Cheating Detection
On Online Tests

1조 B611137 이강복 | B611169 임정민 | B611210 하인혜

Contents

Problems

Types of cheating on online test

Substitute examination

Cheating with blind spot of camera

Web surfing

Limitation of current program

How to prevent substitute examination

OnTest

- Take a picture of one's ID card
- Check the student's ID card with the information

Supervisor must check the identity card against the person.

Monito

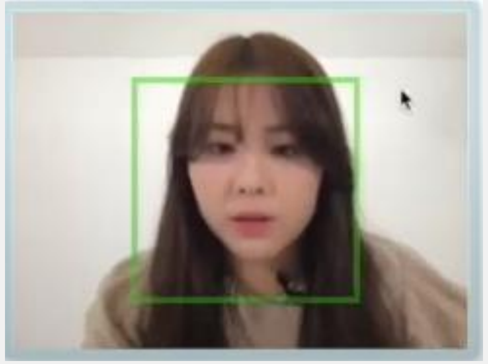
- Take a picture of one's ID card
- Check the birth date and name of the students

Supervisor must check the identity card against the person

Limitation of current program

How to prevent cheating using blind spot of camera

OnTest



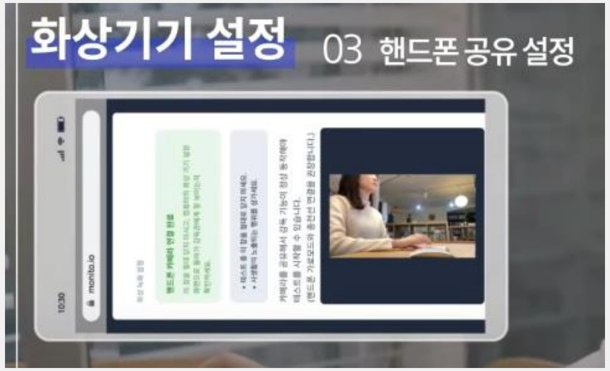
In the square box.

Not in the square box.

- Using a webcam
- A method of warning: when the user's face is out of the bounding box

Cheating is possible in the bounding box

Monito



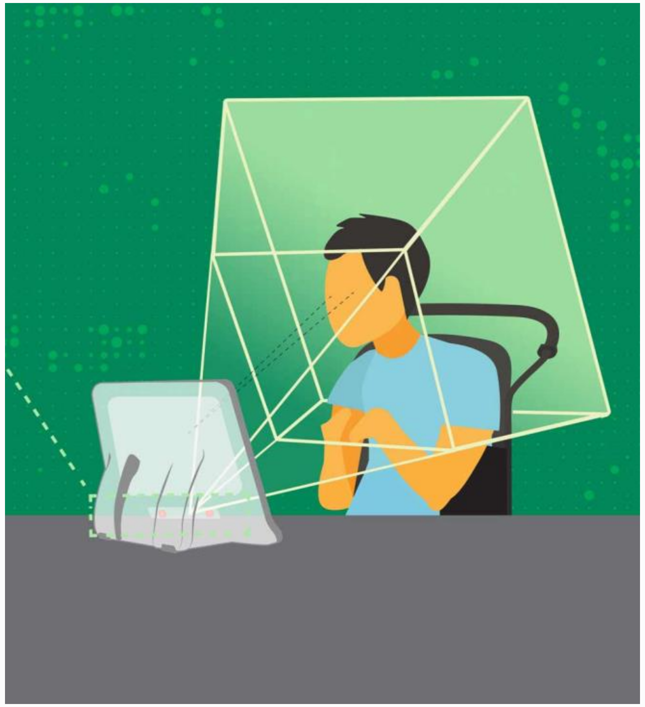
- Use both webcam and cell phone camera to reduce blind spots

A blind spot that can't be captured by two cameras

Solutions

- **Self-certification**

- **Eye-Tracking**

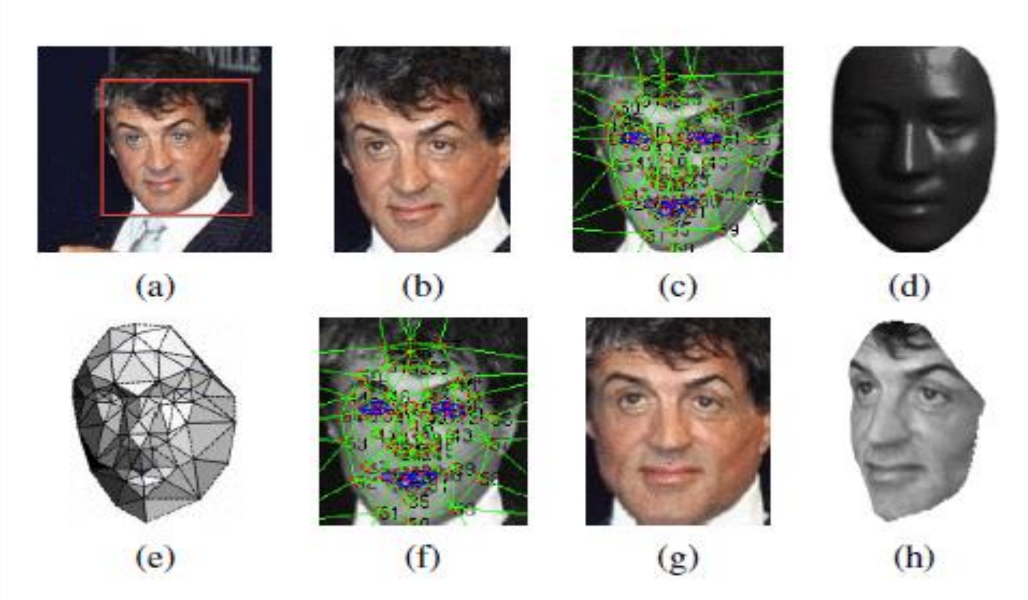


Related Work: DeepFace

Face Recognition PipeLine

: Detectation → Alignment → Representation → Classification

1. Alignment



- 3D Modeling

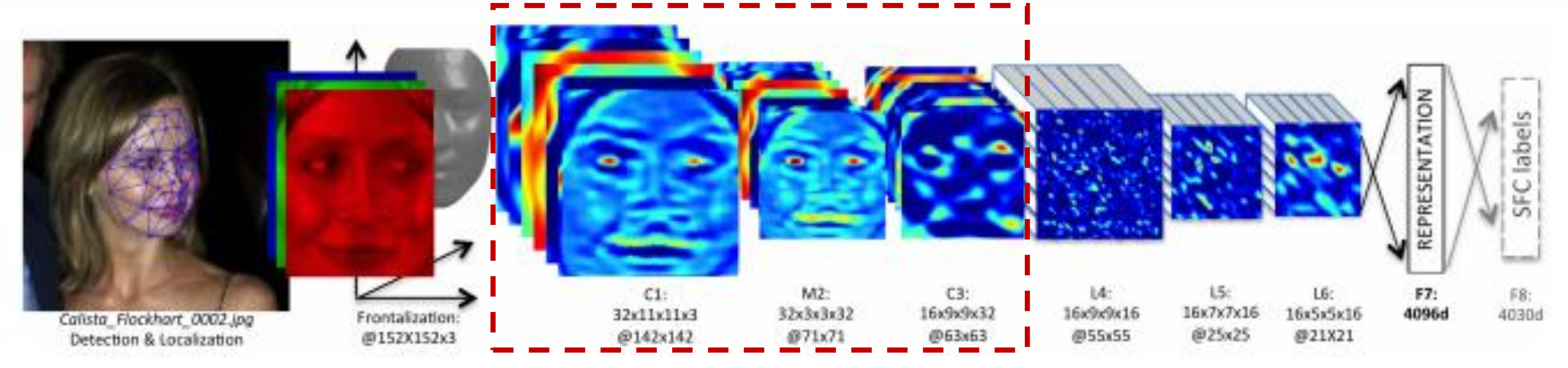
  : Landmark extraction using pre-learned 3D face model

- Frontalization

  : using piecewise affine transformations on each part of the image.

Related Work: DeepFace

2. Representation



- **Front-end adaptive pre-processing part (C1, M2, C3)**

  : Use Convolutional layer , a max-pooling layer, and a convolutional layer.

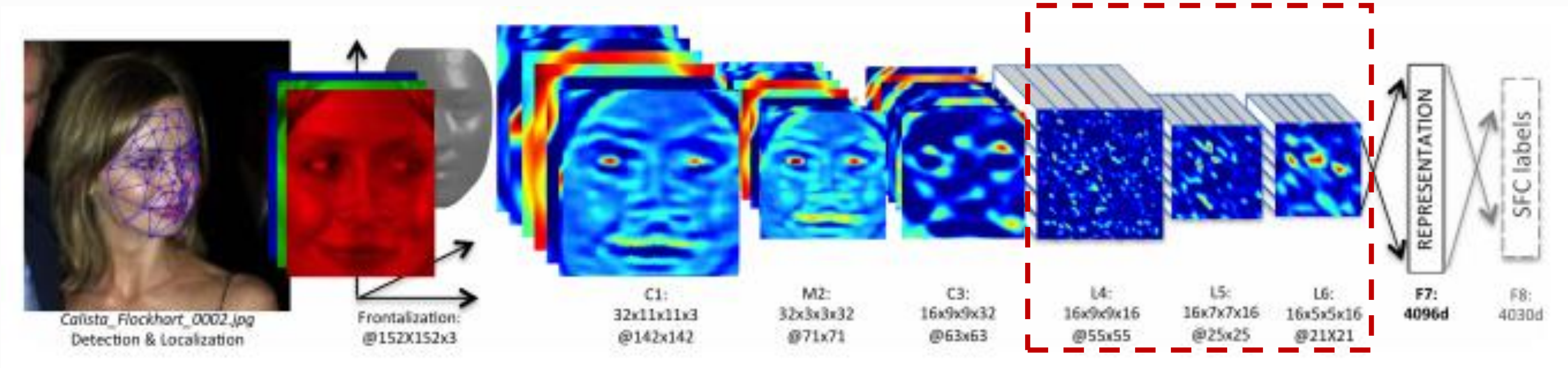
- **Three Locally - connected layers (L4, L5)**

  : Locally connected layer use differently learned weights for all pixels.

- **Two Fully - connected layers (F7, F8)**

Related Work: DeepFace

2. Representation



- **Front-end adaptive pre-processing part (C1, M2, C3)**

  : Use Convolutional layer , a max-pooling layer, and a convolutional layer.

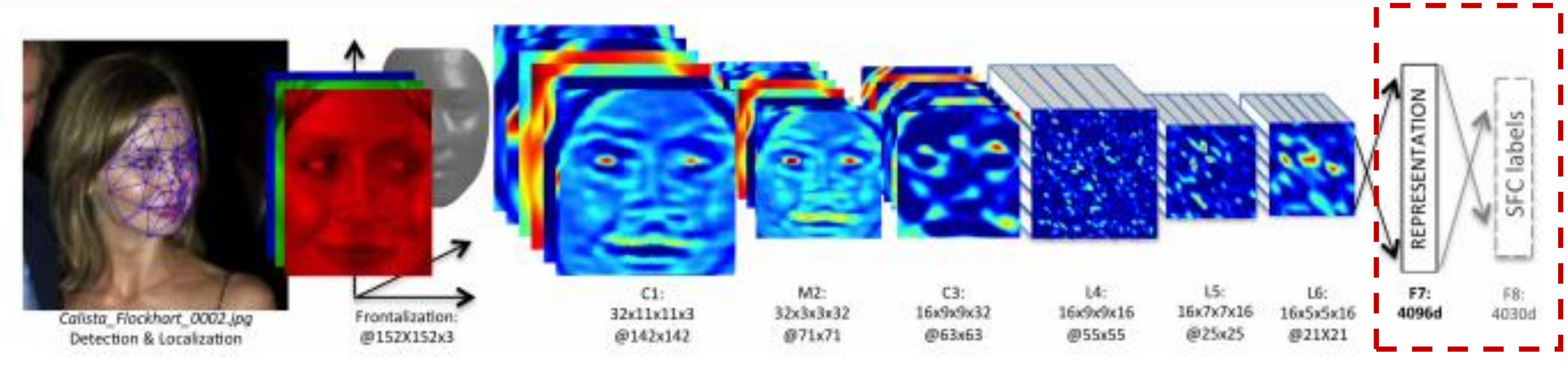
- **Three Locally - connected layers (L4, L5)**

  : Locally connected layer use differently learned weights for all pixels.

- **Two Fully - connected layers (F7, F8)**

Related Work: DeepFace

2. Representation



- **Front-end adaptive pre-processing part (C1, M2, C3)**

  : Use Convolutional layer , a max-pooling layer, and a convolutional layer.

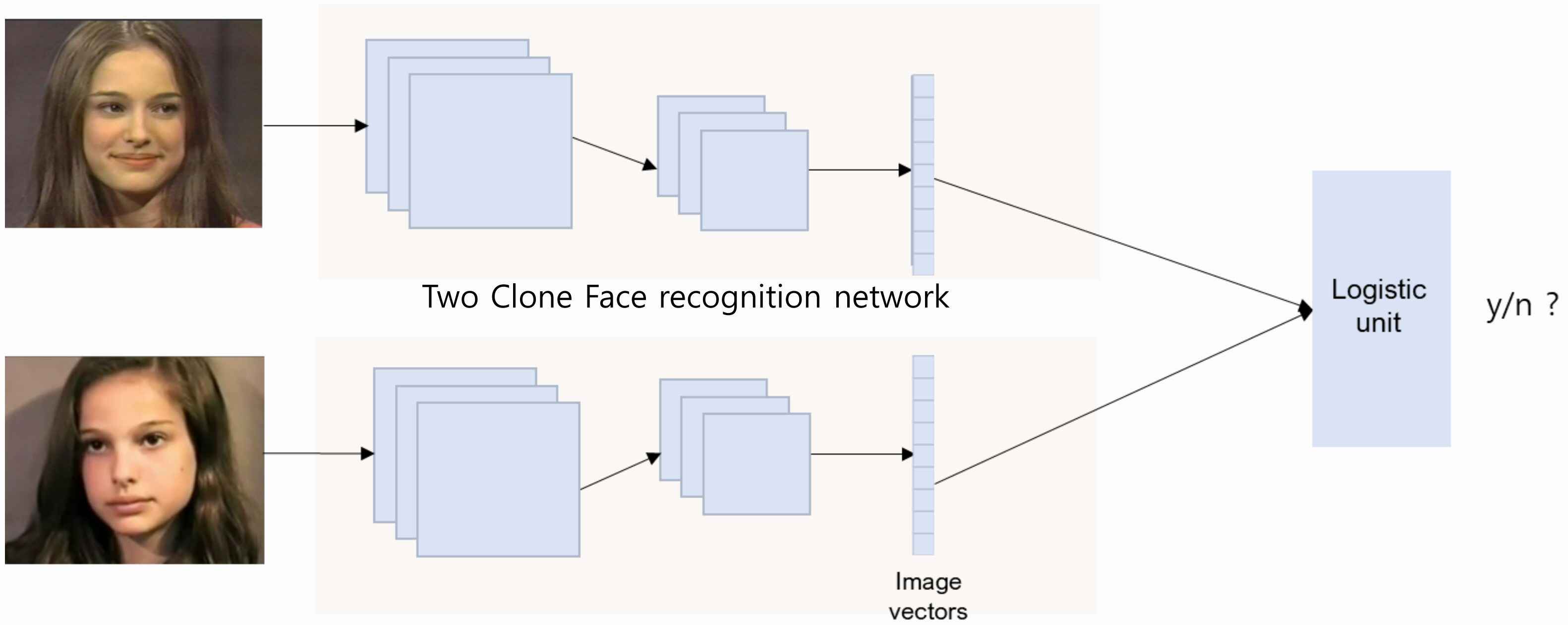
- **Three Locally - connected layers (L4, L5)**

  : Locally connected layer use differently learned weights for all pixels.

- **Two Fully - connected layers (F7, F8)**

Related Work: DeepFace

3. Verification: Siamese Network



- Calculate the distance between two images as an output vector of two images.

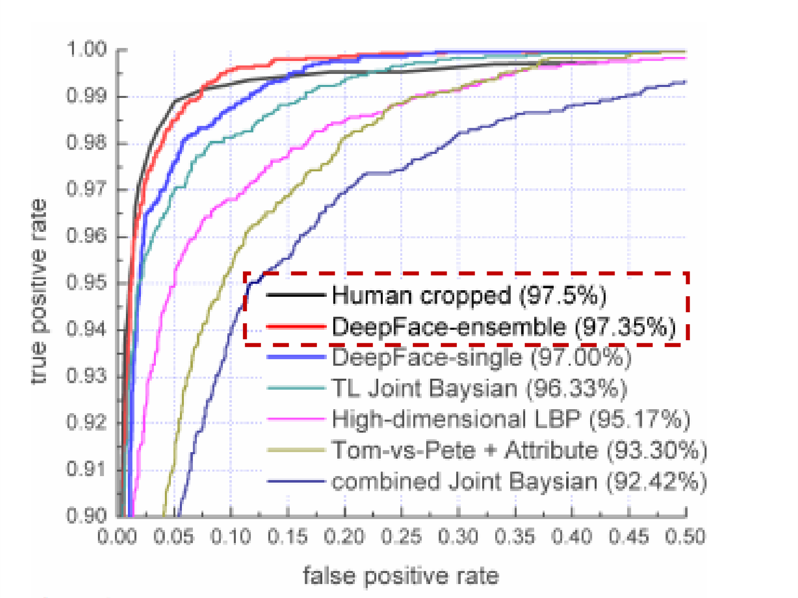
$$d(f1, f2) = \sum_i \alpha_i |f_1[i] - f_2[i]|$$

- A different person if the distance is far and the same person if it is close

Related Work: DeepFace

Result of Experiments

1. **LFW** (the Labeled Faces in the Wild) **Dataset**



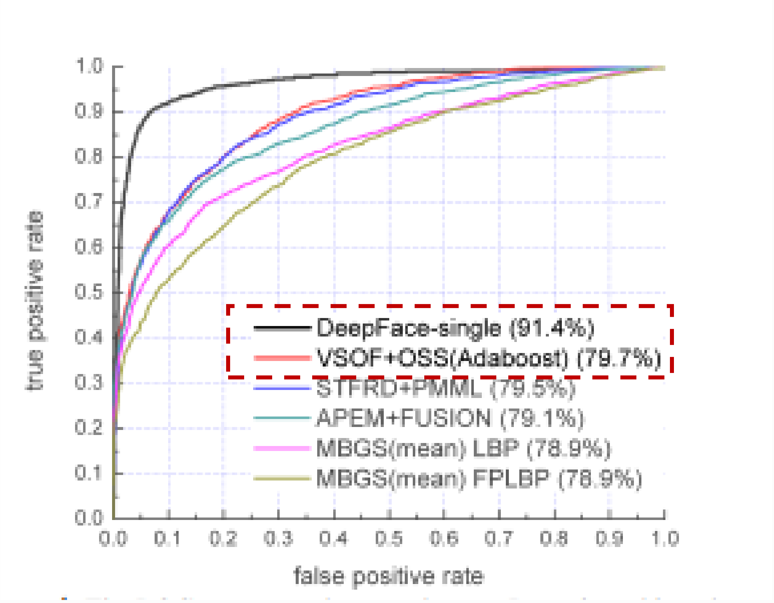
Method	Accuracy ± SE	Protocol
Joint Bayesian [6]	0.9242 ±0.0108	restricted
Tom-vs-Pete [4]	0.9330 ±0.0128	restricted
High-dim LBP [7]	0.9517 ±0.0113	restricted
TL Joint Bayesian [5]	0.9633 ±0.0108	restricted
DeepFace-ensemble	**0.9715** ±0.0027	restricted
DeepFace-ensemble	**0.9735** ±0.0025	unrestricted
Human, cropped	0.9753	

- DeepFace achieve **97.35%** (close to human levels)
- Better performance than Accuracy of the most successful recent study (96.33%)

Related Work: DeepFace

Result of Experiments

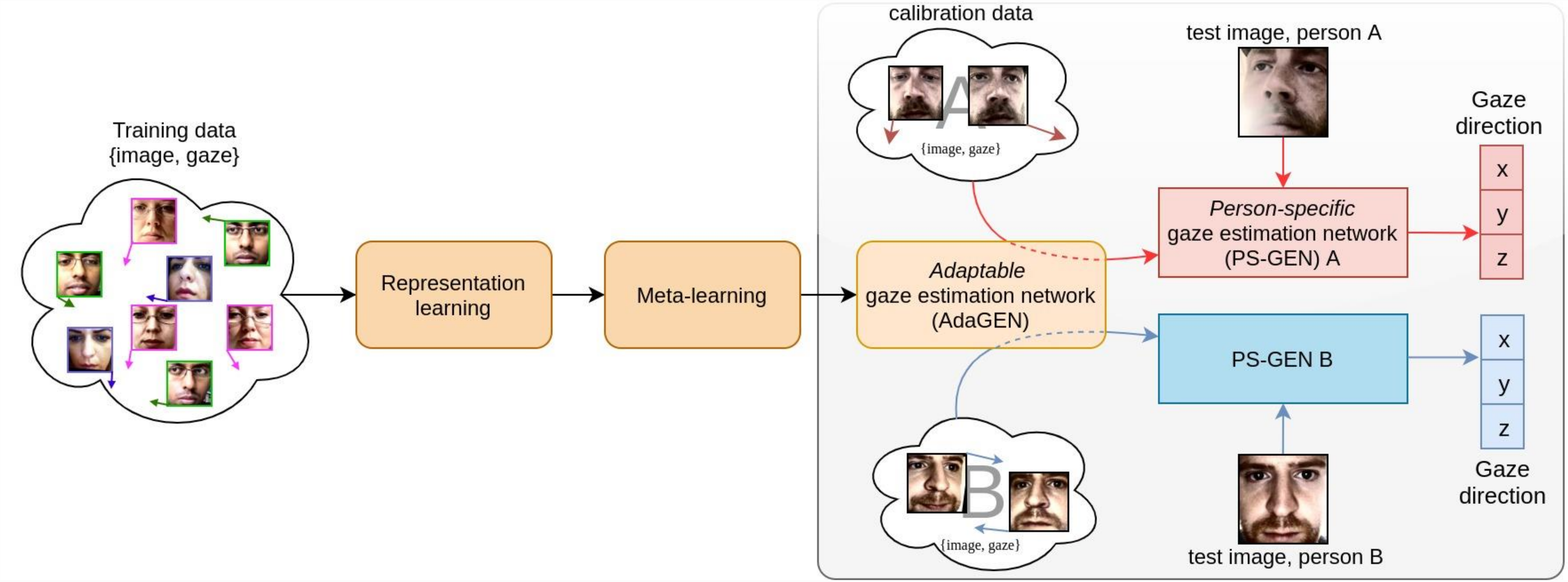
2. WTF (the YouTube Faces : video fame) Dataset



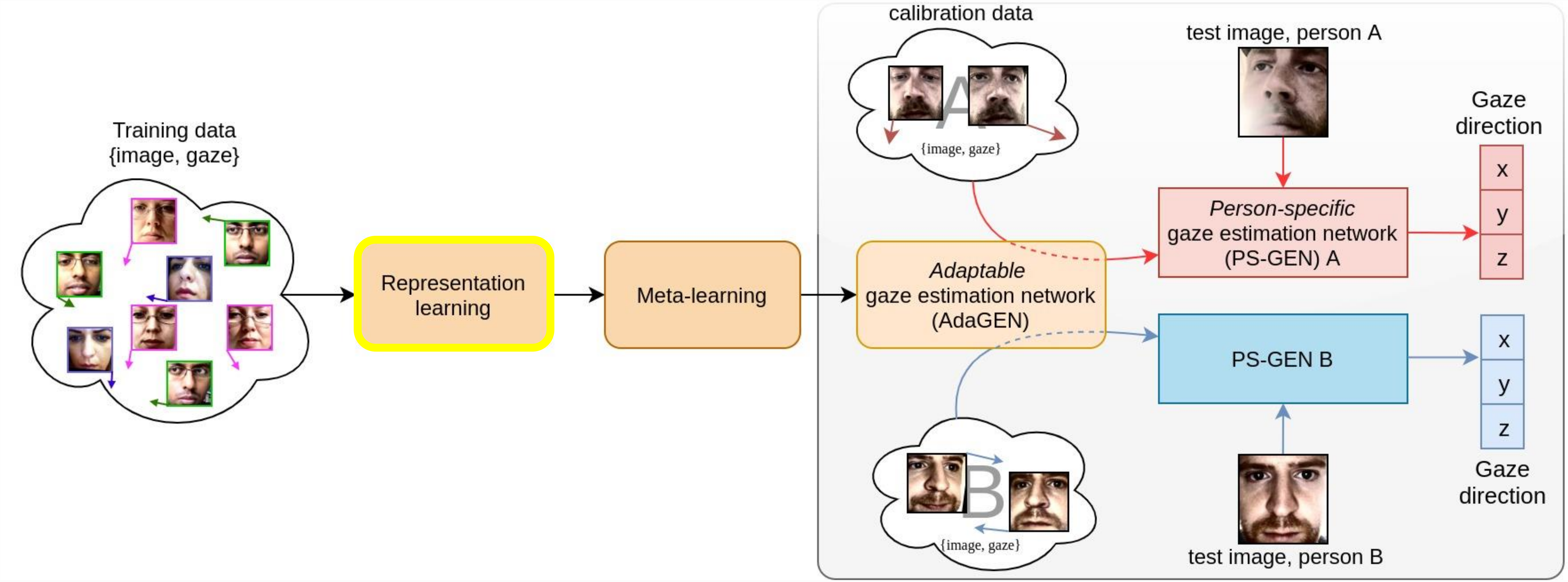
Method	Accuracy (%)	AUC	EER
MBGS+SVM- [31]	78.9 ±1.9	86.9	21.2
APEM+FUSION [22]	79.1 ±1.5	86.6	21.4
STFRD+PMML [9]	79.5 ±2.5	88.6	19.9
VSOF+OSS [23]	79.7 ±1.8	89.4	20.0
DeepFace-single	**91.4** ±1.1	96.3	8.6

- DeepFace achieve **91.4%**
- Much better performance than Accuracy of the most successful recent study (79.7%)
- **Relatively weak performance in video frames** compared to web images

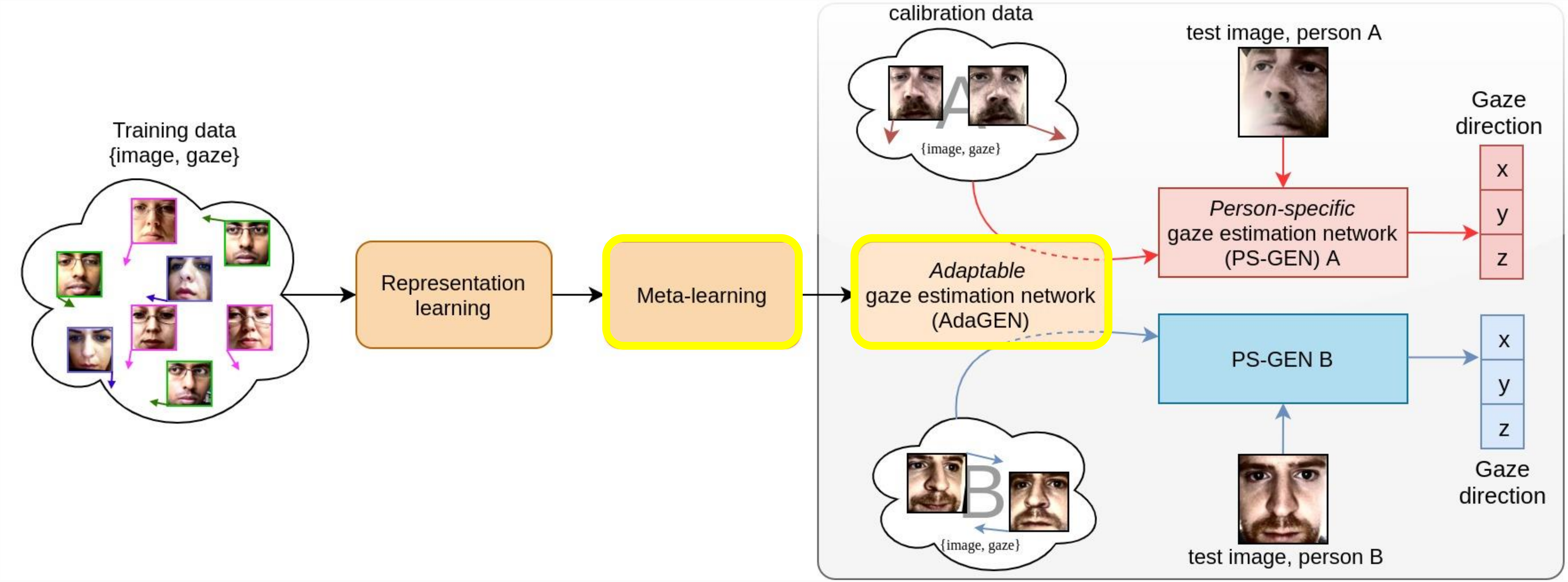
Related Work: FAZE (Few-Shot Adaptive Gaze Estimation)



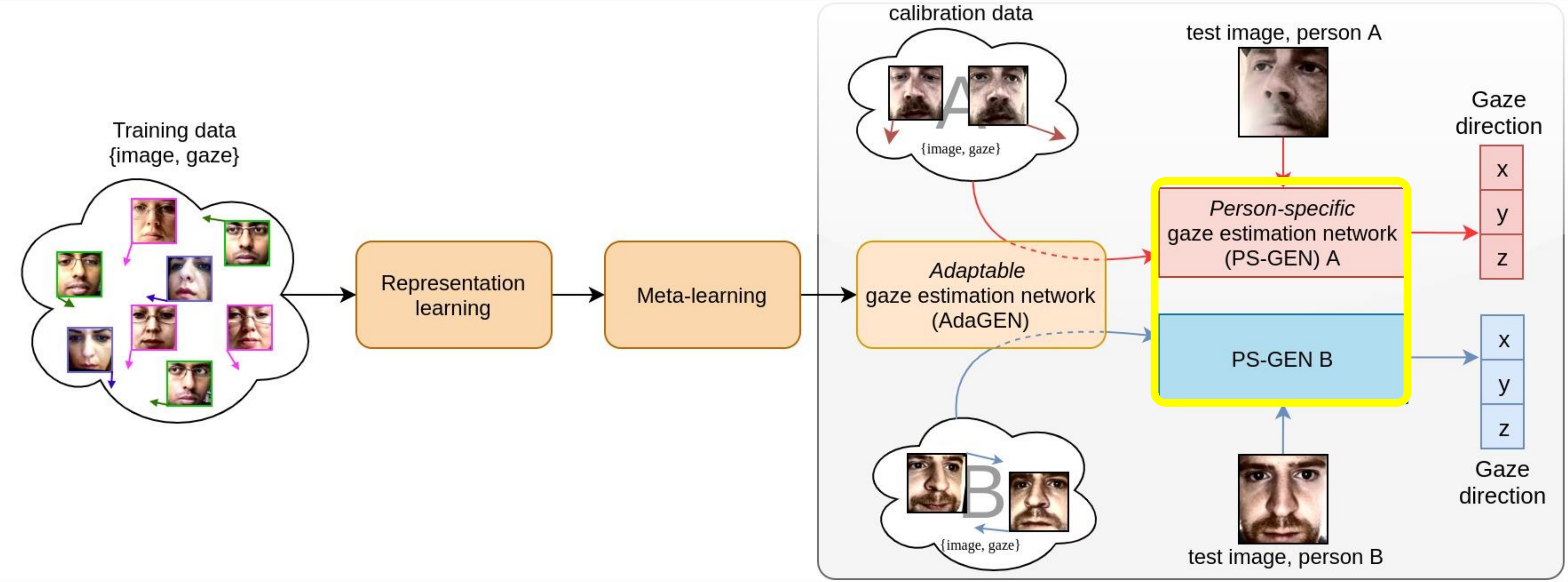
Related Work: FAZE (Few-Shot Adaptive Gaze Estimation)



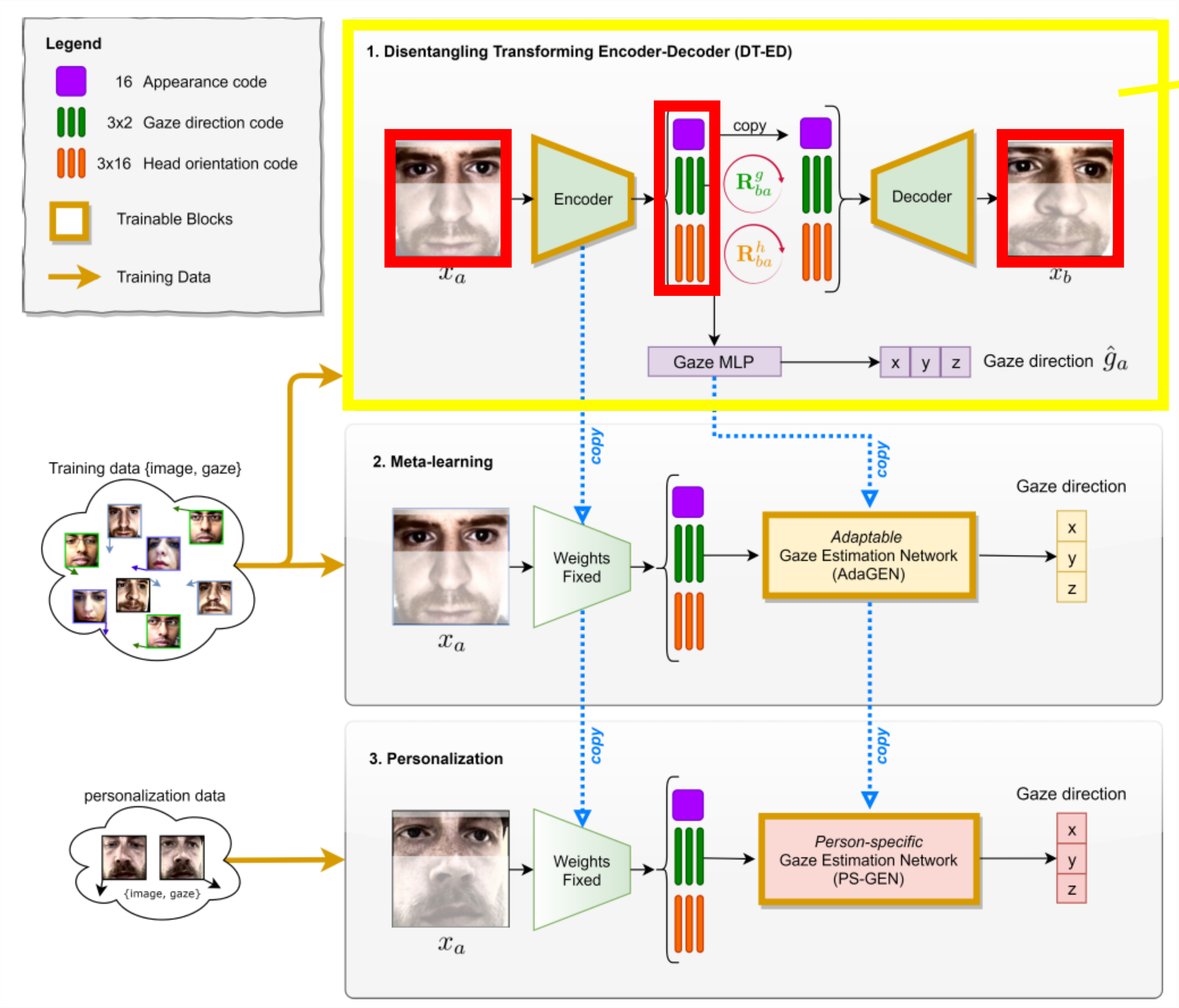
Related Work: FAZE (Few-Shot Adaptive Gaze Estimation)



Related Work: FAZE (Few-Shot Adaptive Gaze Estimation)



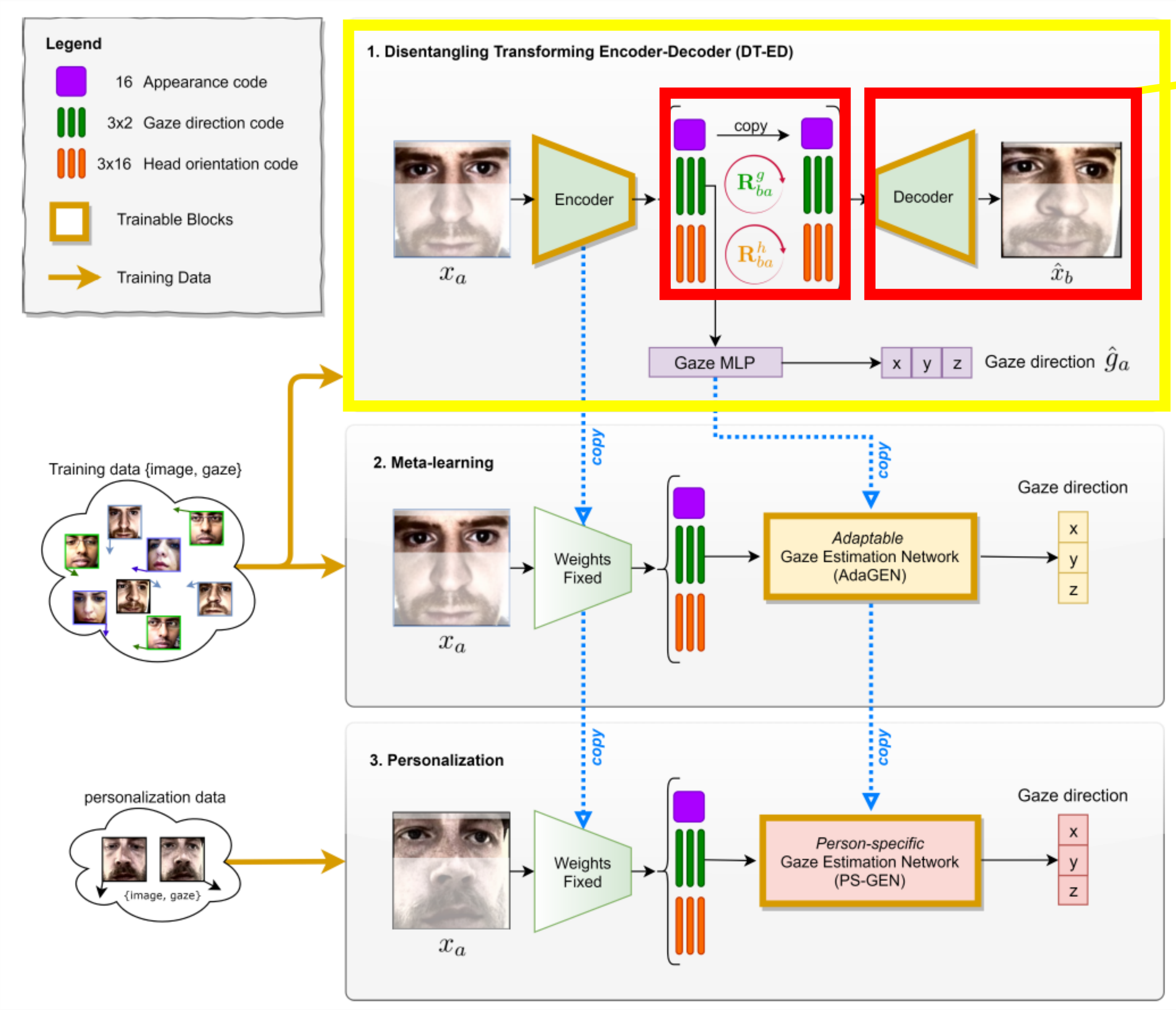
Related Work: FAZE (Few-Shot Adaptive Gaze Estimation)



Representation Learning Part

- Input is a pair of images from the same person
- Image is embedded into 3 latent spaces : **Appearance, Gaze and Head pose**

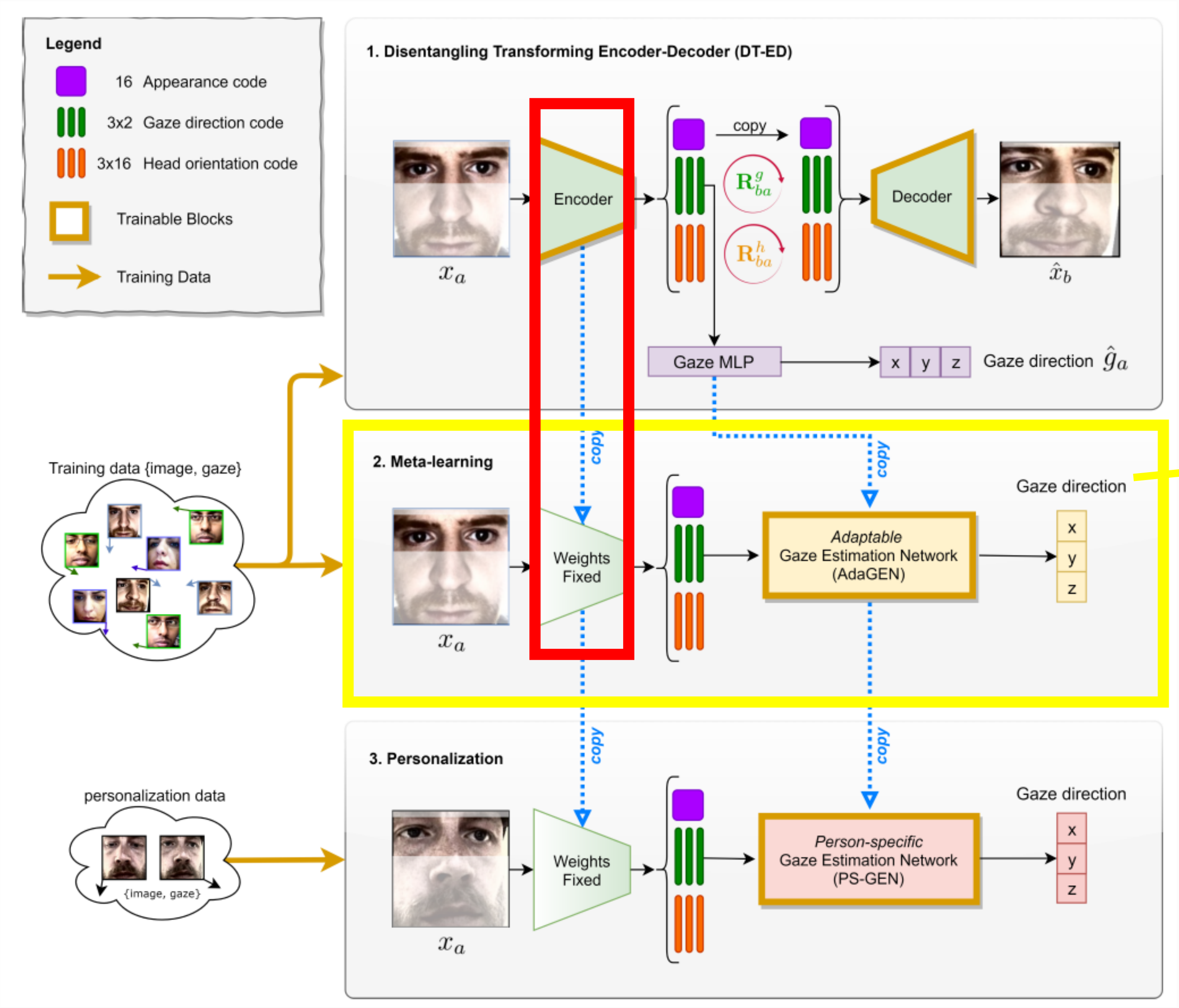
Related Work: FAZE (Few-Shot Adaptive Gaze Estimation)



Representation Learning Part

- Gaze and Head pose codes are **rotated based on known differences.**
- Appearance code passes through.
- Decoder ouput is guided with an L1 loss with the additional loss terms.

Related Work: FAZE (Few-Shot Adaptive Gaze Estimation)



Meta Learning Part

- Learn to learn
- Gaze Embedding can be used.

Related Work: FAZE (Few-Shot Adaptive Gaze Estimation)

θ_n : initial weights of the network

Randomly Selected P_{train}

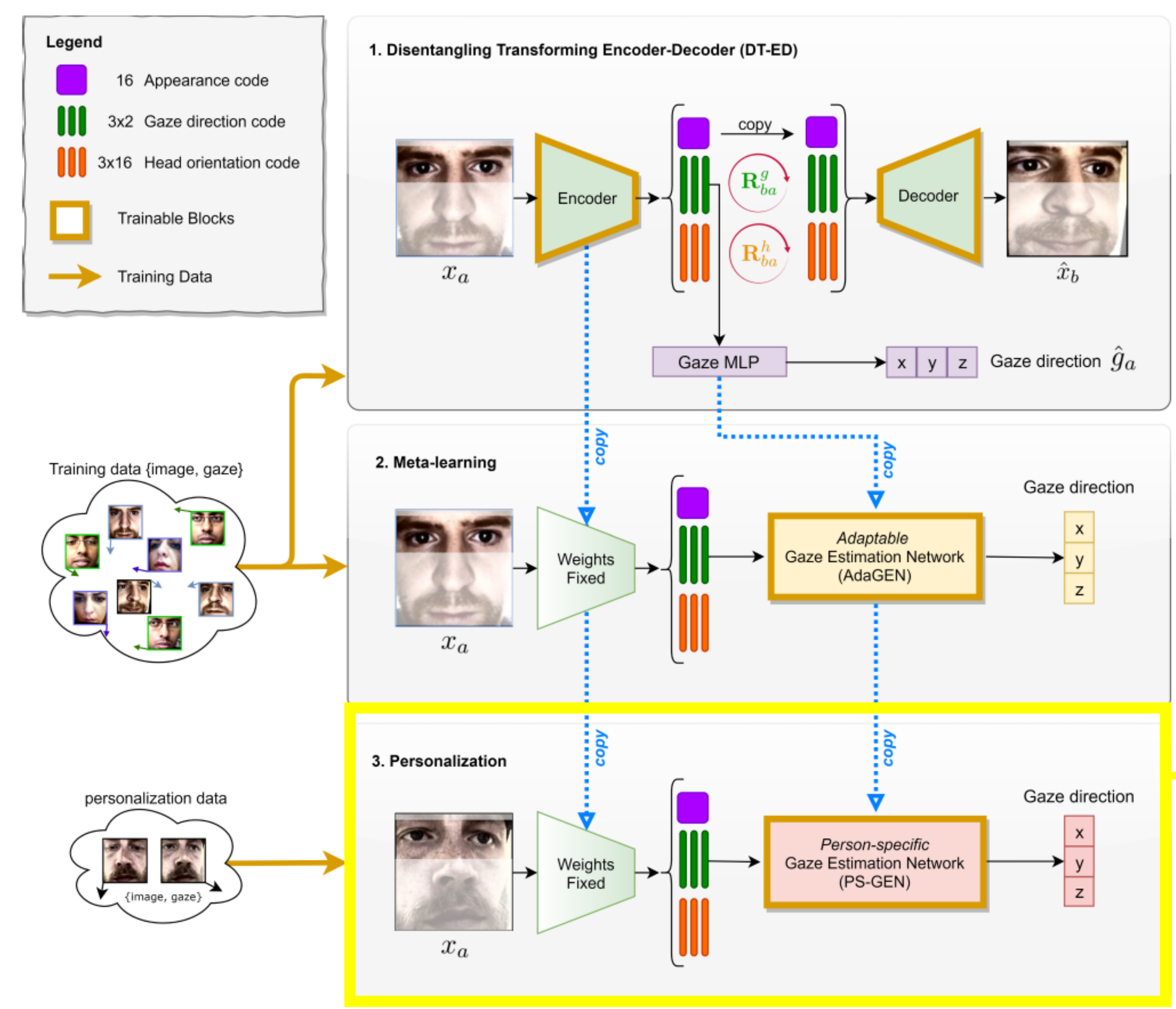
Few-Shot D_{train}

Update : $\theta'_n = f(\theta_n) = \theta_n - \alpha \nabla L^t_{P_{train}}(\theta_n)$

D_{valid}

Update : $\theta_{n+1} = \boldsymbol{\theta_n} - \beta \nabla L^v_{P_{train}}(\boldsymbol{f}(\boldsymbol{\theta_n}))$

Related Work: FAZE (Few-Shot Adaptive Gaze Estimation)



Personalization

- Similar to Meta-Learning
- **Using unseen person's data to yield final person specific model.**

Related Work: FAZE (Few-Shot Adaptive Gaze Estimation)

θ^* : optimal weights of the network

Unseen Person P_{test}

D_{test}

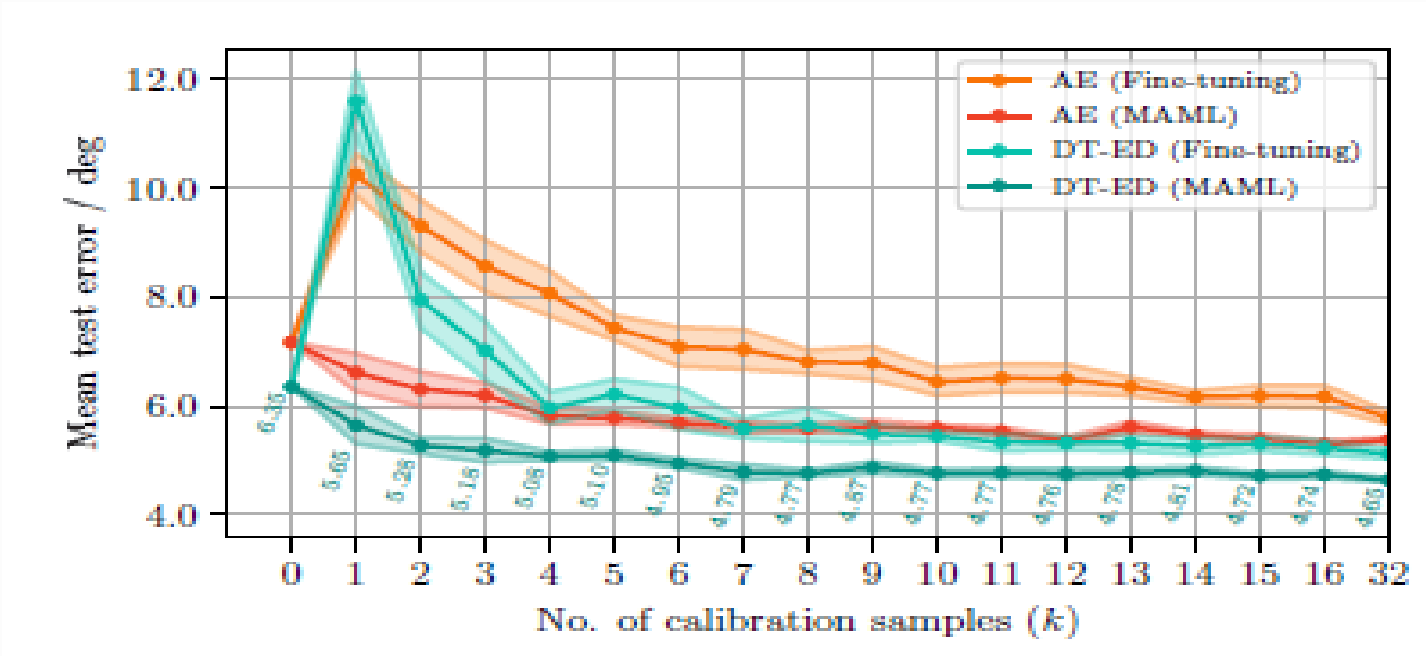
Update : $\theta_{P_{test}} = \theta^* - \alpha \nabla L^t_{P_{test}}(\theta^*)$

D_{valid}

Evaluate

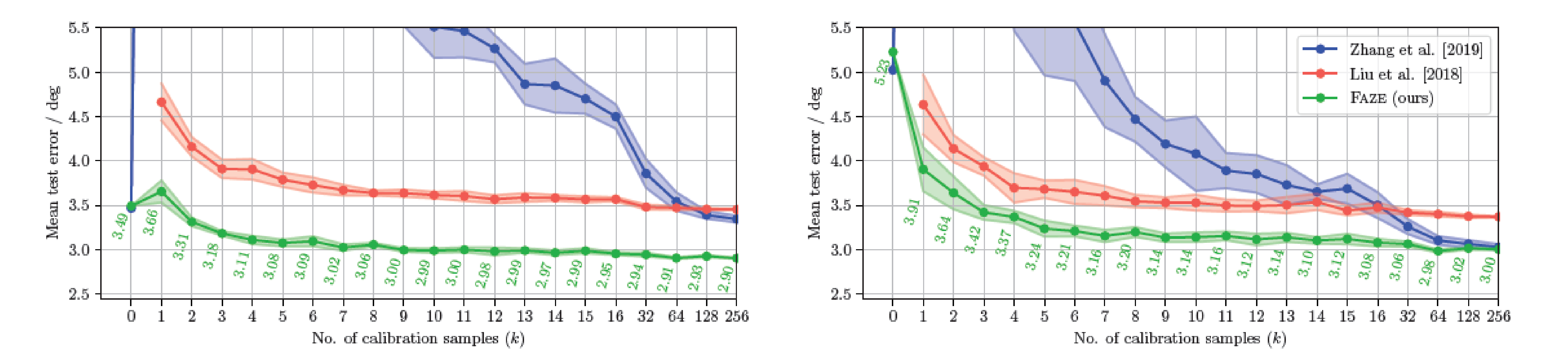
Related Work: FAZE (Few-Shot Adaptive Gaze Estimation)

Result of Experiments



Related Work: FAZE (Few-Shot Adaptive Gaze Estimation)

Result of Experiments



(a) GazeCapture (test) (b) MPIIGaze

Limitation: DeepFace

1. 120 million too many parameters due to Locallly - connected layer

2. Relatively low recognition accuracy for video frames

3. Applicability to low-end camera environment of a laptop computer

Limitation: FAZE

1. Application to Real-Time Video

2. Using a few Sample

3. Person-specific model

Thank you!