ETTAC 2020

1st Workshop on Eye Tracking Techniques, Applications and Challenges

https://vision.unipv.it/ettac2020/

10 January 2021

In conjunction with

ICPR 2020

Introduction

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

- Speaker: Gonzalo Garde Lecumberri
- Ph. D. Student in Public University of Navarre (UPNA), Spain
- Member of the Gaze Interaction for Everyone (GI4E) research group in the same university

Synthetic gaze data augmentation for improved user calibration

Gonzalo Garde, Andoni Larumber-Bergera, Sonia Porta, Rafael Cabeza, Arantxa Villanueva

Objectives

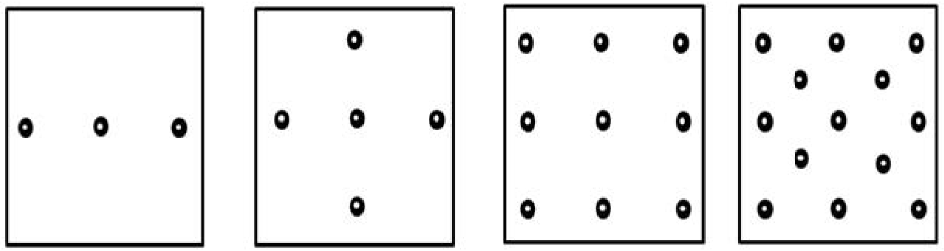
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

- To question the impact of the balance between generalization and personalization in Deep Learning-based Gaze Estimation solutions

- To study the effect of pretraining the model over different domains before training our system for our final application.

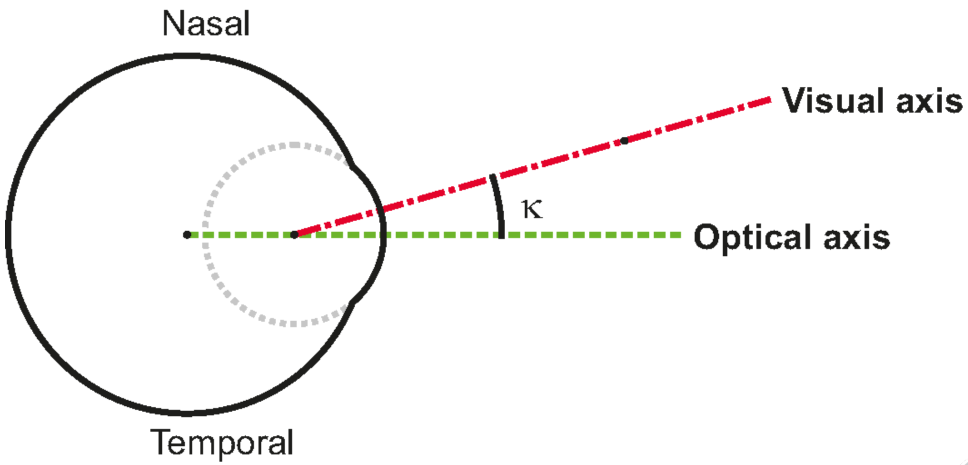
Research introduction: calibration

In Gaze Estimation, personalization is achieved through calibration



Examples of grid of calibration based on points

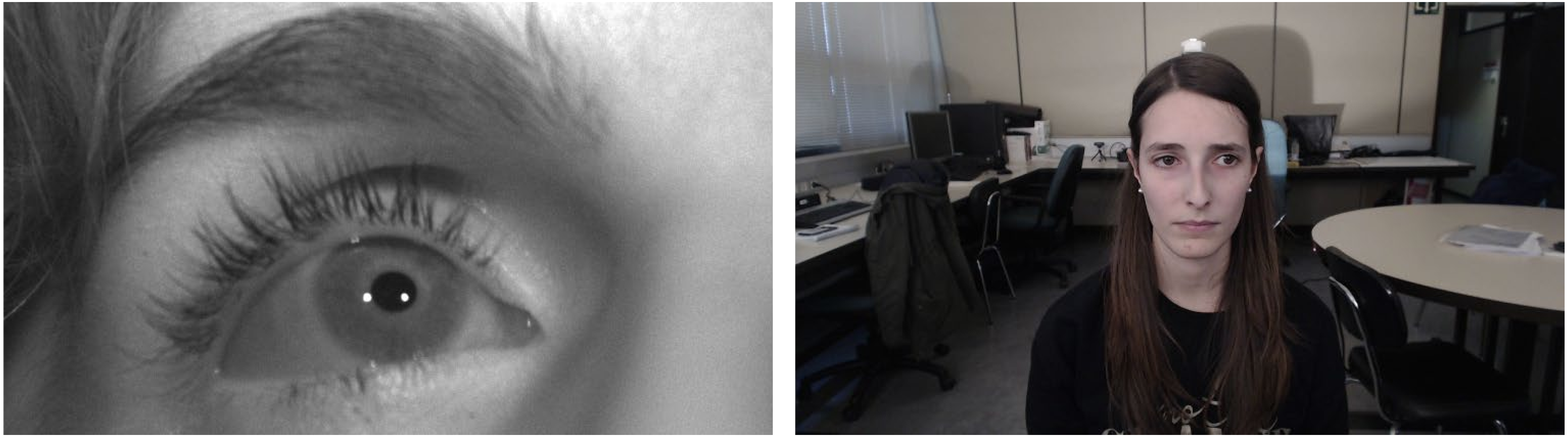
To calibrate: adjust system using individual's intrinsic aspects



Simplified model of the eye

Research introduction: high resolution – low resolution systems

upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa



High resolution vs low resolution

In both type of systems, calibration has probed to improve the obtained accuracy

Research introduction: size limitations

Requisite to implement a Deep Learning solution: large enough databases

If learning in a supervised manner…

Requisite to implement a Deep Learning solution: large enough databases correctly labelled

Two posible solutions:

- Using synthetic databases or enviroments
- Pretraining a model in a different topic and using it as starting point (“transfer-learning”)

Research introduction: transfer-learning

Basis: if a network learns to extract features from images to solve a task, part of these features can be handy to solve others image-related problems

The closer the original domain to the new one, the better the learning transference

Pretrained models in frameworks:

TensorFlow PyTorch

Research design: databases

3 databases were needed, 2 for the pretraining of the model and 1 as target.

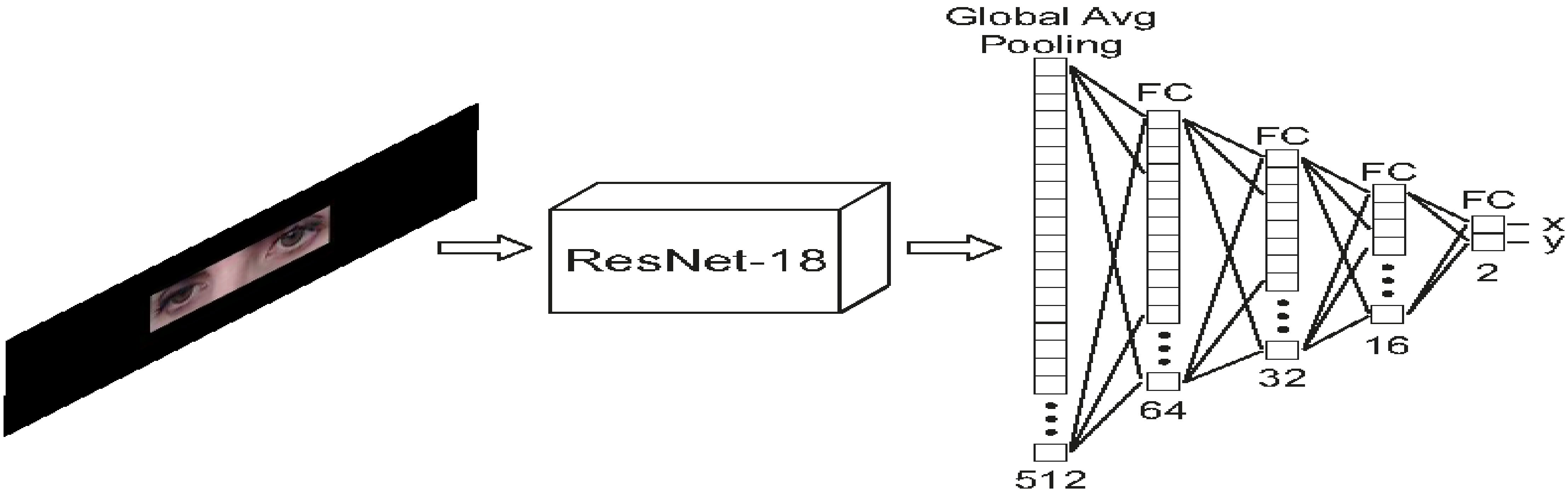
IMAGENET	U2Eyes	I2Head
General-purpose database	Close domain database	Target database

They were not used in this project directly. Our starting models were pre-trained over them

Research design: images preprocessing

Image preprocessing

Research design: network architecture



Network architecture

Research design: training

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

The training process could be divided in two parts:

Starting point → 12 trainings

Select among two models*:

- Model pretrained with U2Eyes
- Model with Resnet-18 weights from Imagenet

This part allow us to study the effect of transfer learning in different domains

*https://github.com/GonzaloGardeL/Synthetic-gaze-data-augmentation-for-improved-user-calibration

Research design: training

The training process could be divided in two parts:

Starting point → 12 trainings

The model is trained over images from I2Head. In these 12 variations, we change the nature of the training dataset from only calibration to calibration + generalization

With this block, we study the balance between generalization and personalization

Research design: training

Model	Train # Users/Images	Calibration # Users/Images	Total Train # Images	Test # Users/Images
ImageNet	K/K*130	1/34	K*130+34	1/130
U2Eyes	K/K*130	1/34	K*130+34	1/130

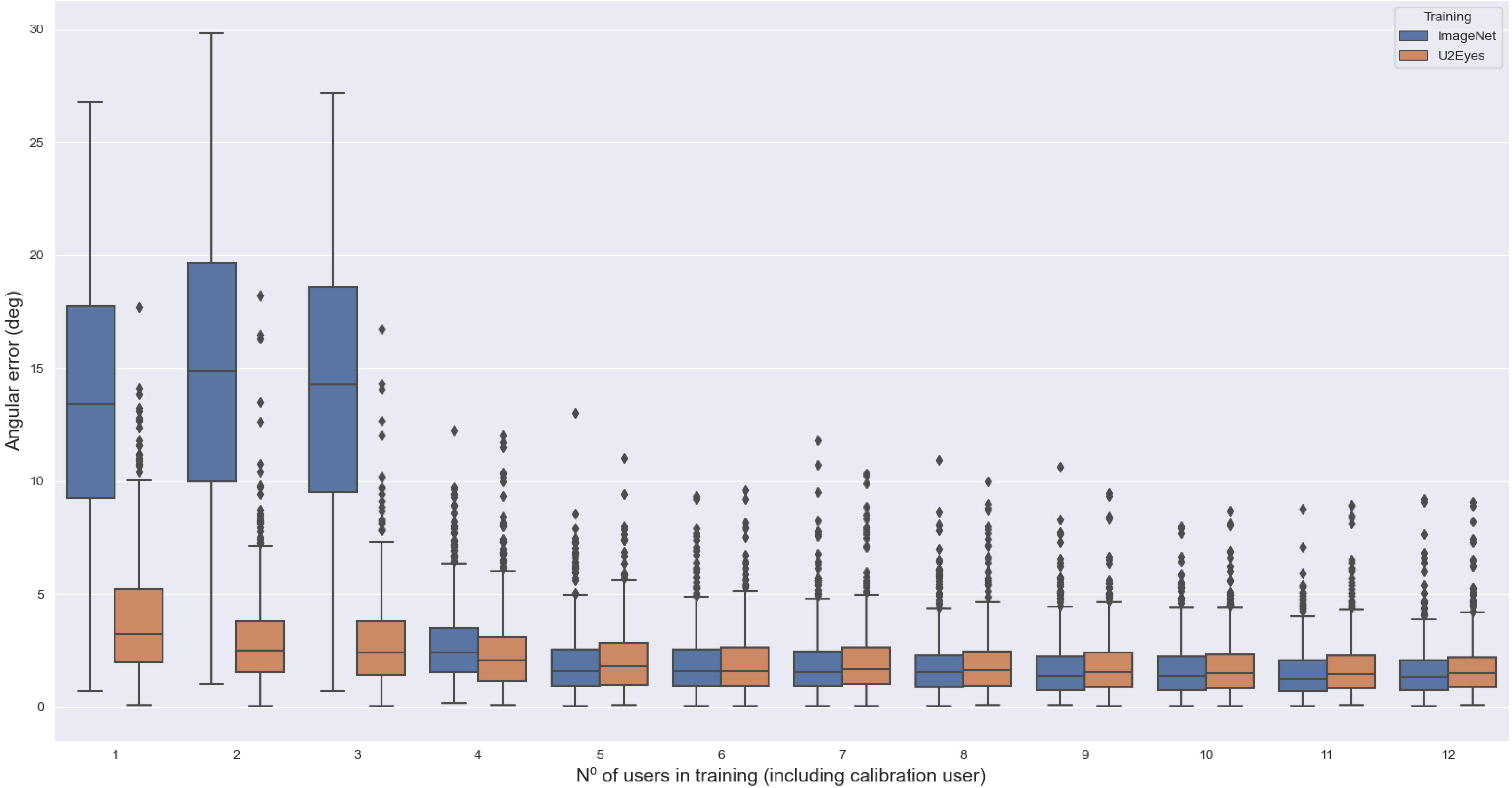
K varies from 0 (only calibration images in the training dataset) to 11 (calibration images + 11 additional users)

- Trained over 240 epochs
- Adam optimizer
- Euclidean distance of the real look-at-point and the estimated look-at-point

A learning rate Schedule based in Cyclic learning rate was used, although other could be used

Research results

upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa



Research results

Users in training	Mean (°)		Median (°)	
	Imagenet	U2Eyes	Imagenet	U2Eyes
1	13.615	**3.891**	13.401	**3.243**
2	**14.812**	2.967	**14.880**	2.522
3	14.069	2.861	14.255	2.404
4	2.867	2.488	2.567	2.149
5	2.004	2.149	1.631	1.867
6	2.028	1.965	1.675	1.667
7	1.987	2.039	1.617	1.746
8	1.877	1.968	1.639	1.724
9	1.758	1.860	1.471	1.611
10	1.681	1.818	1.412	1.588
11	1.615	1.777	**1.334**	**1.485**
12	**1.559**	**1.714**	1.344	1.486

Research results: discusion

From the obtained results, we highlight two cases to study:

- Correlation number of training images – accuracy
- Comparison of the results obtained for each of the pretraining methods

Research results: discusion

upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

From the obtained results, we highlight two cases to study:

- Correlation number of training images – accuracy

Users in training	Mean (°)		Median (°)	
	Imagenet	U2Eyes	Imagenet	U2Eyes
1	13.615	**3.891**	13.401	**3.243**
2	**14.812**	2.967	**14.880**	2.522
3	14.069	2.861	14.255	2.404
4	2.867	2.488	2.567	2.149
5	2.004	2.149	1.631	1.867
6	2.028	1.965	1.675	1.667
7	1.987	2.039	1.617	1.746
8	1.877	1.968	1.639	1.724
9	1.758	1.860	1.471	1.611
10	1.681	1.818	1.412	1.588
11	1.615	1.777	**1.334**	**1.485**
12	**1.559**	**1.714**	1.344	1.486

As the number of users decreases, the accuracy tend to worse

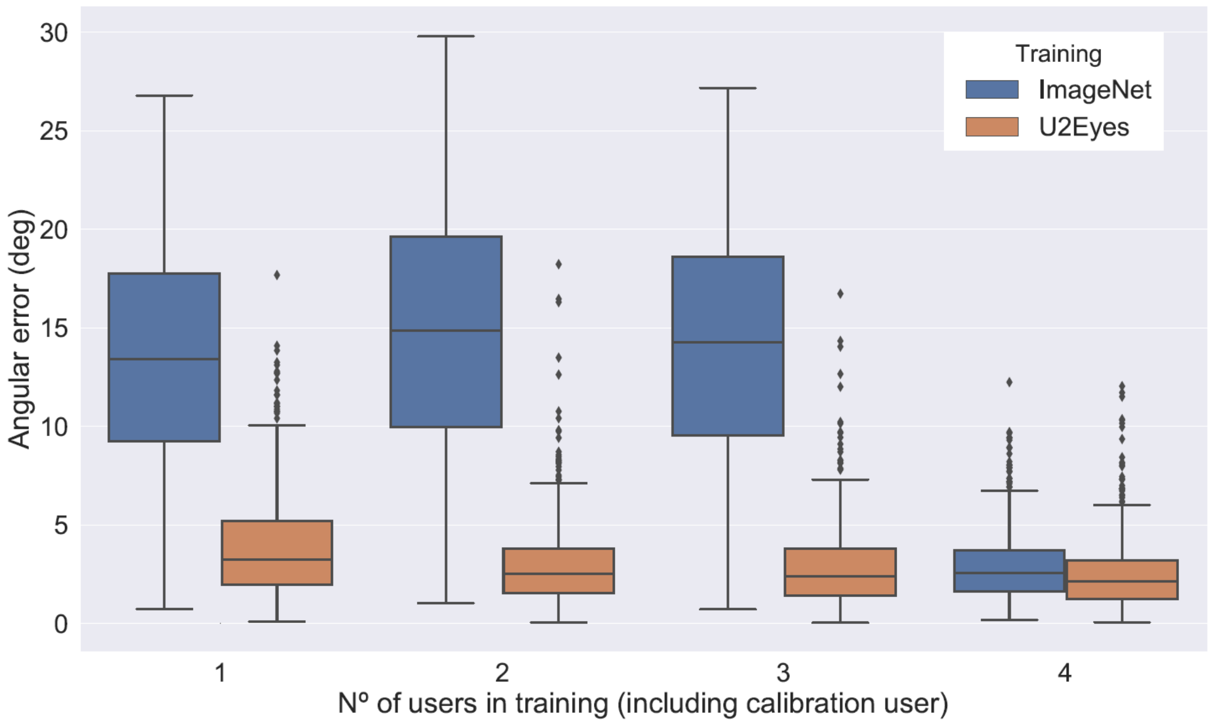
For 12 users, accuracy close to high-resolution systems is obtained (**1.5º**)

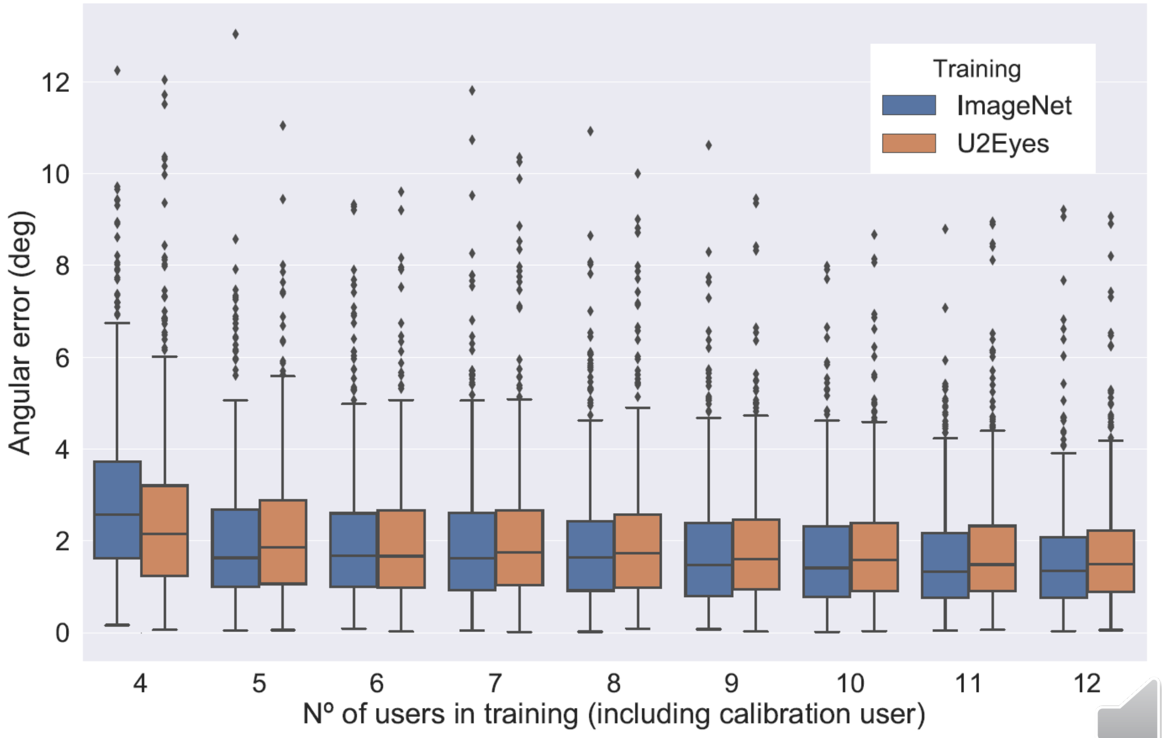
Research results: discusion

upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

From the obtained results, we highlight two cases to study:

- Comparison of the results obtained for each of the pretraining methods



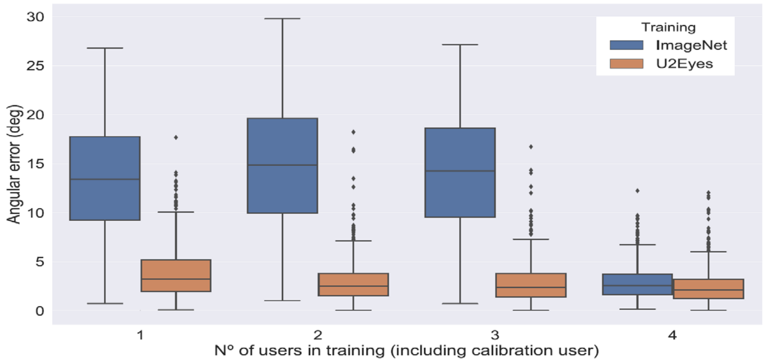


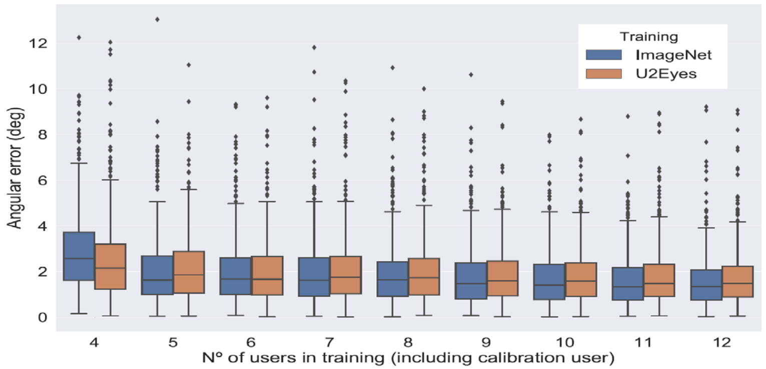
Research results: discusion

upna
Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

From the obtained results, we highlight two cases to study:

- Comparison of the results obtained for each of the pretraining methods





Hypothesis:

- If training dataset is long enough for the Imagenet model to learn, it benefits from real-world variability
- If not, the U2Eyes model is more robust because, as it was trained in a similar domain, is more capable of continuing learning

Conclusions

From the obtained results, we conclude that:

- Results close to high-resolution systems can be obtained (1.5º) by using calibration in low-resolution systems based on Deep Learning
- The importance of providing domain images during the training process has been confirmed, and also the benefits of pre-training the network in a closer domain instead of in a more general dataset, to compensate the lack of useful gaze data images that characterize gaze estimation problems

Universidad Pública de Navarra
Nafarroako Unibertsitate Publikoa

Thank you!

gonzalogarde3@gmail.com | gonzalo.garde@unavarra.es