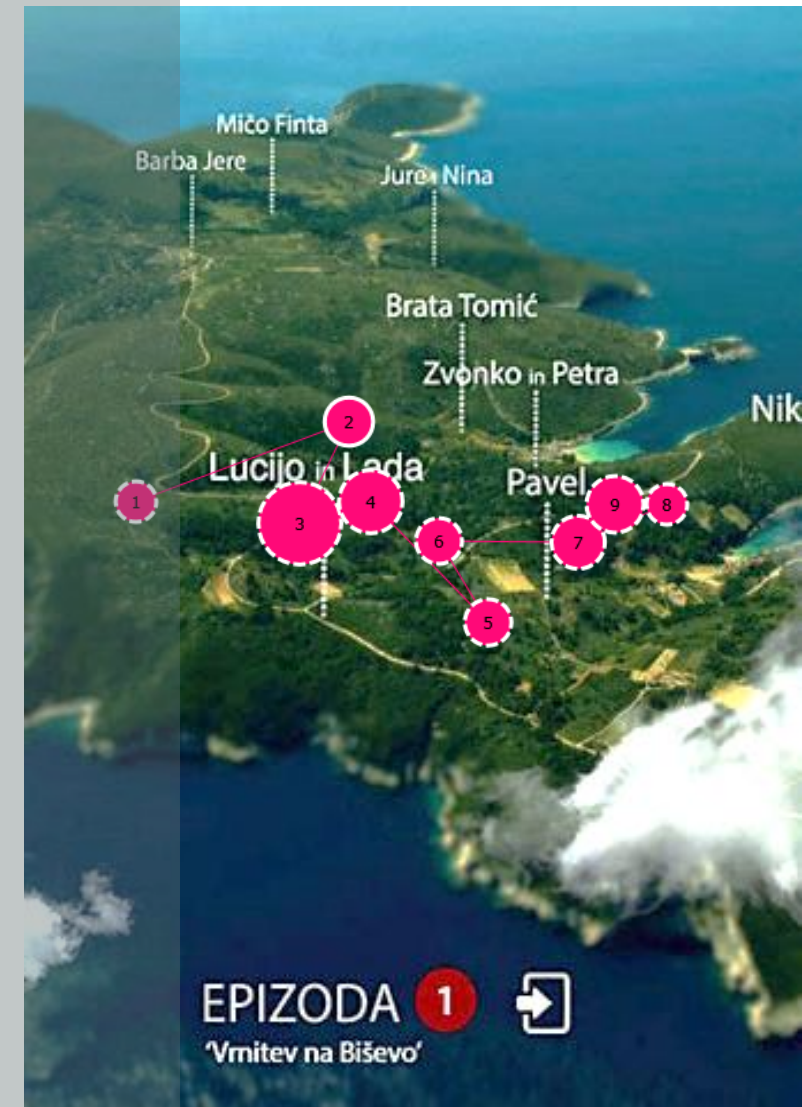


Gaze2AOI: Harnessing Deep-Neural Networks for Area of Interest Annotation in Eye Tracking Research



Presenter: Karolina Trajkovska

Mentors: dr. Klen Čopič Pucihar, dr. Matjaž Kljun



OVERVIEW

EYE-TRACKING AND AOI

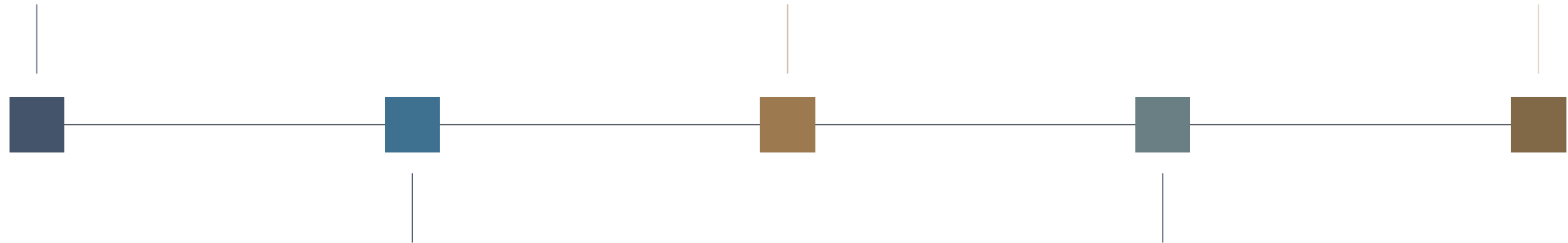
Background

GAZE2AOI

System Architecture,
Functionalities, Interface

FUTURE WORK

Space for Improvement
and Limitations



AUTOMATION OF AOI

Related Work and
Gap in Literature

APPLICATION & IOTOK

Using Gaze2AOI for
Interactive Documentaries

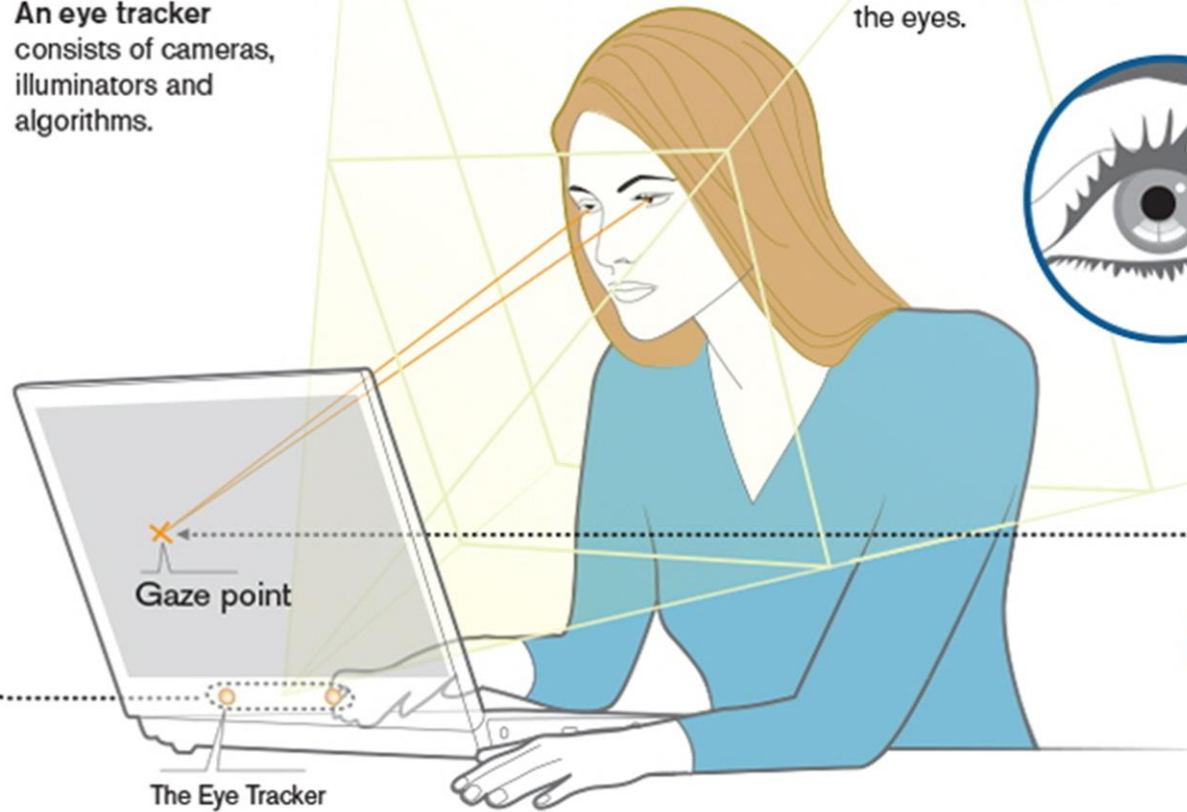
1 An eye tracker consists of cameras, illuminators and algorithms.

2 The illuminators create a pattern of near-infrared light on the eyes.

3 The cameras take high-resolution images of the user's eyes and the patterns.

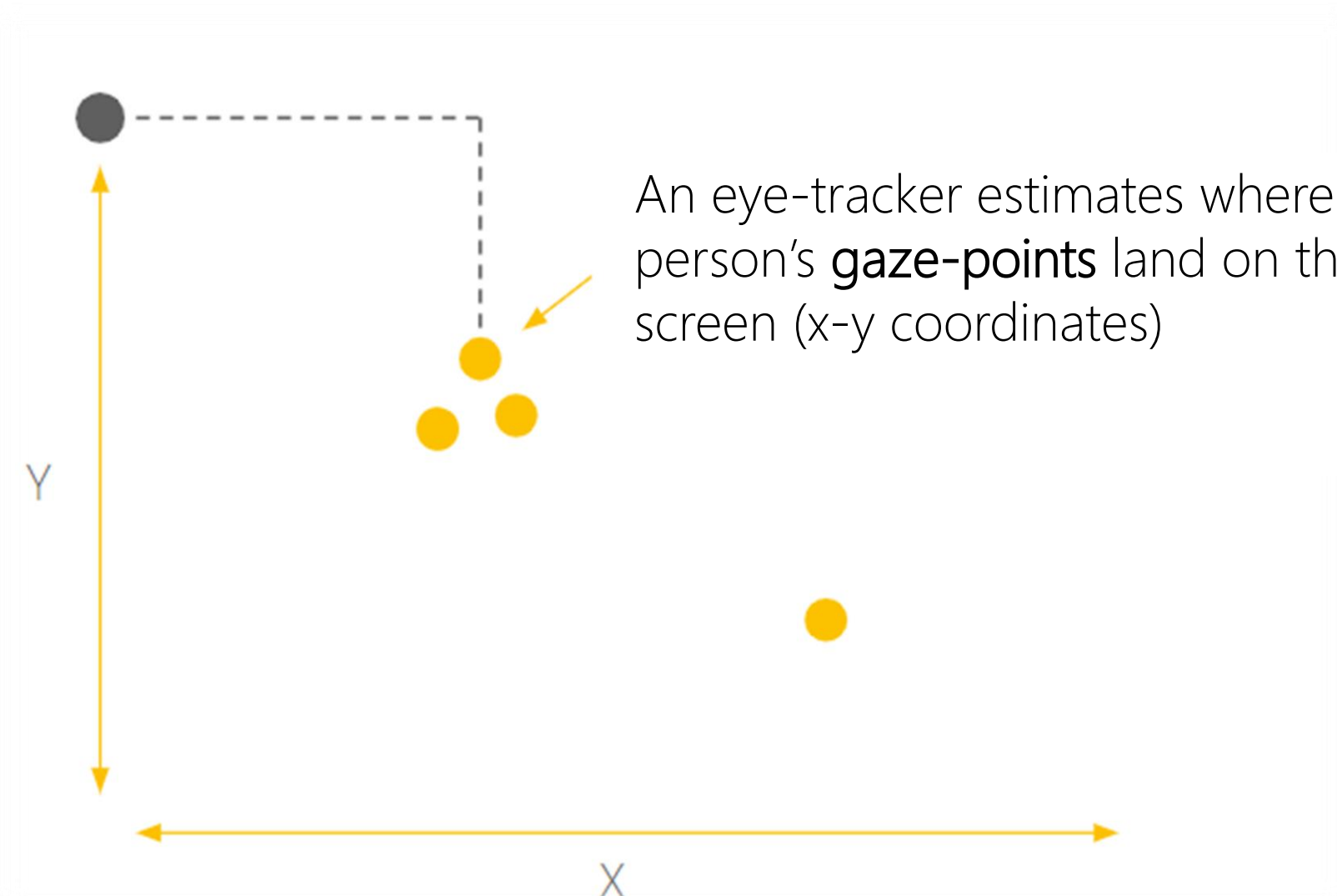
4 The image processing algorithms find specific details in the user's eyes and reflections patterns.

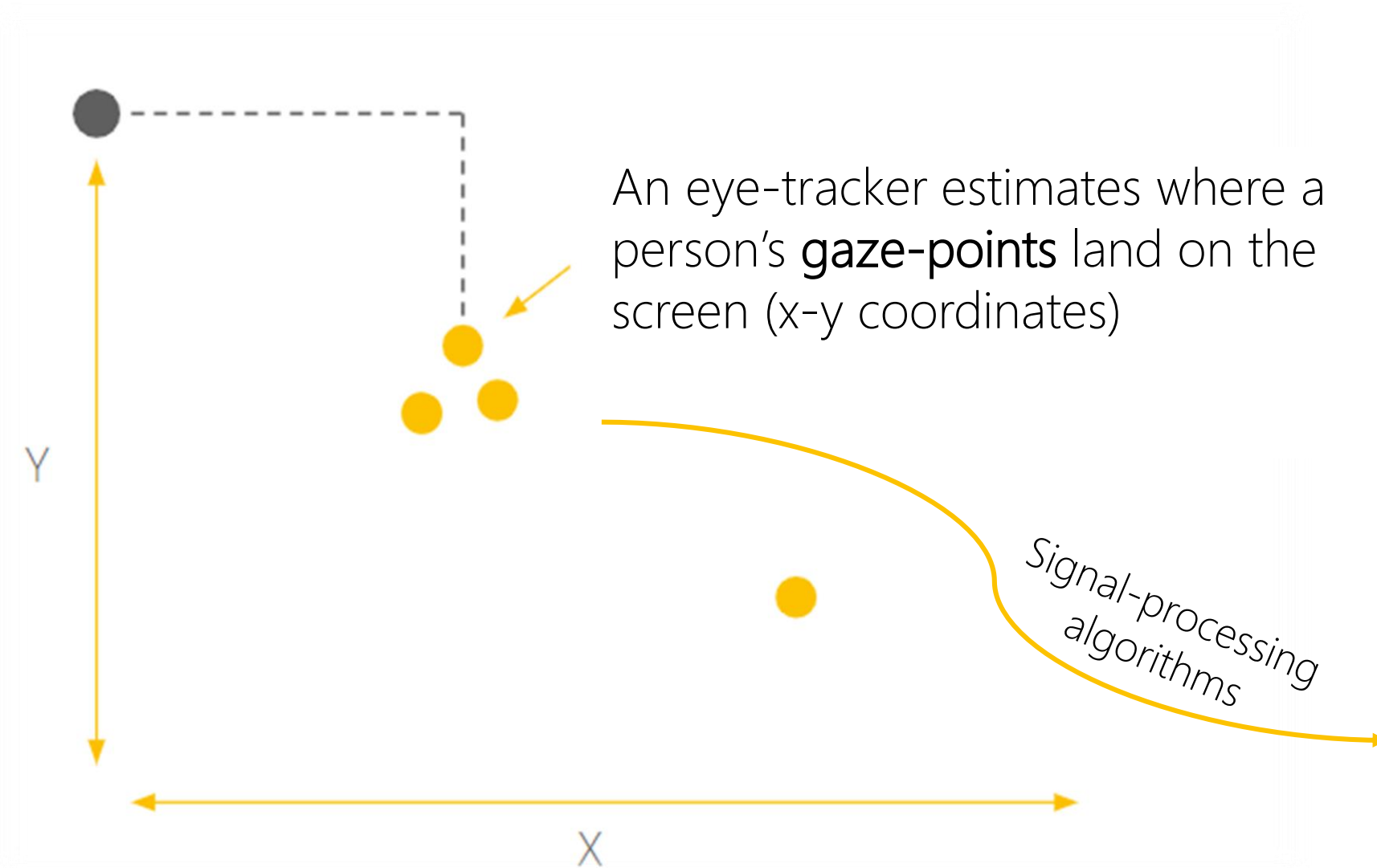
5 Based on these details the eyes' position and gaze point are calculated, for instance on a computer monitor, using a sophisticated 3D eye model algorithm.



WHAT IS EYE TRACKING?

Source: www.tobiipro.com/learn-and-support

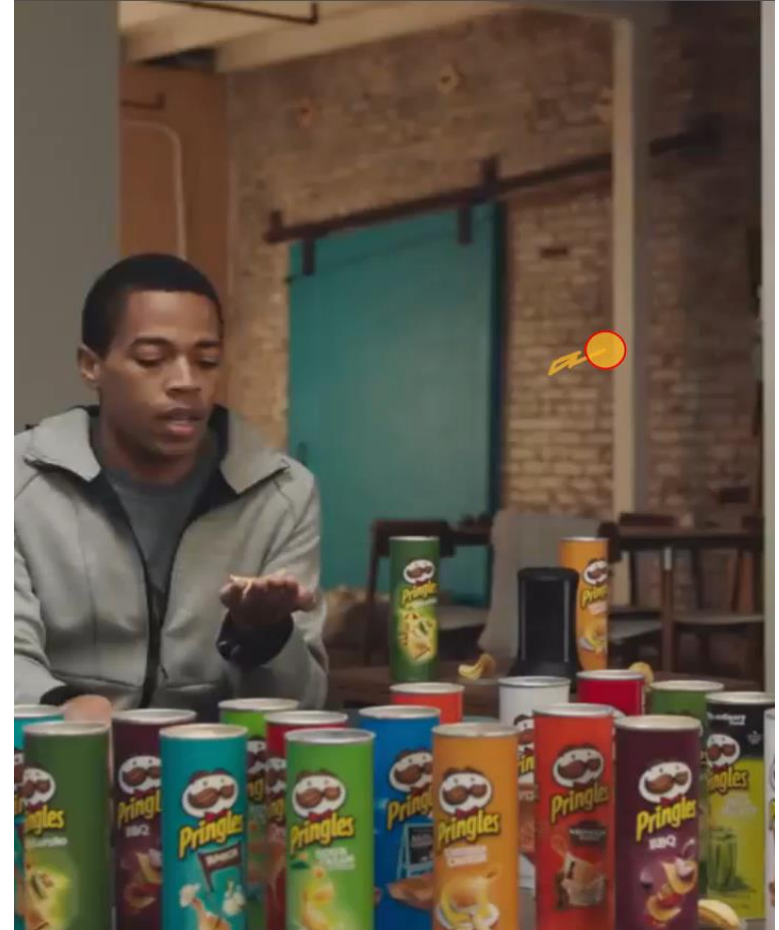




A fixation

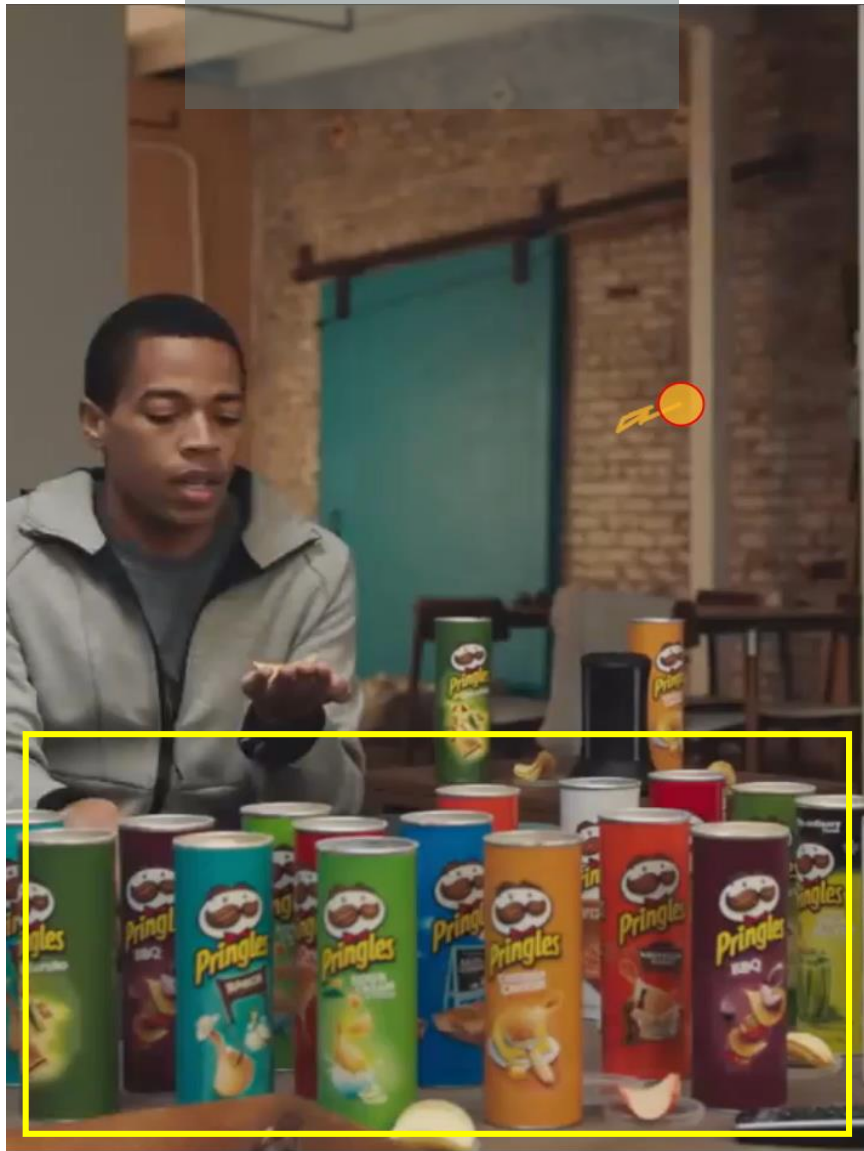
WHY EYE-TRACKING DATA?

- What part of the visual information was noticed first?
- Was an important part of the visual information overlooked?
- What part of the visual information was looked at longest?
- Etc.



AREA OF INTEREST(AOI)

- Give **semantic meaning** to the gaze data collected
- Manual annotation and analysis is **time consuming**



Neutral products

TTF:	3.8s
Time spent:	1.7s
Ratio:	4/4

UGENS HUG

Neutral info

TTF:	6.4s
Time spent:	1.1s
Ratio:	4/4

UGENS HUG

Toothpaste products

TTF:	4.4s
Time spent:	4.6s
Ratio:	4/4

Neutral price

TTF:	10.6s
Time spent:	0.4s
Ratio:	3/4

Toothpaste info

TTF:	9.8s
Time spent:	0.8s
Ratio:	3/4

Toothpaste price

TTF:	10.7s
Time spent:	0.5s
Ratio:	4/4

1
2
3
4
5
6

Frit valg

10,-

19,-

AUTOMATED AOI

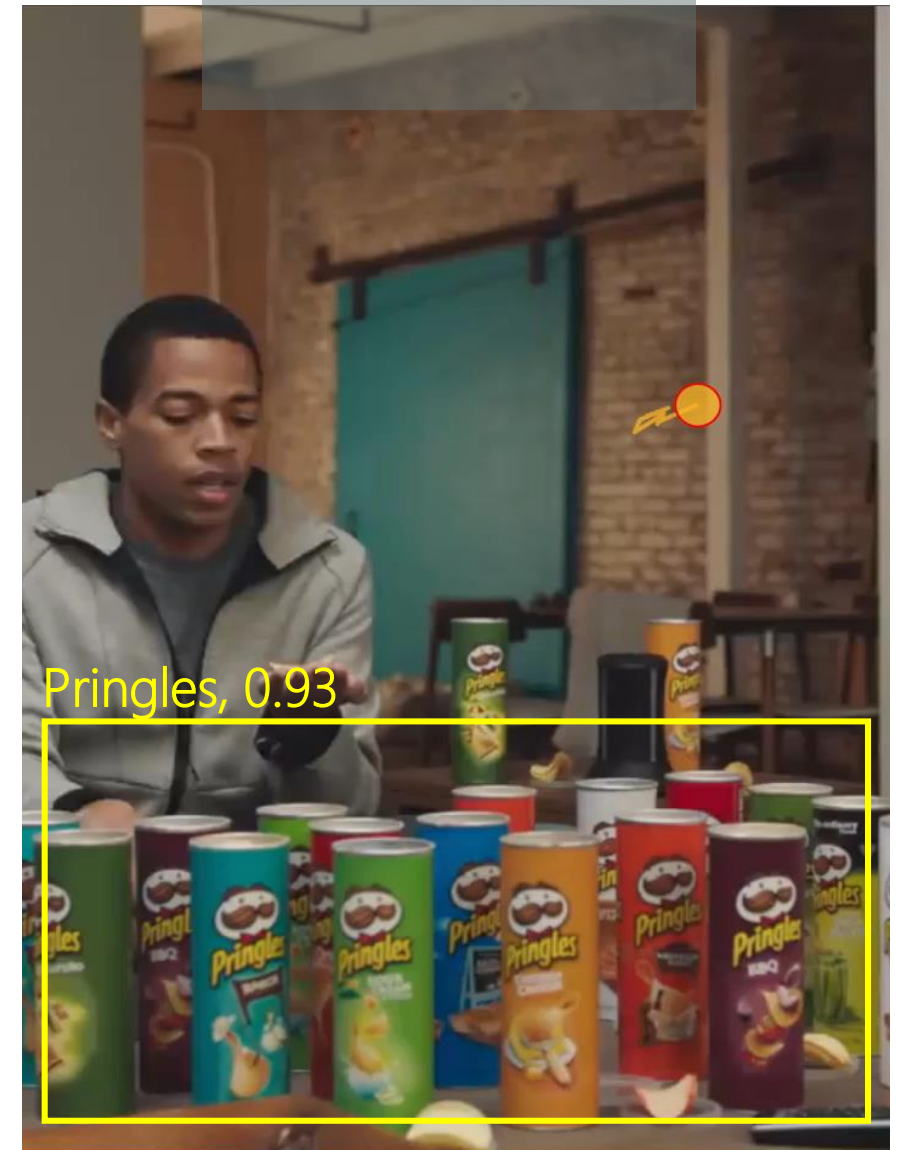
- Use of Object Detection, Segmentation or Image Classification to **automatically define the AOI** and check whether an AOI is fixated or not
- DNNs era sparked this area of research

AOI Bounding Box:
(x, y, w, h)

Object Detection algorithm

Fixation Point:
(x, y)

Software like iMotions



RELATED WORK

Exploring and comparing methods and algorithms [2, 6, 7, 8, 9, 10]

Image Processing techniques (ex. Canny edge detection, Haar Features)

Region based Convolution Neural Network (R-CNNs)

Fast Segmentation Convolutional Neural Networks (Fast – SCNN)

You Only Look Once (YOLO) v2, v3, v4

Implemented algorithm and applied to specific research

- Analyzing students' attention [12]
- Artistic heuristics for face detection: tracking gaze when looking at faces [3]
- Augmented Reality warning systems for onsite construction workers [11]

Publicly available tools

VLEYE: A Complete System for Analysis of Video Lecture Based on Eye Tracking (2018) [1]

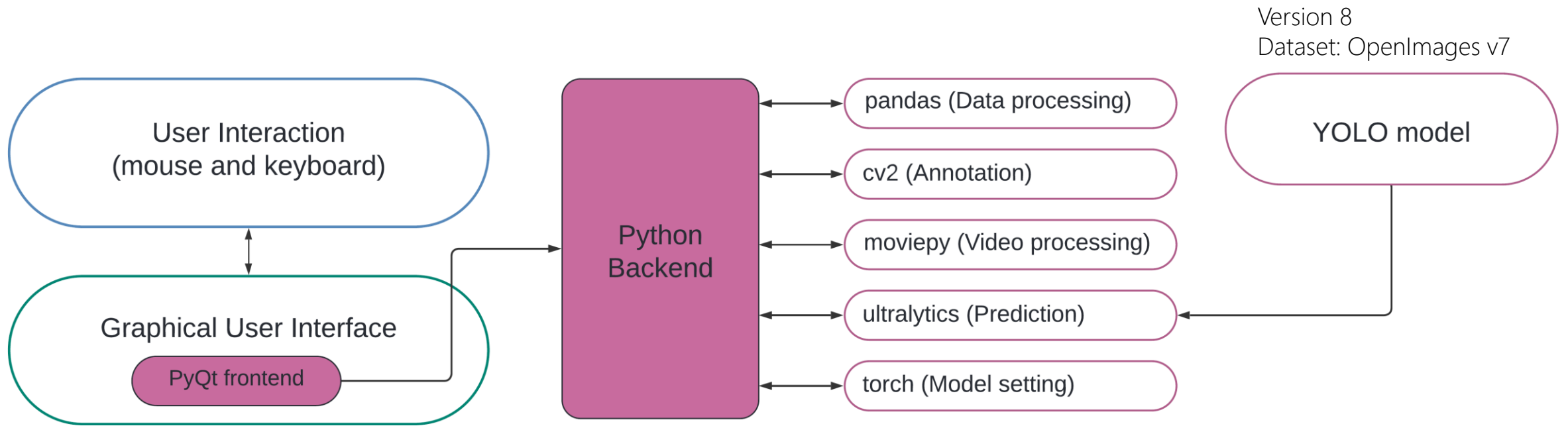
- Open-source graphical user interface with many features, but *automatic AOI only for faces (using Haar Features)*

EyeNotate Tool: Interactive Fixation-to-AOI Mapping for Mobile Eye Tracking Data based on Few-Shot Image Classification (2023) [13]

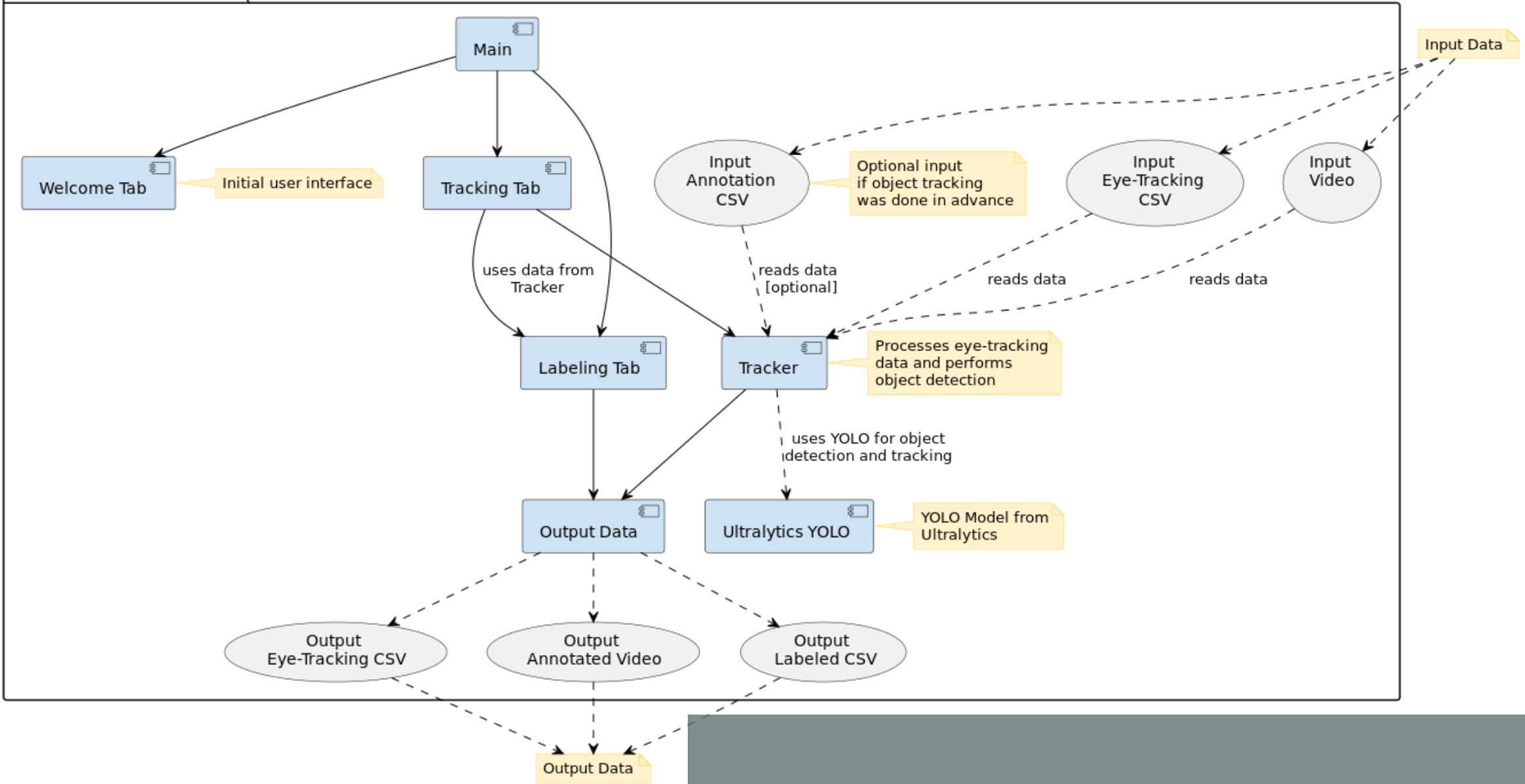
- Web application for automatically AOI detection that cuts the patch around the fixation point and runs Image Classification (*lost information on what is NOT seen*)

Gaze2AOI

- Automatic AOI for 600 different object classes for the whole image / frame
- Additional customized labeling
- Calculation of metrics like time to first fixation, AOI revisits, total time per AOI
- Open source



SYSTEM ARCHITECTURE



UML COMPONENT DIAGRAM

INTERFACE AND FUNCTIONALITIES

File START TRACKING

Home
Tracking
Labeling

- Select All
- Cabbage
- Cabinetry
- Cake
- Cake stand
- Calculator
- Camel
- Camera
- Can opener
- Canary
- Candle
- Candy
- Cannon
- Canoe
- Cantaloupe
- Car
- Carnivore
- Carrot
- Cart
- Cassette deck
- Castle
- Cat
- Cat furniture
- Caterpillar
- Cattle
- Ceiling fan
- Cello
- Centipede
- Chainsaw
- Chair
- Cheese
- Cheetah
- Chest of drawers
- Chicken

All
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z

00:00

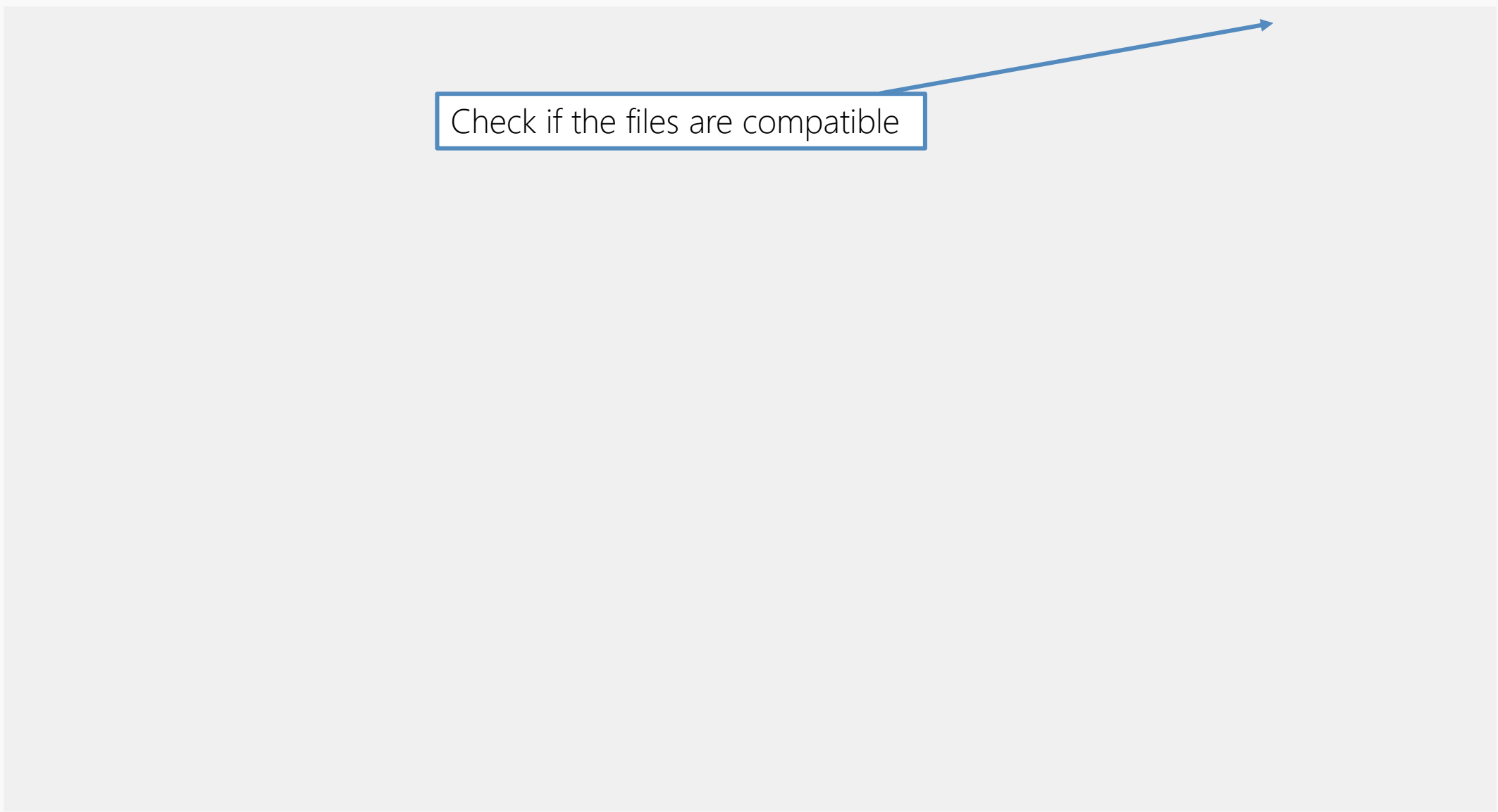
00:00

Select CSV file of eye tracking data and the according video of the same subject.

If predictions were run in advance and now you want to annotate for specific classes, select CSV file with annotations (predictions)

- Select All
- Cabbage
- Cabinetry
- Cake
- Cake stand
- Calculator
- Camel
- Camera
- Can opener
- Canary
- Candle
- Candy
- Cannon
- Canoe
- Cantaloupe
- Car
- Carnivore
- Carrot
- Cart
- Cassette deck
- Castle
- Cat
- Cat furniture
- Caterpillar
- Cattle
- Ceiling fan
- Cello
- Centipede
- Chainsaw
- Chair
- Cheese
- Cheetah
- Chest of drawers
- Chicken

- All
- A
- B
- C
- D
- E
- F
- G
- H
- I
- J
- K
- L
- M
- N
- O
- P
- Q
- R
- S
- T
- U
- V
- W
- X
- Y
- Z

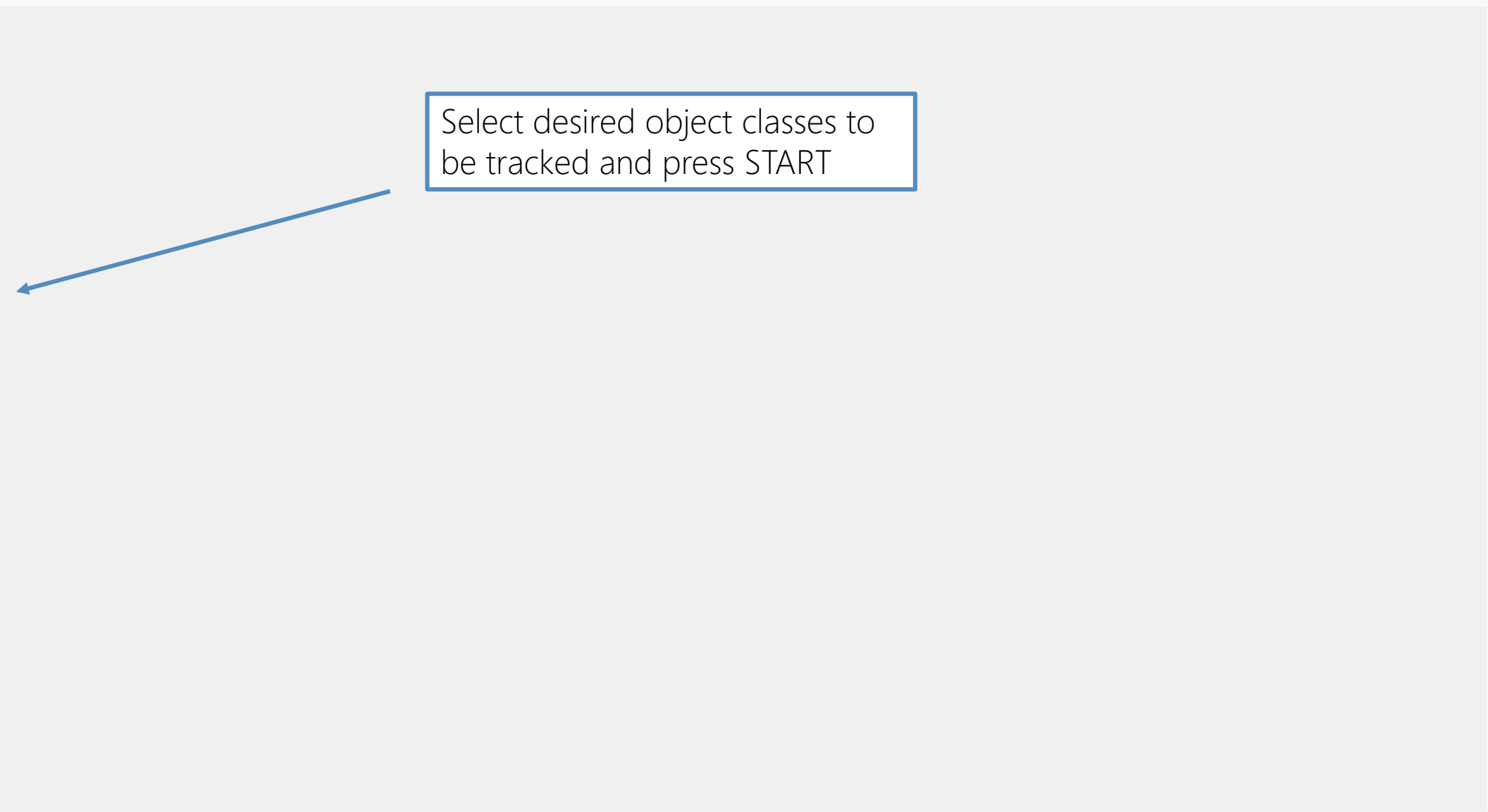


Check if the files are compatible



- Select All
- Cabbage
- Cabinetry
- Cake
- Cake stand
- Calculator
- Camel
- Camera
- Can opener
- Canary
- Candle
- Candy
- Cannon
- Canoe
- Cantaloupe
- Car
- Carnivore
- Carrot
- Cart
- Cassette deck
- Castle
- Cat
- Cat furniture
- Caterpillar
- Cattle
- Ceiling fan
- Cello
- Centipede
- Chainsaw
- Chair
- Cheese
- Cheetah
- Chest of drawers
- Chicken

- All
- A
- B
- C
- D
- E
- F
- G
- H
- I
- J
- K
- L
- M
- N
- O
- P
- Q
- R
- S
- T
- U
- V
- W
- X
- Y
- Z



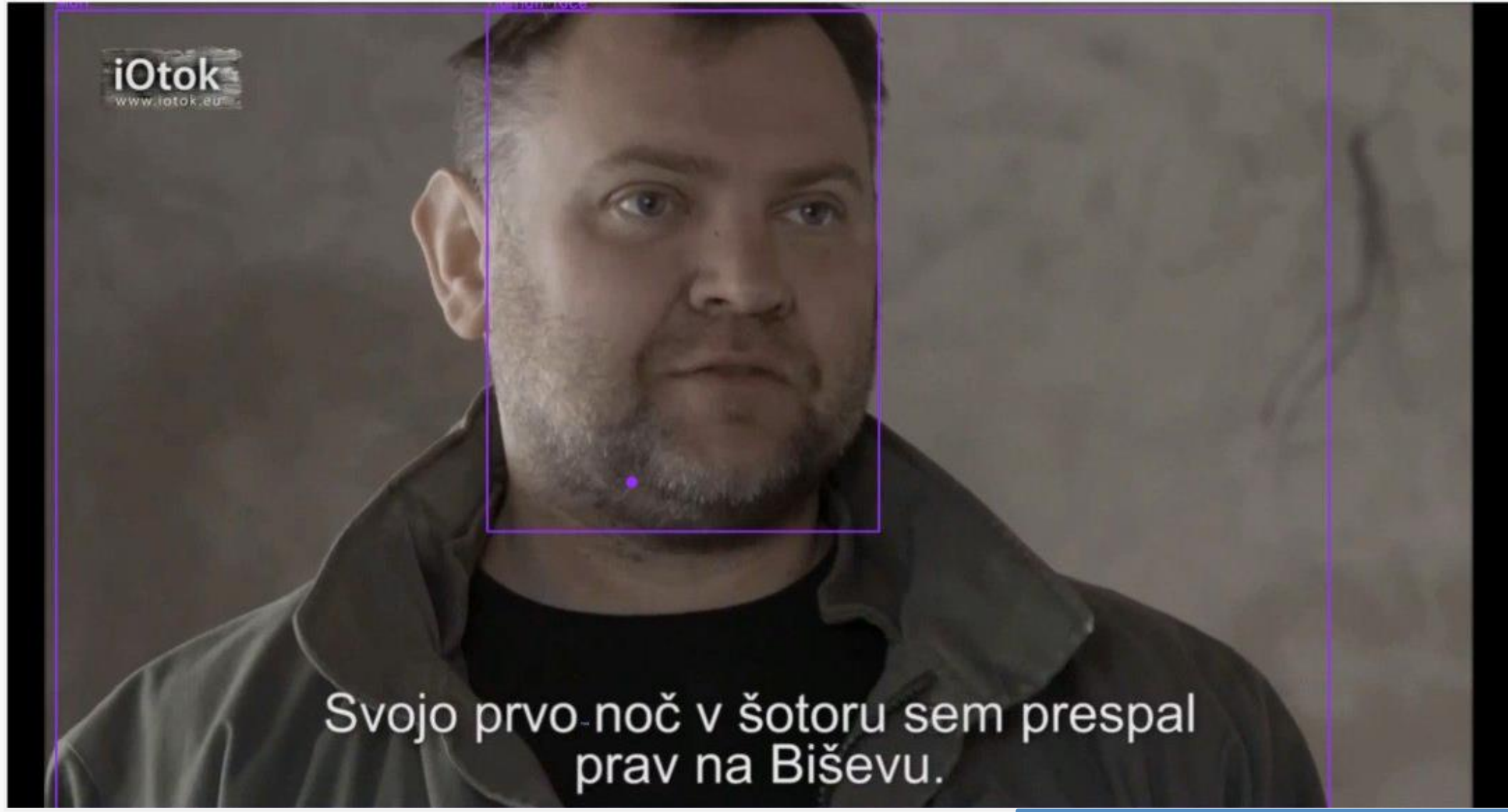
Select desired object classes to be tracked and press START



Home
Tracking
Labeling

- Select All
- Cabbage
- Cabinetry
- Cake
- Cake stand
- Calculator
- Camel
- Camera
- Can opener
- Canary
- Candle
- Candy
- Cannon
- Canoe
- Cantaloupe
- Car
- Carnivore
- Carrot
- Cart
- Cassette deck
- Castle
- Cat
- Cat furniture
- Caterpillar
- Cattle
- Ceiling fan
- Cello
- Centipede
- Chainsaw
- Chair
- Cheese
- Cheetah
- Chest of drawers
- Chicken

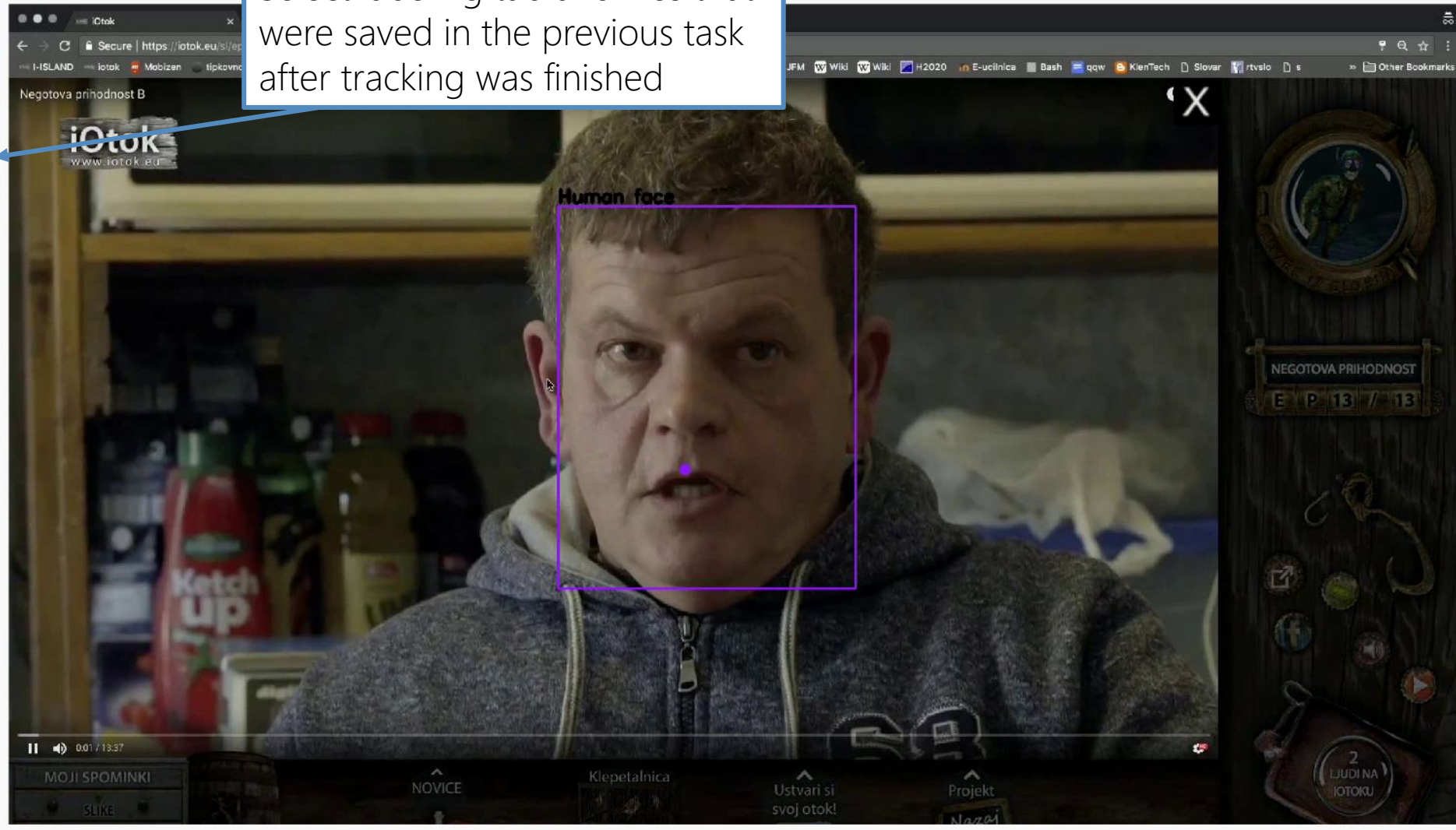
All
A
B
C
D
E
F
G
H
I
J
K
L
M
N
O
P
Q
R
S
T
U
V
W
X
Y
Z



Preview the annotated video including the fixation point

File
Home
Tracking
Labeling

Select labeling tab and files that were saved in the previous task after tracking was finished



Human face

Previous

Next

Restart

Save

The screenshot displays the Video Annotation App interface. On the left, a video player shows a frame from a video with a purple bounding box around a man's face. The text "Human face" is written above the box. A blue arrow points from the text "Human face" in the top right control area to the label in the video frame. A text box in the center-right explains the workflow: "Detected objects will appear here. Insert the desired label and go to next frame (next fixation). At the last frame, save the file." The control panel on the right includes buttons for "Previous", "Next", "Restart", and "Save". The "Next" and "Save" buttons are highlighted in blue. The video player shows a progress bar at 0:01 / 13:37 and a list of video segments at the bottom.

APPLICATION FOR IOTOK

- o Documentary with interactive elements

Ducasse, J., Kljun, M., & Čopič Pucihar, K. (2020). Interactive Web Documentaries: A Case Study of Audience Reception and User Engagement on iOtok. *International Journal of Human-Computer Interaction*, 36(16), 1558–1584. <https://doi.org/10.1080/10447318.2020.1757255>

- o Some people quit the web portal before reaching the core video: was it a lack of interest or was the interface not intuitive enough?
- o Currently running predictions with Gaze2AOI
- o Calculate: TTFF, AOI revisits, gazing time spent on interactive elements before leaving the site

iOtok / island

S tehnologijo si bomo življenje uredili tako, da nam ga na koncu ne bo več potrebno živeti.

Scenarist, režiser in producent / Scriptwriter, director and producer **MIHA ČELAR**

Koproducenti / Co-producers: **NENAD PUHOVSKI** in **MARTA ZACCARON** Izvorna producentka in ko-scenaristka / Executive producer and co-scriptwriter: **TAMARA BABUN**
Direktor fotografije / DOF: **TOMISLAV KRNIC** Tonki snemalci / Sound recordist: **VINCENT LAURENCE** Podvodni snemalci / Underwater cameraman: **MARINO BRZAC**
Montažerji / Editors: **DAREJ ŠÖMEN** and **FABIO TOICH** Barvne korekcije / Color grading: **ANDREA GUARASCIO** Grafično oblikovanje / Graphics: **REMIGIO GUADAGNINI**
Avtor glasbe, aranžmaja in dirigent / Original music composed, arranged and directed: **LUCA CIUT** Oblikovalec zvoka / Sound editing and mixing: **FRANCESCO MOROSINI**
Vodji projekta RTV SLO / Head of the project RTV SLO: **MATEJA ERIKA SMISL** in **MAJA PROMOŽIČ** Izvšna producentka FACTUM / Executive producer FACTUM: **NEDA FRANK**
Vodja projekta SLOVENSKI FILMSKI CENTER / Head of the project SLOVENIAN FILM CENTRE: **LIDIJA ZAJEC** Vodja projekta VIBA FILM / Head of the project VIBA FILM: **ANDREJ RUSTEC**
Avtor spletnih aplikacij: iOtok.si / Author of the iOtok application: **KLEN ČOPIČ PUCIHAR** Avtor island AR aplikacije / Author of the island AR app: dr. **RAHUL SWAMINATHAN**
Urednik dok. prog. RTV SLO / Chief Editor RTV SLO Docu. Prog.: **ANDRAZ POSCHL** Asistentki produkcije / Production assistants: **NADINA STEFANIČ** in **KRISTINA MATIČIČ**

V produkciji / Produced by **ASTRAL FILM** Factum Documentary Film Project **QUASAR MULTIMEDIA** V koprodukciji / In coproduction with **RADIO TELEVIZIJA SLOVENIJA**

S finančno in tehnično podporo / With a financial and technical support of **REPUBLIKA SLOVENIJA** **EVROPSKA UNIJA** **EVROPSKI SKLAD** **ERASMUS+2** **REACT** **AM** V sodelovanju z / In partnership with **TELECOM SLOVENIA** **REACT** **AM** **TELECOM SLOVENIA** **REACT** **AM**

FUTURE WORK

- Complete calculations for IOTOK
- Try YOLO WORLD – prompt to detection
- Similarity networks
- Segmentation instead detection

REFERENCES

- [1] Barz M, Sonntag D. Automatic Visual Attention Detection for Mobile Eye Tracking Using Pre-Trained Computer Vision Models and Human Gaze. *Sensors*. 2021; 21(12):4143.
- [2] Batliner, Martin, et al. "Automated areas of interest analysis for usability studies of tangible screen-based user interfaces using mobile eye tracking." *AI EDAM* 34.4 (2020): 505-514.
- [3] Duchowski, Andrew T., et al. "Art facing science: Artistic heuristics for face detection: tracking gaze when looking at faces." *Proceedings of the 11th ACM symposium on eye tracking research & applications*. 2019.
- [4] Ducasse, J., Kljun, M., Attygalle, N. T., & Pucihar, K. Č. (2022). Interactive Web Documentaries: A Case Study of Video Viewing Behaviour on iOtok. *International Journal of Human–Computer Interaction*, 38(10), 949–972.
- [5] Ducasse, J., Kljun, M., & Čopič Pucihar, K. (2020). Interactive Web Documentaries: A Case Study of Audience Reception and User Engagement on iOtok. *International Journal of Human–Computer Interaction*, 36(16), 1558–1584.
- [6] Jayawardena, Gavindya, and Sampath Jayarathna. "Automated filtering of eye gaze metrics from dynamic areas of interest." *2020 IEEE 21st International Conference on Information Reuse and Integration for Data Science (IRI)*. IEEE, 2020.
- [7] Kumari N, Ruf V, Mukhametov S, Schmidt A, Kuhn J, Küchemann S. Mobile Eye-Tracking Data Analysis Using Object Detection via YOLO v4. *Sensors*. 2021; 21(22):7668.
- [8] Lagmay, Ezekiel Adriel D., and Maria Mercedes T. Rodrigo. "Enhanced Automatic Areas of Interest (AOI) Bounding Boxes Estimation Algorithm for Dynamic Eye-Tracking Stimuli." *APSIPA Transactions on Signal and Information Processing* 11.1 (2022).²⁴

- [9] Ryabinin, Konstantin, Svetlana Alexeeva, and Tatiana Petrova. "Automatic Areas of Interest Detector for Mobile Eye Trackers." *Graphicon-Conference on Computer Graphics and Vision*. Vol. 32. 2022.
- [10] Wang, Felix Sihan. *Algorithmic Gaze Measure Development for Application in Real-World Scenarios*. Diss. ETH Zurich, 2023.
- [11] Wu, Shaoze, et al. "Measuring the impact of Augmented Reality warning systems on onsite construction workers using object detection and eye-tracking." *Proceedings of the 28th EG-ICE International Workshop on Intelligent Computing in Engineering, EG-ICE 2021*. 2021.
- [12] Xu, Hui, et al. "Analyzing students' attention by gaze tracking and object detection in classroom teaching." *Data Technologies and Applications* 57.5 (2023): 643-667.
- [13] Zhang, Xuebai, et al. "A complete system for analysis of video lecture based on eye tracking." *Ieee Access* 6 (2018): 49056-49066.

THANK YOU!

Questions?

