Hybrid eye center localization using cascaded regression and robust circle fitting

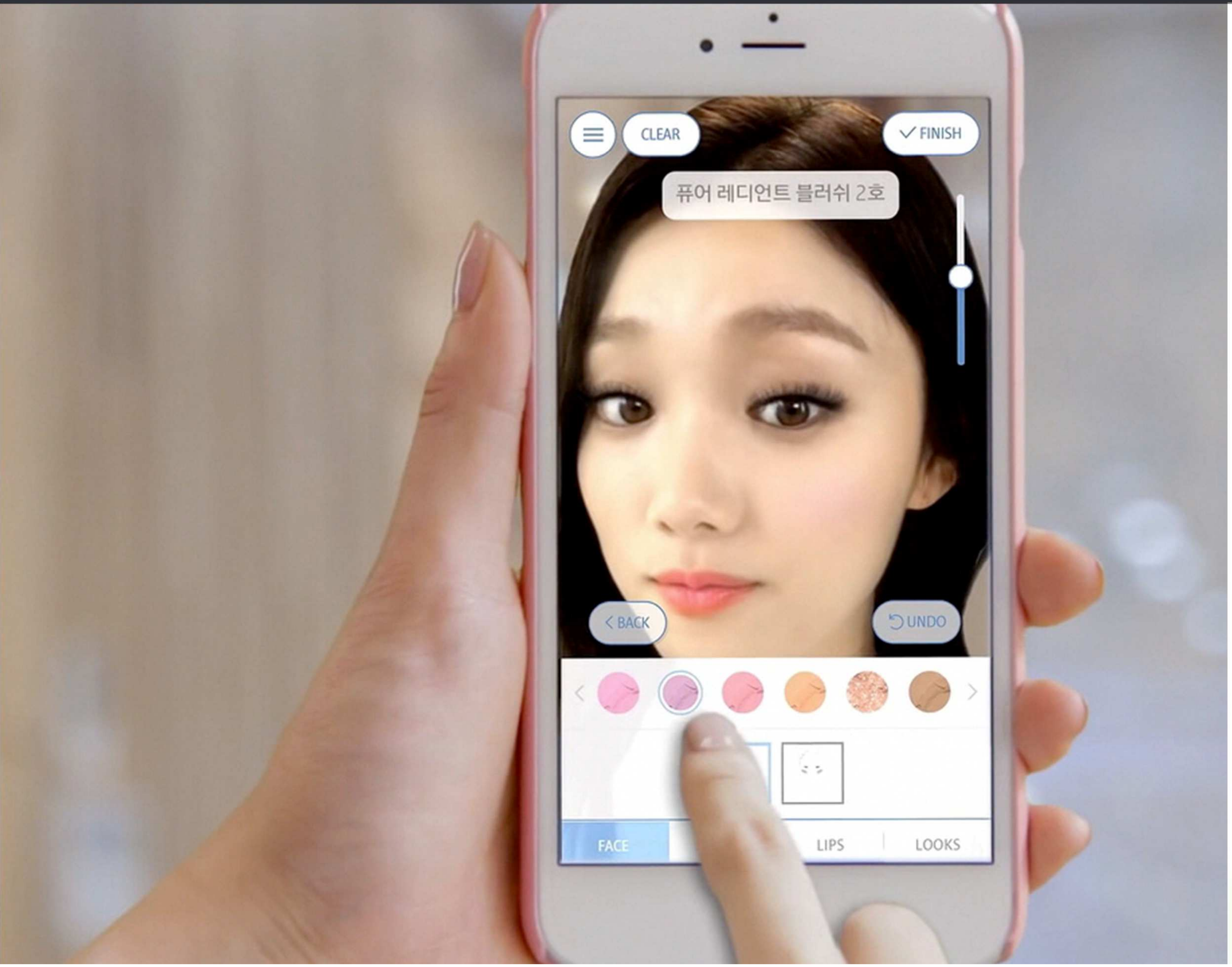
Alex Levinshtein   Edmund Phung   Parham Aarabi

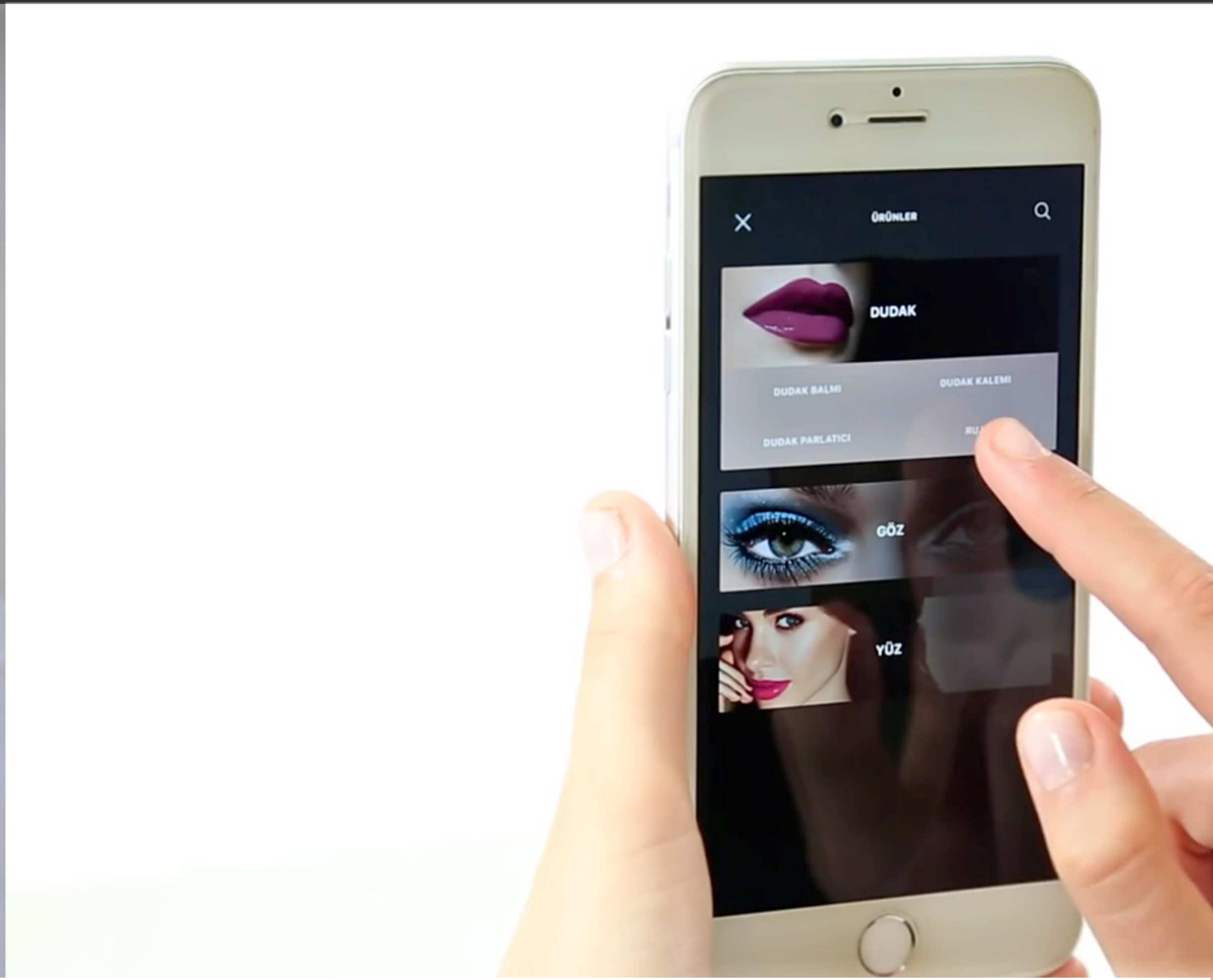
MODIFACE

UNIVERSITY OF TORONTO

LANEÍGE

AVON



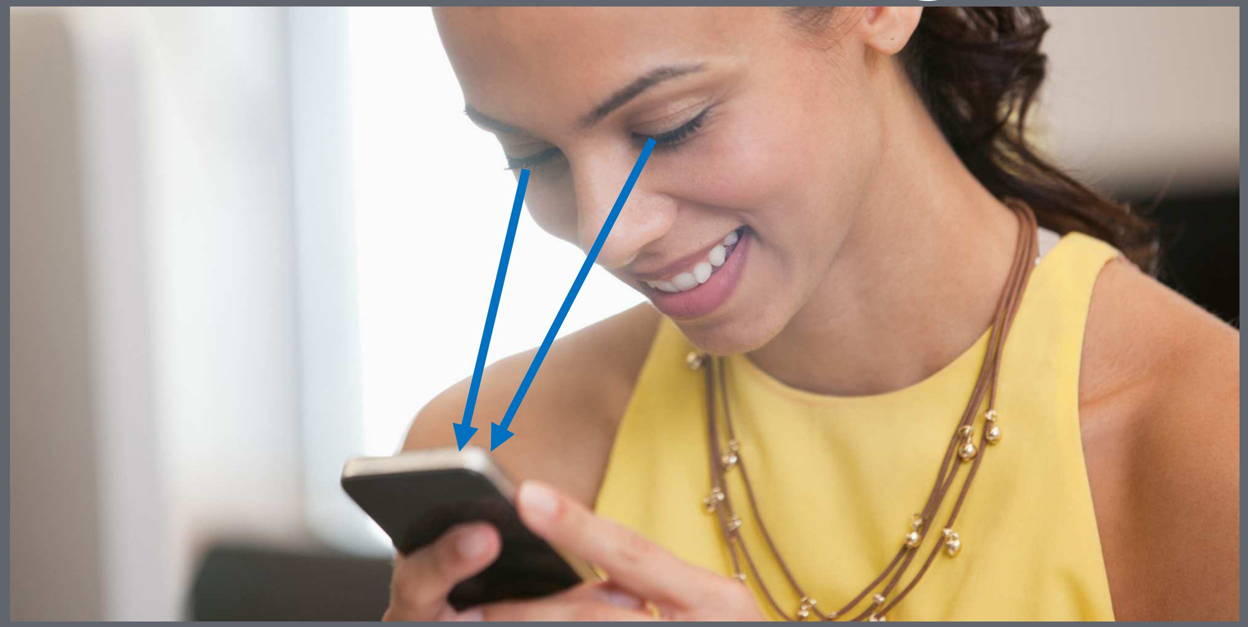


Customized Native iOS/Android/Web Apps

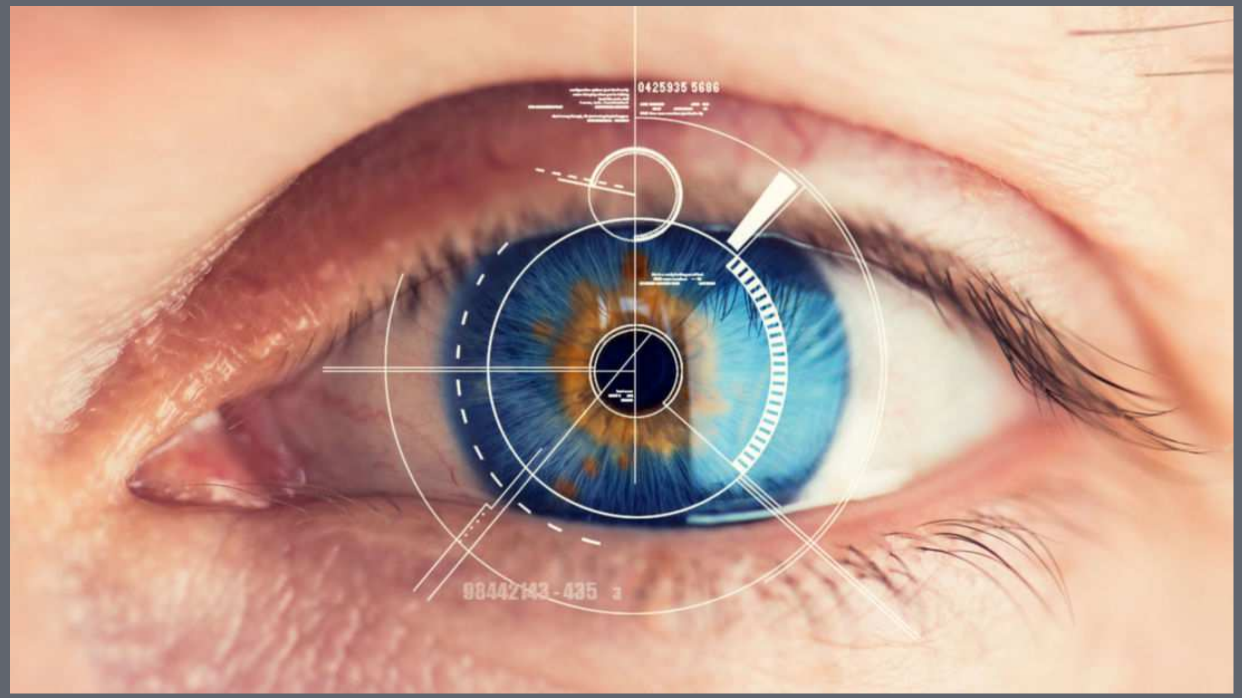
The need for iris detection

Virtual contact lenses try-on

Gaze tracking

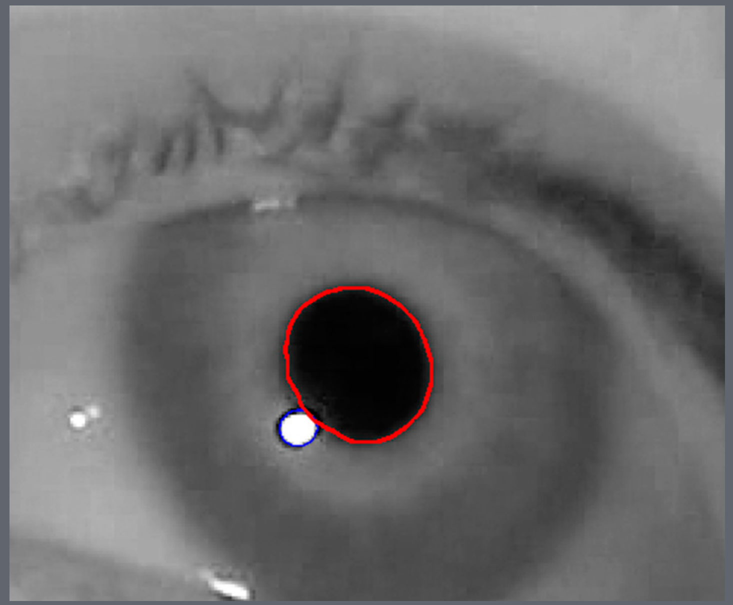


Iris recognition

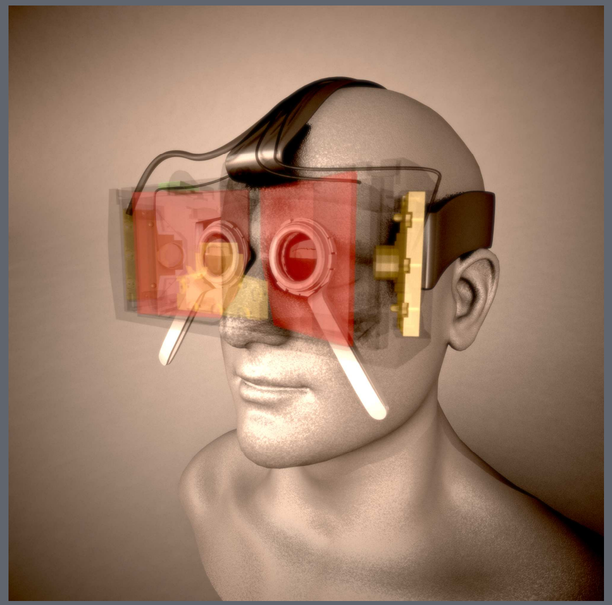


In-the-wild iris detection

Most methods use specialized devices



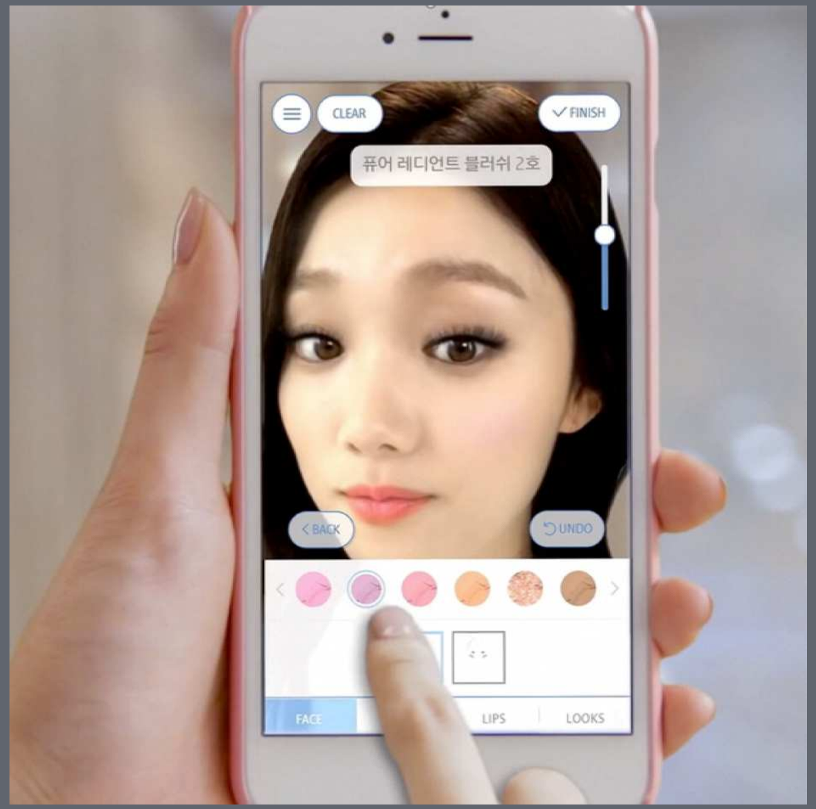
NIR images

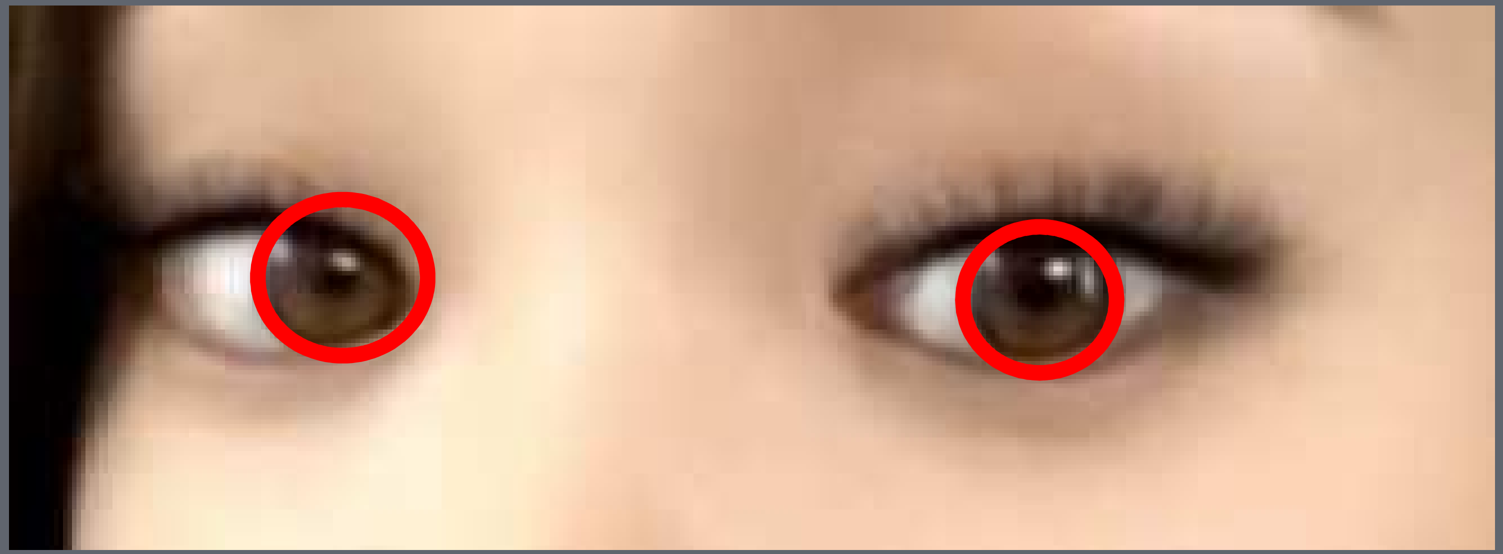


Head mounts

Iris scanners

Our focus: real-time detection on mobile devices using a standard camera





Related Work

Hand-crafted methods

Ahuja et al. 2016
Fuhl et al. 2015,2016
George and Routray 2016
Li et al. 2005
Skodras and Fakotakis 2015
Swirski et al. 2012
Timm and Barth 2011
Valenti and Gevers 2012
Wood and Bulling 2014

Accurate but not robust

Learning-based methods

Markus et al. 2014
Tian et al. 2016
→ Powerful regressors, Simple features

Zhou et al. 2015
→ Simple regressors, Powerful features

More robust but not accurate

Our contributions

1. Learning-based method using powerful regressors **and** powerful features

2. Increased accuracy by combining #1 with robust circle fitting

Method Overview

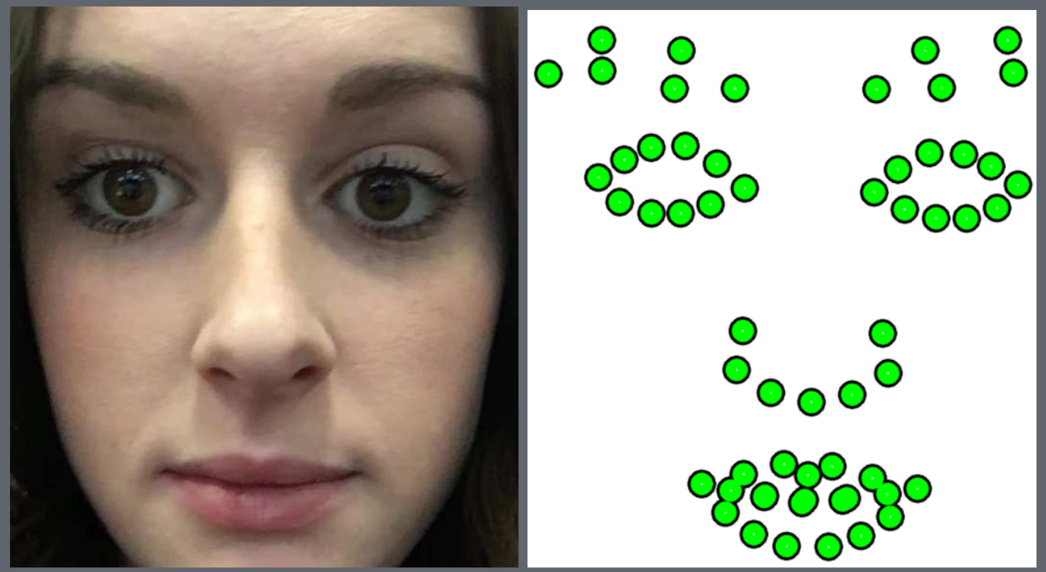
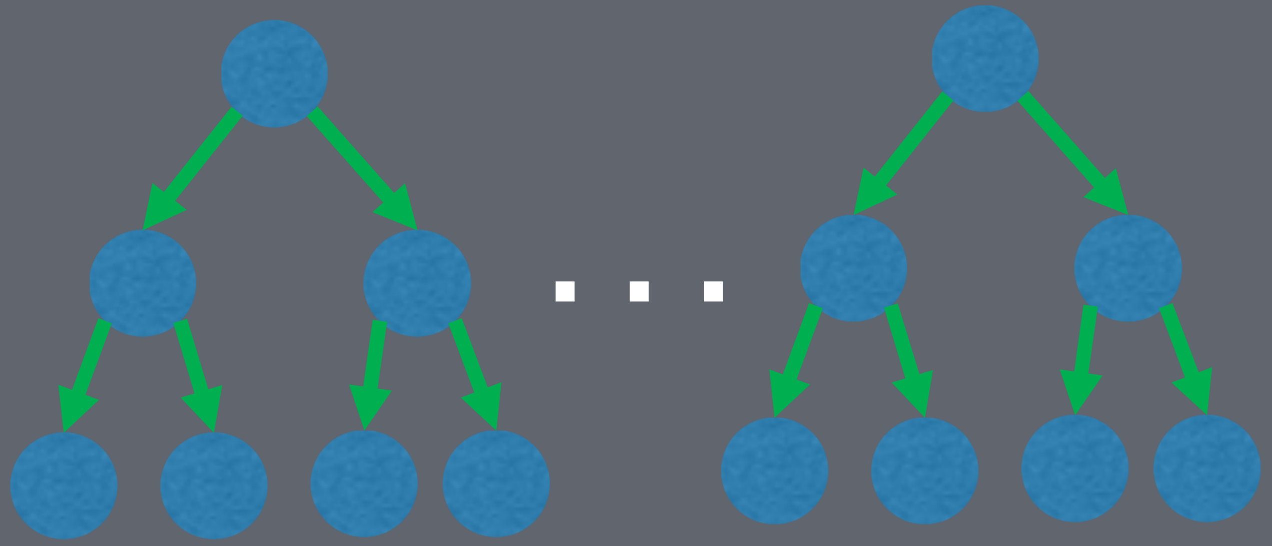
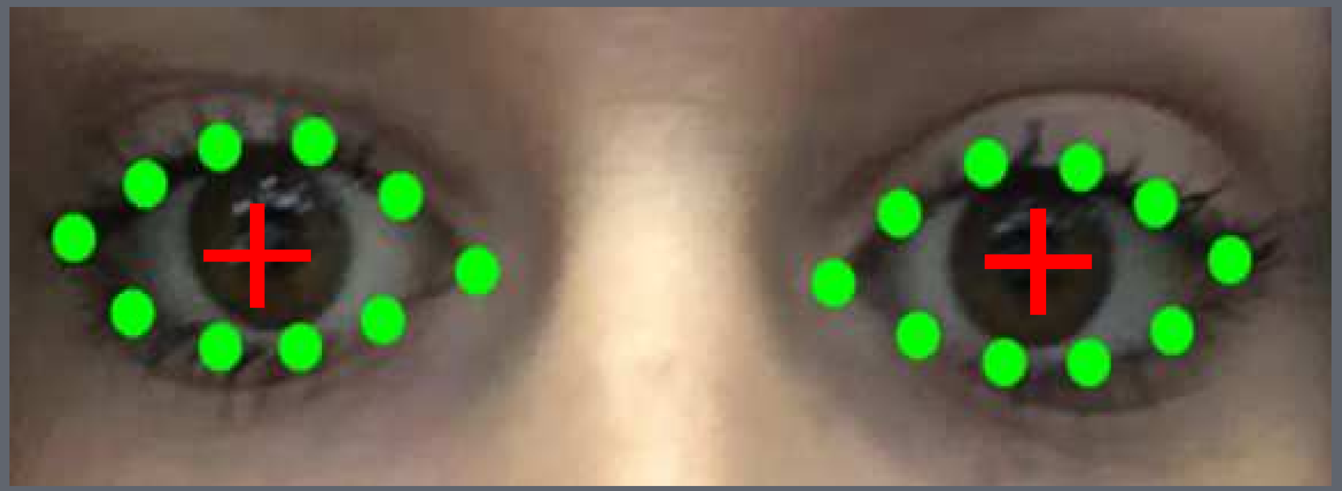


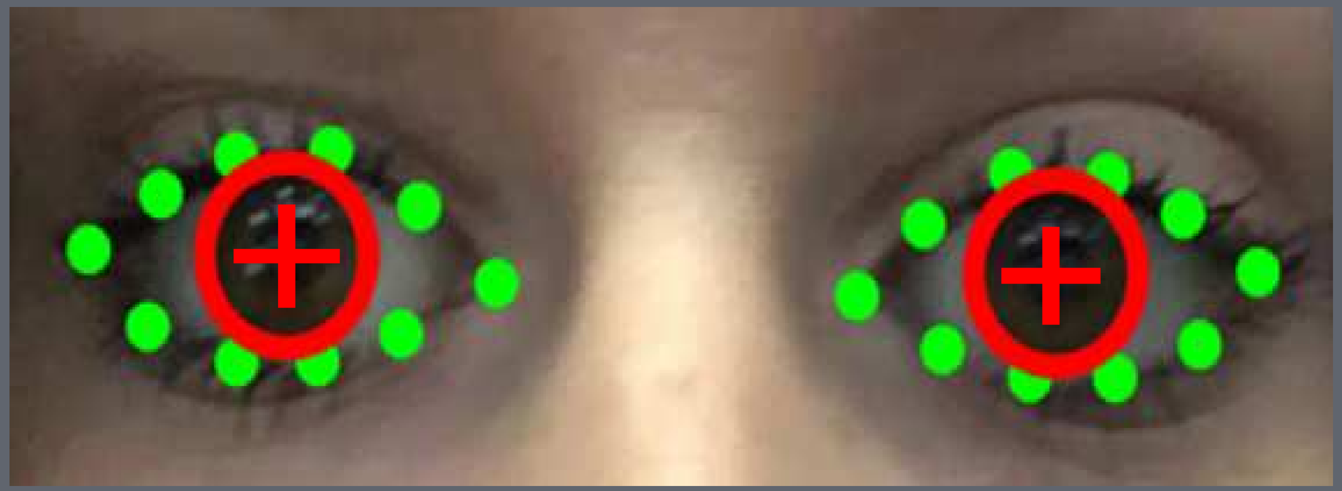
Image + Facial landmarks



Cascade of regression forests



Rough eye center estimates



Accurate eye center and iris radii estimates after circle fitting

Rough eye center detection

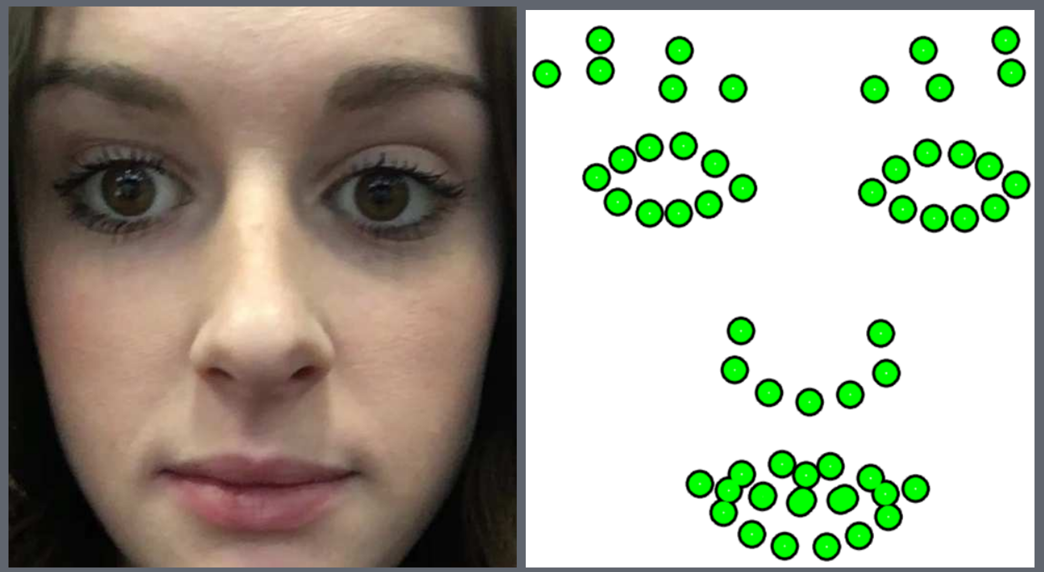
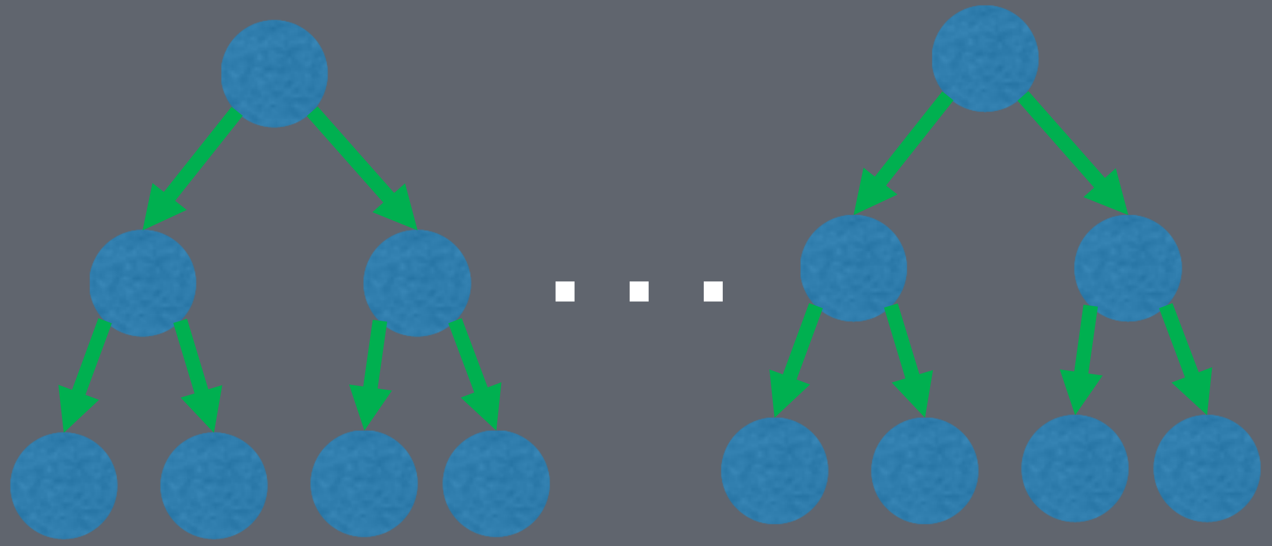
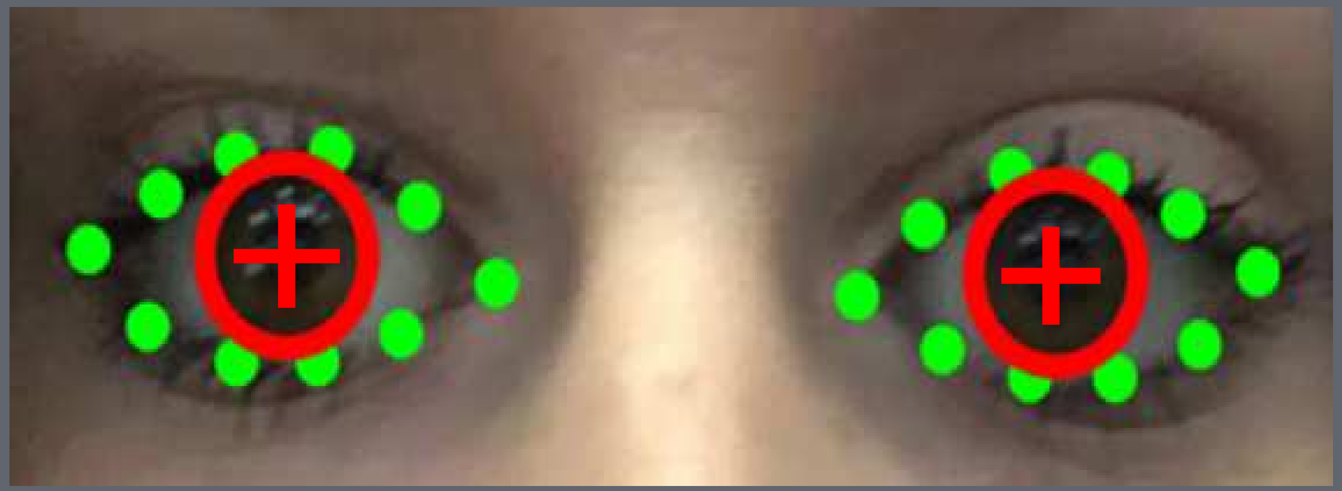


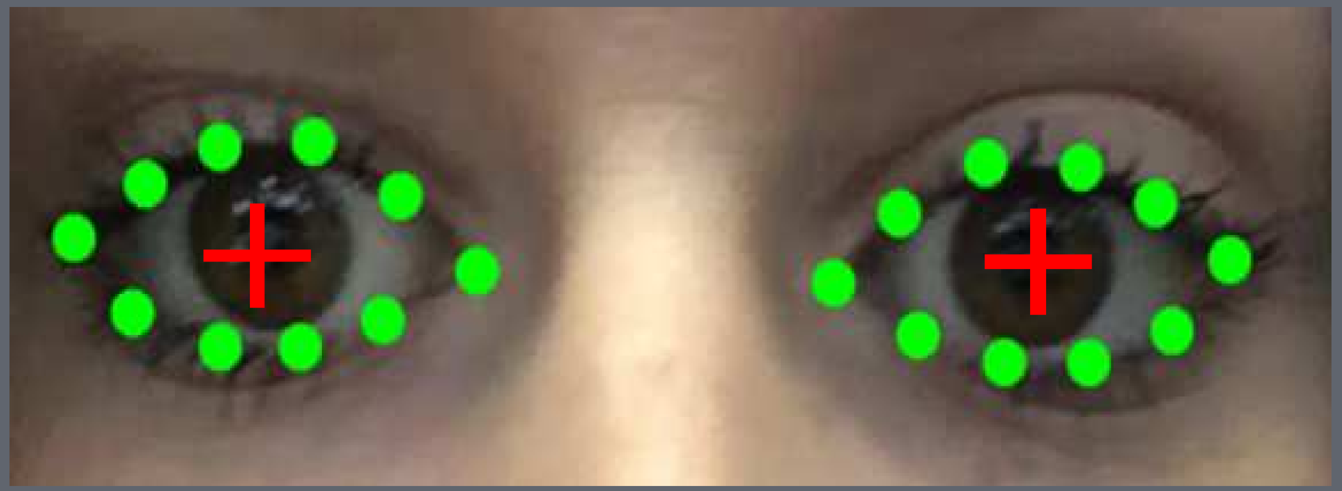
Image + Facial landmarks



Cascade of regression forests



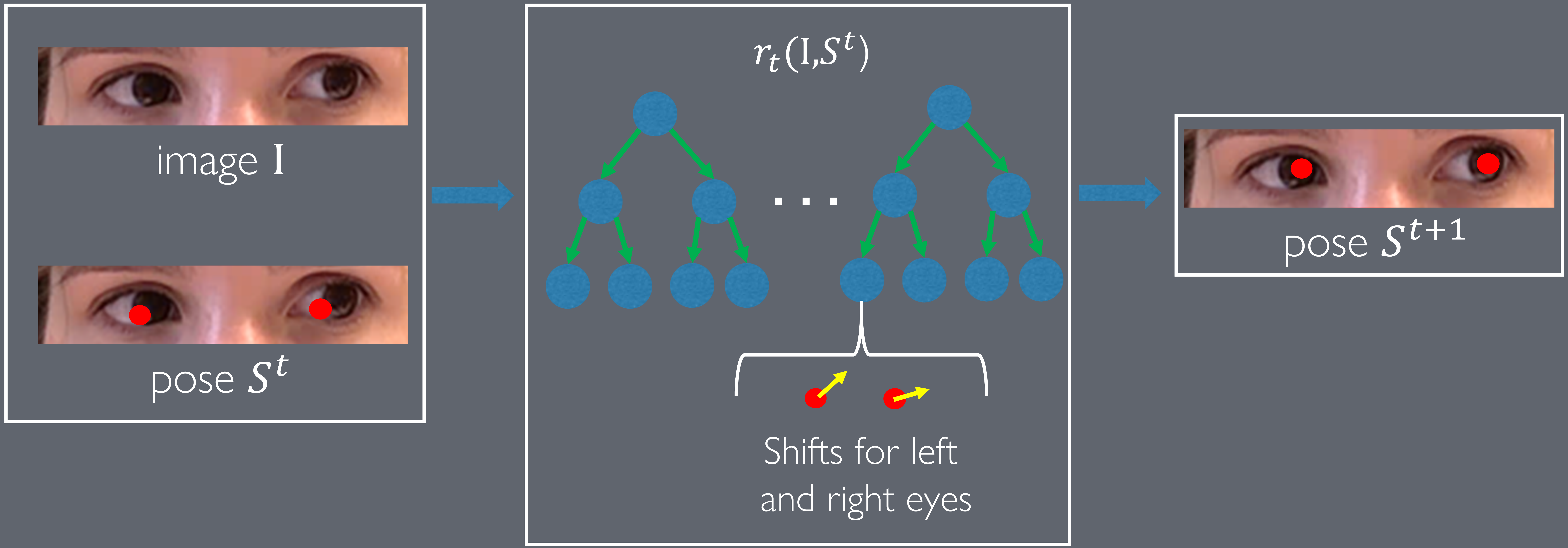
Accurate eye center and iris radii estimates after circle fitting



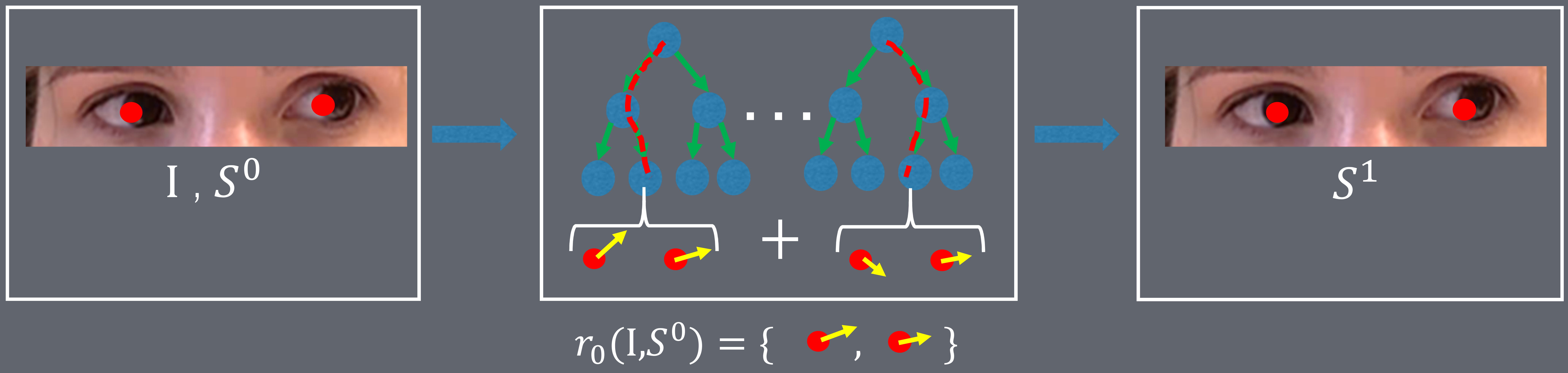
Rough eye center estimates

Cascade of gradient boosted regression forests

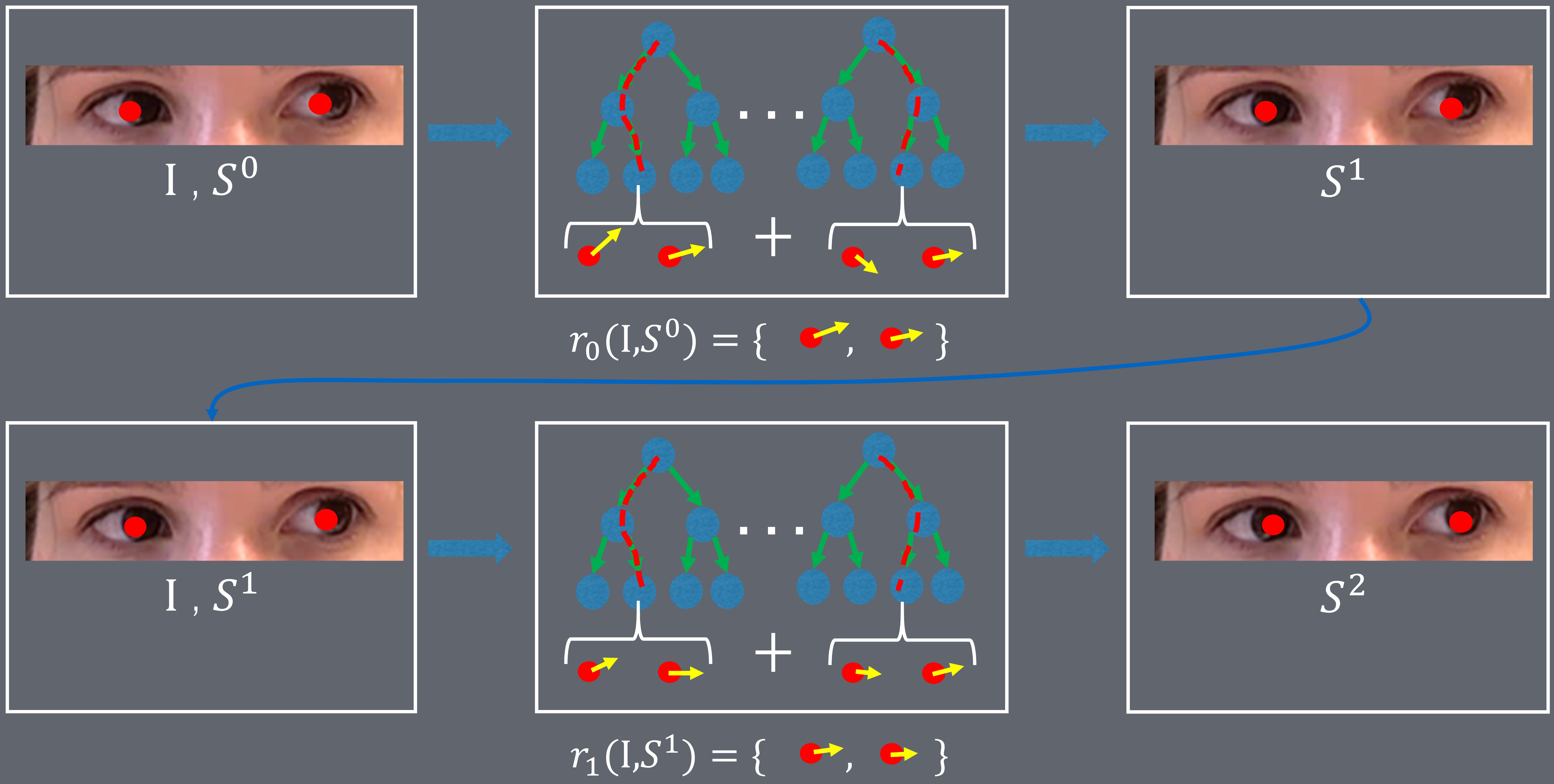
$$S^{t+1} = S^{t} + r_t(\mathrm{I}, S^{t})$$



Cascade of gradient boosted regression forests



Cascade of gradient boosted regression forests

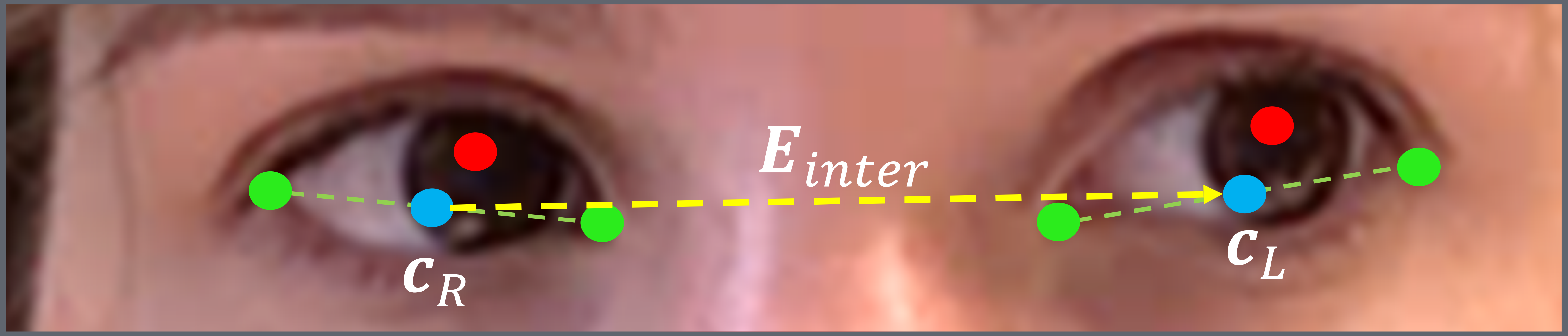


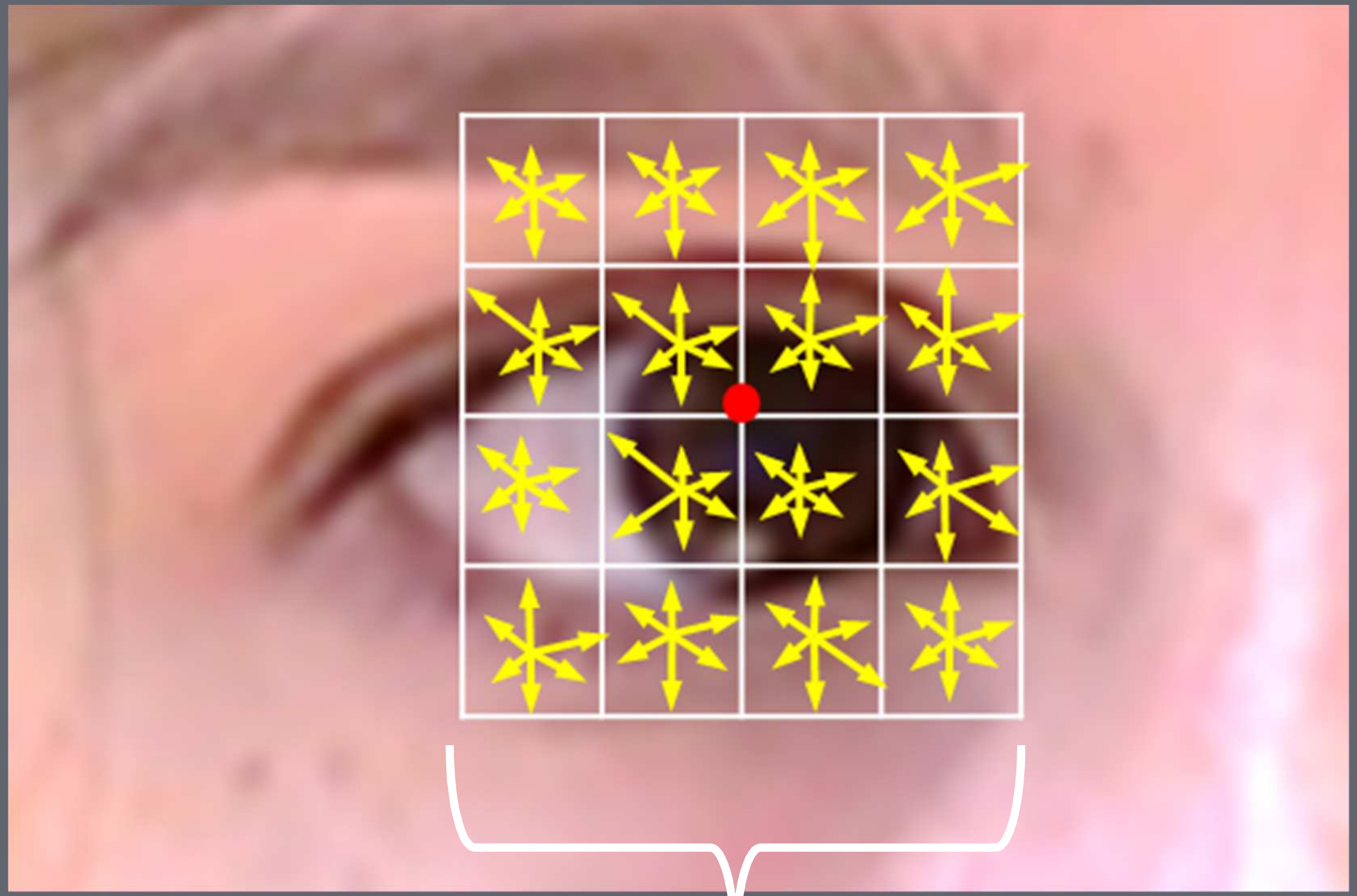
Beyond raw pixels in $r_t(\mathrm{I}, S^t)$

1. Use powerful features, instead of raw pixels

2. Anchor the features to current eye center estimates

Normalized HoG features for iris detection





$0.4 \cdot E_{inter}$

Initial pose

$$S^0 = \{\binom{0}{0}, \binom{1}{0}\}$$

Features:

Pairwise differences of HoG

Iris pose refinement by circle fitting

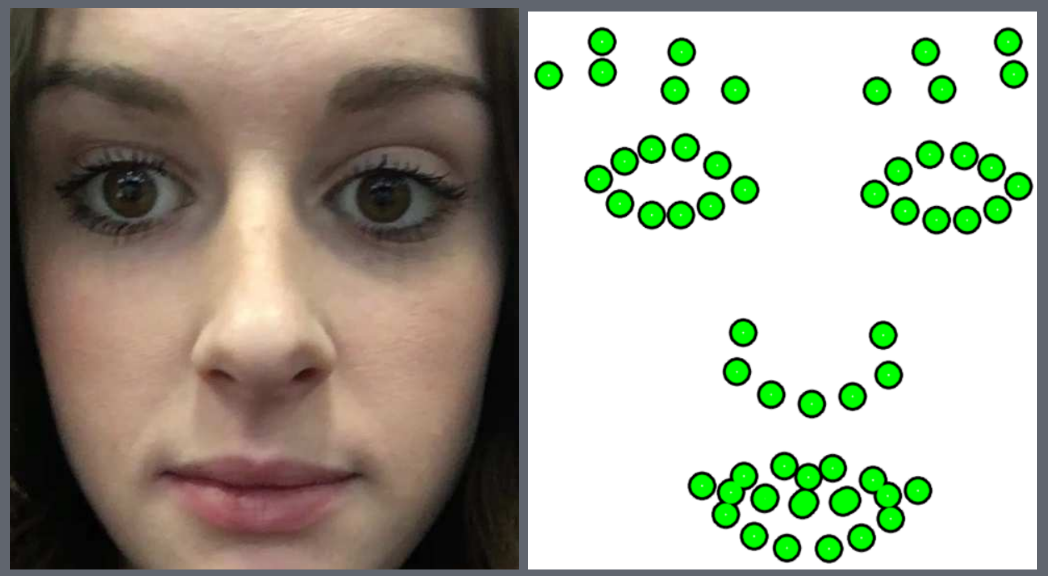
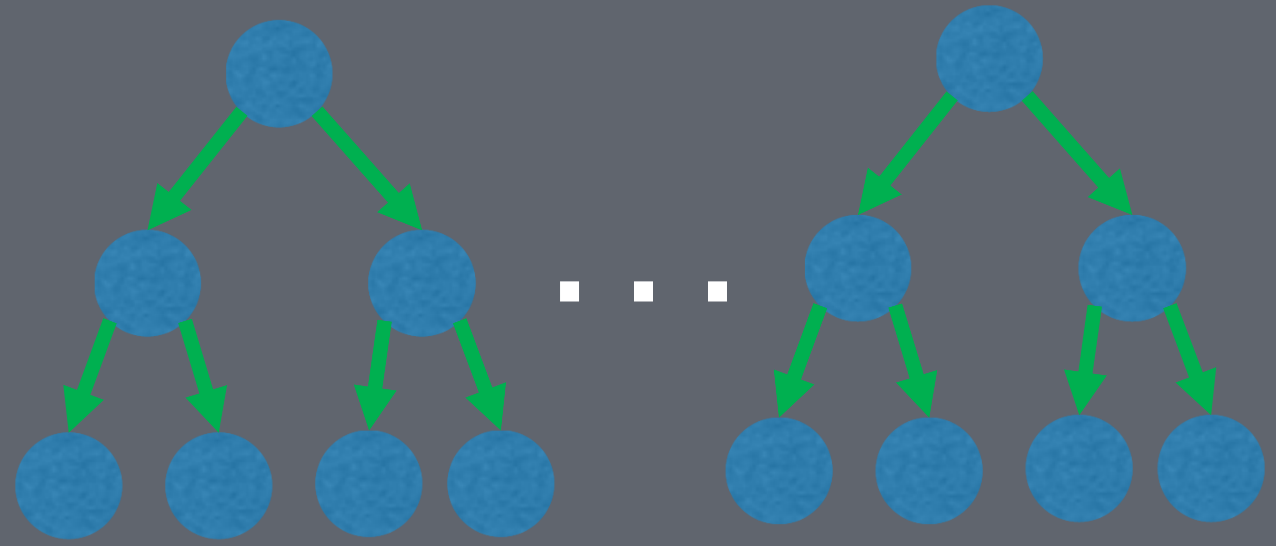
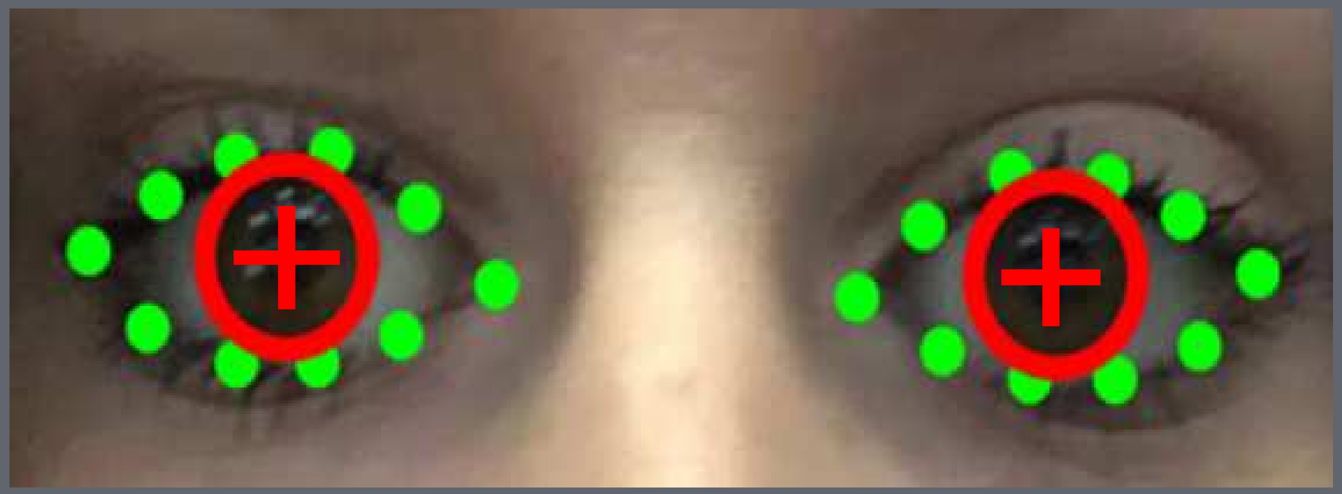


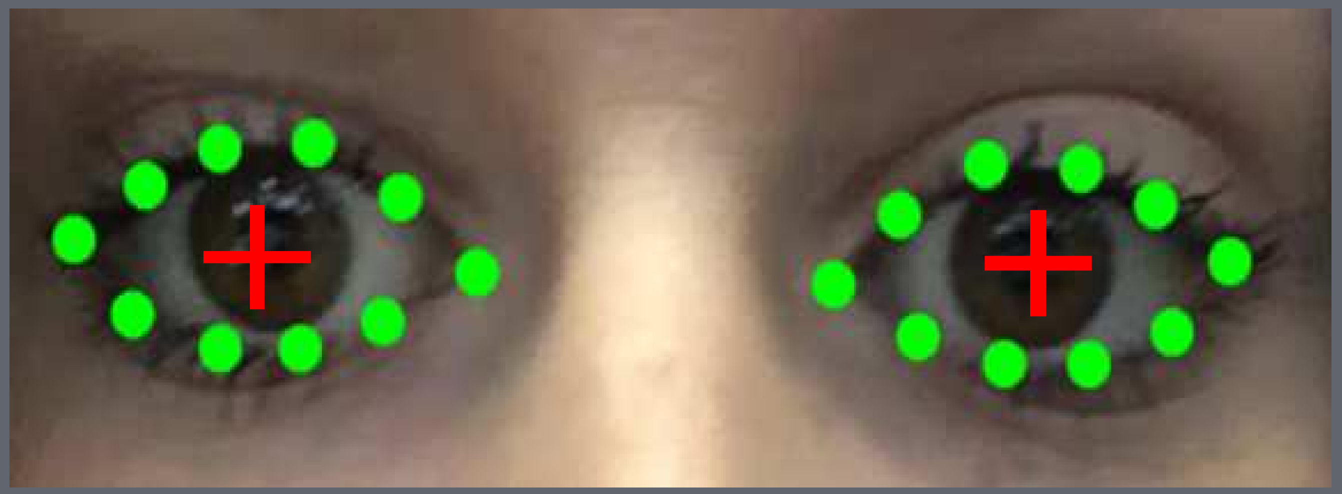
Image + Facial landmarks



Cascade of regression forests

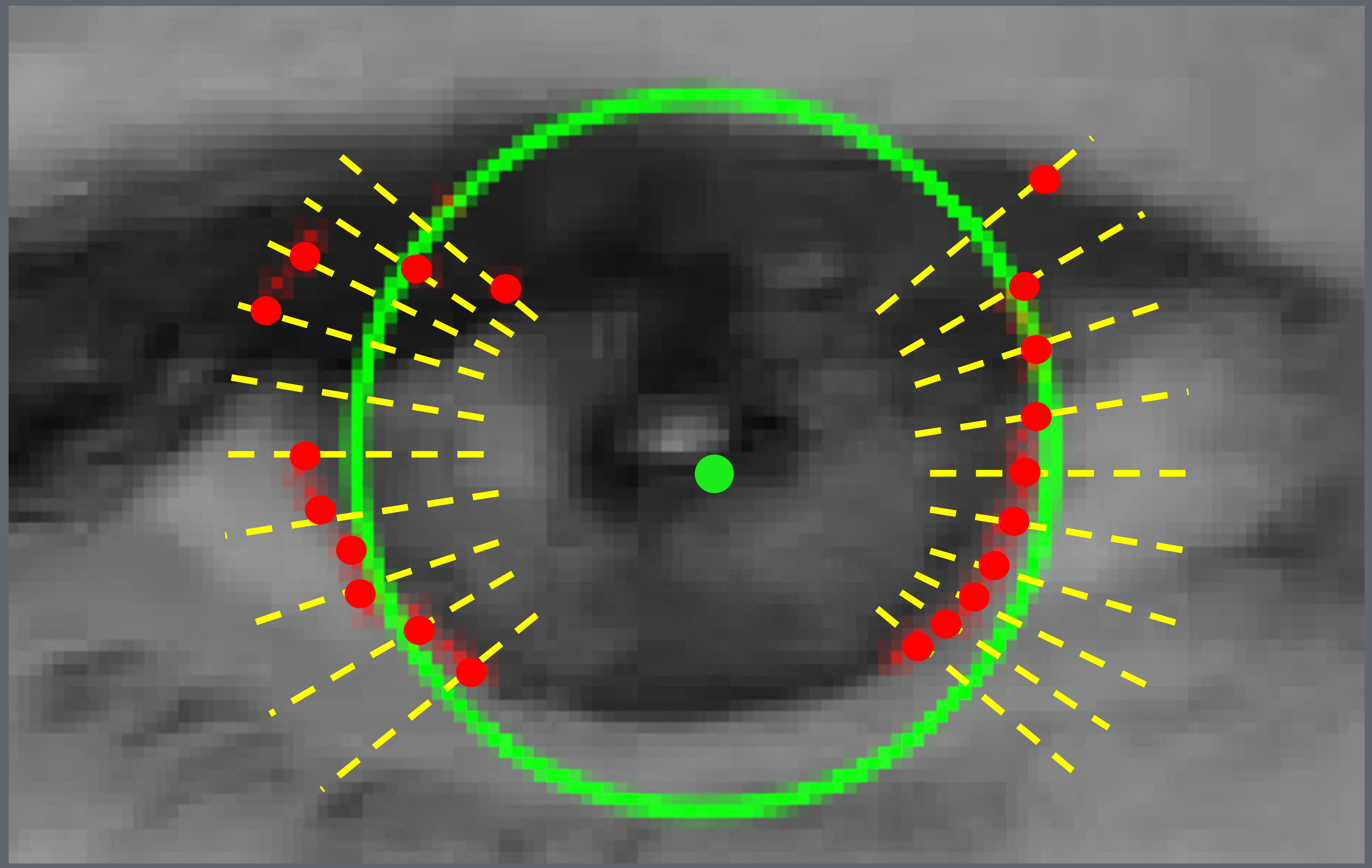


Accurate eye center and iris radii estimates after circle fitting



Rough eye center estimates

Robust circle fitting

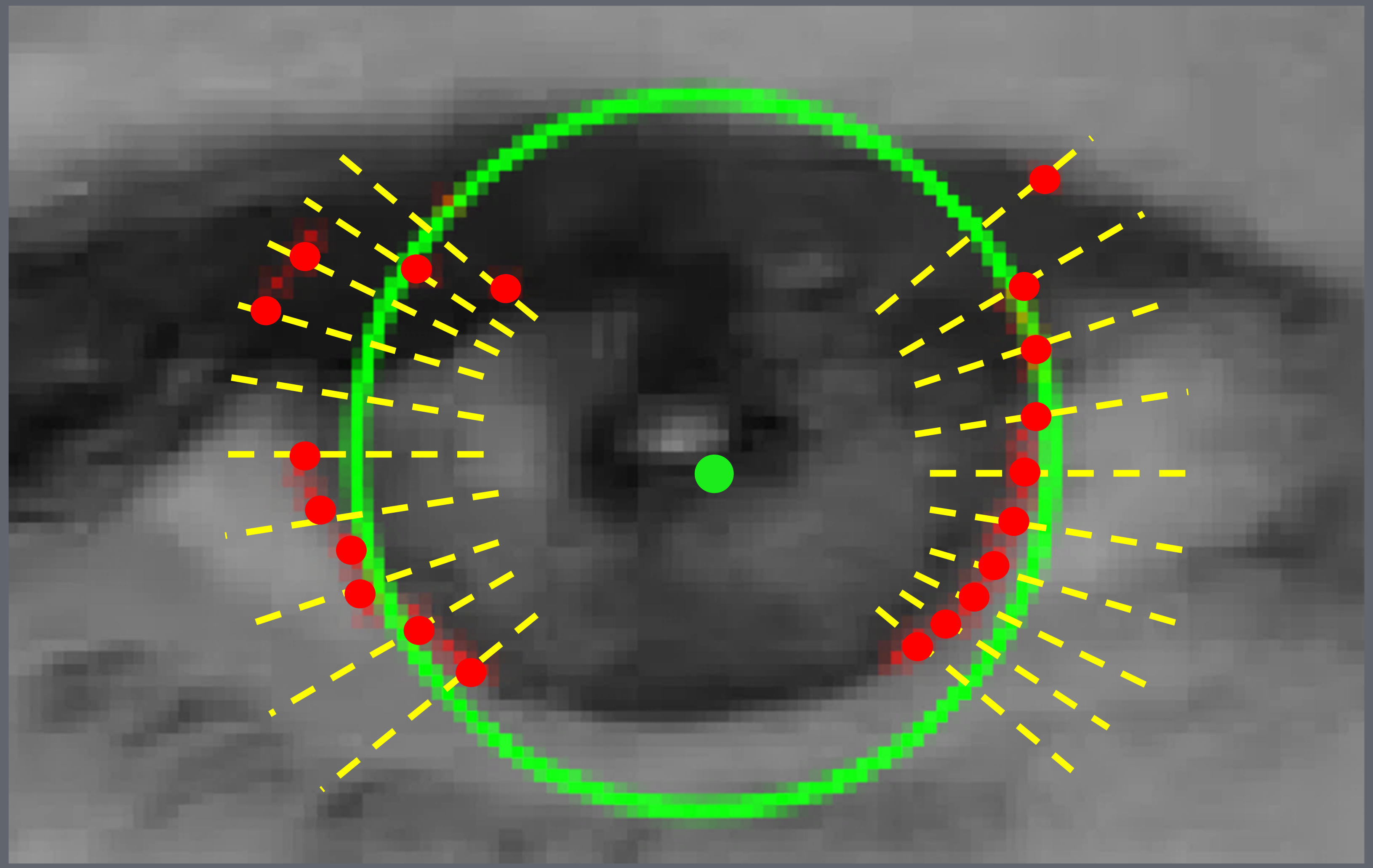


Fit iris using circle refinement

Start with (a, b) set to regressor output and r set to 10% of interocular distance

$$C_1 = \sum_{i=1}^{N} \left(\sqrt{(e_{ix} - a)^2 + (e_{iy} - b)^2} - r \right)^2$$

Robust circle fitting

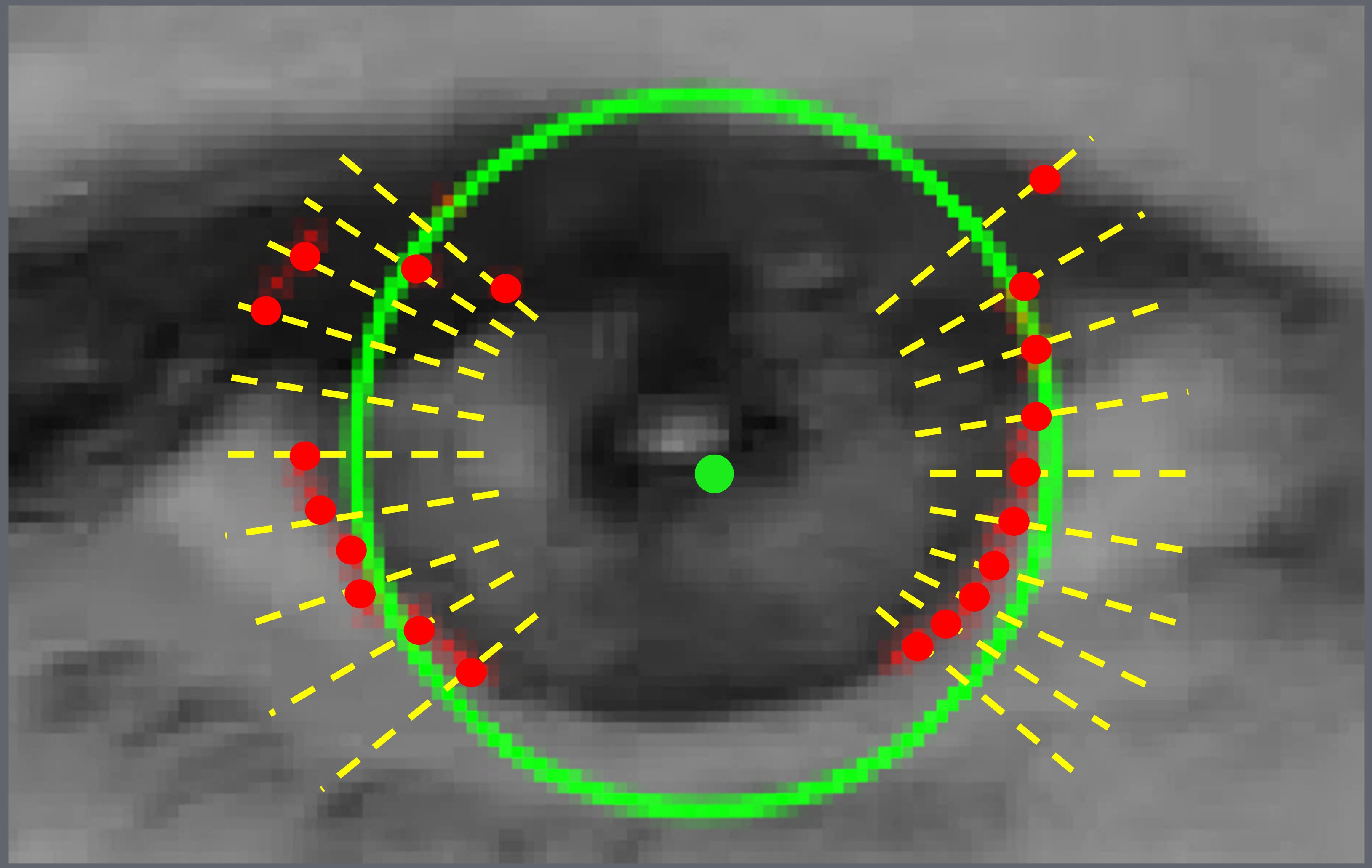


Issues:

1. Not robust to outliers
2. Only uses regressor output as initial guess

$$C_1 = \sum_{i=1}^{N} \left(\sqrt{(e_{ix} - a)^2 + (e_{iy} - b)^2} - r \right)^2$$

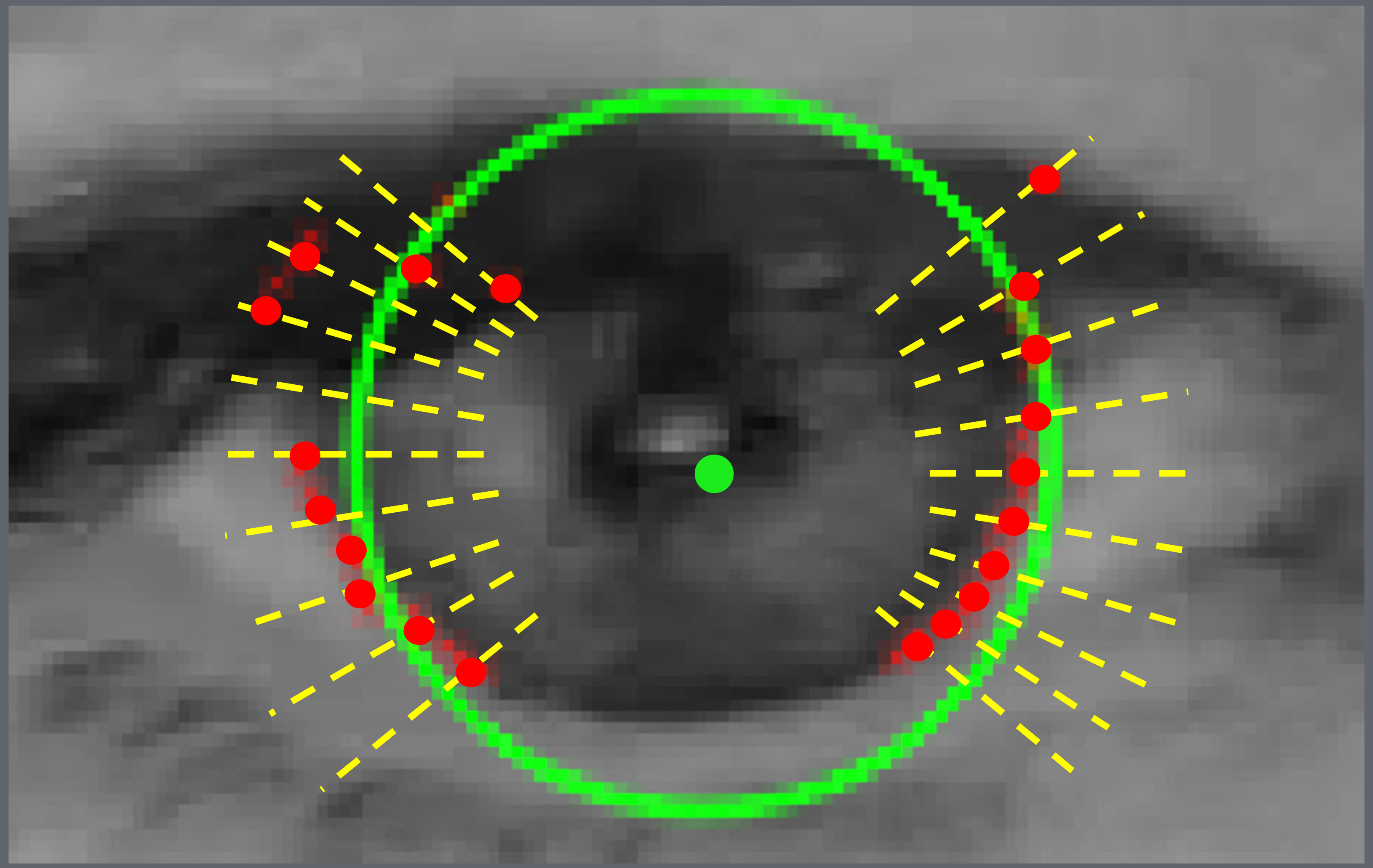
Robust circle fitting



$$C_2 = w_1 \frac{1}{N} \sum_{i=1}^{N} \rho\left(\sqrt{(e_{ix} - a)^2 + \left(e_{iy} - b\right)^2} - r\right) +$$

$$w_2(a - a_0)^2 + w_2(b - b_0)^2 + w_3\left(r - r_{default}\right)^2$$

Robust circle fitting



1. Robust to outliers
2. Uses regressor output as initial guess and prior

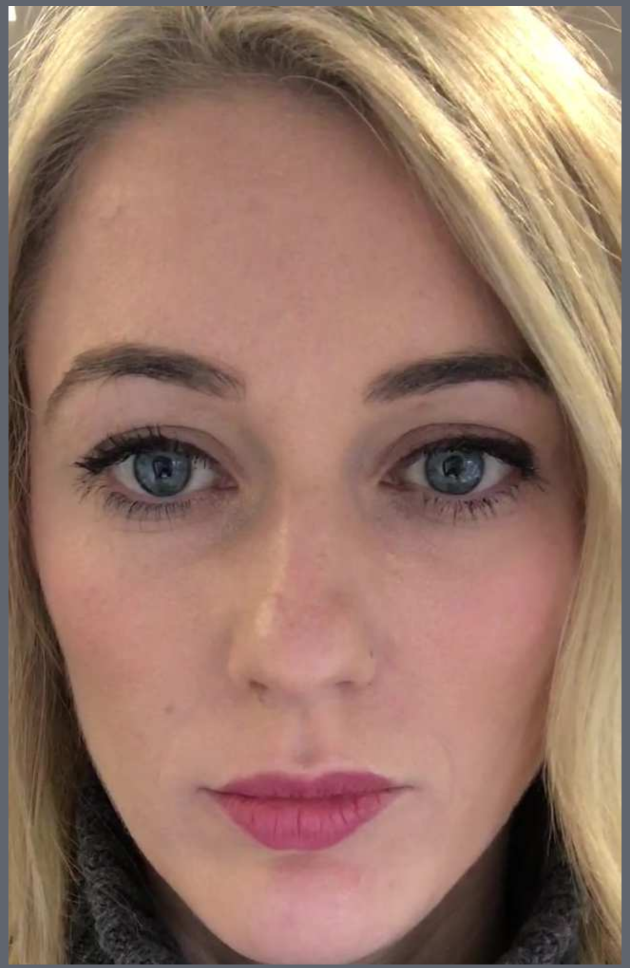
$$C_2 = w_1 \frac{1}{N}\sum_{i=1}^{N} \rho\left(\sqrt{(e_{ix}-a)^2 + (e_{iy}-b)^2} - r\right) +$$

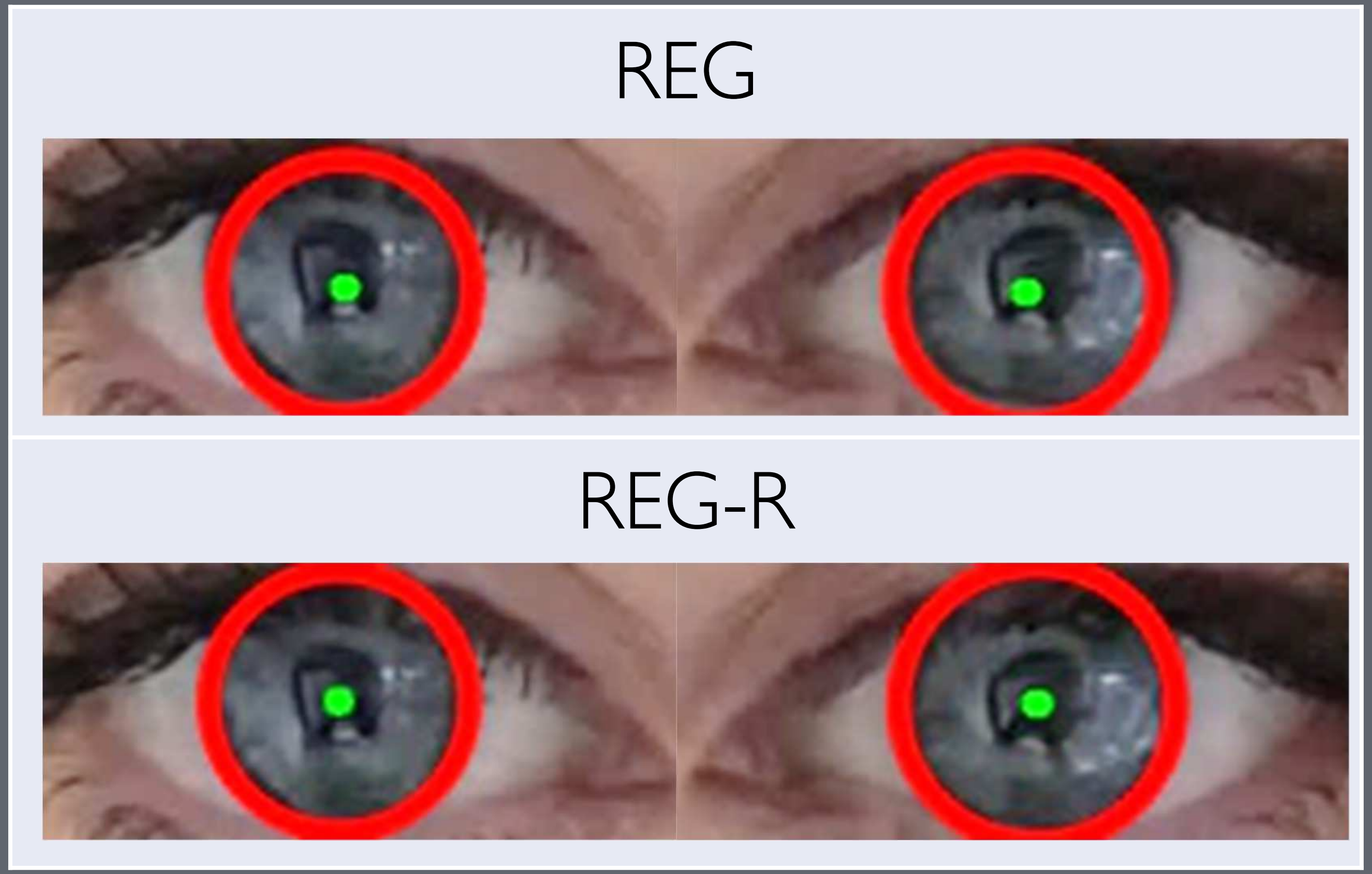
→ Robust circle fitting cost

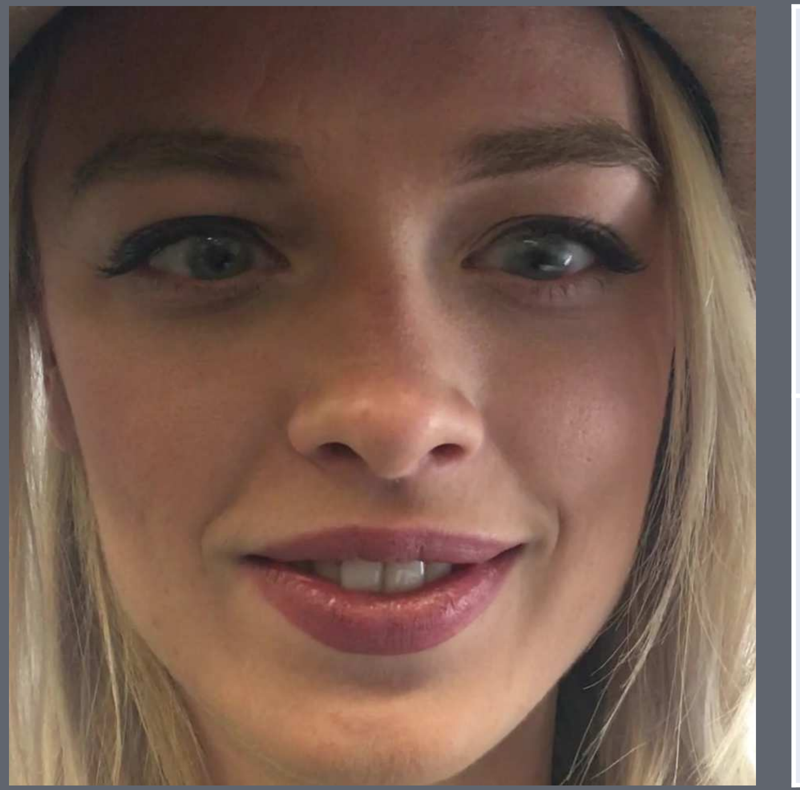
$$w_2(a-a_0)^2 + w_2(b-b_0)^2 + w_3(r - r_{default})^2$$

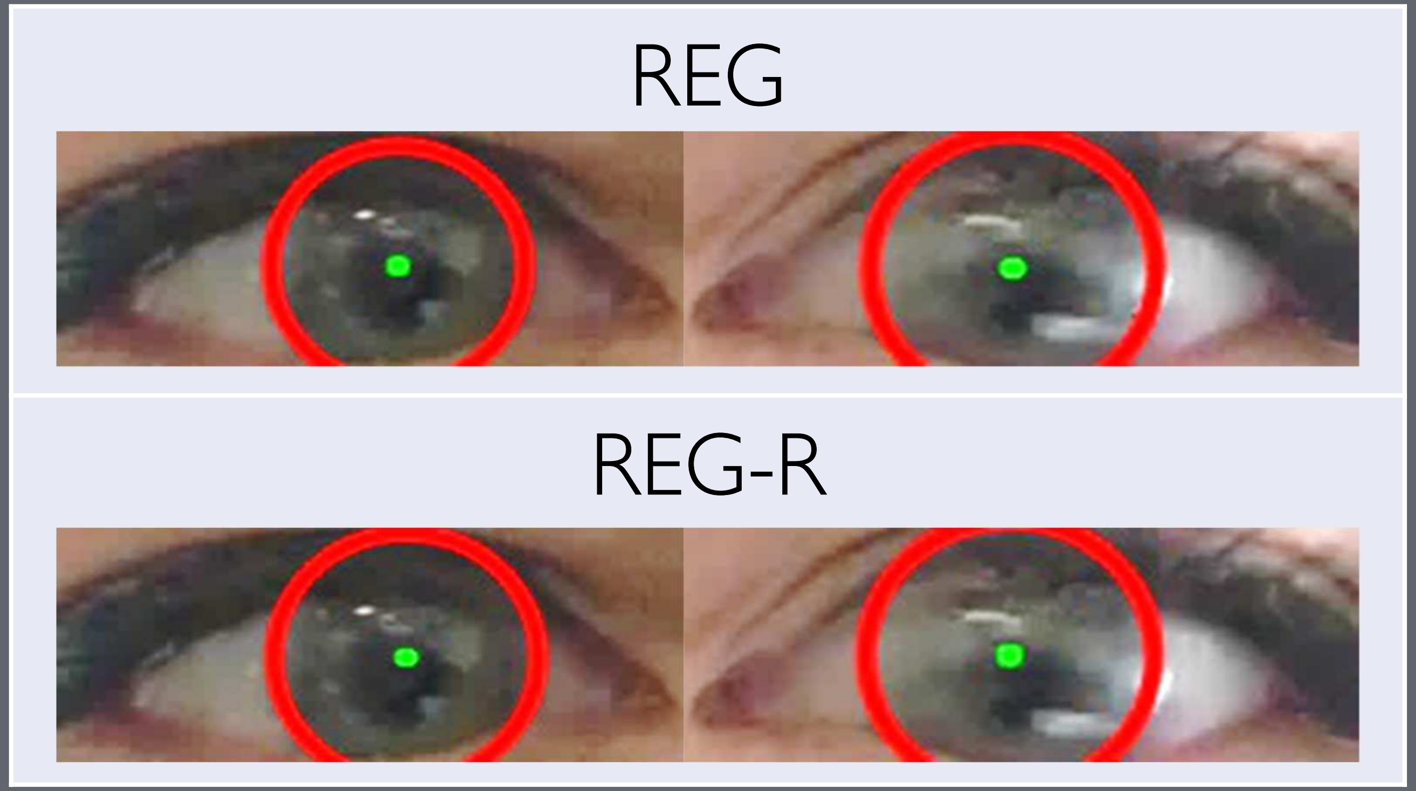
→ Prior terms

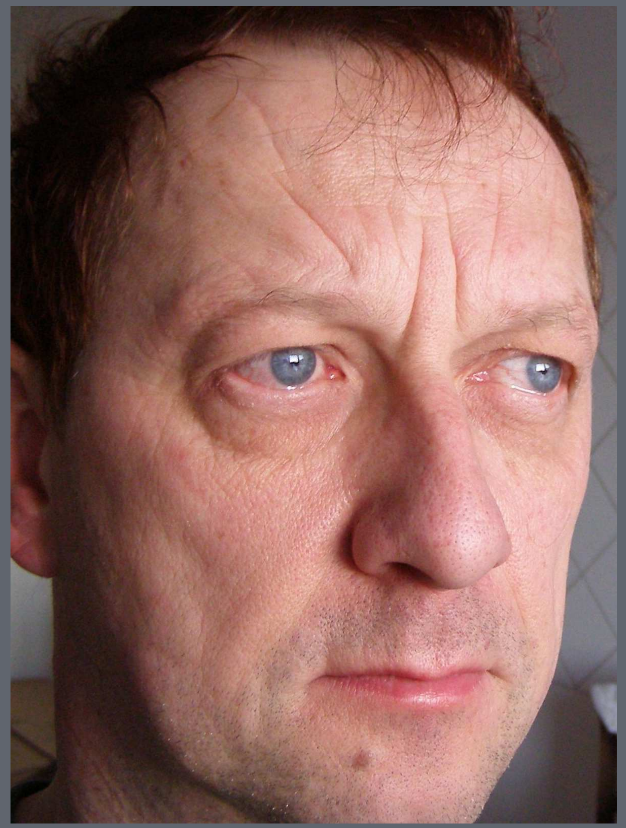
Results

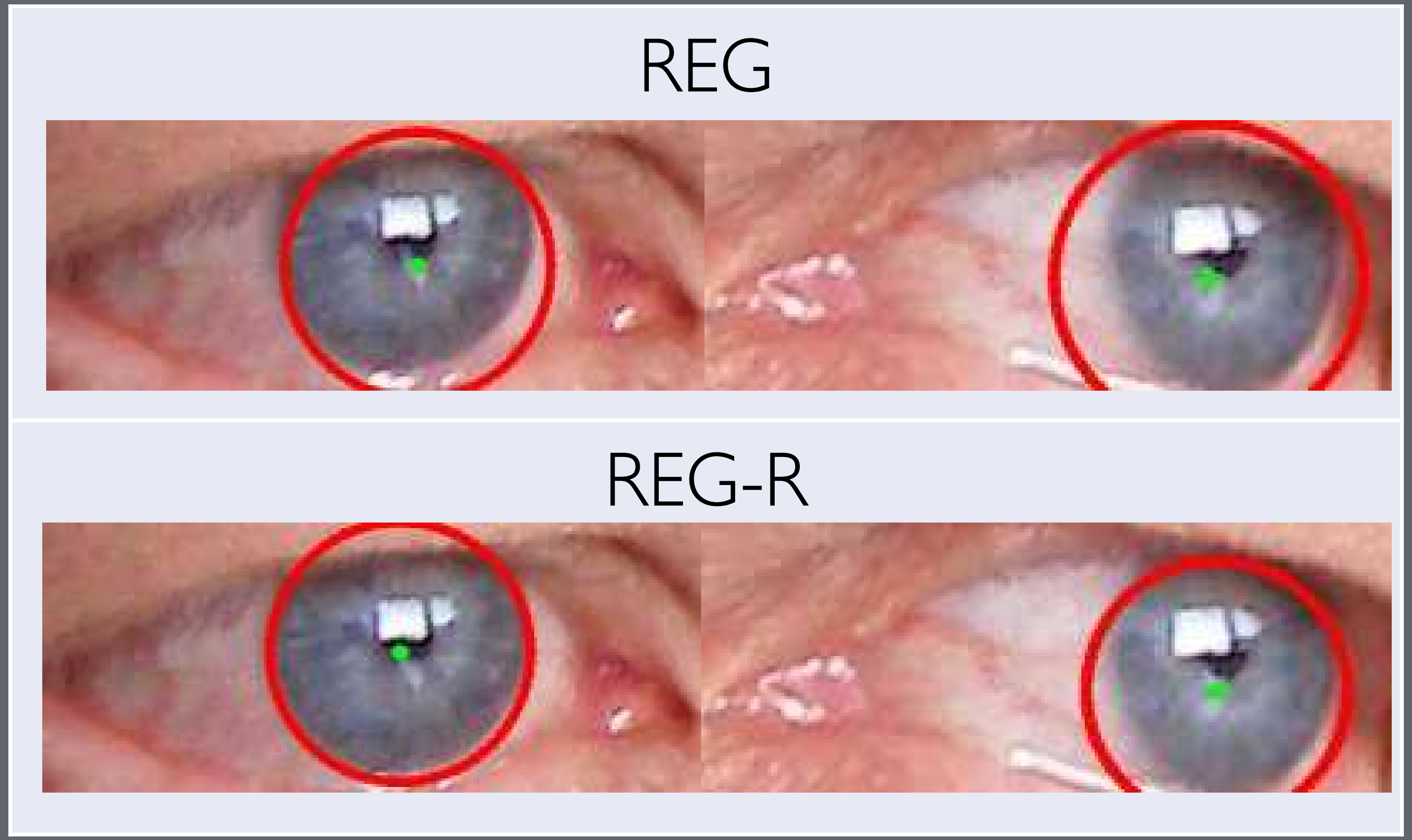








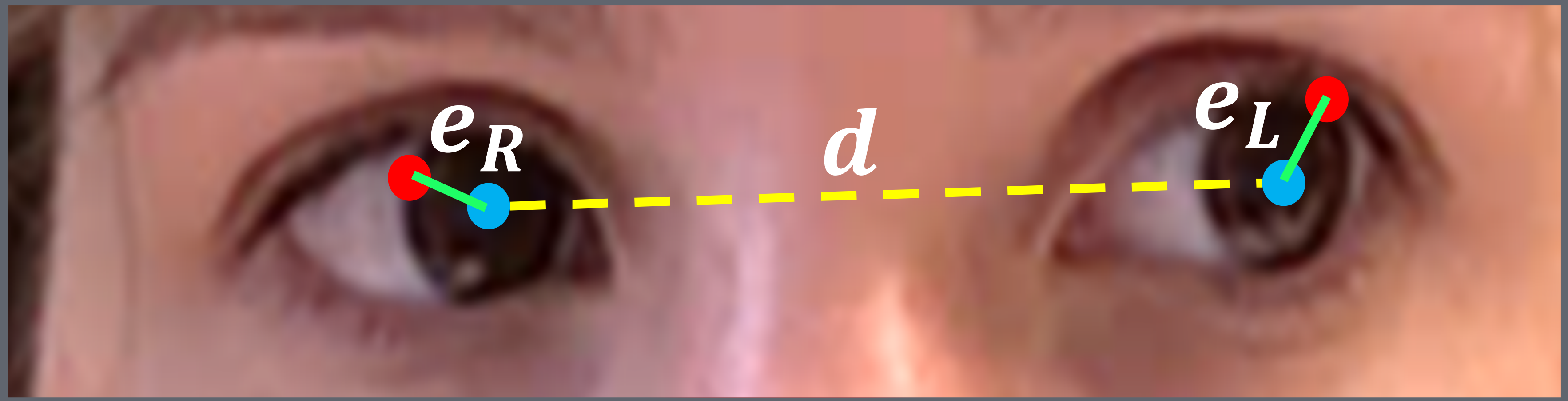




Results

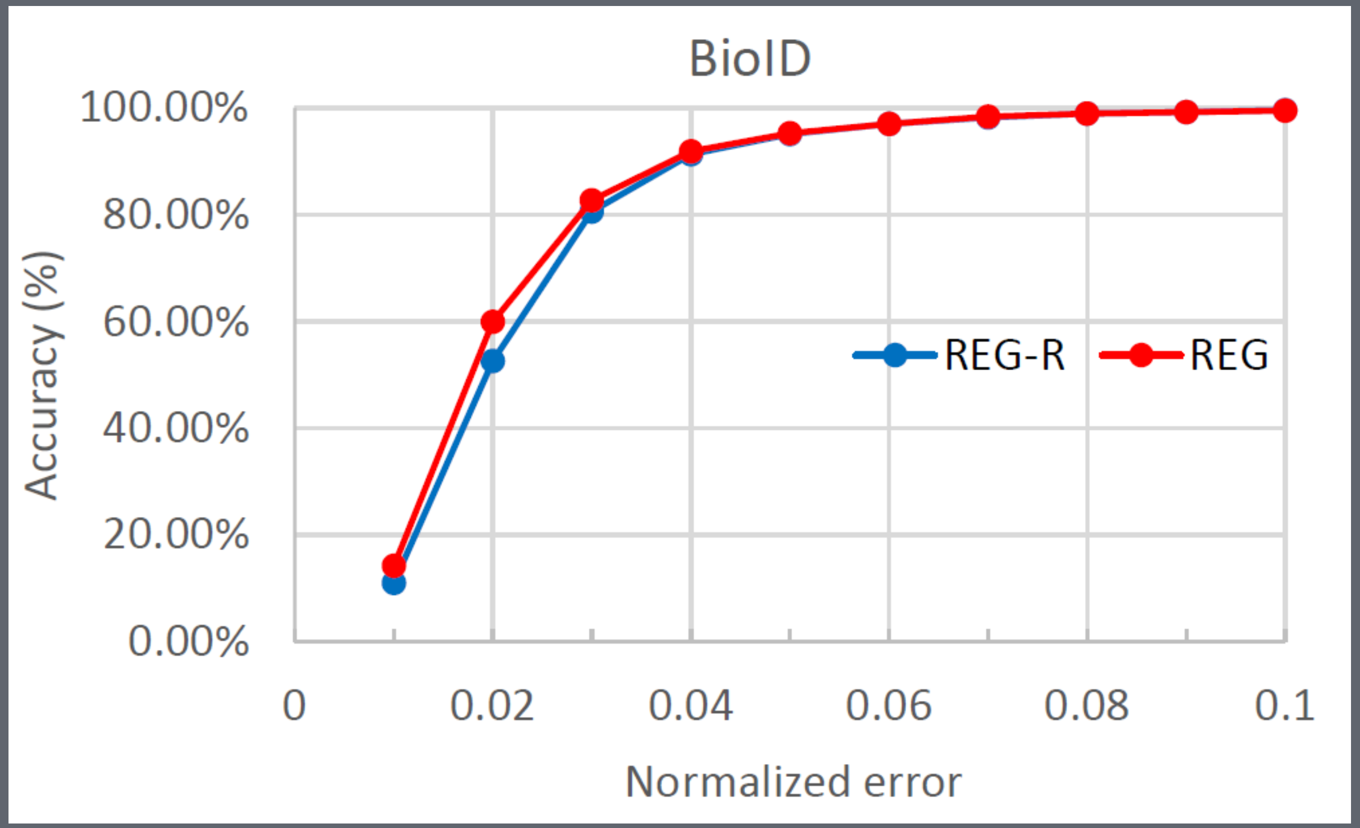
Within the
pupil iris

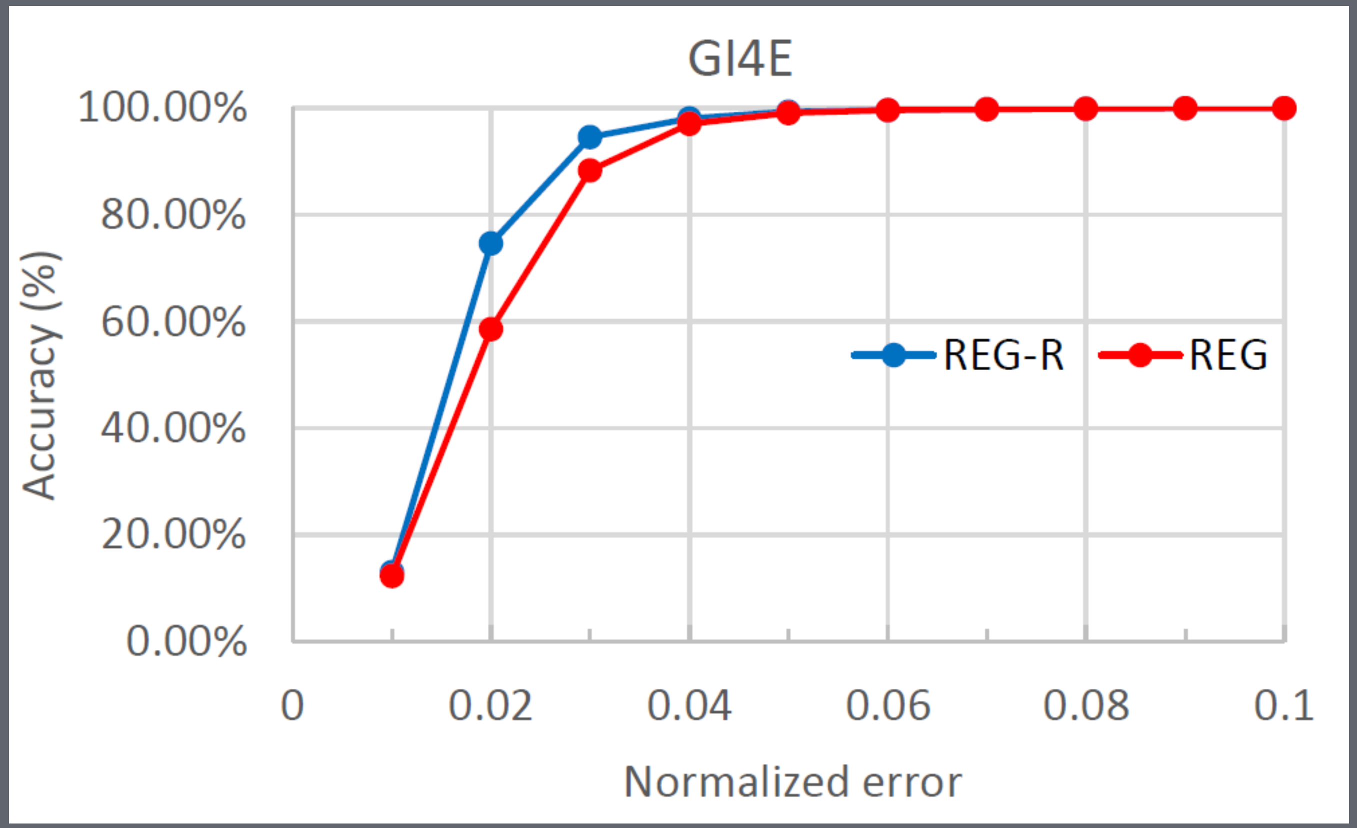
Method	$e < 0.025$	$e < 0.05$	$e < 0.1$	$e < 0.25$
BioID Dataset				
REG-R	68.13%	95.07%	99.59%	100%
REG	74.3%	95.27%	99.52%	100%
Timm [8]	38%*	82.5%	93.4%	98%
Valenti [9]	55%*	86.1%	91.7%	97.9%
Zhou [17]	50%*	93.8%	99.8%	99.9%
Ahuja [1]	NA	92.06%	97.96%	100%
Markuš [15]	61%*	89.9%	97.1%	99.7%
GI4E Dataset				
REG-R	88.34%	99.27%	99.92%	100%
REG	77.57%	99.03%	99.92%	100%
ELSE [3]	49.8%	91.5%	97.17%	99.51%
Anjith [4]	NA	89.28%	92.3%	NA
Talking Face Dataset				
REG-R	65.78%	95.68%	99.88%	99.98%
REG	18.7%	95.62%	99.88%	99.98%
ELSE [3]	59.26%	92%	98.98%	99.94%
Ahuja [1]	NA	94.78%	99%	99.42%

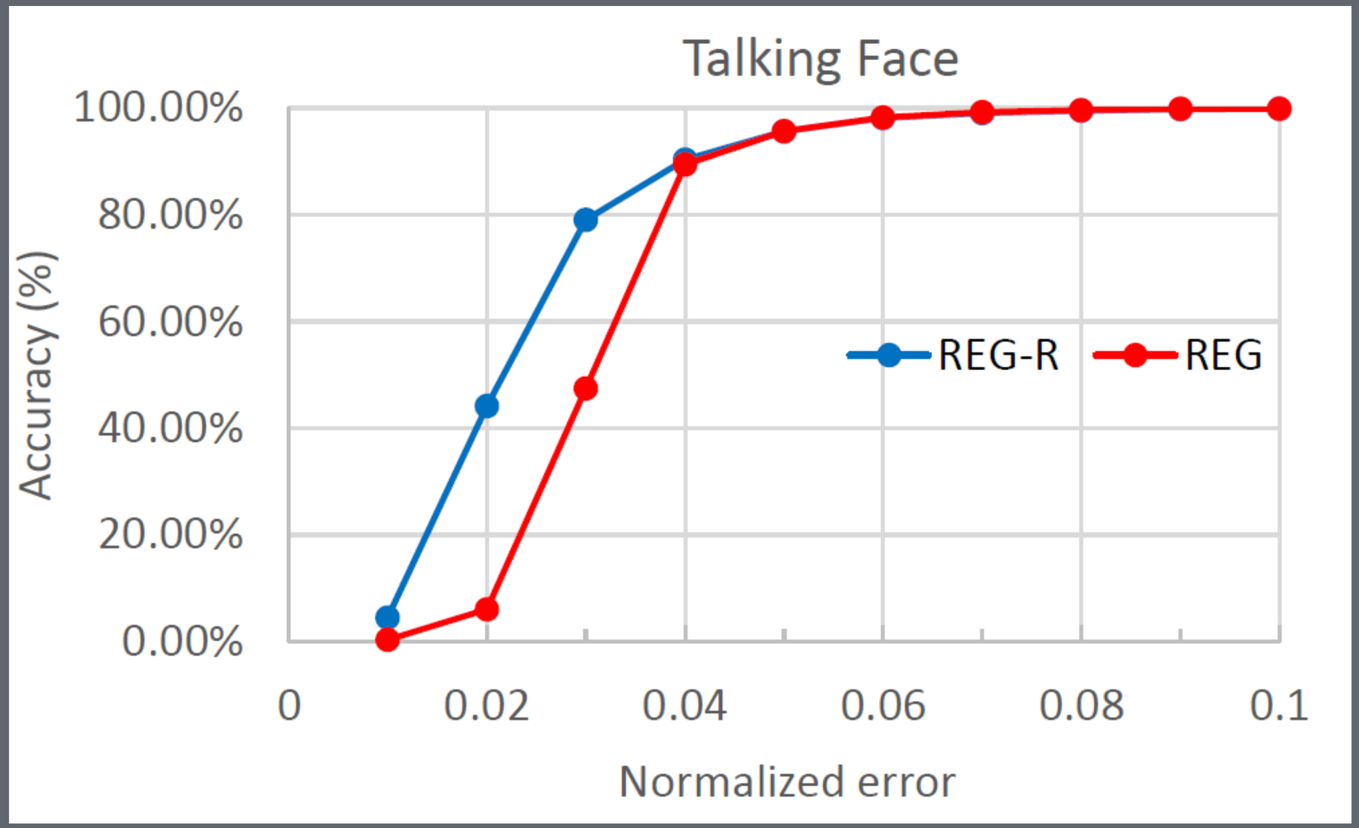


$$e = \frac{1}{d}\max\{e_R, e_L\}$$

Results







Increased fine-level accuracy due to circle refinement

Majority of methods are not even evaluated at $\boldsymbol{e < 0.025}$

Summary

1. Real-time iris detection system from standard images

2. Eye center localization using cascades of boosted regression trees with HoG features

3. Accurate iris localization using robust circle fitting

4. State-of-the-art performance on multiple datasets

Questions?