Contents

List of Figures **vi**

1 Introduction **1**
1.1 Motivation . . . 1
1.2 Hypotheses . . . 2
1.3 Objectives . . . 2
1.3.1 Main Objectives . . . 2
1.3.2 Specific Objectives . . . 2
1.4 Thesis Structure . . . 3

2 Scanpath Ocular **4**
2.1 Eye Movements . . . 4
2.2 Brain Selection . . . 5
2.3 Spatial Selection . . . 5
2.4 Fixation Duration . . . 6
2.5 Saliency and Selective Order . . . 6
2.6 Calculating saliency . . . 6
2.7 Fovea centralis . . . 8
2.8 Scanpath Comparison Metrics . . . 9

3 Theoretical background for Scanpath Prediction **12**
3.1 Models for Scanpath Prediction . . . 12
3.2 Summary of Artificial Neural Models . . . 13
3.2.1 Theoretical definition . . . 14
3.2.2 Supervised Learning . . . 15
3.2.2.1 Stochastic Gradient Descent . . . 15
3.2.2.2 Backpropagration . . . 17
3.2.3 Neural networks . . . 17
3.2.3.1 Convolutional neural network . . . 17
3.2.3.2 Recurrent neural network . . . 18
3.2.3.3 Attention neural network . . . 20
3.2.4 Regularization methods . . . 21
3.2.4.1 Dropout . . . 21
3.2.4.2 Early stop . . . 21
3.3 Auto-regressive models for multi-step-ahead forecast in time series . . . 22
3.3.1 Recursive forecast . . . 22
3.3.1.1 Training phase . . . 22

3.3.1.2 Inference . . . 23
3.3.2 Direct forecast . . . 24
3.3.3 Forecast strategy selection . . . 24
3.4 Modelling uncertainty via MC-Dropout . . . 25
3.4.1 Approximate Variational Inference . . . 25

4 Applications of Artificial Neural Networks for Scanpath Prediction 27
4.1 Methods . . . 28
4.1.1 Dataset . . . 28
4.1.2 Modelling Procedure . . . 29
4.2 Modelling scanpath with a recurrent neural model using positional information . . . 31
4.2.1 Positional Scanpath and LSTM model . . . 31
4.2.2 Analysis of PosScan model . . . 32
4.2.2.1 PosScan results grouped by train image type . . . 33
4.2.2.2 PosScan results grouped by predicted image type . . . 41
4.2.2.3 Prediction in other subjects rather than the trained one . . . 46
4.2.3 Remarks . . . 49
4.3 Selecting features to enhance the model . . . 50
4.4 Modelling scanpath with an attention neural model using positional and spatial information though time . . . 55
4.4.1 Saliency maps from foveated images and Attention model . . . 55
4.4.2 Analysis of FovSOS-FS model . . . 57
4.4.2.1 FovSOS-FS results grouped by predicted image type . . . 58
4.4.3 Analysis of FovSOS-FSD . . . 62
4.4.3.1 FovSOS-FSD results grouped by train image . . . 63
4.4.3.2 FovSOS-FSD results grouped by predicted image . . . 66
4.4.4 Remarks . . . 69
4.5 Comparative discussion of models . . . 69

5 Conclusions and Future Work 74

Bibliography 77

ANNEXES 86

Annexed A PosScan architecture 87
A.1 Architecture parameters . . . 87
A.2 Metric results grouped by train image type . . . 88
A.3 Metric results grouped by predicted image type . . . 92

Annexed B FovSOS-FS architecture 96
B.1 Architecture parameters . . . 96
B.2 Metric results grouped by predicted image type . . . 97

Annexed C FovSOS-FSD architecture 101
C.1 Architecture parameters . . . 101
C.2 Metric results grouped by train image type . . . 102
C.3 Metric results grouped by predicted image type . . . 106