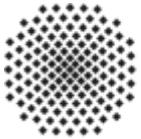


University of Stuttgart
Germany

Predicting Visual Search Targets

Ramvinojen Narayana Perumal
Shyamnarayan Sankarasubramanian
Vijeth Kumar

Institute for Visualisation and
Interactive Systems

Motivation

- Is it possible to infer the persons goals by analyzing their eye fixations?
- Eye movements can reveal complex cognitive states of mind
- Analyse Visual Behavior
- Closed World Vs Open World Settings
- Prediction of Search Targets using Stationary Eye Tracker

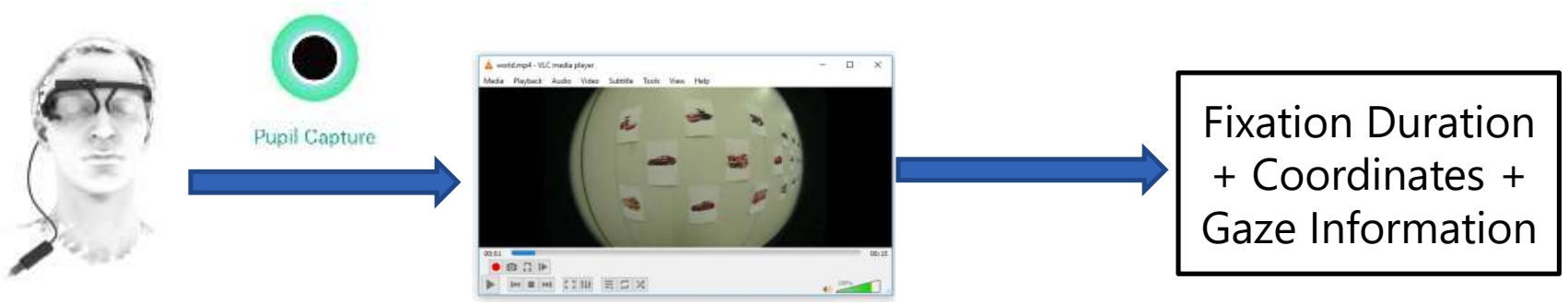
Can we predict Search Targets using Mobile Eye Tracker?

Agenda

1. Hardware
2. Data Collection
3. Preprocessing Data
4. Feature Extraction
5. Models
6. Applications
7. References

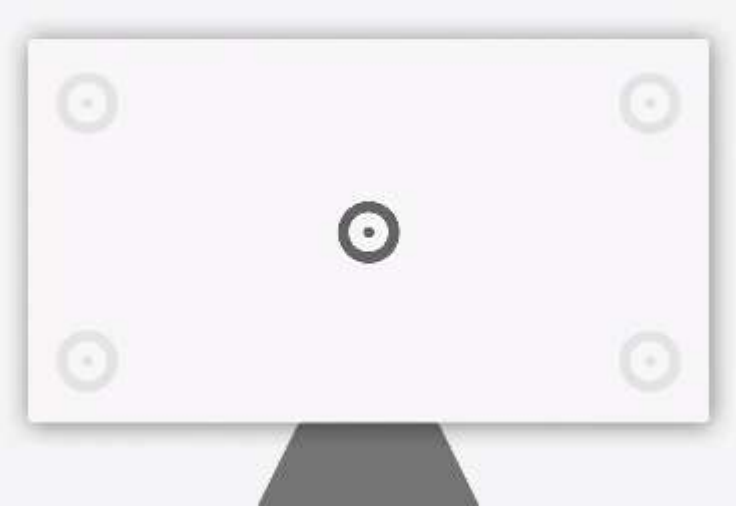
Mobile Eye Tracker

- Hardware:
  - Mobile Eye Tracker – World Camera and 2 Eye Cameras
  - Sampling Frequency Eye Camera – 200Hz
  - Sampling Frequency World Camera – 30Hz
  - Recording - 1920X1080p HD



Calibration

- Screen Marker Calibration - Small field of vision
- Manual Marker Calibration - Low Accuracy & Time Consuming
- Natural Feature Calibration - Better Approach



Screen Marker Calibration

Natural Feature Calibration

Data Collection

Object Type	Pros	Cons
Real Object 	1) Mimics Real World	1) Maximum 3 classes 2) Lower Distractors options 3) Costly 4) Complex Image Processing
Virtual Object 	1) More classes 2) Better Dataset 3) More Distractors 4) Economically feasible	1) Deviation from Real life situation

Data Collection

Class	Sub Class
Car	Sedan
	Hatchback
	Pickup
Truck	Fire Engine
	Cement Truck
	Road Roller
Bus	Public Bus
	School Bus
Distractors	Cement Trailer
	Crane
	Tractor

Description	Number
Participants	17
Dataset	286
Training dataset	190(**~70%**)
Validation dataset	48(**~15%**)
Testing dataset	48(**~15%**)

Preprocessing Data

Video

Laplacian of Gaussian

Best Frame

Gaze Information

Best Frame

Extracting Image Patch

Fixation Coordinate

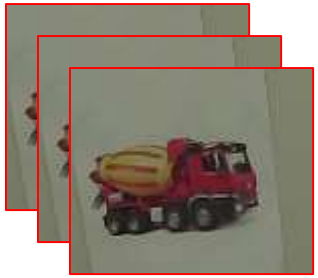


Image Enhancer + Reshape

224X224

Feature Extraction

Image Patches

ResNet50

Pre-trained Network

2048 Features per Fixation

Feature Vector1

Normalized Fixation Duration1

Feature Vector N

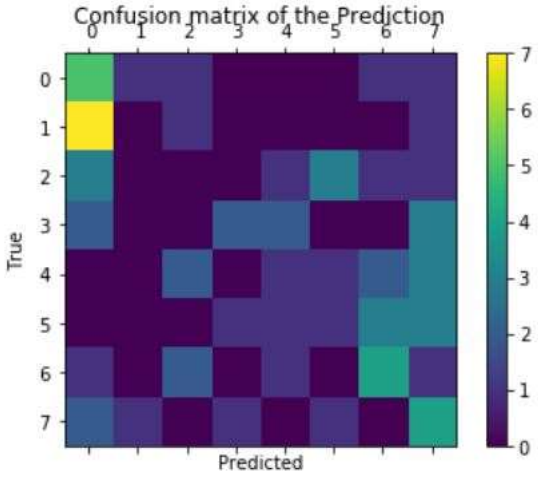
Normalized Fixation Duration N

Combined Fixation Feature Vector

Approach I - Regression & Classifier

- Linear Regression
- Loss function - Least Squares

Classifier Trained with Augmented Image Patches

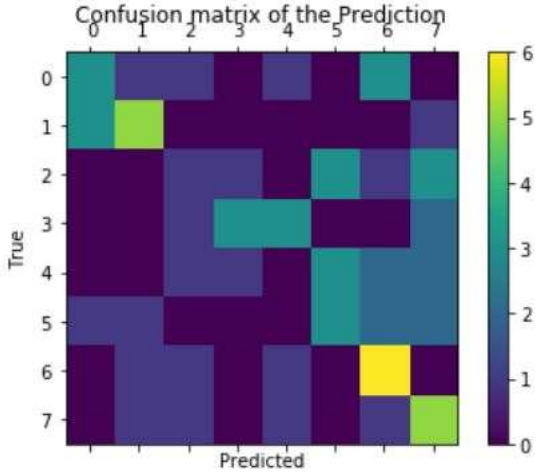


Summary for each class:	precision	recall	f1-score	support
0	0.25	0.56	0.34	9
1	0.00	0.00	0.00	9
2	0.00	0.00	0.00	9
3	0.50	0.22	0.31	9
4	0.17	0.11	0.13	9
5	0.17	0.11	0.13	9
6	0.36	0.44	0.40	9
7	0.24	0.44	0.31	9
avg / total	0.21	0.24	0.20	72

Approach I - Regression & Classifier

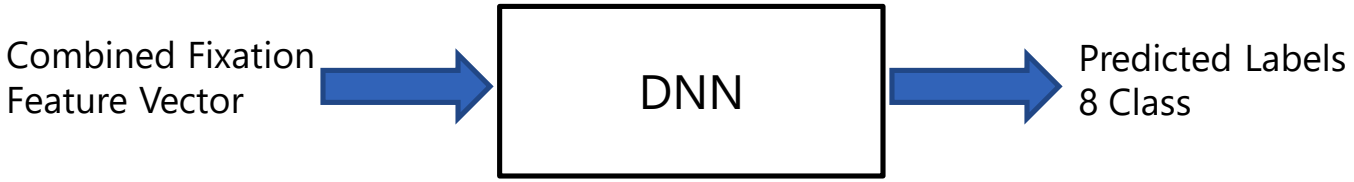
- Linear Regression
- Loss function - Least Squares

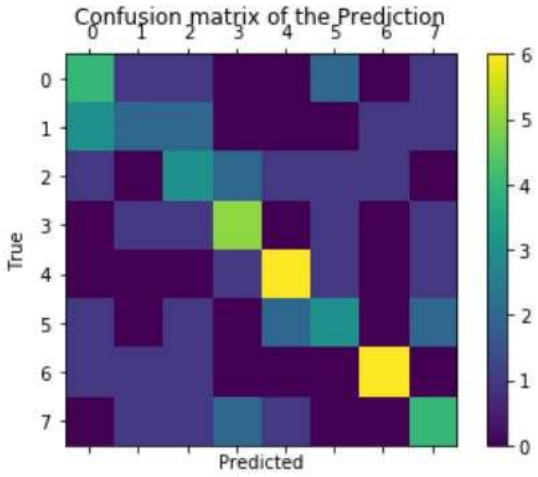
Classifier Trained with Augmented Image Patches



	precision	recall	f1-score	support
0	0.43	0.33	0.38	9
1	0.56	0.56	0.56	9
2	0.17	0.11	0.13	9
3	0.60	0.33	0.43	9
4	0.00	0.00	0.00	9
5	0.33	0.33	0.33	9
6	0.40	0.67	0.50	9
7	0.33	0.56	0.42	9
avg / total	0.35	0.36	0.34	72

Approach II - DNN

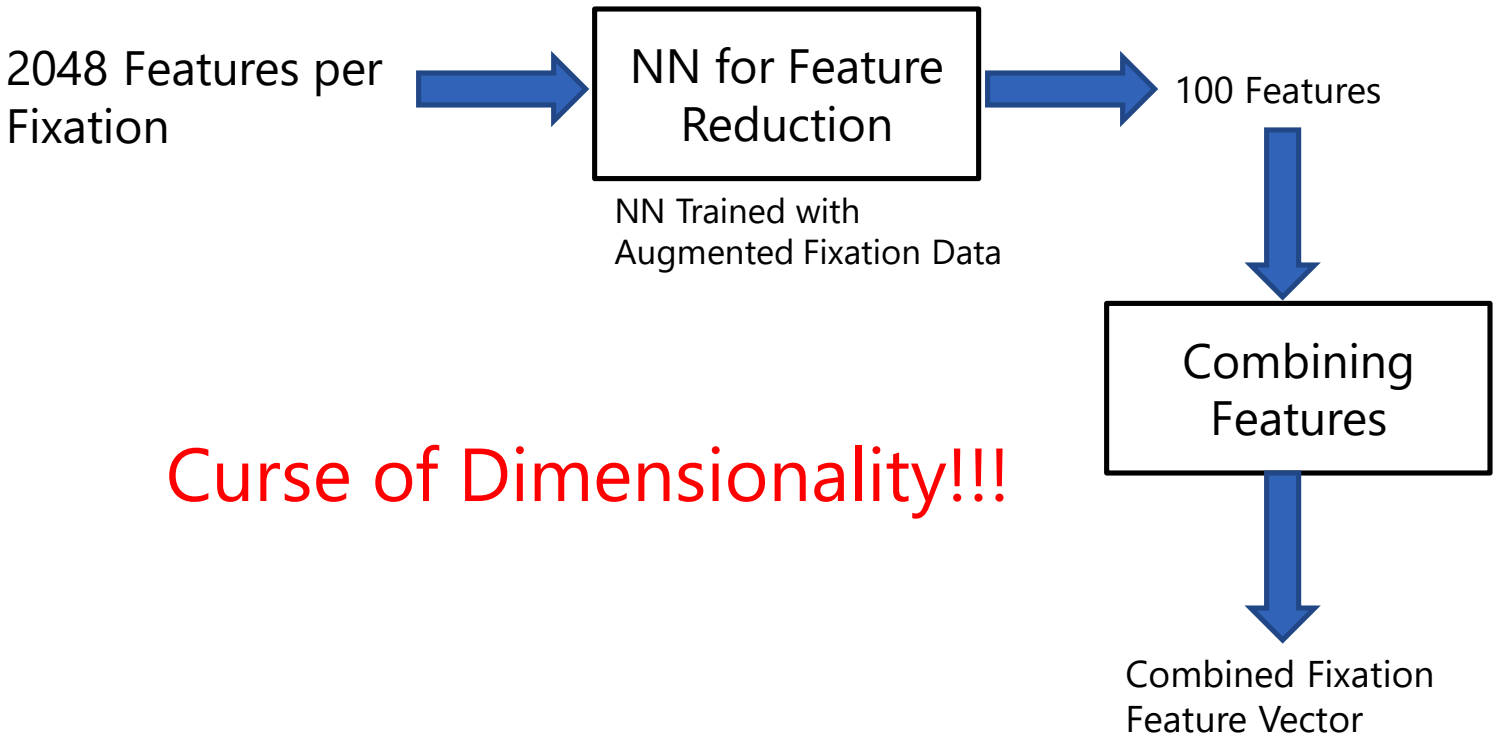




Summary for each class:

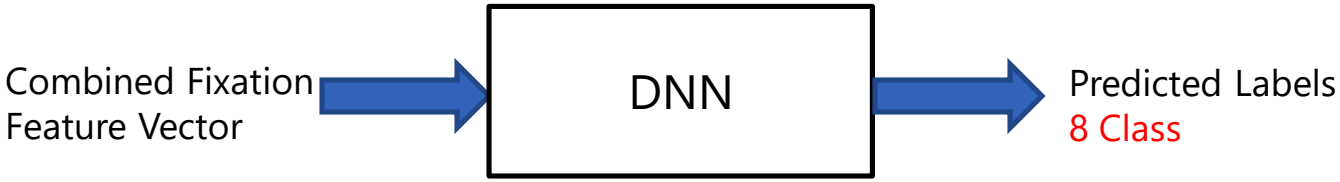
	precision	recall	f1-score	support
0	0.40	0.44	0.42	9
1	0.33	0.22	0.27	9
2	0.30	0.33	0.32	9
3	0.50	0.56	0.53	9
4	0.60	0.67	0.63	9
5	0.38	0.33	0.35	9
6	0.75	0.67	0.71	9
7	0.40	0.44	0.42	9
avg / total	0.46	0.46	0.46	72

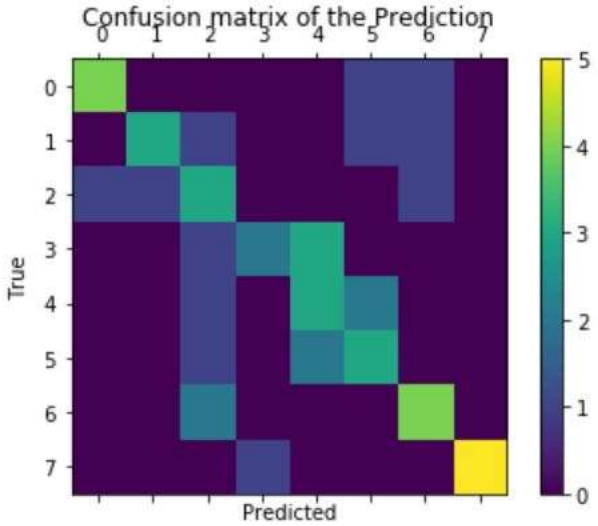
Approach III – Feature Reduction



Curse of Dimensionality!!!

Approach III – Feature Reduction

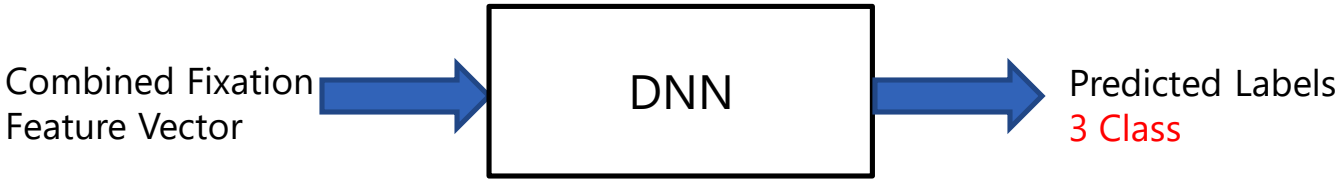


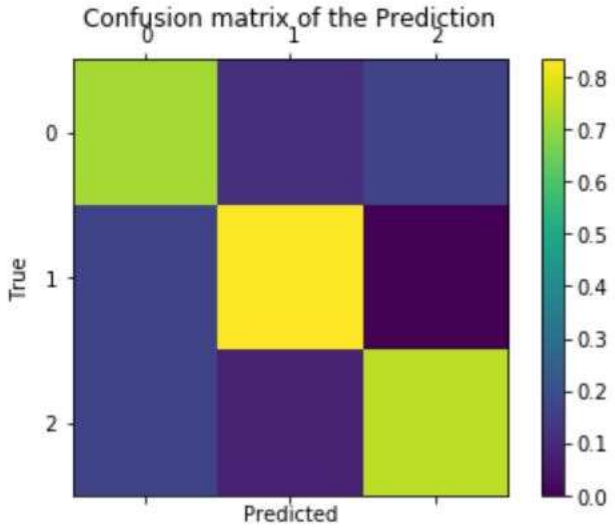


Summary for each class:

	precision	recall	f1-score	support
0	0.80	0.67	0.73	6
1	0.75	0.50	0.60	6
2	0.33	0.50	0.40	6
3	0.67	0.33	0.44	6
4	0.38	0.50	0.43	6
5	0.43	0.50	0.46	6
6	0.57	0.67	0.62	6
7	1.00	0.83	0.91	6
avg / total	0.62	0.56	0.57	48

Approach III – Feature Reduction

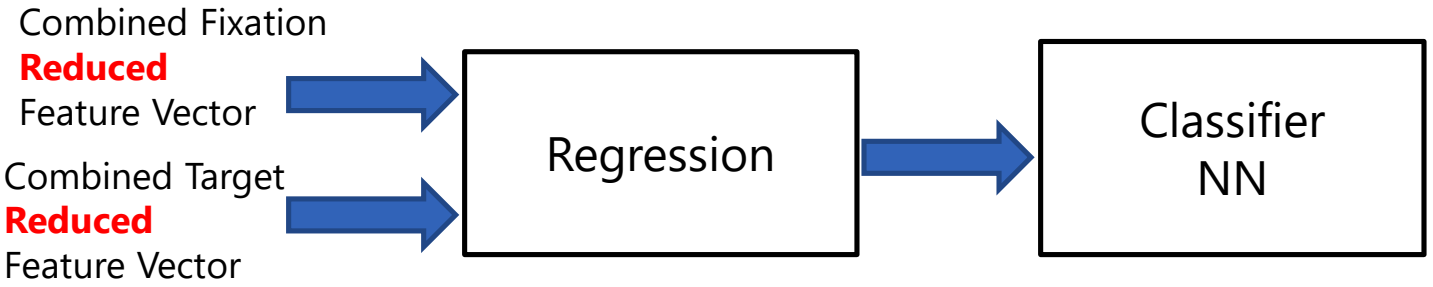


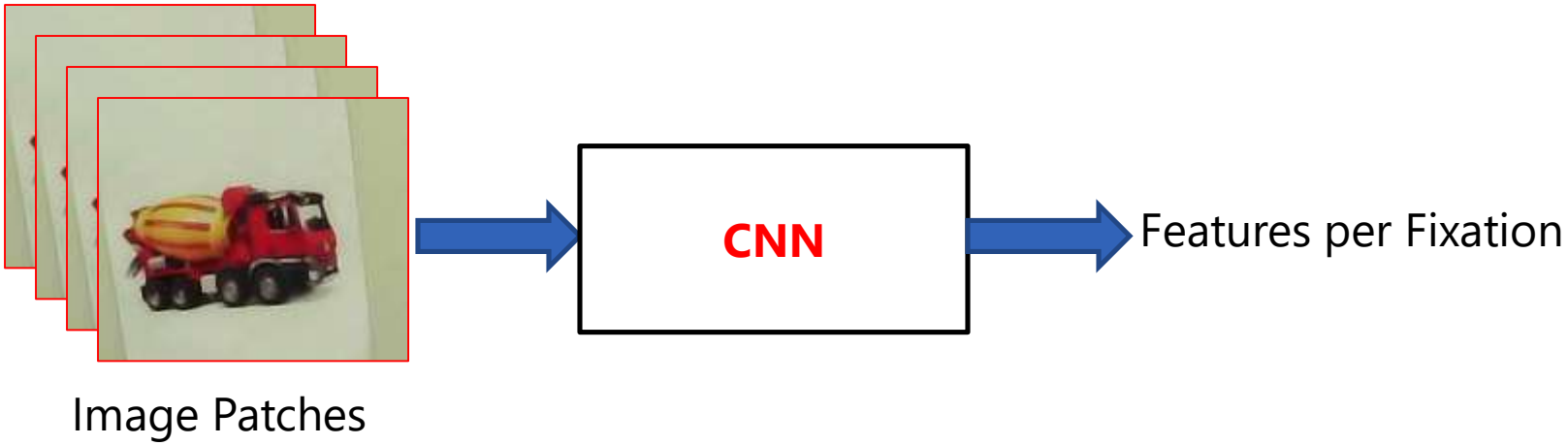


Summary for each class:

	precision	recall	f1-score	support
0	0.72	0.72	0.72	18
1	0.83	0.83	0.83	18
2	0.75	0.75	0.75	12
avg / total	0.77	0.77	0.77	48

Other Approaches





Below Par Results!!!

Applications

Virtual Search Assistant in Grocery Store

Virtual Search Assistant for Drivers

Enhances Human – Machine Interaction

References

1. Prediction of Search Targets From Fixations in Open-World Settings -Hosnieh Sattar, Sabine Mueller, Mario Fritz, Andreas Bulling
2. Predicting the Category and Attributes of Visual Search Targets Using DeepGaze Pooling - Hosnieh Sattar, Andreas Bulling, Mario Fritz
3. Fixation Detection for Head-Mounted Eye Tracking Based on Visual Similarity of Gaze Targets – Julian Steil, Michael Xuelin Huang, Andreas Bulling
4. What do eyes reveal about the mind? Algorithmic inference of search targets from fixations - Ali Borji, Andreas Lennartz, Marc Pomplun
5. Eye can read your mind: Decoding gaze fixations to reveal categorical search targets - Gregory J. Zelinsky, Yifan Peng, Dimitris Samaras