rEPORT

DIGITS RECOGNIZER KAGGLE

INTRODUCTION

The project features an ensemble of different neural network architectures which are designed for handwritten digits recognition. By combining different architectures which includes - Residual Neural Network(ResNet), Dense Neural Network(DenseNet), Vision Transformer(ViT) and Hybrid Vision Transformer(Hybrid ViT) along with data augmentation yields an accuracy of 99.62% on Kaggles Public Score for the test data.

TRAINING DATA

The digits dataset used for this project consists of grayscale images provided
in kaggle train dataset representing numerical digits. The dataset was preprocessed and augmented to increase the diversity of the training data and to improve the generalizability of the models.
Data augmentation techniques include:

- Random rotations – 10 degrees
- Horizontal and vertical shift – 0.1*
- Scaling and cropping – 0.1*

Data augmentation applied over 33600 train images (42000*0.2) with apprx. 50-100 epochs allows each model to be train on over 1.6-3.3 million unique digit images.!
Training the models along with the ReduceLROnPlateau learning rate scheduler for better convergence of the loss function.

MODEL ARCHITECTURE & ENSEMBLING

The Architecture involves Custom Implementation and independent training of ResNet, DenseNet, ViT and Hybrid ViT followed by the Ensembling procedure - weighted summation of predictions from the models.

KEY FEATURES

1. ResNet
    - Single Layer Skip Block
2. DenseNet
    - Dense Block Layers: 4
    - Growth Rate : 32 channels/layer
3. ViT
    - Transformer Blocks w/ Skip Connections : 3
    - Positional + Patch Embeddings included in Custom Tensorflow Layer

4 Hybrid ViT
    - Feature Extractor Model : pre-made custom DenseNet Model
    - Positional + CNN Feature Embeddings
    - Transformer Blocks : 1 w/ 8 Attention Heads

Results

Model	Accuracy
ResNet	99.01%
DenseNet	98.56%
ViT	96.85%
Hybrid	99.54%
Ensemble	99.63%

Precision, Recall, F1 Score : 0.99% for all Models

Better Accuracy for predictions can be obtained by increasing the models in the ensemble procedure