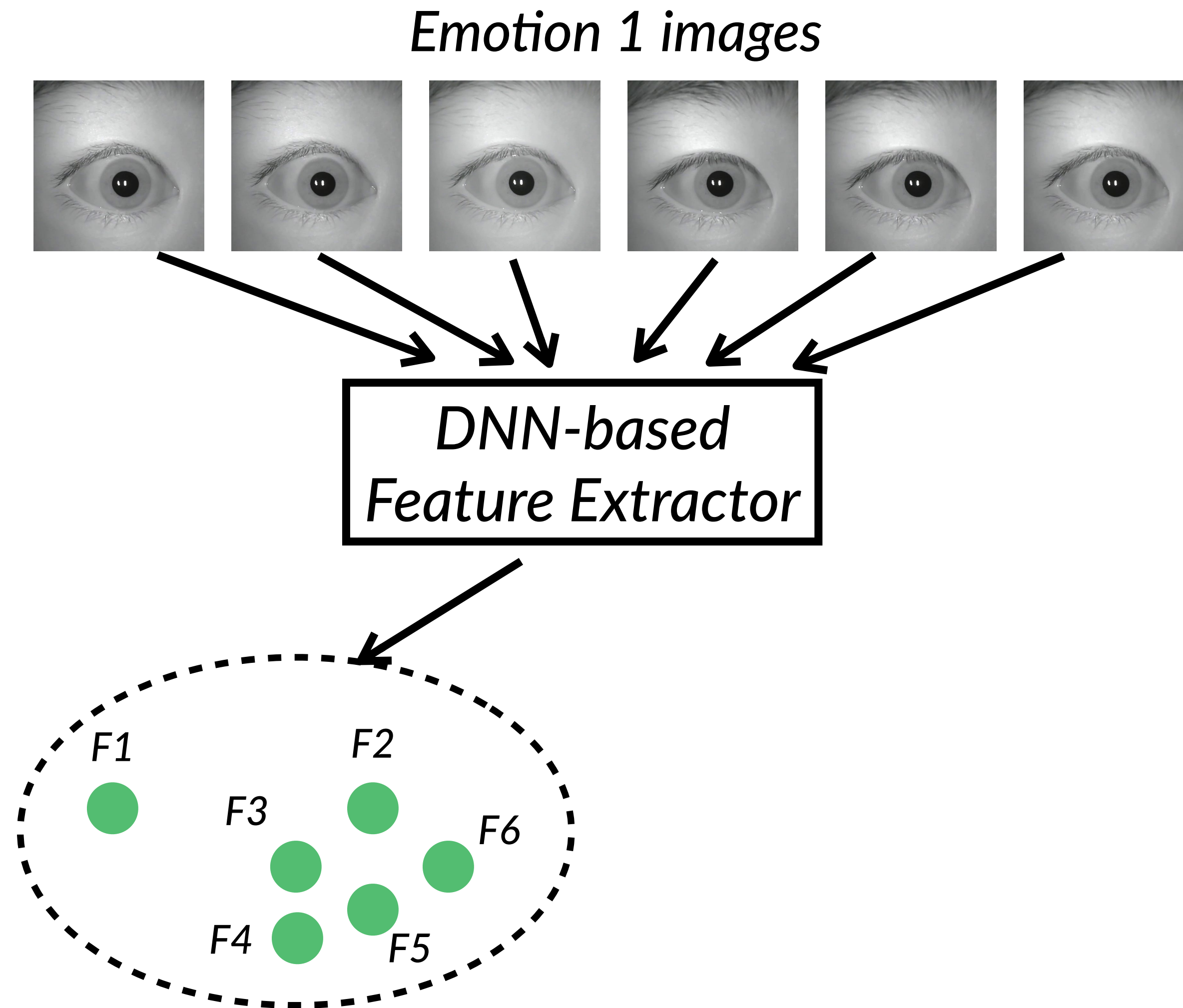
Emotion 1 images



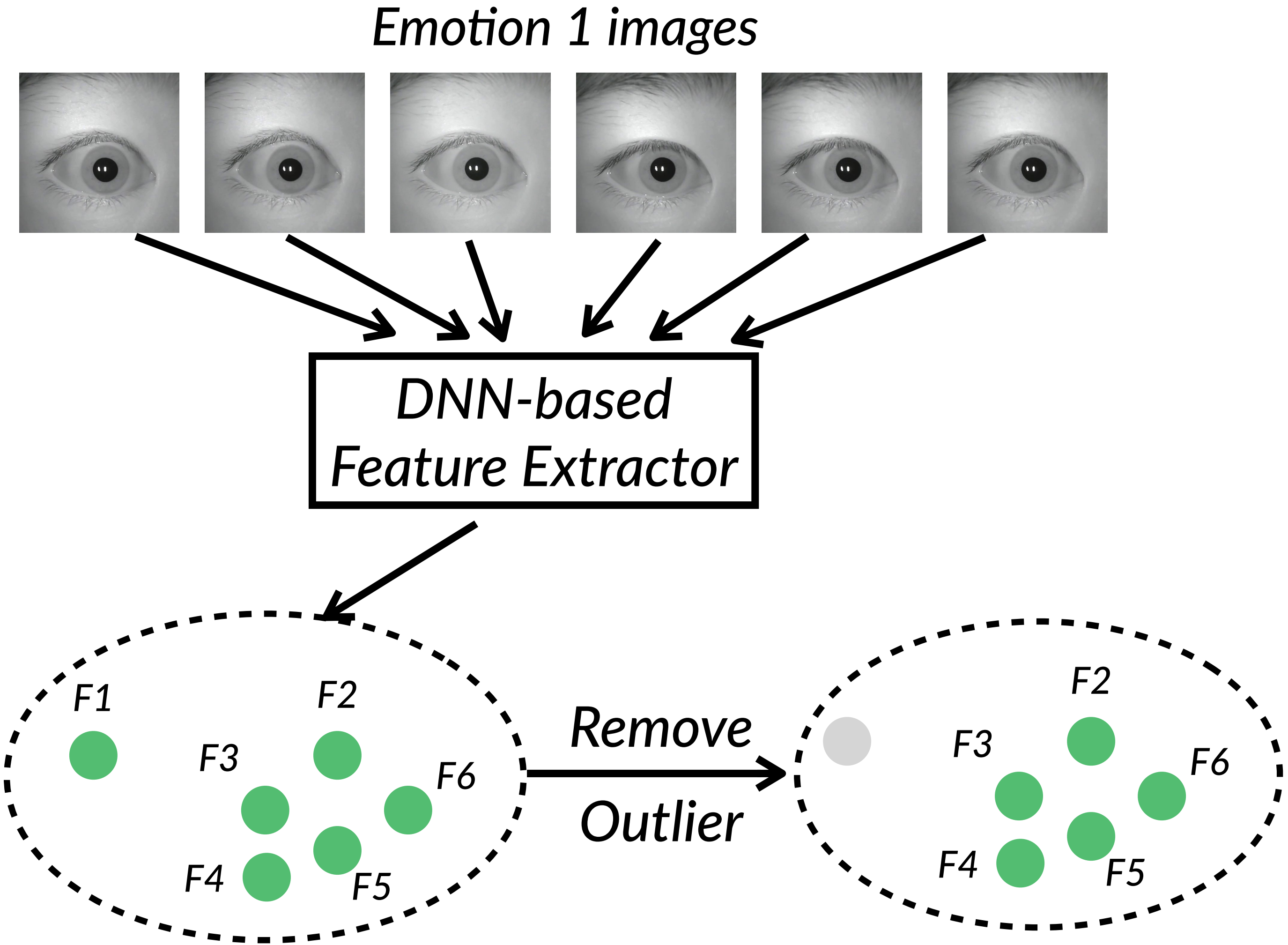
Personalized fine-grained emotion classification



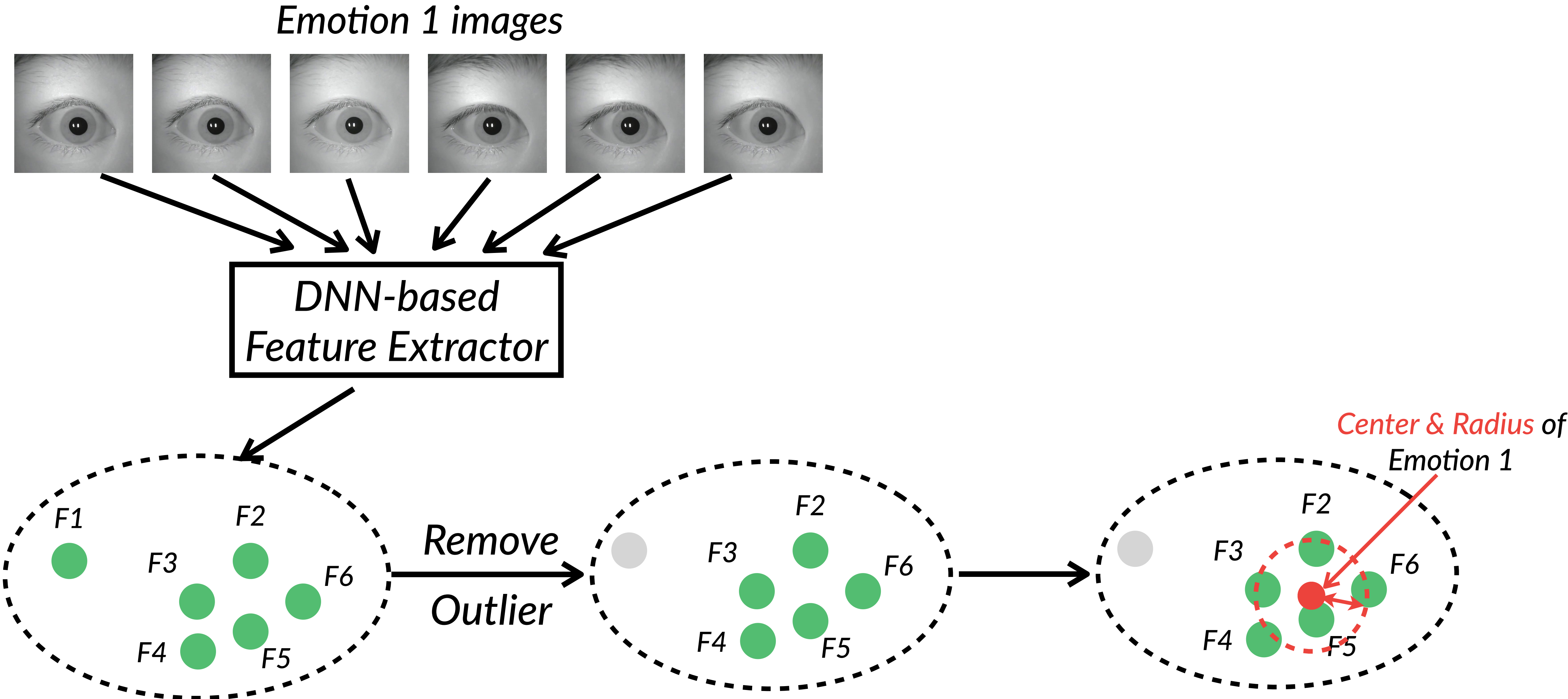
Personalized fine-grained emotion classification



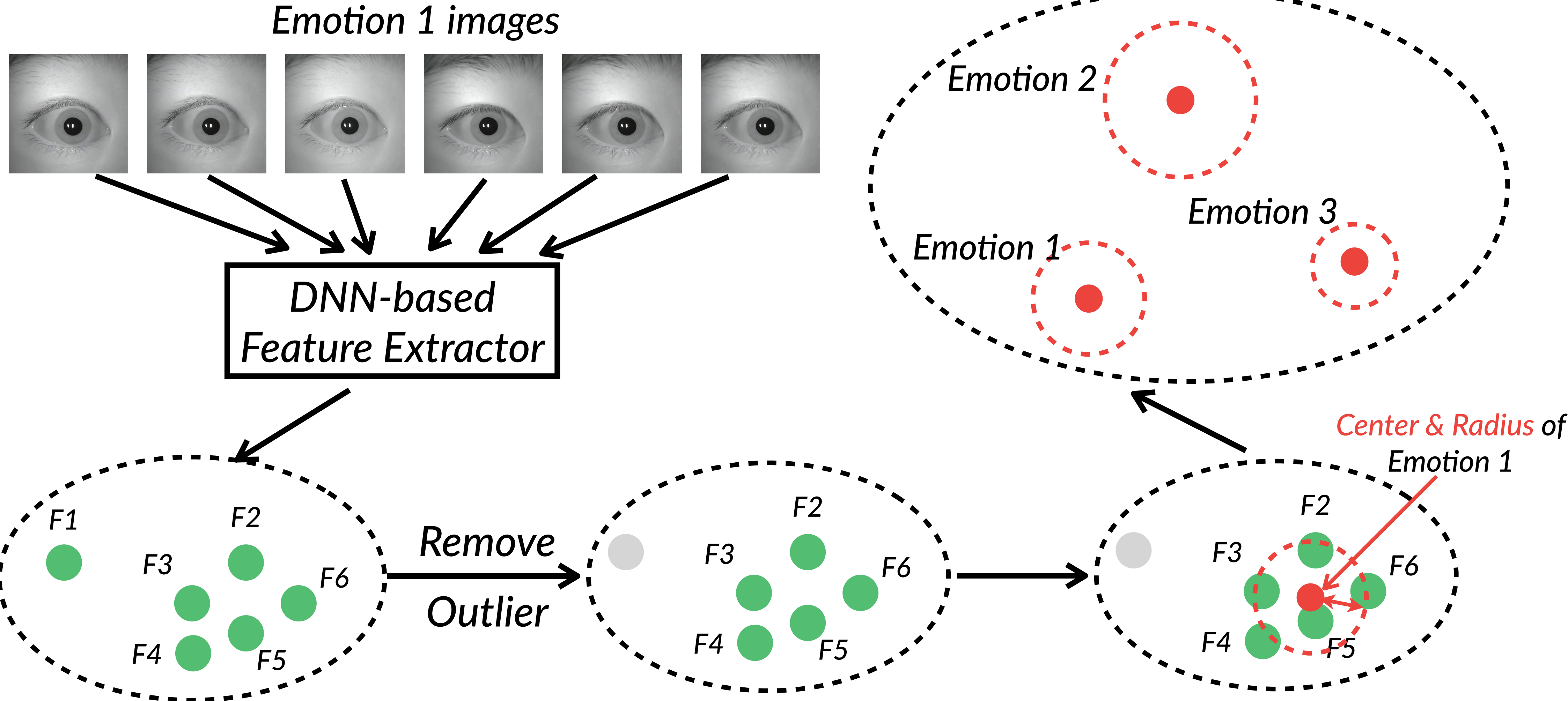
Personalized fine-grained emotion classification



Personalized fine-grained emotion classification

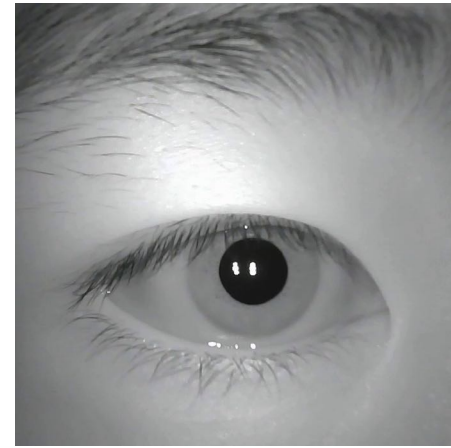


Personalized fine-grained emotion classification

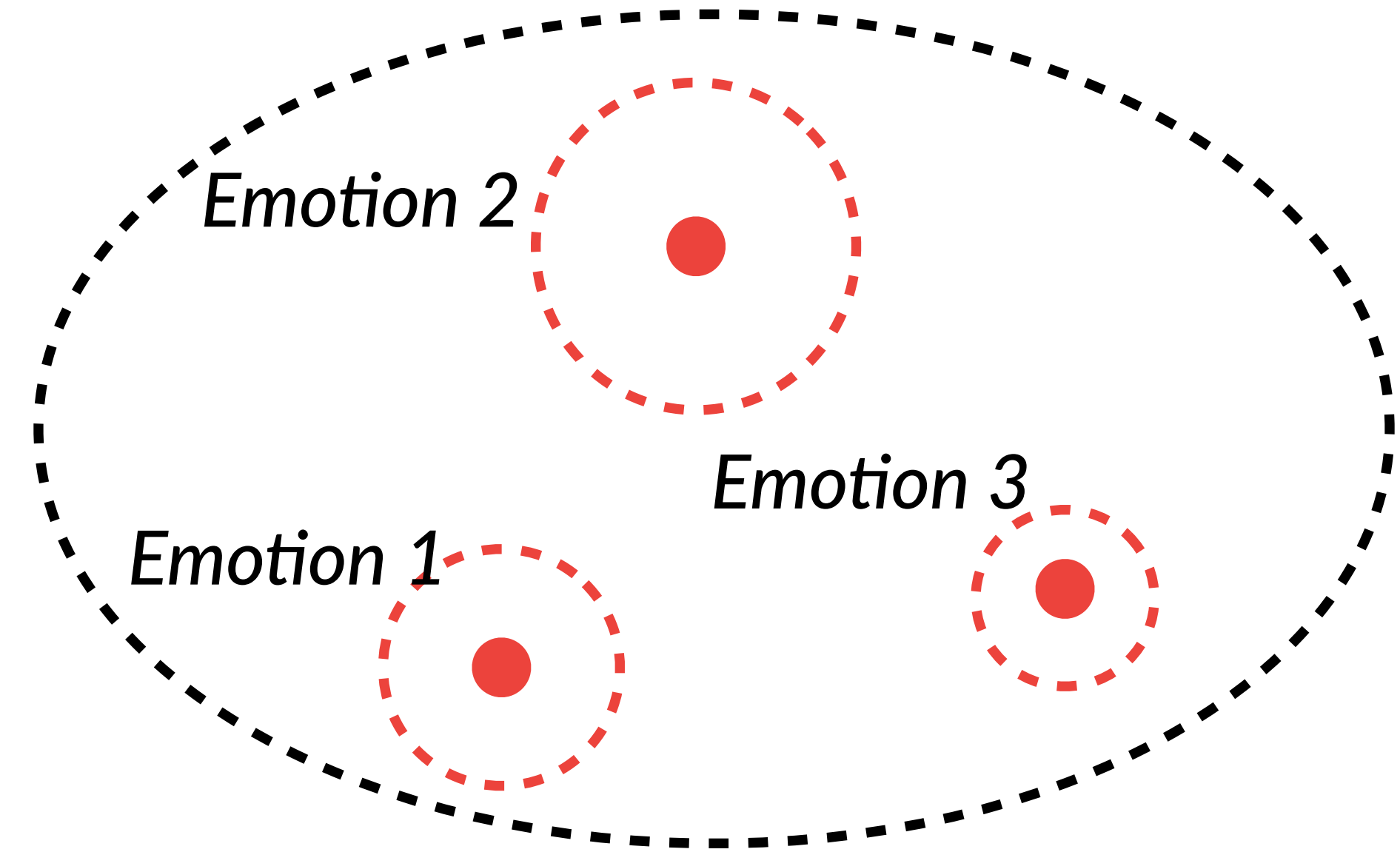


Personalized fine-grained emotion classification

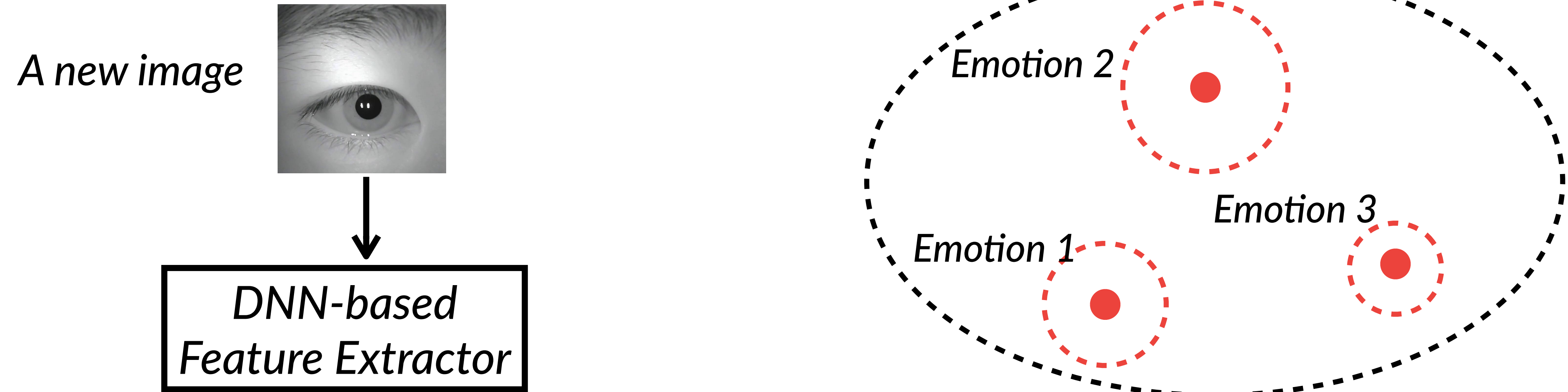
A new image



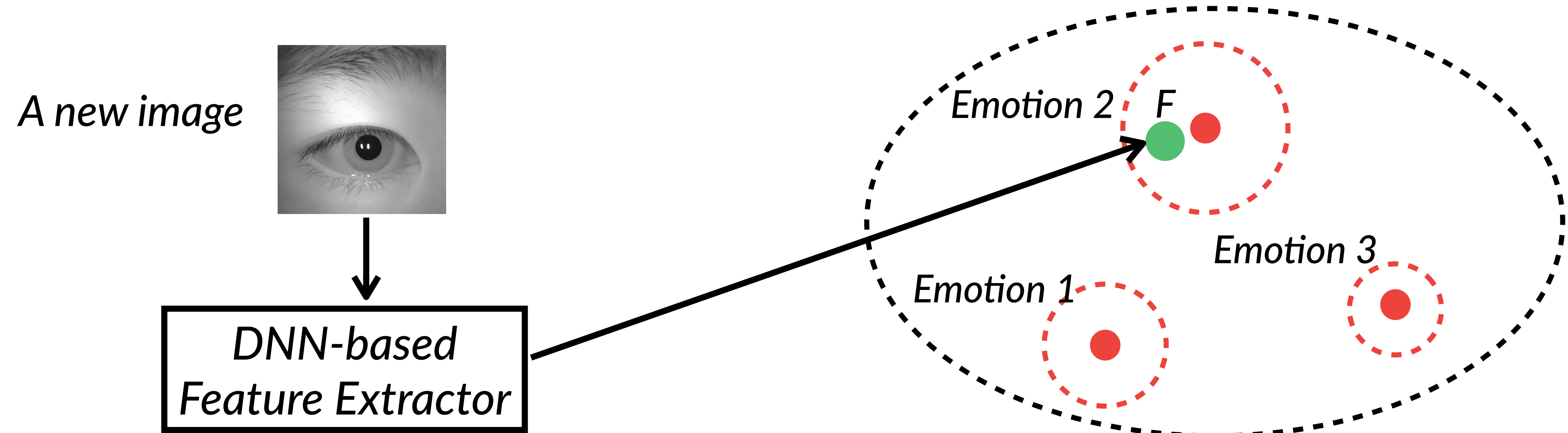
*DNN-based
Feature Extractor*



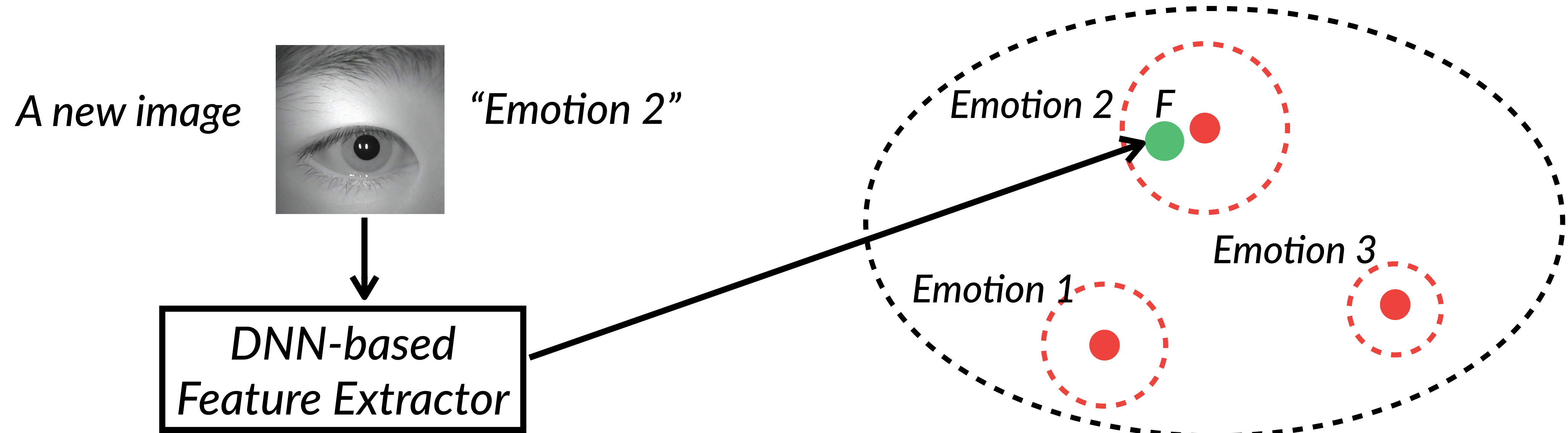
Personalized fine-grained emotion classification



Personalized fine-grained emotion classification

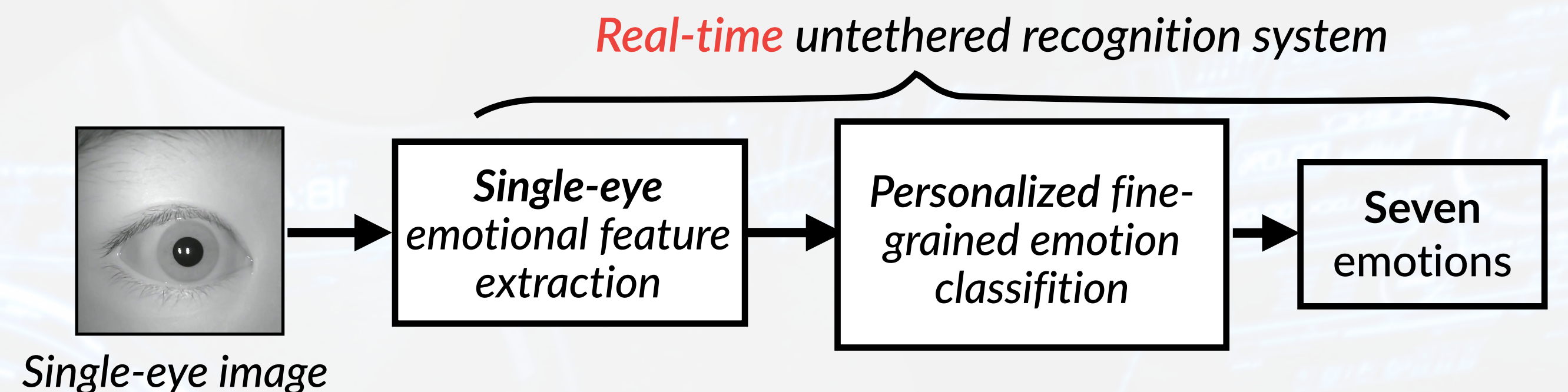


Personalized fine-grained emotion classification

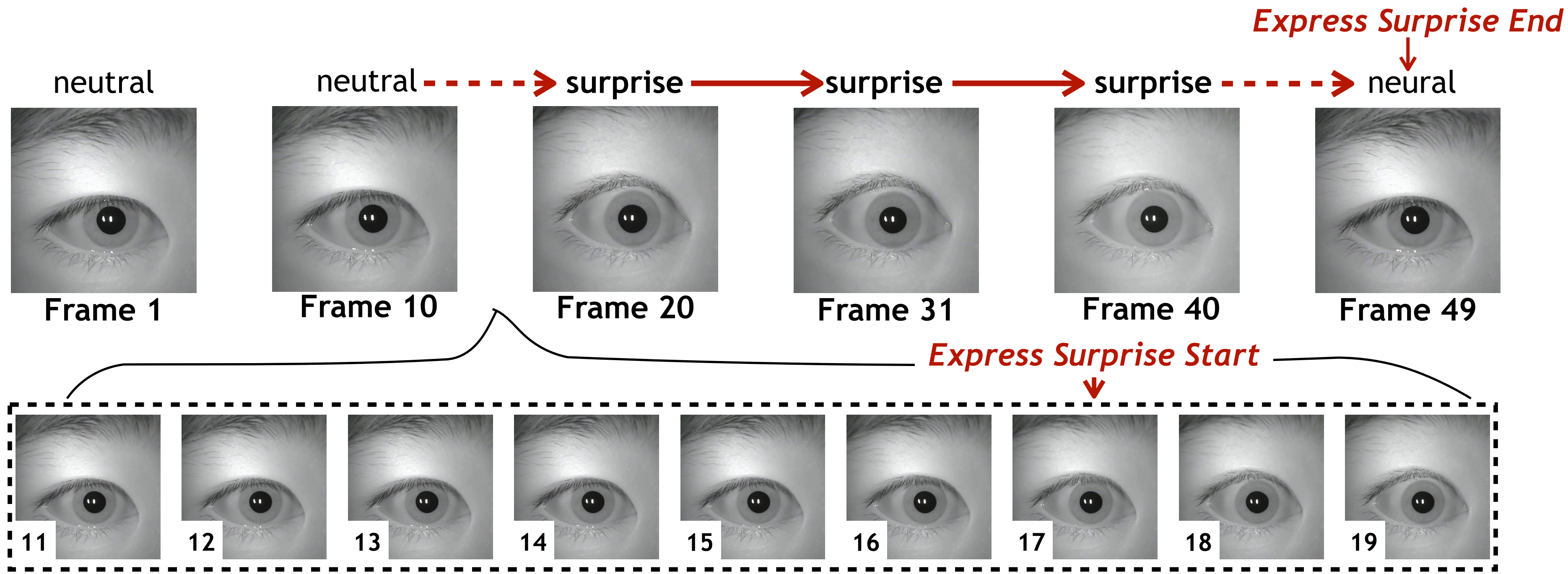


Today's talk

1. *Single-eye emotional feature extraction.*
2. *Personalized fine-grained emotion classification.*
3. *Real-time untethered recognition system.*
4. *The prototype and evaluation.*

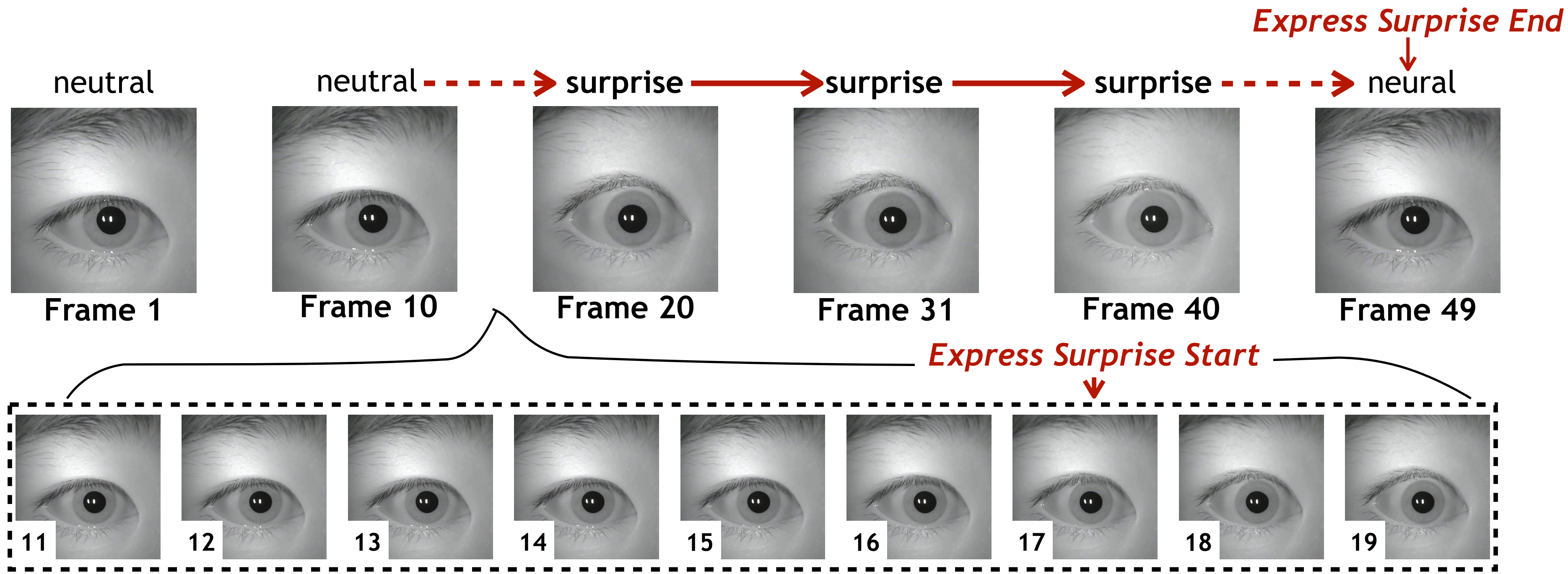


Limited device resources for real time computing



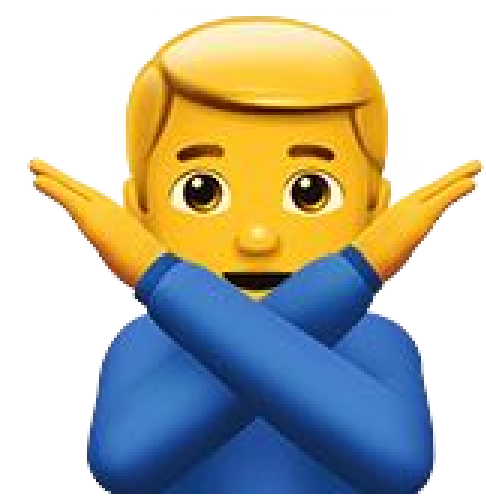
Eye-area images captured at 25 fps. The "surprise" emotion lasts 22 frames.

Limited device resources for real time computing



Eye-area images captured at 25 fps. The “surprise” emotion lasts 22 frames.

Slow recognition leads emotion
Missed or Inconsistent



Real-time untethered recognition system

Temporal Locality:

Fast Forwarder: Two consecutive frames are usually same -> **Quickly** judge the similarity

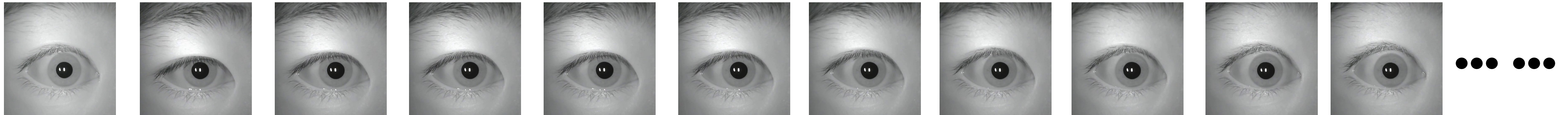
Frame Sampler: Stagnation immediately following a change -> Label some frames *immediately*

Real-time untethered recognition system

Temporal Locality:

Fast Forwarder: Two consecutive frames are usually same -> **Quickly** judge the similarity

Frame Sampler: Stagnation immediately following a change -> Label some frames *immediately*



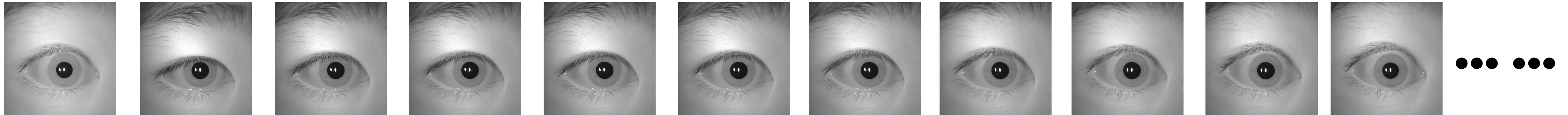
Real-time untethered recognition system

Temporal Locality:

Fast Forwarder: Two consecutive frames are usually same -> **Quickly** judge the similarity

Frame Sampler: Stagnation immediately following a change -> Label some frames **immediately**

surprise



Real-time untethered recognition system

Temporal Locality:

Fast Forwarder: Two consecutive frames are usually same -> **Quickly** judge the similarity

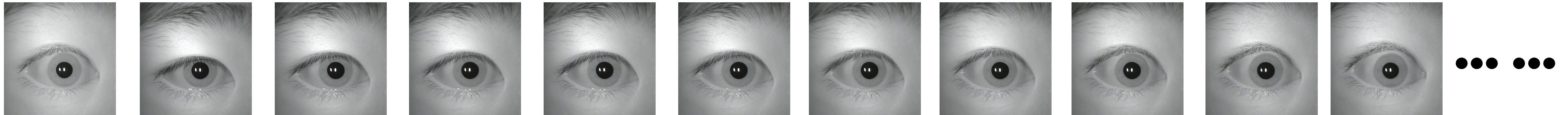
Frame Sampler: Stagnation immediately following a change -> Label some frames *immediately*

Personalized Classifier



surprise

neutral

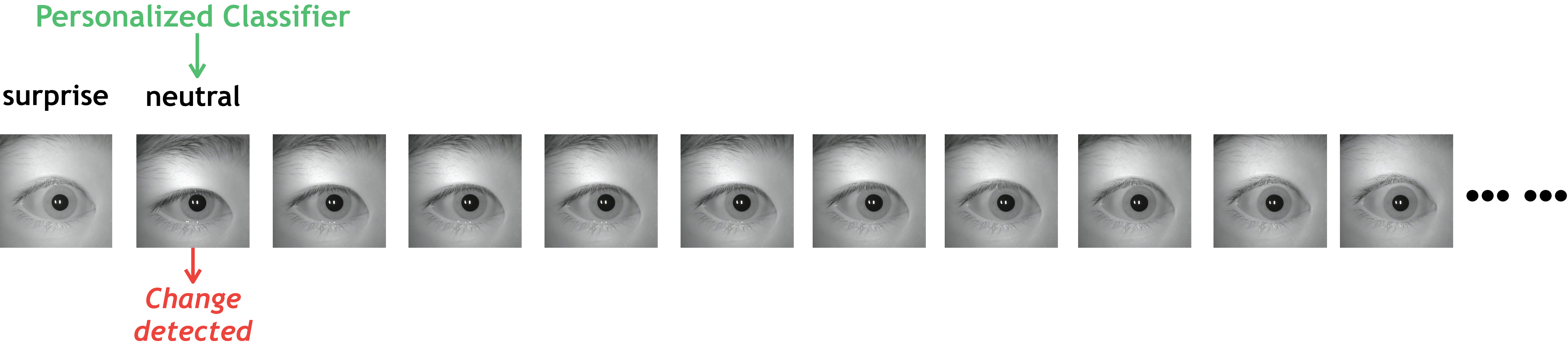


Real-time untethered recognition system

Temporal Locality:

Fast Forwarder: Two consecutive frames are usually same -> **Quickly** judge the similarity

Frame Sampler: Stagnation immediately following a change -> Label some frames **immediately**

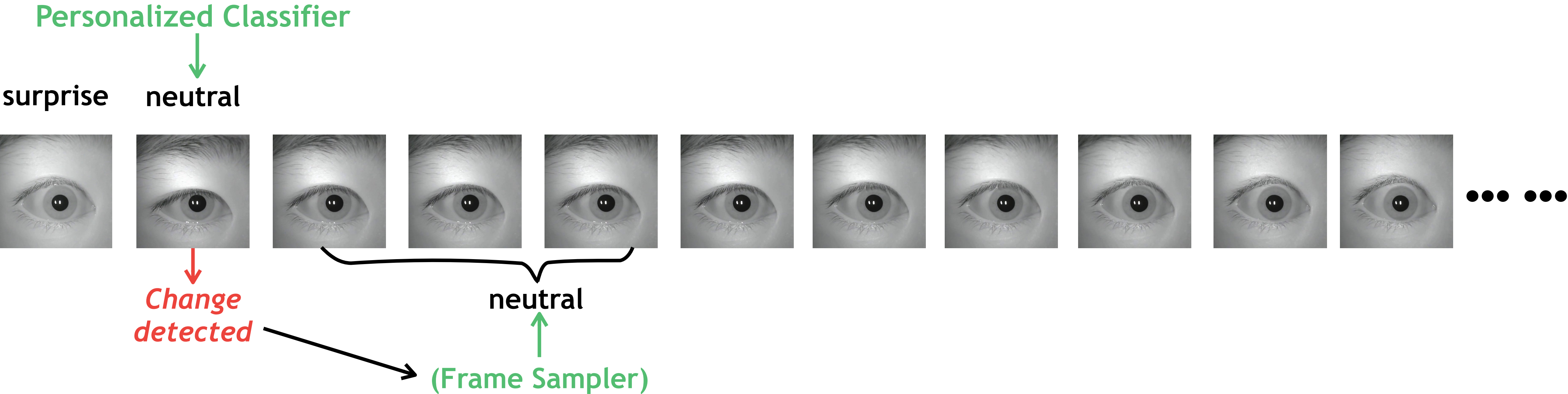


Real-time untethered recognition system

Temporal Locality:

Fast Forwarder: Two consecutive frames are usually same -> **Quickly** judge the similarity

Frame Sampler: Stagnation immediately following a change -> Label some frames **immediately**

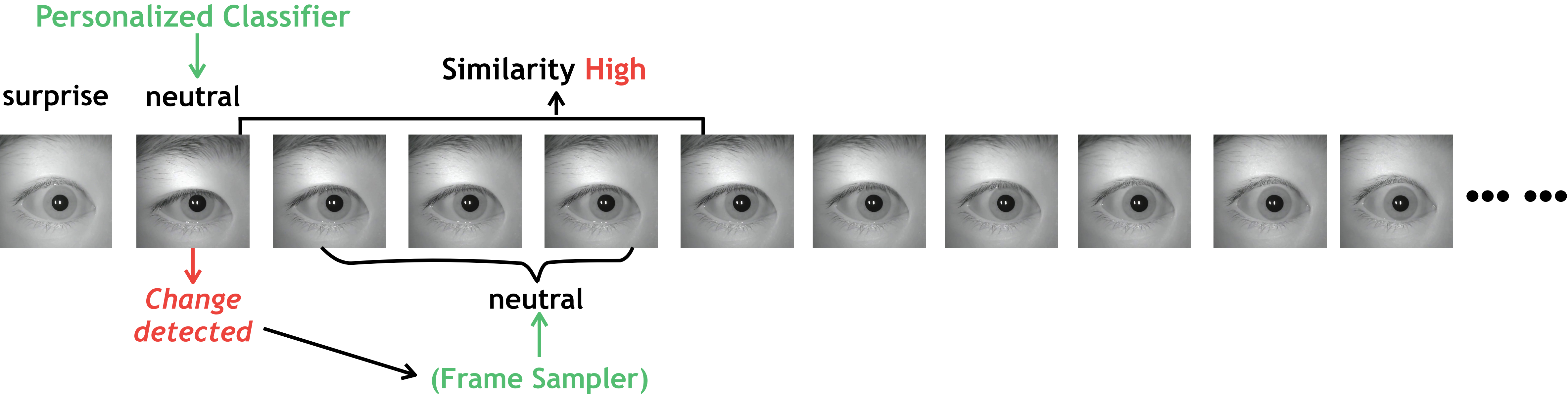


Real-time untethered recognition system

Temporal Locality:

Fast Forwarder: Two consecutive frames are usually same -> **Quickly** judge the similarity

Frame Sampler: Stagnation immediately following a change -> Label some frames **immediately**

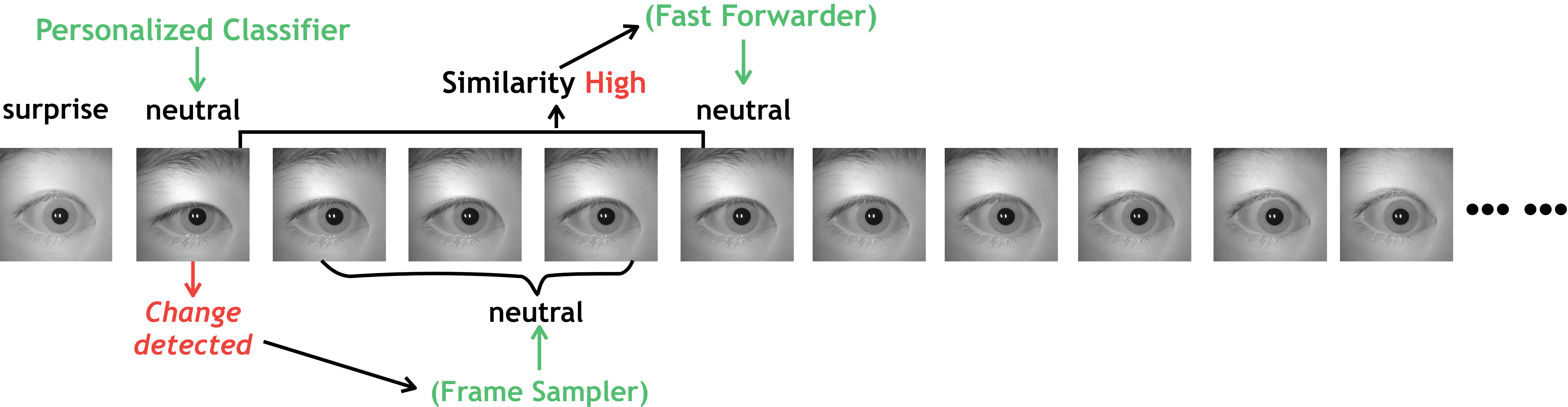


Real-time untethered recognition system

Temporal Locality:

Fast Forwarder: Two consecutive frames are usually same -> Quickly judge the similarity

Frame Sampler: Stagnation immediately following a change -> Label some frames *immediately*

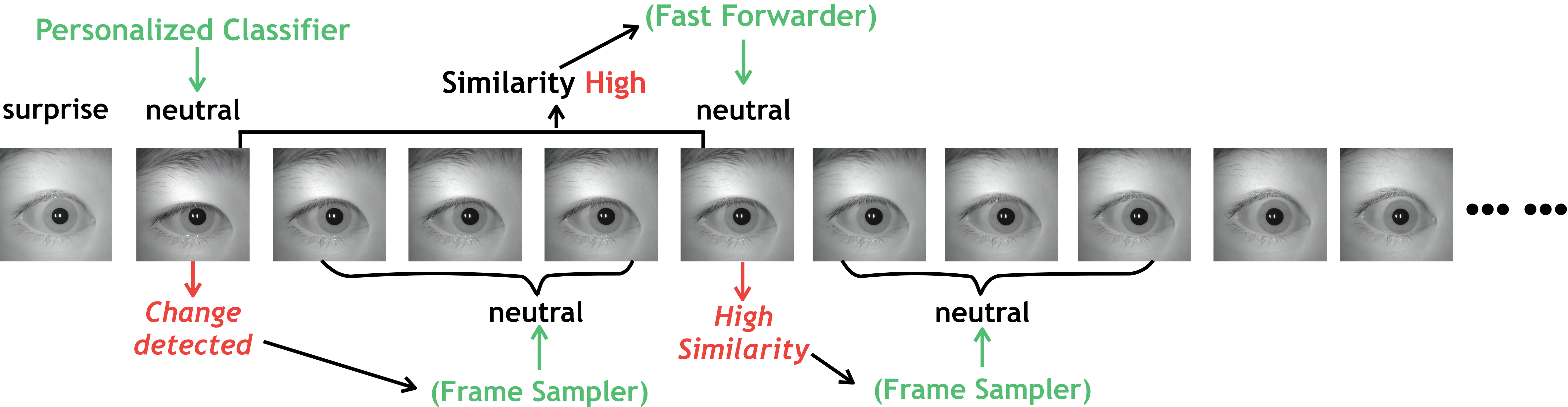


Real-time untethered recognition system

Temporal Locality:

Fast Forwarder: Two consecutive frames are usually same -> Quickly judge the similarity

Frame Sampler: Stagnation immediately following a change -> Label some frames *immediately*

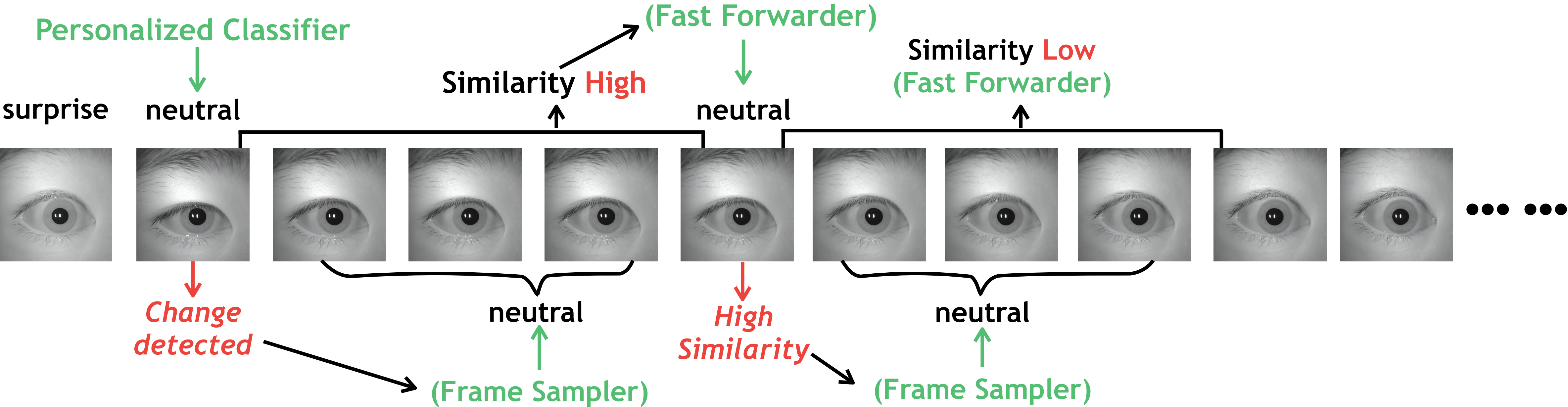


Real-time untethered recognition system

Temporal Locality:

Fast Forwarder: Two consecutive frames are usually same -> Quickly judge the similarity

Frame Sampler: Stagnation immediately following a change -> Label some frames *immediately*

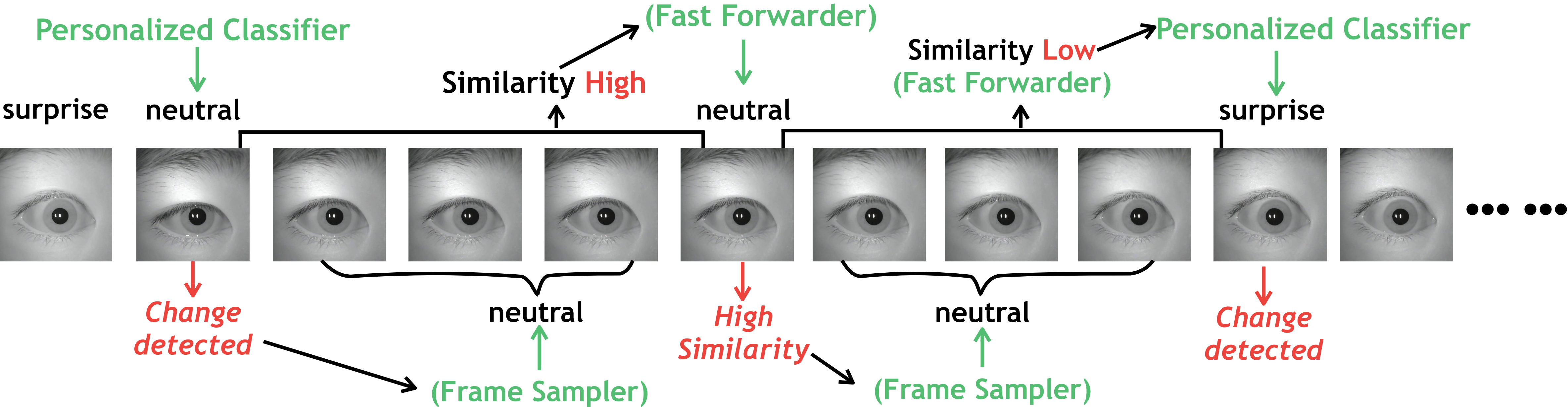


Real-time untethered recognition system

Temporal Locality:

Fast Forwarder: Two consecutive frames are usually same -> Quickly judge the similarity

Frame Sampler: Stagnation immediately following a change -> Label some frames *immediately*

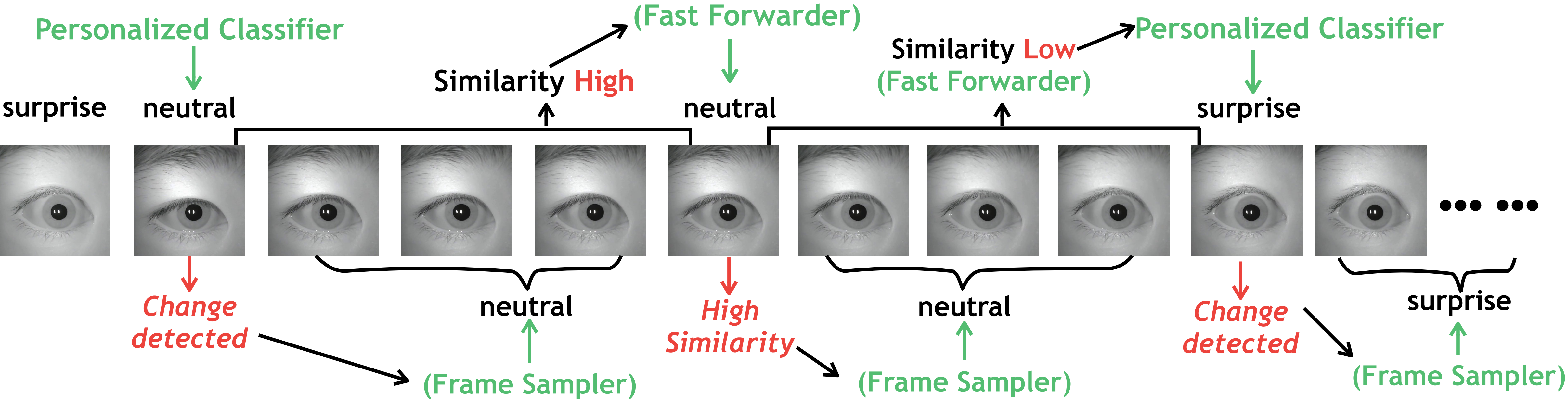


Real-time untethered recognition system

Temporal Locality:

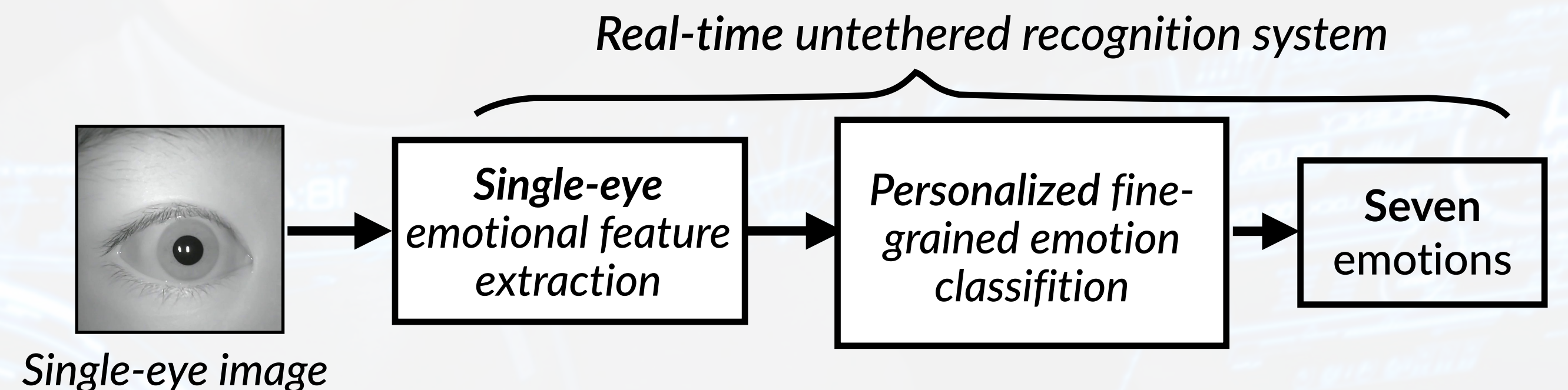
Fast Forwarder: Two consecutive frames are usually same -> Quickly judge the similarity

Frame Sampler: Stagnation immediately following a change -> Label some frames *immediately*

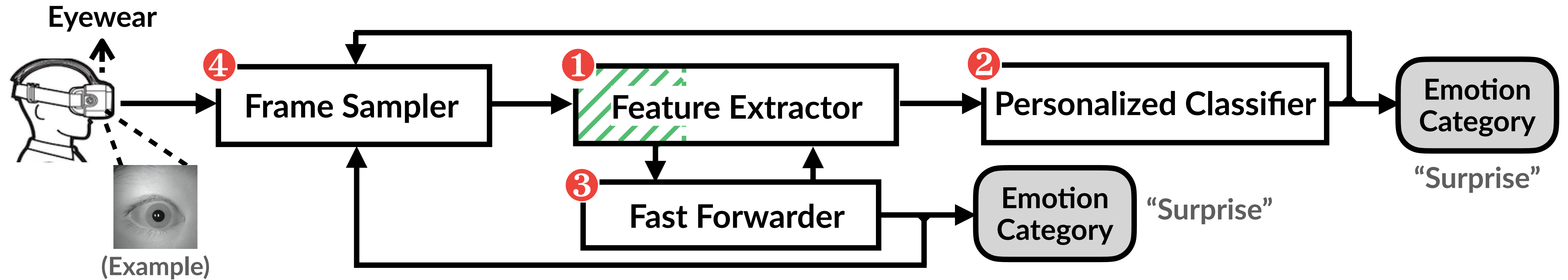


Today's talk

1. *Single-eye emotional feature extraction.*
2. *Personalized fine-grained emotion classification.*
3. *Real-time untethered recognition system.*
4. *The prototype and evaluation.*

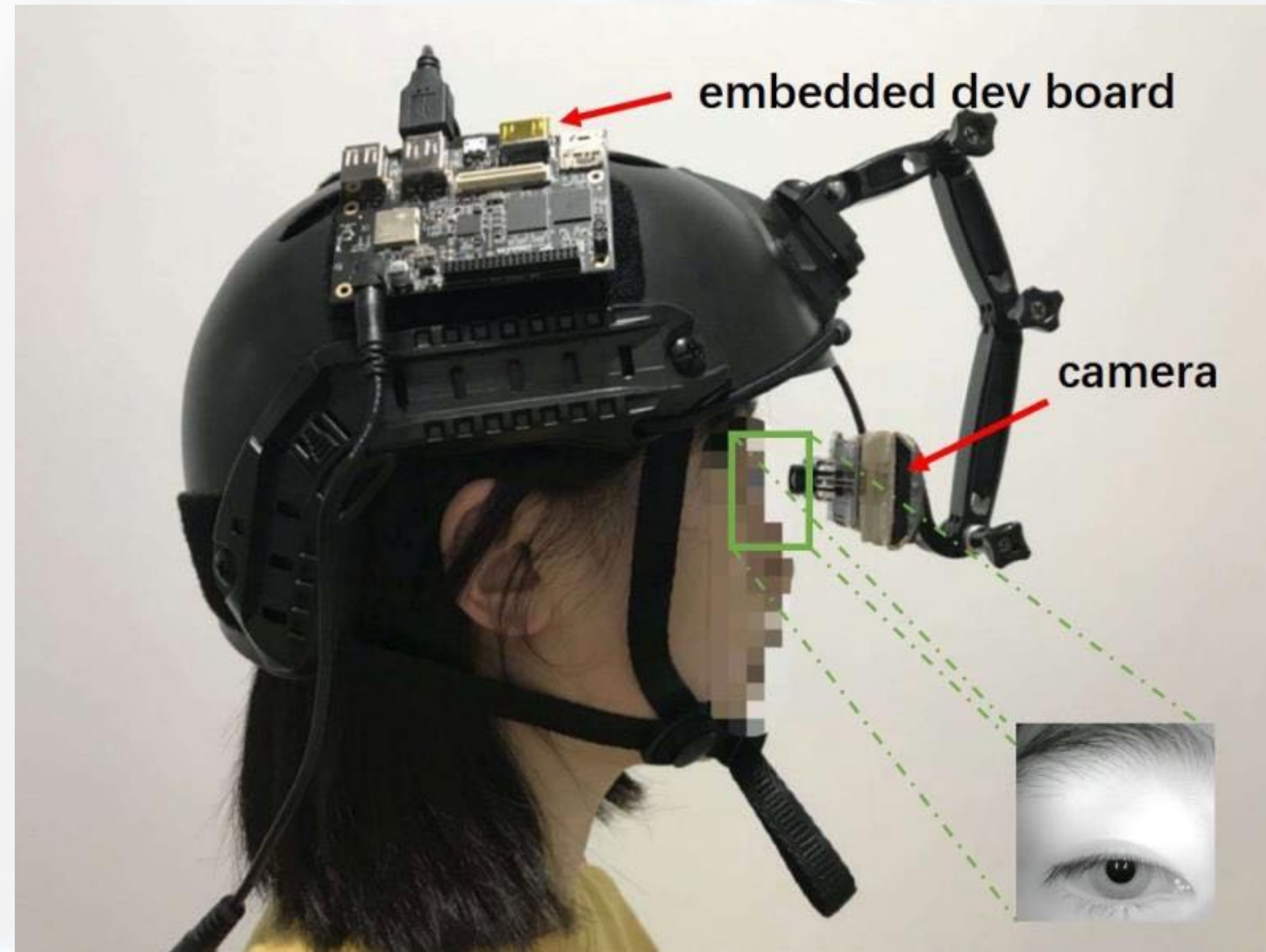


System Workflow



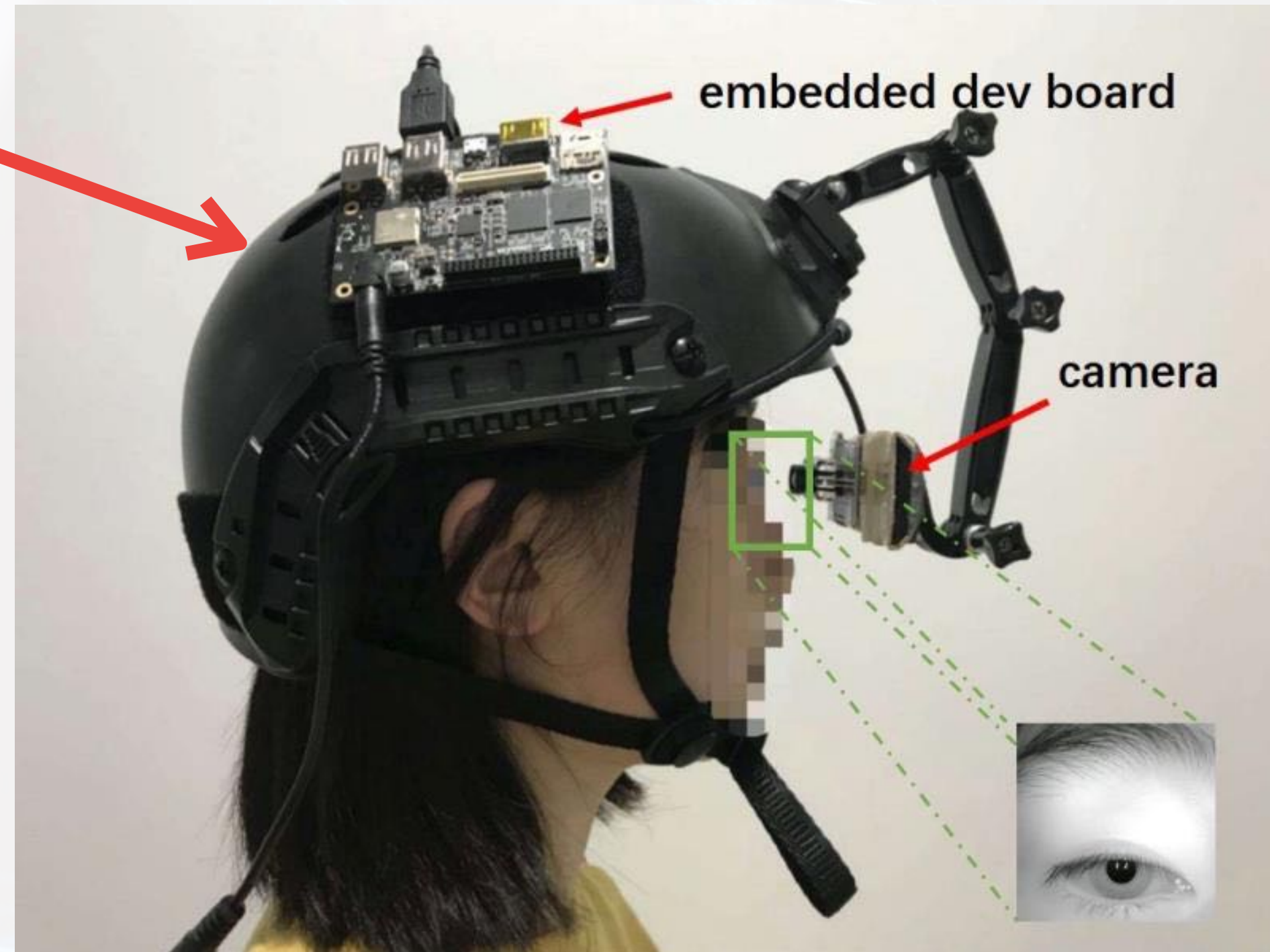
The system workflow of EMO.

Prototype



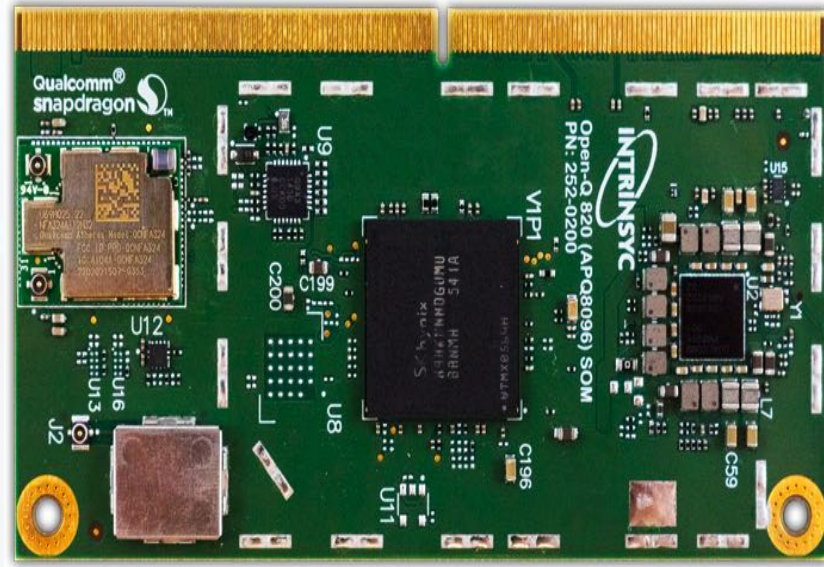
Prototype

GoPro Helmet



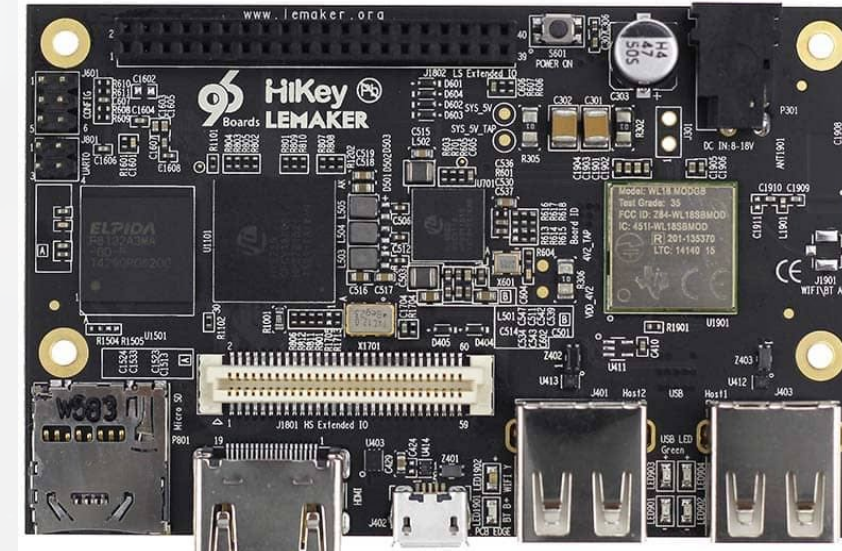
Prototype

Open Q-820

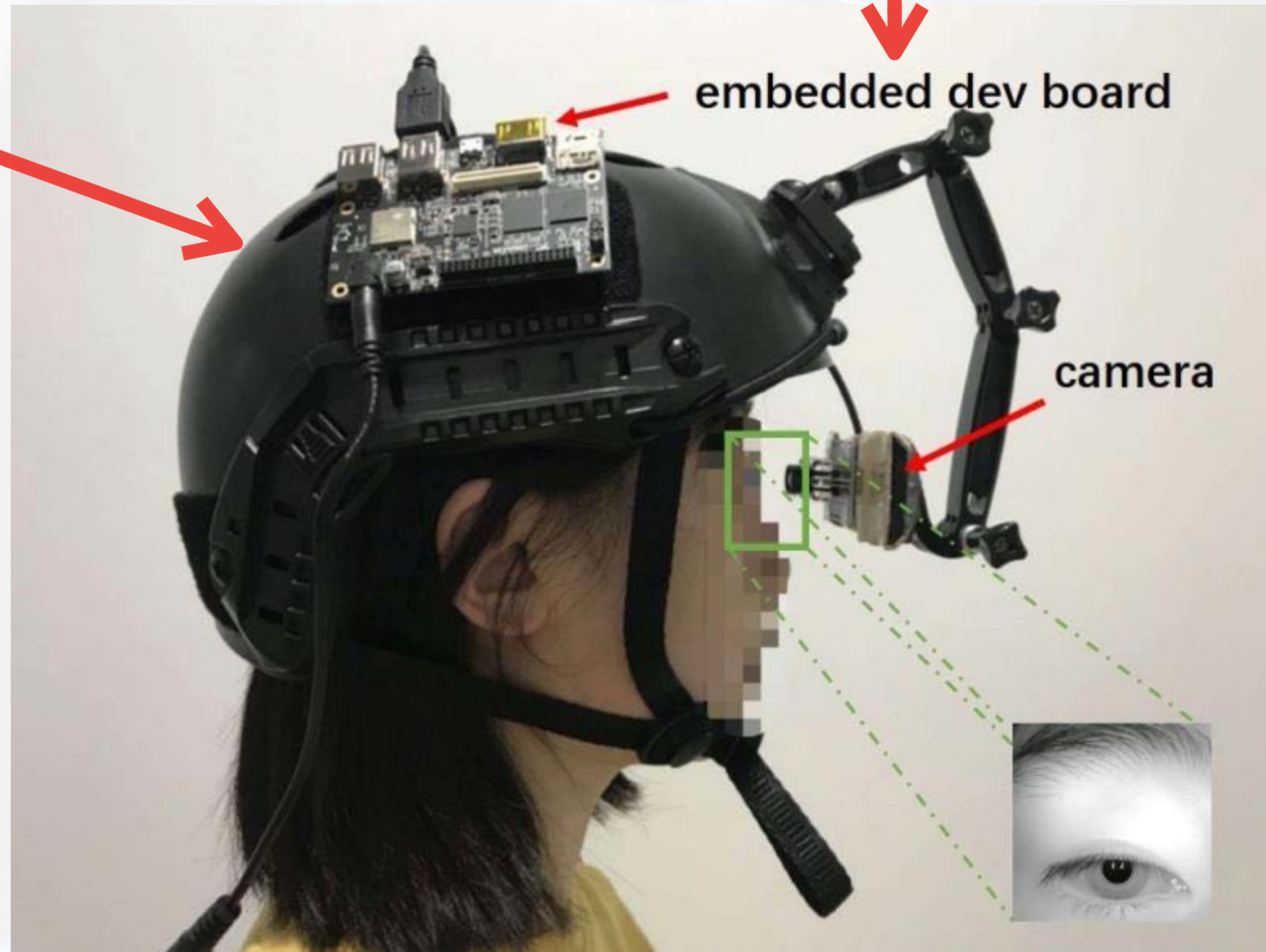


or

Hikey 620

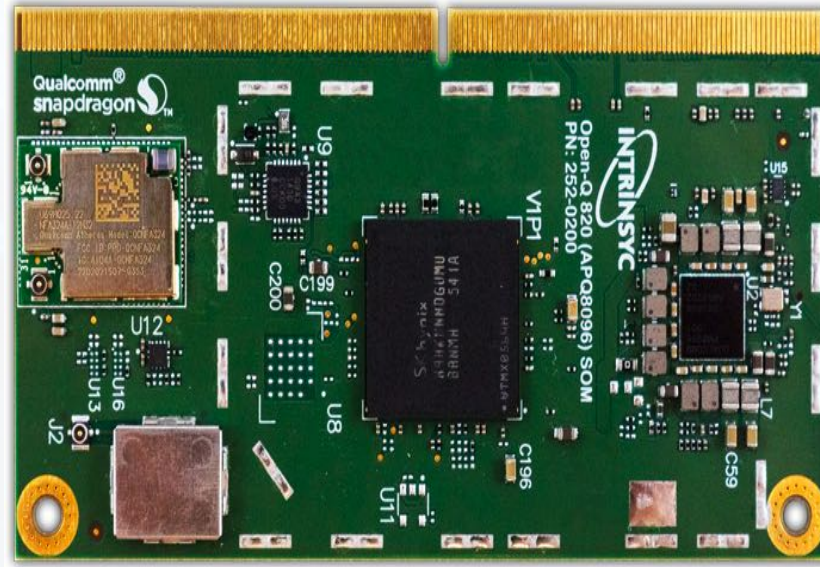


GoPro Helmet



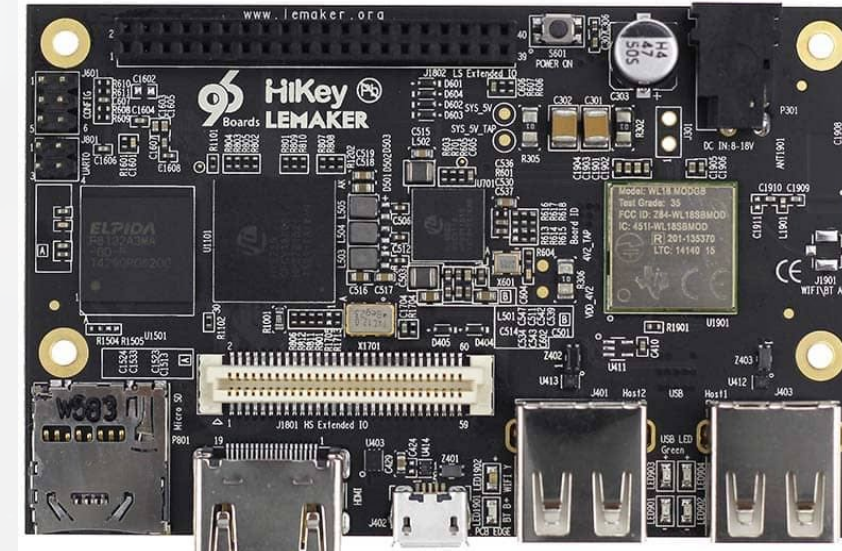
Prototype

Open Q-820

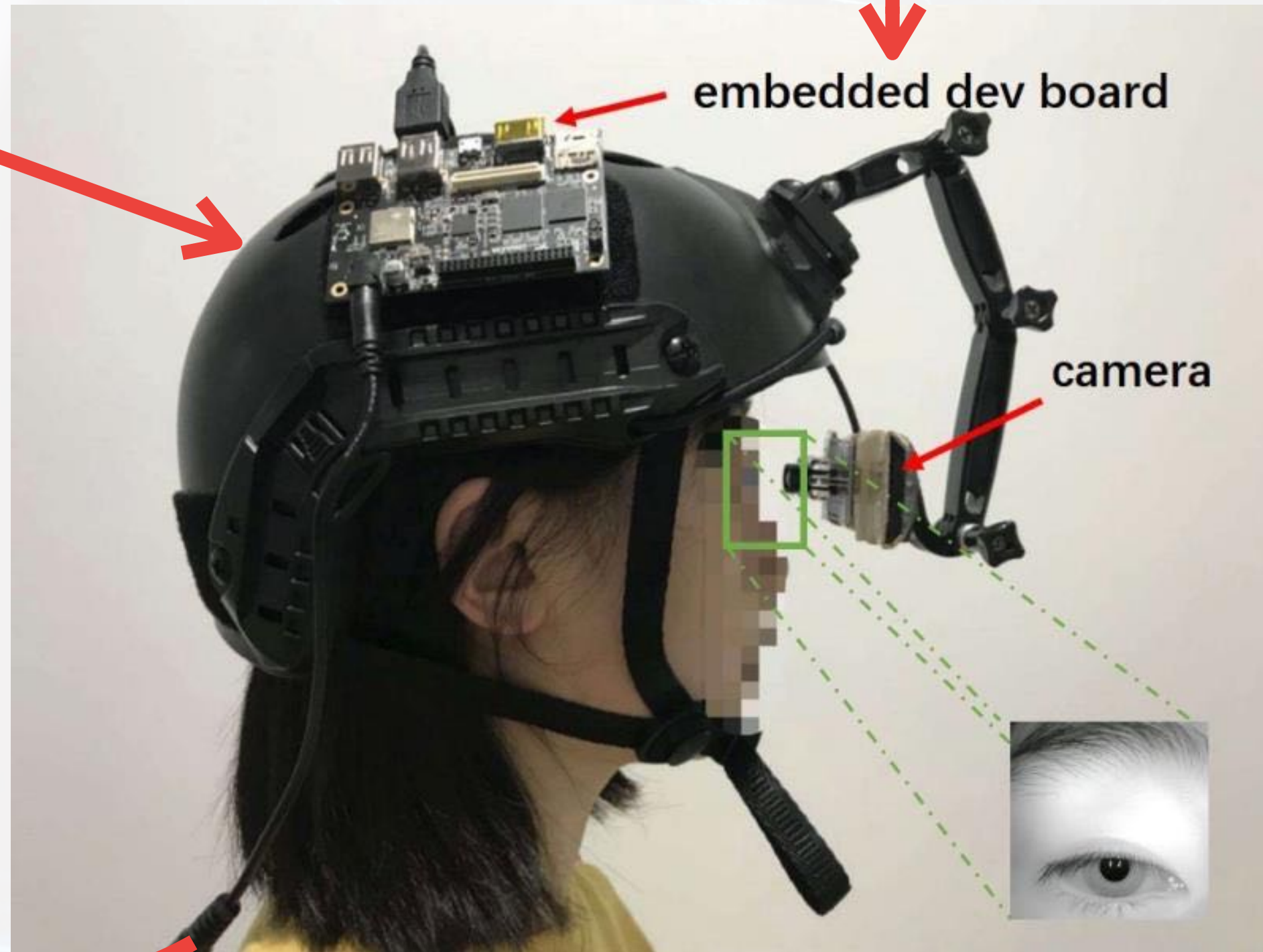


or

Hikey 620



GoPro Helmet



embedded dev board

camera

A portable 185Wh battery

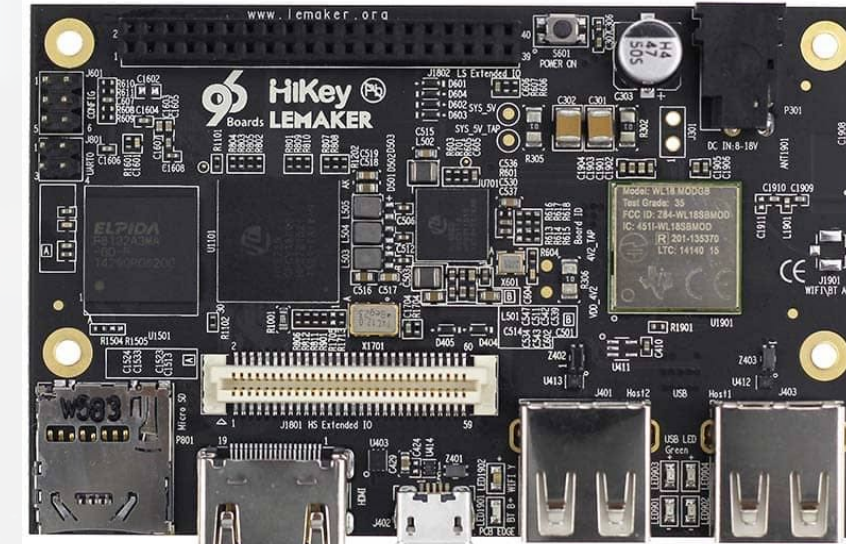


Prototype

Open Q-820

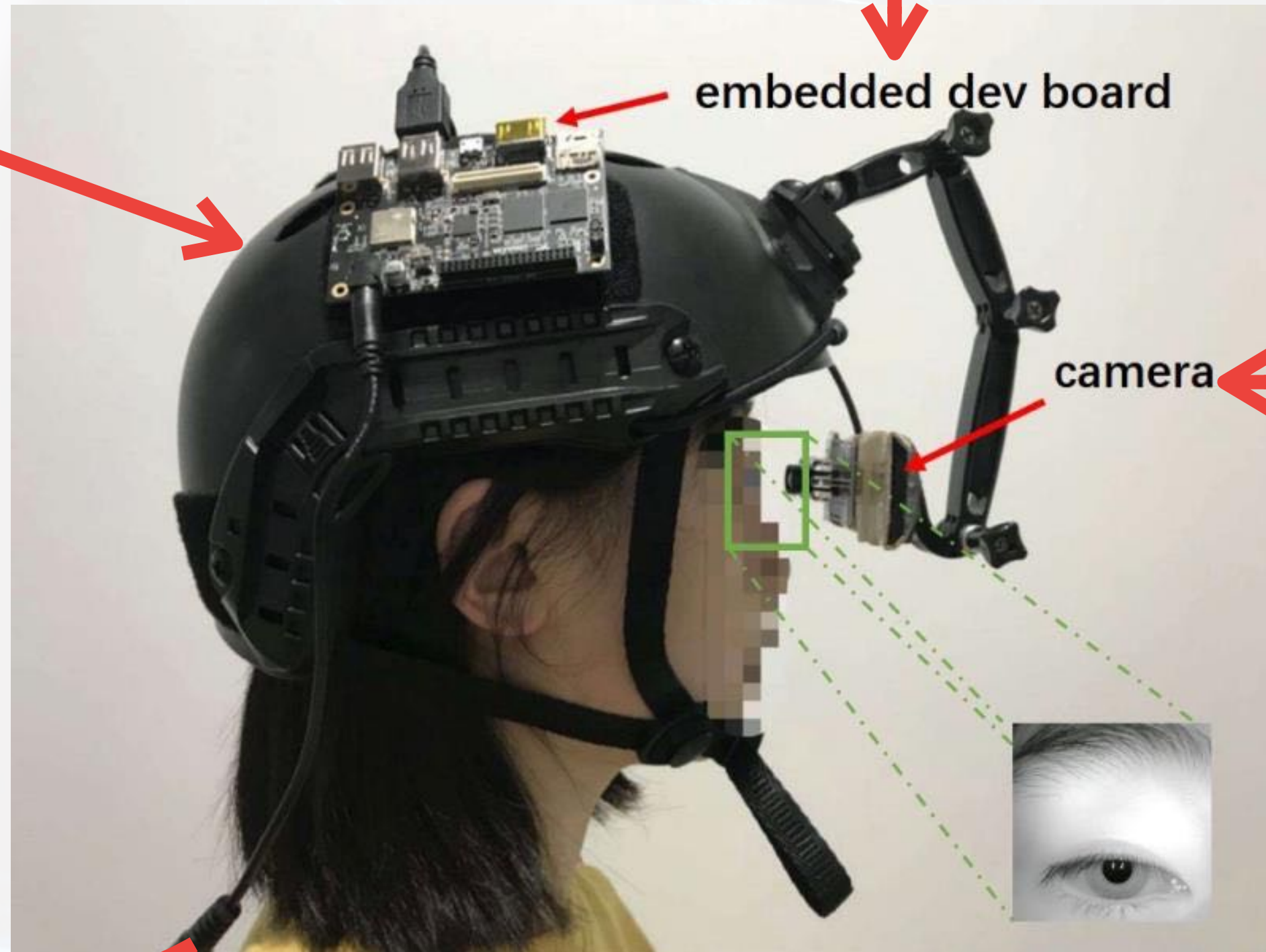


Hikey 620



or

GoPro Helmet



embedded dev board

camera

3.6mm focal length IR camera

A portable 185Wh battery

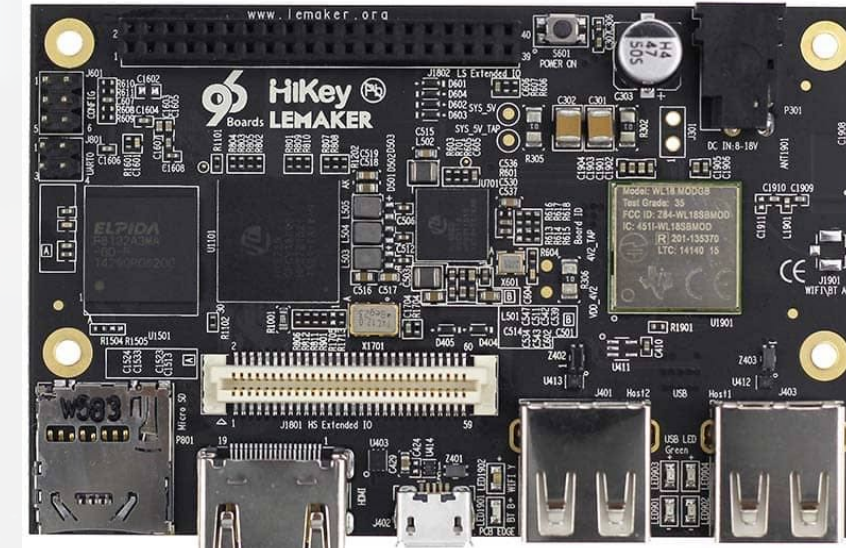


Prototype

Open Q-820

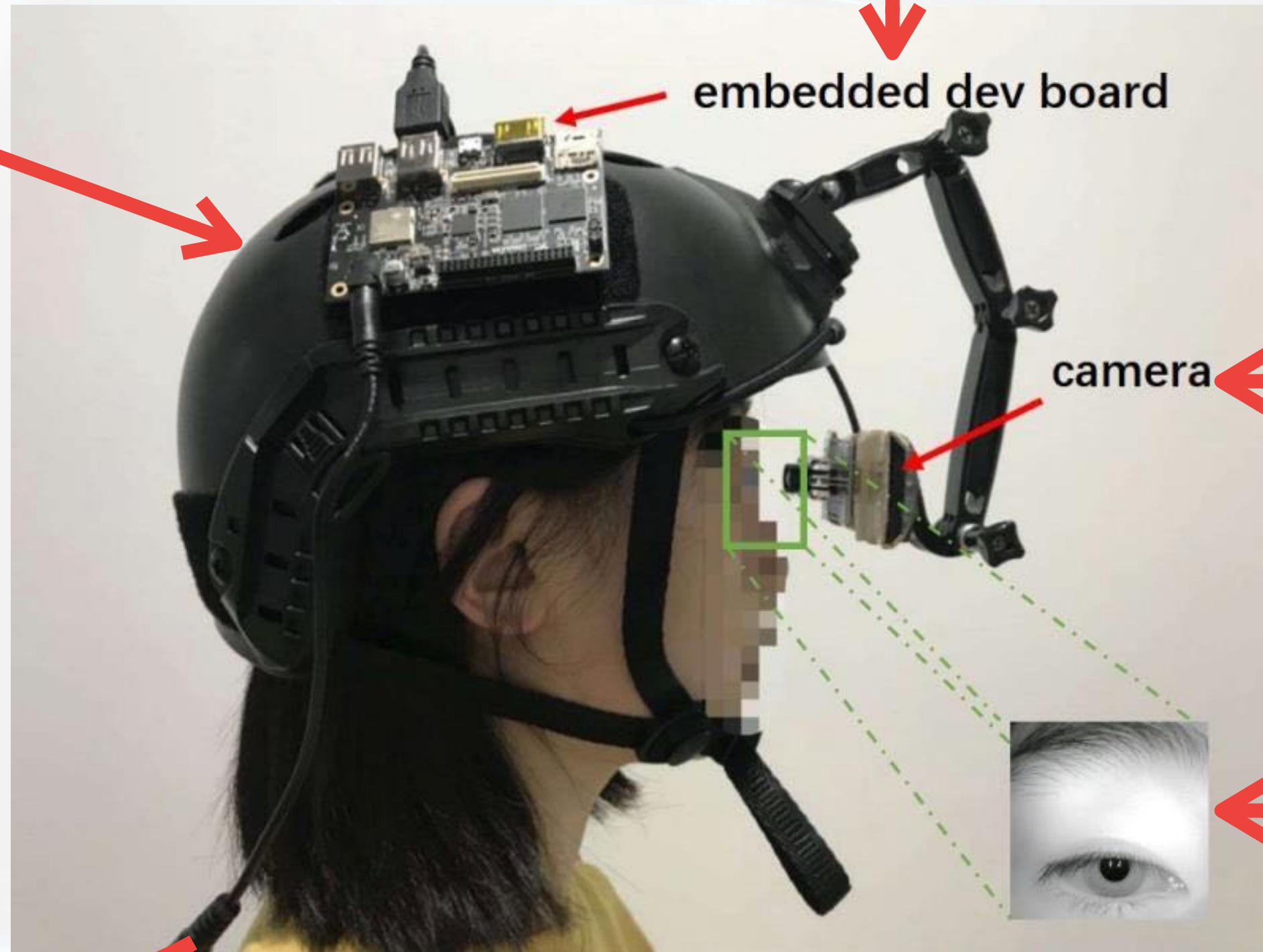


Hikey 620



or

GoPro Helmet



embedded dev board

camera

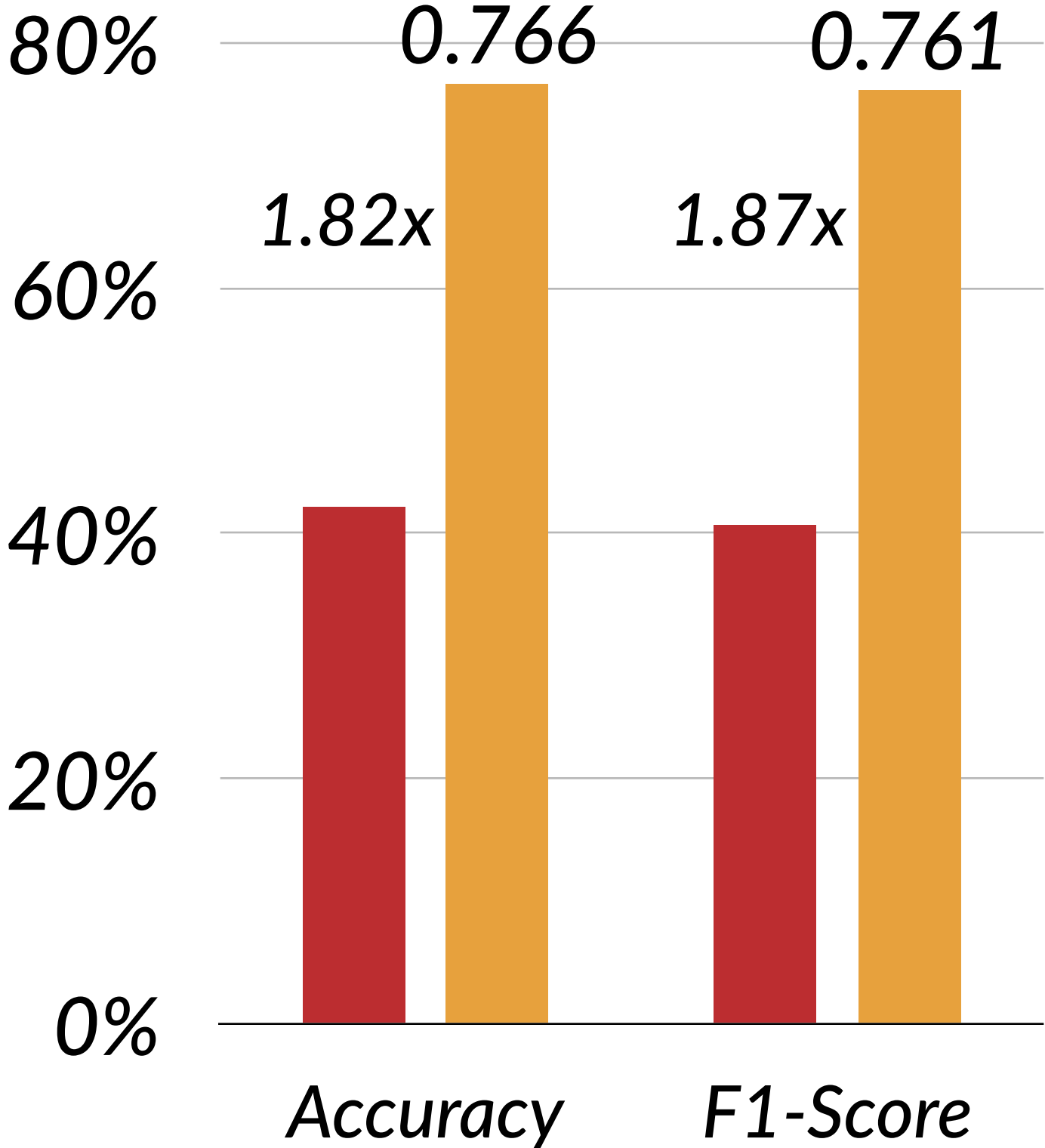
3.6mm focal length IR camera

Single-eye Image

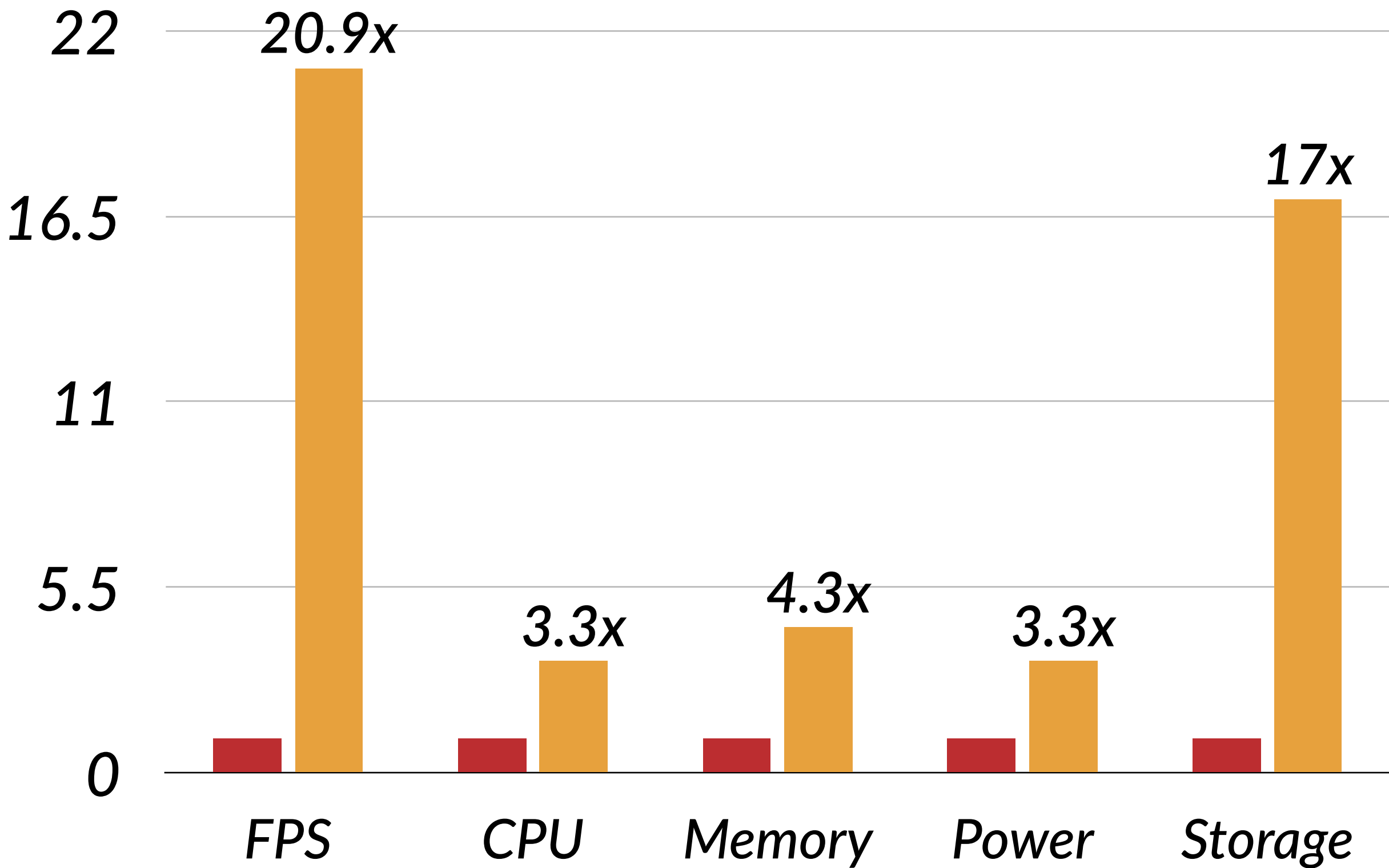
A portable 185Wh battery

Evaluation

■ *Baseline* ■ *EMO (our)*



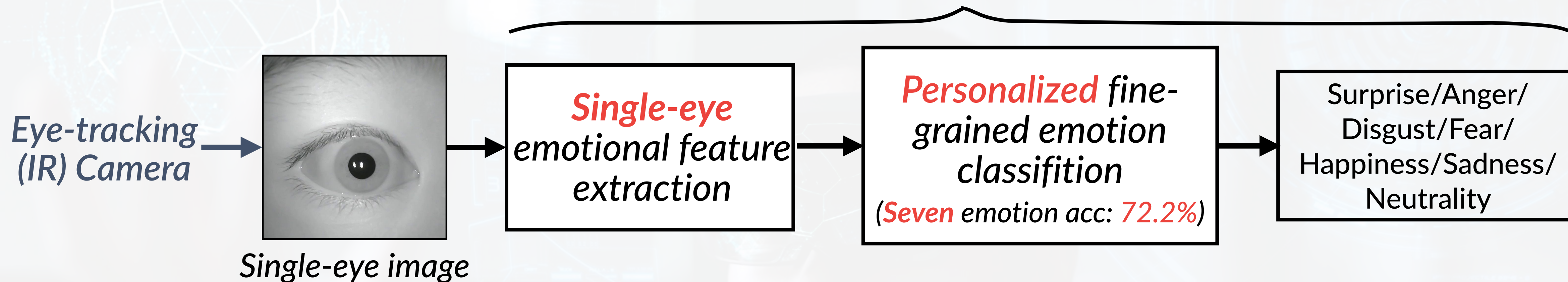
Recognition Performance.



System Performance.

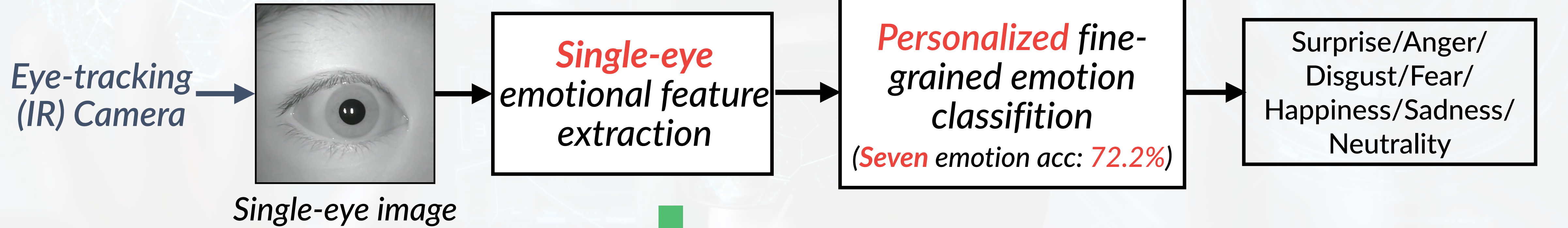
Take away

Real-time untethered recognition system.
(CPU: *Open-Q820/12%*, Memory: *73MB*, Speed: *12.8fps*)

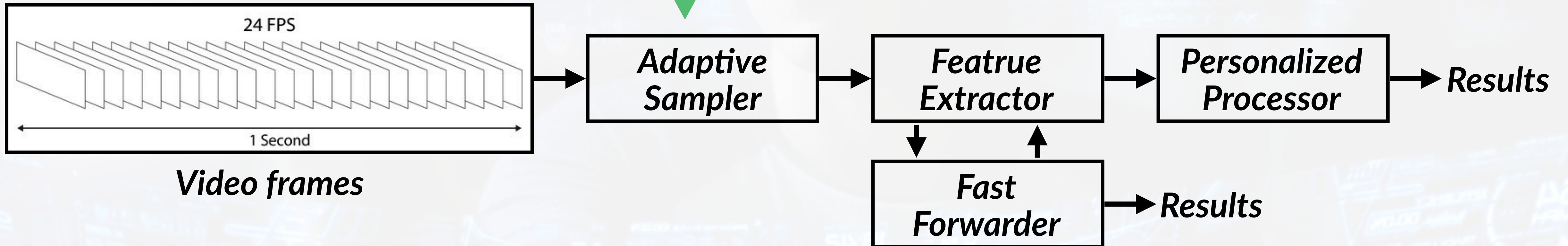


Take away

Real-time untethered recognition system.
(CPU: *Open-Q820/12%*, Memory: *73MB*, Speed: *12.8fps*)

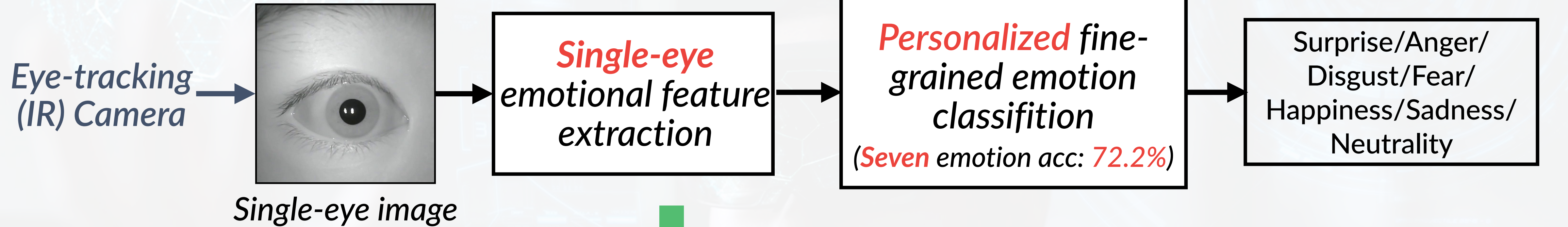


Personalized Intelligent (Video) Processing

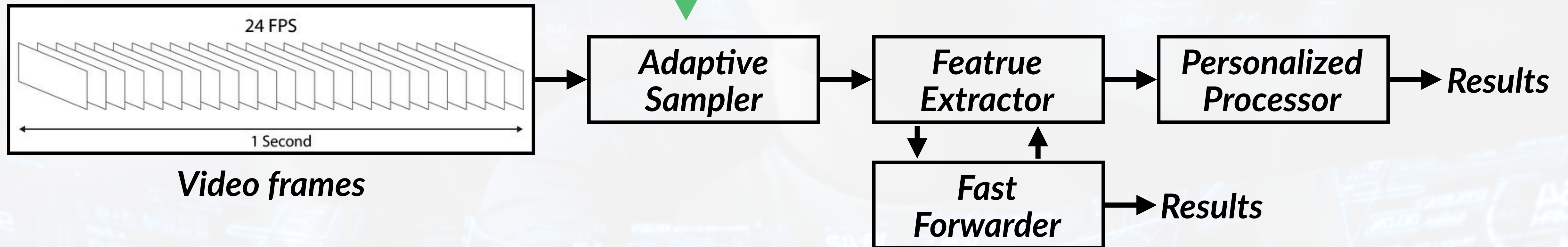


Take away

Real-time untethered recognition system.
(CPU: *Open-Q820/12%*, Memory: *73MB*, Speed: *12.8fps*)



Personalized Intelligent (Video) Processing



Thank you for attention! Questions?