Neural-Assisted Feature Matching

Internship Report

Author: EL OUARRAT Haytam
Internship Period: Mars – August 2025
Location: SteelSeries, Lille, France

Advisors: Pierre Biret, Damien Granger, Raphaël Greff
University Supervisor: Phillipe Joly

Engineering Degree in Robotics and Interactive Systems
UPSSITECH
University of Toulouse

June 24, 2025

FOR GLORY

Contents

FOR GLORY

List of Figures

Chapter 1

Acknowledgements

Chapter 2

Introduction

2.1 Host Organism

2.1.1 Nahimic

2.1.2 SteelSeries

2.1.3 GN Group

2.1.4 Mission

2.2 Context & Motivation

2.2.1 Role of Feature Matching in Computer Vision

2.2.2 Challenges in Gaming Applications

2.2.3 Limitations of Traditional Feature Matching Techniques

2.3 Project Objectives

2.3.1 Reprodusing Feature Matching Techniques with Neural Networks

2.3.2 Improving computational efficiency

2.3.3 Ensure matching accuracy for gaming footage

2.4 Industrial Relevance

2.4.1 Integration with SteelSeries Moments Software

2.4.2 Real-time performance constraints

Chapter 3

Literature Review

3.1 Traditional Feature Matching

3.1.1 Overview of SIFT, ORB, FAST

3.1.2 Comparative Strengths, Weaknesses, and Computational Costs

3.2 Neural Feature Matching

3.2.1 Review of Recent Methods: LoFTR, ALIKE, LightGlue, XFeat

3.3 Knowledge Distillation

3.3.1 Distillation Types: Response-Based, Feature-Based, Relation-Based

3.3.2 Applications in Model Compression and Matching Tasks

3.4 Lightweight Architectures for Edge Deployment

3.4.1 MobileNet, ShuffleNet, XFeat-Style Networks

3.4.2 Trade-Offs Between Efficiency and Accuracy

3.5 Gaps and Opportunities

3.5.1 Where Traditional Methods Fall Short

3.5.2 Where Neural Methods Remain Overkill for Real-Time CPU Usage

3.5.3 Motivation for a Hybrid/Distilled Approach

Chapter 4

Methodology

4.1 Problem Formulation

4.1.1 Define Feature Matching as Correspondence Prediction

4.1.2 Objectives in Terms of Speed, Accuracy, and Robustness

4.2 Baseline Selection

4.2.1 Justification for Using ORB or FAST as Teacher Models

4.2.2 Benchmark Datasets (e.g., HPatches, Gaming Clips)

4.3 Neural Architecture Design

4.3.1 Choice of Lightweight CNN or Transformer Backbone

4.3.2 Feature Extraction vs. Matching Separation

4.4 Distillation Strategy

4.4.1 Design of Teacher-Student Framework

4.4.2 Distillation Losses (e.g., L2 on Descriptors, Cross-Entropy on Match Maps)

4.5 Evaluation Metrics

4.5.1 Matching Precision, Recall, Repeatability

4.5.2 Runtime (FPS), Memory Footprint, CPU Load

Chapter 5

Implementation

5.1 Dataset Preparation

5.1.1 Gaming Video Frame Extraction

5.1.2 Synthetic Transformation Generation for Ground Truth Correspondences

5.2 Training Pipeline

5.2.1 Data Augmentation Strategies

5.2.2 Loss Function Components and Training Schedule

5.3 Model Optimization

5.3.1 Quantization, Pruning, or ONNX Export (if applicable)

5.3.2 Inference Optimization for CPU

5.4 Integration with SteelSeries Pipeline

5.4.1 Data Flow Alignment with Moments Software (if available)

5.4.2 Latency Tracking and Bottleneck Identification

Chapter 6

Results & Analysis

6.1 Matching Quality

6.1.1 Quantitative Comparison with ORB, SIFT, and XFeat

6.1.2 Visual Results on Gaming Footage

6.2 Computational Efficiency

6.2.1 FPS and Latency Benchmarks

6.2.2 Memory and CPU Usage Profiles

6.3 Ablation Studies

6.3.1 Effect of Different Distillation Losses

6.3.2 Model Depth vs. Performance Trade-Offs

6.4 Real-Time Viability

6.4.1 End-to-End Latency Breakdown

6.4.2 Suitability for Gaming Hardware

Chapter 7

Conclusion and Future Work

7.1 Summary of Contributions

7.2 Limitations

7.2.1 Domain Generalization

7.2.2 Extreme Low-Light or High-Motion Scenes

7.3 Future Work

7.3.1 Self-Distillation or Online Distillation Strategies

7.3.2 Hardware-Specific Optimizations (e.g., ARM CPU Tuning)

7.3.3 Real-Time Deployment on End-User Devices

Chapter 8

References

Chapter 9

Appendices

9.1 Additional Figures

9.2 Code Snippets

9.3 Hyperparameter Tables

9.4 Hardware Specifications