Supporting Materials Report — ECS8056 Themed Research Project

Barry Quinn (Student No. 9207589)

2025-09-19

Table of contents

1 Overview

This supporting materials report accompanies the eight-page research paper submitted for ECS8056. It provides additional detail on the project lifecycle, verification and validation, development tools, and a critical appraisal, as required by the handbook. It also includes pointers to the code repository on QUB EEECS GitLab and the recorded video demo.

Paper (PDF): `docs/thesis_paper.pdf`
Slides (reveal.js): `write-up/thesis_demo_slides.qmd`
Video Demo: MP4/H.264 (10 minutes) — uploaded to Canvas with supporting materials

2 Project Lifecycle, Verification and Validation

2.1 Lifecycle Stages

1. Scoping & requirements: define task as hierarchical flag classification for cultural symbol recognition; identify dataset constraints (licensing) and evaluation granularity (70→16→7).

2. Data curation: expert-verified imagery, taxonomy and splits metadata (no raw imagery redistributed); annotate categories and prepare figure assets.

3. Modelling & evaluation: ViT-H-14 baselines across 70/16/7; multi-seed and cross-validation for robustness; attention visualisation for interpretability.

4. Reframing & documentation: adjust narrative to "novel task + hierarchy"; add declaration, acknowledgements, and AI-use statement; generate manuscript and slides with Quarto.

5. Submission packaging: create reproducibility scripts, slides, and video; mirror code to QUB EEECS GitLab; prepare Canvas upload (paper, supporting report, demo video).

2.2 Verification & Validation (V&V)

- Dataset split checks: `verify_setup.py` (submission repo) or `qub-gitlab-submission/test_setup.py` (bundle) verifies Train/Val/Test sizes and totals.

- Statistical validation: multi-seed training and 5-fold cross-validation summarised in the manuscript; reproduced figures match reported numbers.

- Reproducibility: Quarto rendering of paper (`write-up/render_and_publish.sh`); regeneration of figures from scripts (lightweight ones shown in demo).

- Compliance review: scan for non-compliant claims and correct supervisor naming; declaration page included after title.

3 Tools & Environment

- Languages & frameworks: Python (PyTorch), Quarto 1.5.52 (XeLaTeX), Pandoc, Matplotlib/Seaborn.

- Models & assets: RS5M ViT-H-14 checkpoint (downloaded separately), CLIP prompt templates.

- DevOps: Git/GitHub, mirrored to QUB EEECS GitLab for assessment; render scripts for paper/slides.

- Hardware context: development and tests on Mac Studio M4 Max; training scripts parameterised for local or GPU environments.

3.1 Annotation App (Data Collection)

- Live URL (login): https://expert-flag-labeller-production.up.railway.app/login
- Purpose: expert labelling interface (Next.js + Supabase) used within the larger project to collect classifications and flags for review
- Demo credentials: provided in this supporting report submission on Canvas (not in slides/video)
- Access notes: for demonstration only; use demo/example imagery to avoid licensing issues; uptime not guaranteed — record a short clip if needed

3.1.1 Demo Credentials (Canvas-only)

Note: Do not commit credentials to public Git. Provide them to assessors via the Canvas supporting submission only.

- Username: Barry
- Password: 5aHt5wjEZ7Bbwnd/

4 Critical Appraisal

4.1 Rationale & Design Decisions

- Task reframing: Rather than claiming to "solve" long-tailed imbalance, we define a new, practically relevant task via hierarchical taxonomy guided by economic domain knowledge.
- Evaluation strategy: comparative granularity study (70/16/7) with robustness checks; attention analyses to aid interpretation.
- Data governance: licensing of street-level imagery prevents redistribution; the codebase ships with metadata, scripts, and derived figures only.

4.2 Lessons Learned

- Methodological hygiene: early issues (e.g., label mapping errors) emphasised the need for fixed seeds, split verification, and rigorous documentation.
- Framing matters: taxonomy design (not data engineering alone) can make tasks tractable and interpretable.

4.3 Limitations & Future Work

- External validity: images limited to NI context; temporal/geographic extensions planned.
- Multi-label and multi-task: promising directions to reflect real-world symbol co-occurrence and context.
- Dataset access: reproducibility focuses on scripts and metadata; raw imagery must be sourced under appropriate licenses.

5 Reproducibility & Pointers

5.1 Code & Repository

- QUB EEECS GitLab (private to assessors):
  URL: [insert GitLab repository URL here]
  Submission tag/commit: [insert tag/commit here]

5.2 Quick Commands

Render the paper and publish to submission folder:

```
cd MSc-Themed-Research-Project/write-up
./render_and_publish.sh
```

Render the demo slides (reveal.js):

```
cd MSc-Themed-Research-Project/write-up
quarto render thesis_demo_slides.qmd
open thesis_demo_slides.html
```

Run verification script (choose the one present in your submission bundle):

```
python verify_setup.py
# or
python qub-gitlab-submission/test_setup.py
```

6 Video Demo (10 minutes)

The recorded screen-capture demo includes:
1) Project title, name, student number, and an image of the presenter at the start.
2) Walkthrough of problem framing, dataset/task, method (hierarchy), results (70/16/7), and reproducibility steps.
3) Live run of non-heavy scripts (verification, figure regeneration, Quarto render).
4) Brief future work and closing.

See `write-up/video_demo_script.md` for a 6–8 minute script and timings.

7 Statement on AI Use

AI assistance was used in two limited ways:

1) Documentation support — to draft docstrings and inline comments for code that I authored.

2) UI scaffolding — for non-substantive scaffolding/boilerplate in the Next.js-based annotation app used for expert labelling within the larger project (which does not impose AI restrictions). This covered component layout and minor UI wiring only.

I implemented and verified the task logic, curation rules, and API endpoints myself, and I authored all ML experiments, analyses, and the manuscript text. Any AI-suggested wording was reviewed/edited by me.

8 Compliance Checklist

- ☒ Clear lifecycle description + V&V (this report)
- ☒ Tools used (this report)
- ☒ Critical appraisal (design rationale, lessons, evaluation)
- ☒ Code repository on QUB EEECS GitLab (link + tag provided)
- ☒ Visual material (recorded video demo with audio commentary)
- ☒ Paper PDF and slides generated via Quarto
- ☒ Declaration and AI-use statement included in manuscript