Reflection and Traceability Report on OCRacle

Phillip Tran

1 Changes in Response to Feedback

1.1 SRS and Hazard Analysis

Instructor feedback was addressed in a pull request. Changes include:

- Removed units from unitless symbols.

- Used symbolic constants to define the size of the input image and number of classification classes.

- Introduced name of program in Introduction.

- Added specific coursework necessary for intended reader to understand the document.

- Removed irrelevant technical user responsibilities.

- Clarified user characteristics.

- Removed system constraints.

- Defined all symbols used in the document.

- Clarified confusing sentence in Cross-Entropy Loss Function description.

- Added more specific for usability requirements.

- Revised maintainability requirements.

- Revised rationale to explain why the dataset was chosen.

- Fixed references for Theoretical Models.

- Clarified R2 to explain how the input image would be processed for classification.

- Split functional requirements into Training and Prediction.

- Moved some General Definitions (ReLU, Softmax) to Theoretical Models.

- Modified GD1 (Neural Network) to align with final implementation.

- Added greater details to Instance Models to clarify implementation.

The following peer review feedback was addressed:

- Suggestions for the Model: Issue closed for being out of scope for the project.

- Defining Accuracy: Explicit accuracy calculation was added to the documentation.

- Assumptions and FRs: Issue closed for not being a concern brought up by the instructor.

- SRS Minor Details: Addressed minor details pointed out by the reviewer.

1.2 Design and Design Documentation

Instructor feedback was addressed in a pull request. Changes include:

- MG: Update to the Uses Hierarchy to accurately reflect the architecture of the program.

- MIS: Starting all modules on a new page.

- MIS: Removing names from outputs.

- MIS: Fixing the environment variable for the Input Format Module.

- MIS: Added footnote to explain why the Input Format Module and Image Preprocessing Module are separate.

- MIS: Defined notation used to describe matrices.

- MIS: Changed LABEL from a set to a list to reflect the behavior of the program.

- MIS: Removed local functions from Image Preprocessing Module.

- MIS: Added GUI sketches to define the Graphical User Interface Module.

The following peer review feedback was addressed:

- Labels Sharing Confidence: The behavior of the argmax function was defined.

- Clearer Control: The MG and MIS documents were updated to reflect how the Application Module interacts with the other modules.

- Bicubic Interpolation Exceptions: Issue was closed for being out of scope, since computational errors in bicubic interpolation are not a concern for the project.

- Validating Input Value: The behavior of the normalization function was clarified to show how the program would avoid divide-by-zero errors.

- Centralized Config File: Issue was closed for being out of scope, though it was a great suggestion.

- Decoupling Optimizer and Training: Issue was closed since ADAM optimization and the training routine are too tightly coupled to be separated.

- Normalization Consistency: MIS document was corrected to correspond with the normalization behavior defined in the SRS.

1.3   VnV Plan and Report

Instructor feedback was addressed in a pull request. Changes include:

- Changed citation format from "cite" to "citep" in LaTeX.

- Gave a greater number of citations to existing documentation, including the MG and MIS.

- Added more detailed methods for code walkthroughs.

- Gave more explicit expectations of inputs where a single uppercase Latin alphabet character is expected.

- Defined the dataset to be used during testing.

- Clarified why the sum of the probability vector is equal to 1.

- Clarified T8 to define how the test would be automated.

- Clarified T10 in terms of expected test users and tasks.

- Clarified T13 in terms of what tests would be run to verify cross-platform compatibility.

- Added queries used for ChatGPT table generation.

-

The following peer review feedback was addressed:

- Comparison with OAR: The accuracy metrics used to compare the OCRacle project to OAR have been defined.

- Docker Config Details: Issue was closed since Docker was no longer being used for the project.

- Evaliating Command Line Skills: The MIT Missing Semester was added as the baseline knowledge of "basic command line skills."

- **PyTest in the Python Stlib?**: PyTest is not included in the Python standard library, so this mention was removed.

- **Specifying Known Correct Outputs**: Clarified what the known correct outputs would be for T4 and T9.

2   Challenge Level and Extras

2.1   Challenge Level

The challenge level of this project is considered to be general.

2.2   Extras

The extra for this project was a video code walkthrough. The code walkthrough explained the purpose of each module, briefly ran through the code, and detailed how each component contributes to the overall functionality of the program.