Ahmedabad University

CSE623 Machine Learning

Weekly Report 2(Group 9)

Section 1

Submitted to faculty: Prof. Mehul Raval

Date of Submission: 21/02/2026

Student Details

Roll No.	Name of the Student	Name of the Program
AU2340065	Bhumika Jain	Btech-CSE
AU2340152	Vishwa Patel	Btech-CSE
AU2340161	Devansh Gupta	Btech-CSE
AU2340246	Akkshat Shah	Btech-CSE

1. Detailed Summary of Work Done During the Current Week

The main tasks that were to be accomplished during the current week were to launch the actual implementation of the suggested stone-counting system by compiling a preliminary dataset and carrying out a first training experiment of the model. It was characterized by the purchase of a small collection of embroidery pictures, the implementation of the least amount of labeling, and the confirmation of the potential of a rudimentary end-to-end training system. Despite the fact that the system is currently not as accurate as it should be, these initial steps provided a working prototype and shed some light on the limitations linked to using a highly constrained dataset.

1.1 Dataset Preparation and Organization

The basic data points are gathered, sorted, and then stored in a table format .First, the basic data points are collected, classified, and after that, they are organized in a table format

- Using the sources available, a small collection of embroidery cloth images was built with the use of decorative stones.
- The images were organized in a simple directory tree to make retrieval convenient using preprocessing pipelines and training pipelines.
- The baseline quality assessments were carried out to determine appropriateness to initial experimentation, to evaluate the quality of resolution and clarity.

1.2 Basic Annotation / Labelling

- Not much annotation of the identified pictures was conducted to facilitate downstream model training.

Depending on the training methodology chosen, the stone regions were drawn out, or binary classification (stone / non-stone) was used.

- The annotation was intended to give enough labelled examples to evaluate a prototype as opposed to creating a high-quality, high-fidelity dataset at this point.

1.3 First Training Pipeline Configuration.

● Integrated a basic pipeline that enabled the reception of the refined image dataset and its annotations into the training programme.

● Preprocessing schemes such as resizing and normalisation to normalise the dimensions and intensities of images to be ingested by the model.

● Now known seamless data propagation between raw imagery to model input and, therefore, error-free transitions become affordable.

1.4 Initial Model Training

● Trained a first model case on the constrained labeled corpus, using the chosen machine-learning procedure of stone or non-stone discrimination or counting.

● Ensured that the training pipeline runs in end-to-end mode without performance failure.

● Noticed that due to the lack of data, the learning ability of the model is constrained to a great extent and makes it highly unstable.

1.5 Behaviour and Observations (Undercounting / Overcounting)

● Assessment of the trained model on the same compact data demonstrated poor-optimal accuracy in counting stones.

● The model, in certain specific cases, underestimates the quantity of stones, which ignores a number of valid inclusions in the embroidery.

● On the other hand, some of them also have overcounting with background motifs or embroidery/zari motifs mistakenly considered as stones.

● The results highlight the need to use a larger data sample, better quality annotations, and continuous refinement to improve the results.

1.6 Key Learnings from the Week

● Sparse and annotated data is insufficient to rely on when counting stones.

● Even crude prototypes can be used to shed light on the practical problems, such as misclassifications and inaccuracy of counting, at an initial stage of the project.

● Future initiatives must focus on dataset growth and a greater level of data quality before engaging in a more advanced modelling.

(Underfitting)

(Overfitting)

2. Activities to be undertaken in the Week ahead

The next week will also be the mid-semester examination week hence limiting our capacity to do our project-based activities because of the high examination preparation and other academic tasks.

2.1 The minimal data set maintenance is accomplished

● Prepare a brief listing of the images accrued, with attached basic descriptors of what they are.

2.2 Review and Documentation

● Revise the initial results of training and observation during this week to clarify the under-counting and over-counting trends.

● Save the current prototype pipeline setup and major insights in a README file.

2.3 Planning to go through Week Post-Mid Sem.

● Spell out precise specifications regarding the future larger dataset gathering- such as the number of target images, and annotation procedures.

● Research on quick annotation tools/methodologies that might help tabulate labels faster when examinations are done.

2.4 Light Debugging (When time allows)

● In case there are some short durations between tests, test the current pipeline on one or two more sample images to determine consistency.

2.5 Academic Priority

● Continue to be primarily concerned with mid-semester examination preparation in order to retain academic status.

● Start the complete implementation of projects as soon as examinations are made, with an extended dataset and improved annotation.