Bhav Beri
2021111013
bhav.beri@research.iiit.ac.in

Analog Clock - Time Reader

(A Computer Vision Approach using Neural Networks)

A Project by the CVIT Team

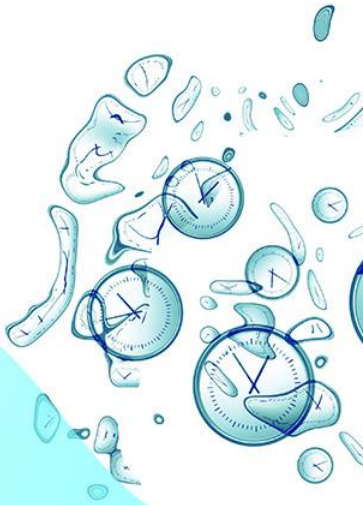
Introduction

Analog clocks continue to be widely used in various settings, necessitating an efficient method for extracting time information from clock images.

Our research project focuses on developing an intelligent system using *computer vision and neural networks* to accurately read time from analog clock images.

Objective

- Extract and interpret time
- Multi-layer Perceptron (MLP) / Convolutional Neural Networks (CNN)
- Different Learning Rates & Batch Sizes
- Aim for Adaptability & Accuracy



Dataset

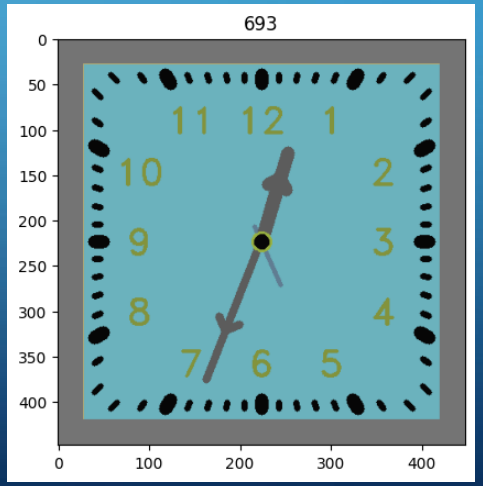
Training

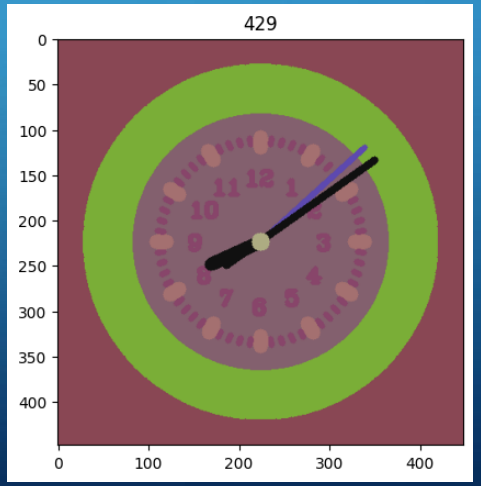
- Synthetic Dataset (Custom Made)
  - Various Backgrounds
  - Different Clock Hands and Text Variants

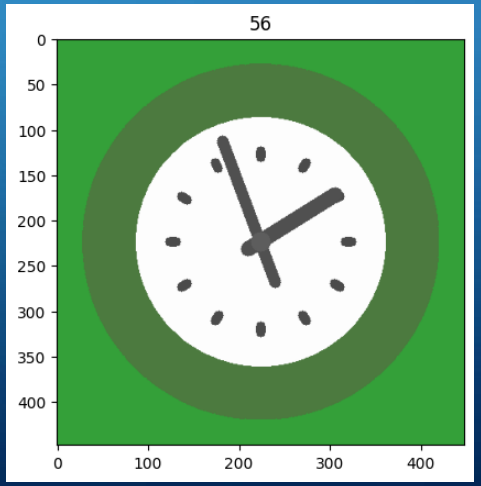
Testing

- Synthetic Dataset
- COCO Dataset
- Open-Images Dataset

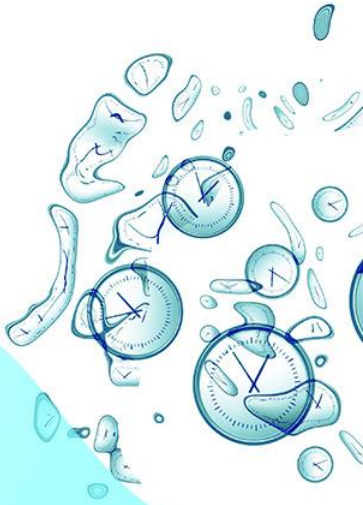
Demo Images







Experimental Setup



- Pre-processing of the images
- Implement the Multi-layer Perceptron (MLP) and Convolutional Neural Networks (CNN) architectures.
- Use different learning rates and no of epochs
- Employ rigorous evaluation metrics

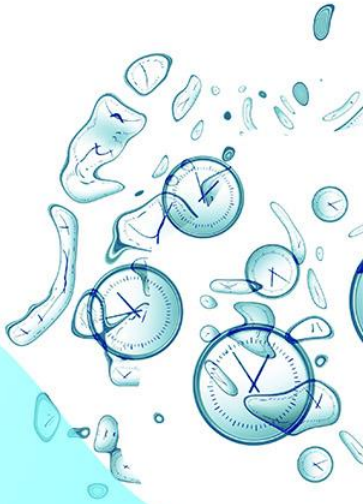
Neural Network Layers Information

- MLP -> `Linear(448 * 448 * 3, 256) + Linear(256, 720)`
- CNN -> `(Conv2D) {224 x 224} + {112 x 112} + {56 x 56} + {8 x 8} + {4 x 4}`

(Optimizer Used -> Adam)

RESULTS

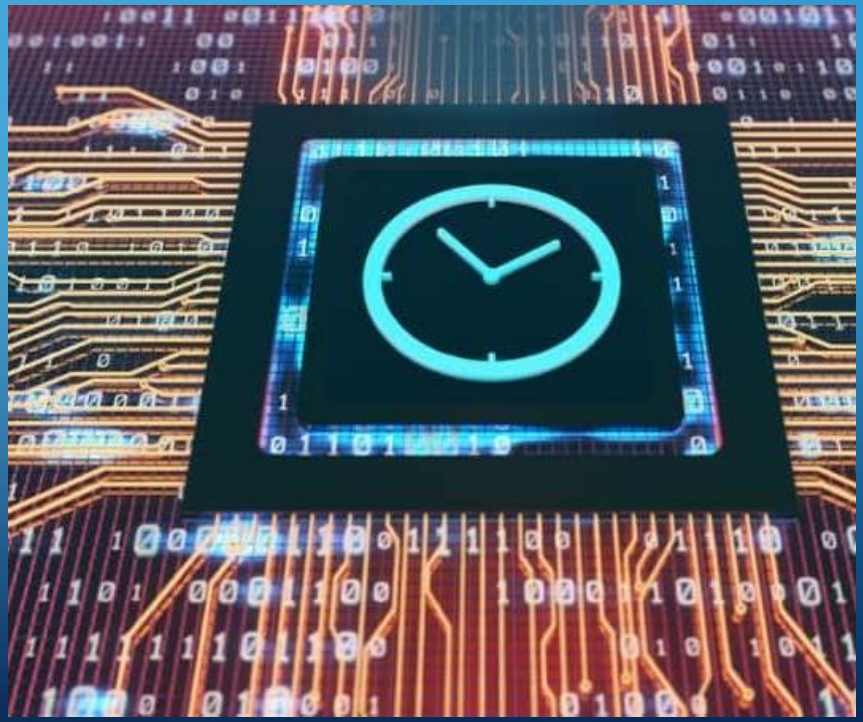
- MLP with almost all the configs (learning rates and epochs) give us low accuracy scores.
- CNN was able to provide good results with batch size of 64 for 20k iterations and learning rate of 1e-3.



DATASET	1	1-H	1-M	2	3
COCO	80.4	86.9	84.4	84.3	86.5
Open-Images	77.3	83.5	81.9	81.5	84.6
Synthetic	79.0	82.3	82.4	83.3	85.5

We report the accuracy on the three datasets. 1, 2, 3 are the top 1, 2, and 3 overall accuracies respectively, whereas 1-H and 1-M are the top-1 hour and minute accuracies.

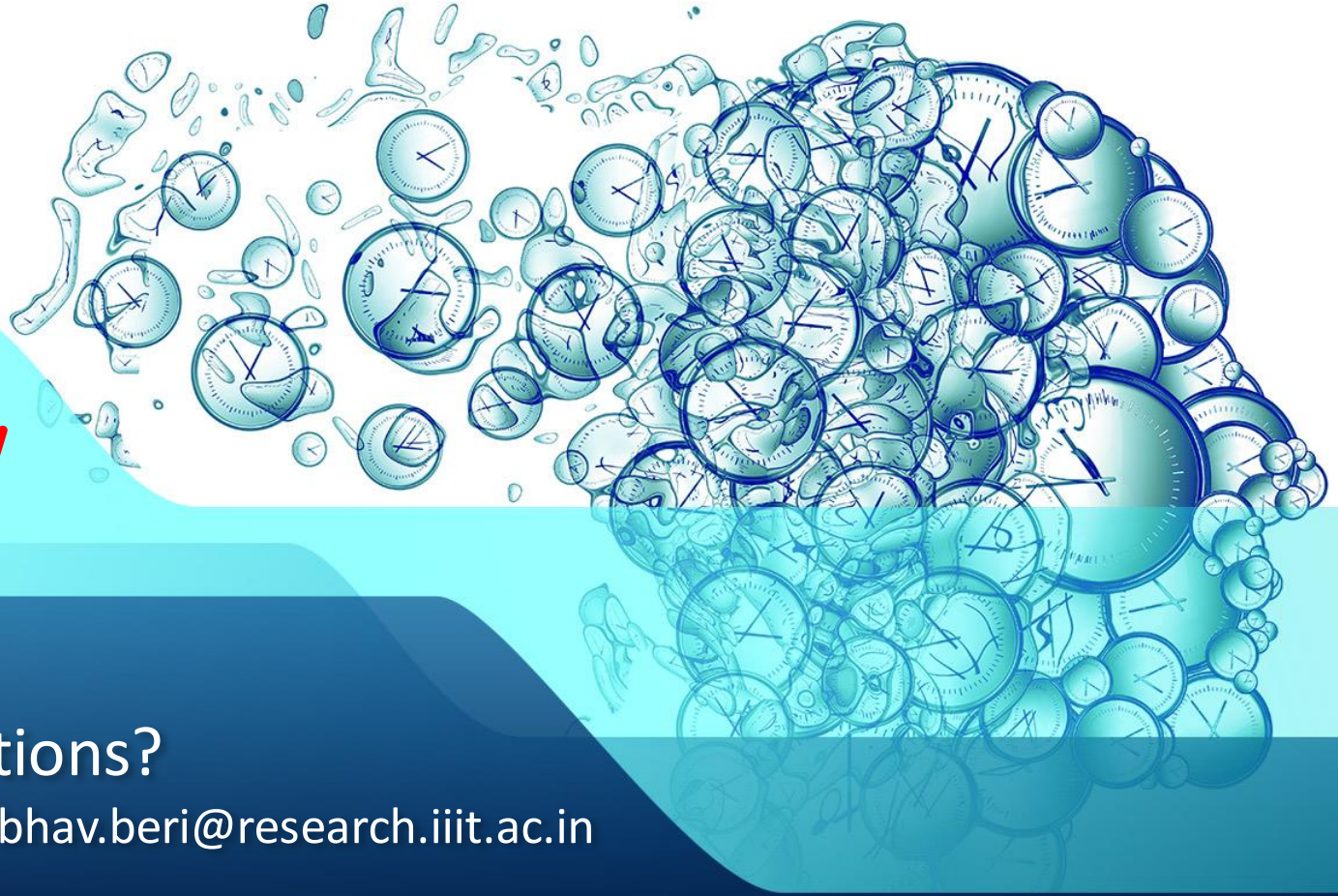
CONCLUSION



- ✓ Introduce a framework for clock reading
- ✓ Make a synthetic dataset generator
- ✓ Compared various configurations for the neural network
- ✓ Benchmarking on three datasets

In Future, we expect this to be taken forward, with clock reading as a pre-processing step in image processing.

Further research can be continued on getting better models with > 90 % accuracy.



THANKS !

Any Questions?

Mail them to bhav.beri@research.iiit.ac.in

PLEASE NOTE

Some data about the results and the setup is based on the original paper which was used for reference for our project, and thus is not exactly what we did in the session for this project (which was pretty-less).
Paper URL -> https://arxiv.org/pdf/2111.09162.pdf