

LIST OF CONTENTS

CHAPTER I INTRODUCTION	1
1.1 Background	1
1.2 Research Problem.....	2
1.3 Research Scope.....	2
1.4 Research Objectives	3
1.5 Research Benefits	3
CHAPTER II LITERATURE REVIEW	4
CHAPTER III THEORETICAL FOUNDATION	8
3.1 Cat Health Monitoring.....	8
3.2 Computer Vision.....	9
3.3 ESP32 - CAM.....	10
3.4 Convolutional Neural Network	11
3.4.1 CNN-Based Object Detection Architectures	13
3.4.2 Speed–Accuracy Trade-offs.....	15
3.4.3 Comparison to Transformers and RNNs.....	16
3.5 YOLO (You Only Look Once).....	18
3.6 FOMO(Faster Object More Object)	21
3.6.1 FOMO Network Architecture	21
3.6.2 FOMO Algorithmic Design and Optimizations.....	22
3.6.3 FOMO Comparison with YOLO, SSD, and EfficientDet	23
3.6.4 FOMO Implementation and Deployment on Edge Devices.....	24
CHAPTER IV SYSTEM ANALYSIS AND DESIGN	26
4.1 System Analysis	26
4.2 Data Collection.....	27
4.3 Data Preprocessing	28
4.4 Yolov8 with Computer Vision System Design Analysis	30
4.4.1 Hardware Architectural Design	33
4.4.2 Hardware Model Design.....	33

4.4.3 Label Studio Label Annotation	34
4.5 FOMO with ESP32-CAM System Design Analysis	36
4.5.1 Hardware Architectural Design	39
4.5.2 Hardware Model Design	39
4.5.3 Label Studio Label Annotation	40
4.6 System Testing and Validation	42
CHAPTER V IMPLEMENTATION	45
5.1 Tools and Materials	45
5.2 Hardware Implementation	47
5.3 Web Live Stream Implementation	48
5.3.1 Camera Configuration.....	49
5.3.2 Live Stream.....	50
5.4 YOLOv8 With Computer Vision	51
5.4.1 Model Training	51
5.4.2 Library Imports and Model Initialization	53
5.4.3 Video Input Configuration	54
5.4.4 YOLOv8 Inference and Bounding Box Processing.....	54
5.4.5 Motion Detection	55
5.4.6 Activity Classification and Time Tracking	57
5.4.7 Behavioural Evaluation and Logging	58
5.4.8 Email Notification System.....	59
5.5 FOMO Model with ESP-32 CAM.....	60
5.5.1 Model Training	62
5.5.2 Library Imports	67
5.5.3 Camera Configuration.....	67
5.5.4 Image Acquisition and Preprocessing.....	68
5.5.5 FOMO Model Inference	70

5.5.6 Activity Tracking	72
5.5.7 Real-Time Feedback and Notification	73
CHAPTER VI RESULTS AND DISCUSSION	75
6.1 YOLOv8 with Computer Vision Result	75
6.1.1 Model Training and Loss Evaluation.....	76
6.1.2 Label Distribution and Bounding Box Analysis	77
6.1.3 Confusion Matrix and Class Performance	78
6.1.4 Precision, Recall, F1, and Matthew Correlation Coefficient Evaluation	80
6.1.5 Real-Time Video Streaming with ESP32-CAM	83
6.1.6 FPS and Model Performance Evaluation.....	83
6.1.7 Detection Accuracy Evaluation	85
6.1.8 Region of Interest (ROI) for Eating Behaviour	87
6.1.9 Failure Cases	89
6.1.10 Detection Accuracy and Error Analysis.....	90
6.1.11 Email Notification System.....	92
6.2 ESP-32 CAM with FOMO Result.....	94
6.2.1 Model Evaluation.....	95
6.2.2 FPS and Inference Evaluation.....	99
6.2.3 Serial Monitor Output Analysis for Real-Time Classification	99
6.2.4 Web Interface for Wireless Behaviour Monitoring.....	101
6.2.5 Failure Cases	102
6.2.6 Accuracy Evaluation: FOMO Detection vs. Ground Truth	103
6.2.7 Email Notification System for Daily Activity Report	105
6.3 Overall Comparison Between Yolov8 vs FOMO for Real Time Cat Health and Behaviour Monitoring	106
CHAPTER VII CONCLUSION	109



Real-Time Monitoring of Cat Health and Activity in a Pet Shop Environment Using ESP-CAM and Computer Vision

Hafizh Ilham, Triyogatama Wahyu Widodo, S.Kom., M.Kom

Universitas Gadjah Mada, 2025 | Diunduh dari <http://etd.repository.ugm.ac.id/>

7.1 Conclusion.....	109
7.2 Recommendations	110
REFERENCES.....	112

LIST OF FIGURES

Figure 3.1 Yolo Model Structure.....	19
Figure 3.2 YOLO Bounding Box.....	19
Figure 3.3 Training YOLO.....	20
Figure 4. 1 Resting Cat Data Collection	28
Figure 4. 2 Sleeping Cat Data Collection.....	28
Figure 4. 3 Standing Cat Data Collection	28
Figure 4. 4 Computation Research Steps	31
Figure 4. 5 PC Computation System Design Flowchart	32
Figure 4. 6 PC Computation Hardware Architectural Design.....	33
Figure 4. 7 ESP-32 CAM with PC Wiring diagram.....	33
Figure 4. 8 Rest Label Annotation with Label Studio.....	35
Figure 4. 9 Sleep Label Annotation with Label Studio.....	36
Figure 4. 10 Stand Label Annotation with Label Studio.....	36
Figure 4. 11 Computation Research Steps	37
Figure 4. 12 ESP-32 CAM Computation System Design Flowchart.....	38
Figure 4.13 ESP-32 CAM Hardware Architectural Design.....	39
Figure 4. 14 ESP-32 CAM Wiring diagram.....	40
Figure 4. 15 Rest Label Annotation with Edge Impulse	41
Figure 4. 16 Sleep Label Annotation with Edge Impulse	42
Figure 4. 17 Stand Label Annotation with Edge Impulse	42
Figure 5. 1 ESP-32 CAM Hardware Implementation	47
Figure 5. 2 Web Live Stream Library	48
Figure 5. 3 Web Live Stream Camera Setup	49
Figure 5. 4 Live Stream Code	50
Figure 5. 5 YOLO Data preparation.....	52
Figure 5. 6 YOLO Model Training	53
Figure 5. 7 Computer Vision Implementation Library.....	53
Figure 5. 8 Video Input Configuration.....	54
Figure 5. 9 Class Index Code	55
Figure 5. 10 Extracting Bounding Box Coordinates.....	55

Figure 5. 11 Motion Detection	56
Figure 5. 12 ROI Result	57
Figure 5. 13 Label stamp.....	57
Figure 5. 14 Timestamp Code	58
Figure 5. 15 Behavioural Evaluation	58
Figure 5. 16 Logging to CSV	59
Figure 5. 17 Email body.....	60
Figure 5. 18 Send Email.....	60
Figure 5. 19 Edge Impulse Data Preprocessing	63
Figure 5. 20 Image Normalization	64
Figure 5. 21 MobileNetV2 Backbone (Source : MobileNet-JDE: a lightweight multi-object tracking model for embedded systems - Scientific Figure on ResearchGate.)	65
Figure 5. 22 Training Configuration	66
Figure 5. 23 Library used for Object Detection with FOMO	67
Figure 5. 24 Camera Configuration for Object Detection with FOMO.....	68
Figure 5. 25 Camera Capturing and preprocessing for Object Detection with FOMO	69
Figure 5. 26 FOMO Model Inference	71
Figure 5. 27 Activity Tracking Code.....	72
Figure 5. 28 Webpage live activity Code	74
Figure 5. 29 Email Code	74
Figure 6. 1 Yolov8 Pipeline.....	76
Figure 6. 2 Model Training and Loss Evaluation.....	76
Figure 6. 3 Label Distribution Evaluation	77
Figure 6. 4 Label Correlogram Evaluation	78
Figure 6. 5 Confusion Matrix Graph.....	79
Figure 6. 6 Normalized Confusion Matrix Graph.....	80
Figure 6. 7 P-R Curve	81
Figure 6. 8 F1 Curve	81
Figure 6. 9 Real Time Video Streaming with ESP-32 CAM	83

Figure 6. 10 GUI Result	84
Figure 6. 11 YOLOv8 Detecting a Cat in a Resting Posture	85
Figure 6. 12 YOLOv8 Detecting a Cat in a Standing Posture	86
Figure 6. 13 YOLOv8 Detecting a Cat in a Sleeping Posture.....	87
Figure 6. 14 Eating detection process.	88
Figure 6. 15 Example of YOLOv8 Failure Cases.....	89
Figure 6. 16 16 Minutes Detection process.....	92
Figure 6. 17 Email Notification System	94
Figure 6. 18 FOMO Model Evaluation	97
Figure 6. 19 FPS and Inference Time Result	99
Figure 6. 20 Serial Monitor Result.....	101
Figure 6. 21 Web Wireless Behaviour Monitoring	102
Figure 6. 22 FOMO Failure Cases	102
Figure 6. 23 Detection Result over ~ 15 minutes.....	105
Figure 6. 24 Email Notifications Result.....	106

LIST OF TABLES

Table 2.1 Research Corelation	6
Table 6. 1 Summary Table of YOLOv8 MCC Values.....	82
Table 6. 2 Actual vs Detection Time	90
Table 6. 3 Summary Table of FOMO MCC Values	98
Table 6. 4 FOMO vs Actual Detected Times	103
Table 6. 5 Comparison Between Yolov8 and FOMO.....	108