Performance Comparison of various Feature Extractors on MNIST dataset

Eshan Gujarathi
Department of Computer Science and Engineering
IIT Gandhinagar

Introduction

The Dataset

- Handwritten digits
- Size normalized and centered
- Training Data: 60,000 images
- Testing Data: 10,000 images
- Image size: 28x28

Applications:
- Bank cheque processing
- Form data entry

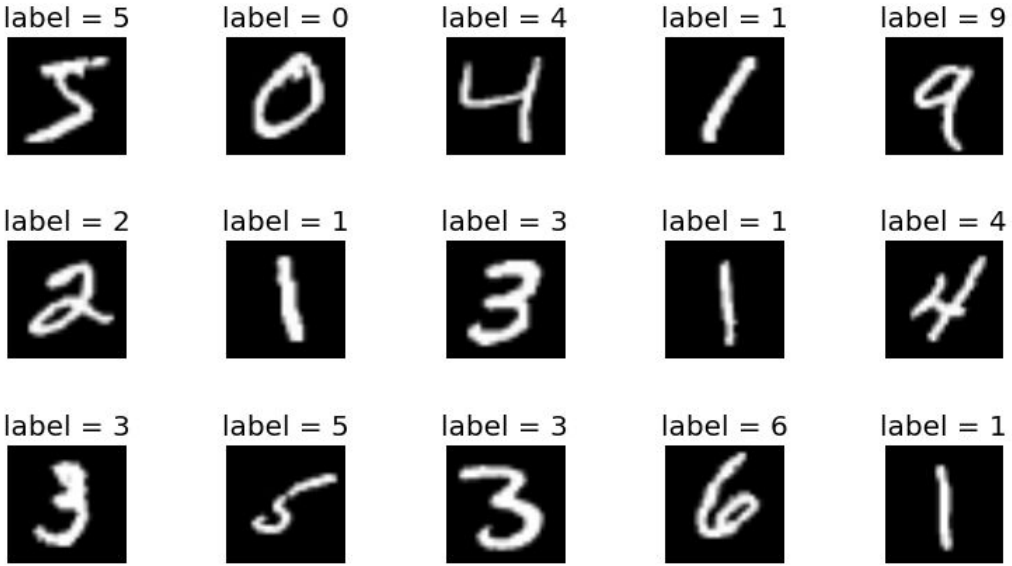


Image Source

Models

1. Fully Connected Network
2. Convolutional Neural Network (ConvNet/CNN)
3. AlexNet (Winner of 2012 ImageNet Challenge)
4. VGG (Runners Up of 2014 ImageNet Challenge)
5. ResNet (Winner of 2015 ImageNet Challenge)

Training

Fully Connected Network

- Images flattened
- Epochs = 30
- Validation split = 0.2

Layer (type)	Output Shape	Param #
dense (Dense)	(None, 1000)	785000
dropout (Dropout)	(None, 1000)	0
dense_1 (Dense)	(None, 784)	784784
dropout_1 (Dropout)	(None, 784)	0
dense_2 (Dense)	(None, 400)	314000
dropout_2 (Dropout)	(None, 400)	0
dense_3 (Dense)	(None, 200)	80200
dropout_3 (Dropout)	(None, 200)	0
dense_4 (Dense)	(None, 10)	2010

Total params: 1,965,994
Trainable params: 1,965,994
Non-trainable params: 0

Convolutional Neural Network (ConvNet)

- Epochs = 30
- Validation split = 0.2

Layer (type)	Output Shape	Param #
conv2d (Conv2D)	(None, 24, 24, 28)	728
max_pooling2d (MaxPooling2D)	(None, 12, 12, 28)	0
conv2d_1 (Conv2D)	(None, 8, 8, 56)	39256
max_pooling2d_1 (MaxPooling2D)	(None, 4, 4, 56)	0
flatten (Flatten)	(None, 896)	0
dense (Dense)	(None, 56)	50232
dense_1 (Dense)	(None, 10)	570

Total params: 90,786
Trainable params: 90,786
Non-trainable params: 0

AlexNet

- Epochs = 10
- Validation split = 0.2
- Images resized to = (32*32*3)
- Output layer with 10 classes

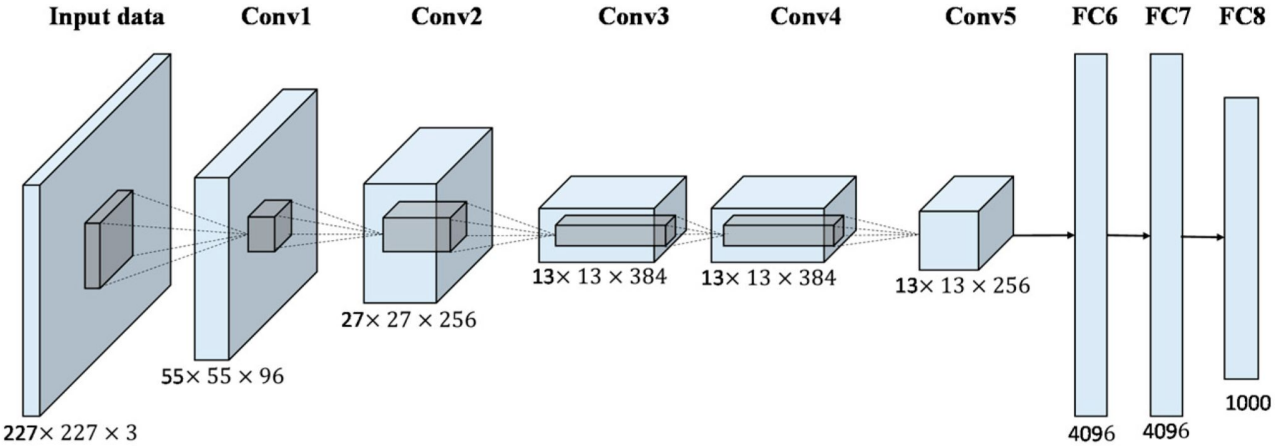


Image source

VGG-16

- Epochs = 10
- Validation split = 0.2
- Images resized to = (48*48*3)
- Used Transfer Learning
- Output layer with 10 classes

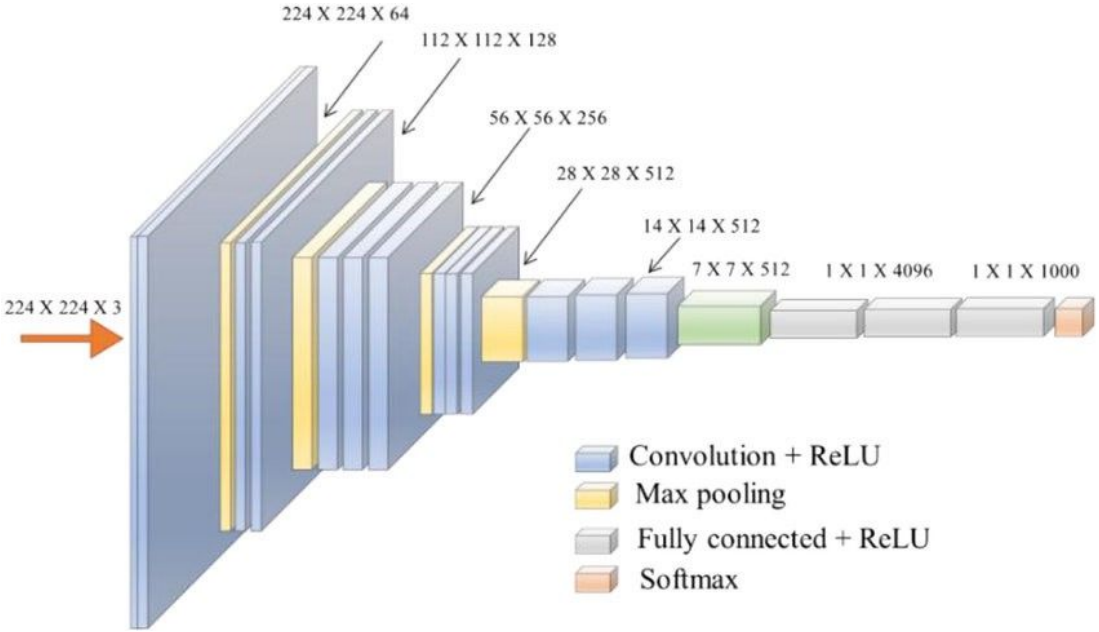


Image source

ResNet-50

- Epochs = 2
- Validation split = 0.2
- Images resized to = (32*32*3)
- Output layer with 10 classes

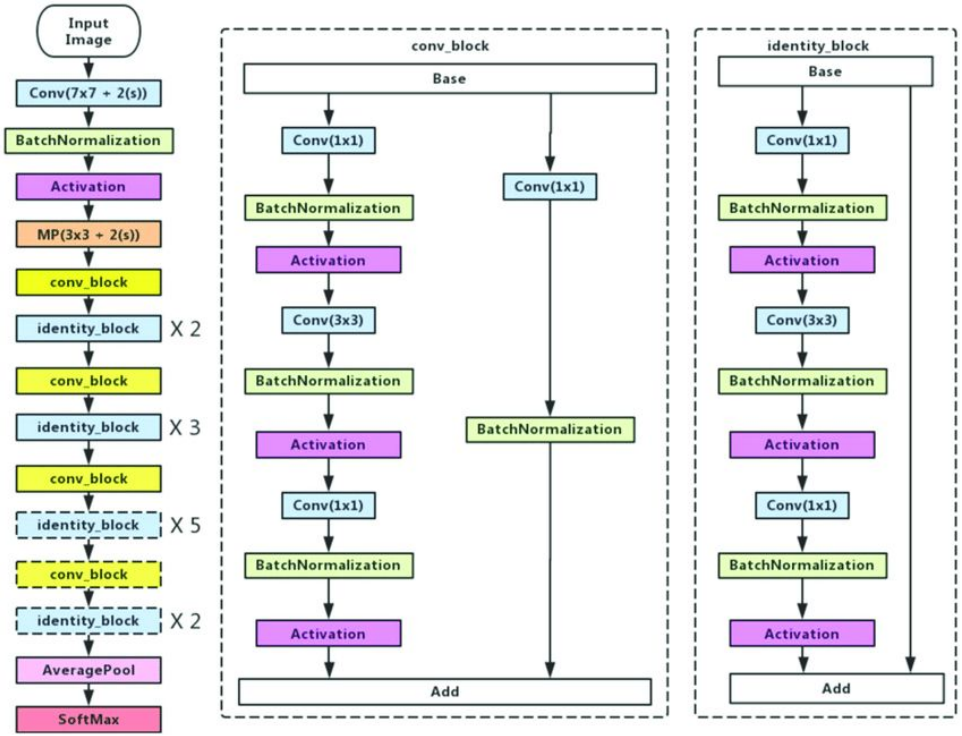


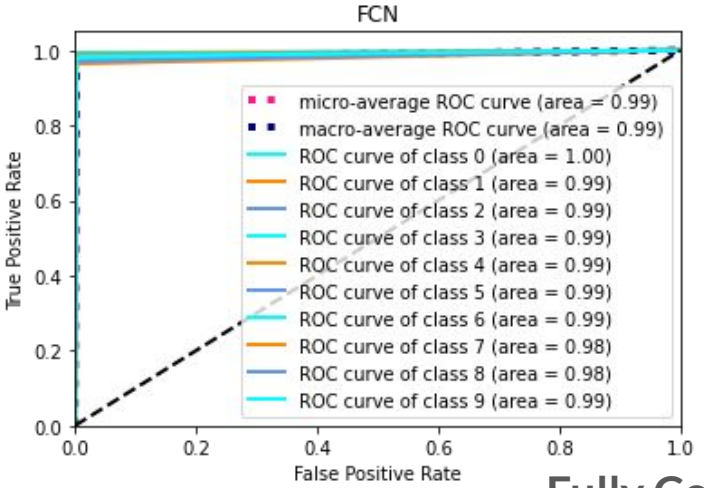
Image source

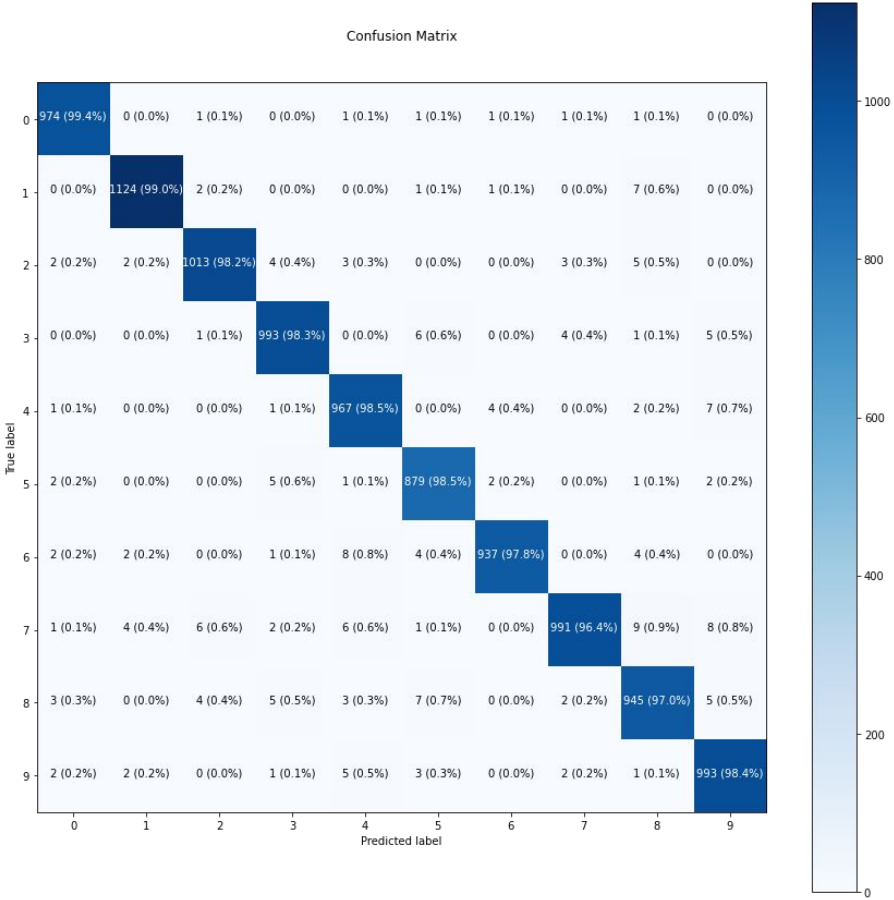
Performance Results

Evaluation Metrics

- Accuracy
- Confusion Matrix
- Precision, Recall and F1 score for each class
- ROC curve and AUC

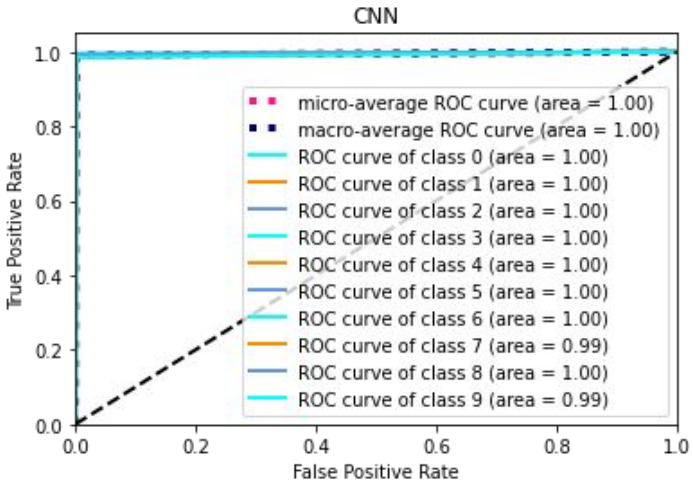
	precision	recall	f1-score	support
0	0.99	0.99	0.99	980
1	0.99	0.99	0.99	1135
2	0.99	0.98	0.98	1032
3	0.98	0.98	0.98	1010
4	0.97	0.98	0.98	982
5	0.97	0.99	0.98	892
6	0.99	0.98	0.98	958
7	0.99	0.96	0.98	1028
8	0.97	0.97	0.97	974
9	0.97	0.98	0.98	1009
accuracy			0.98	10000

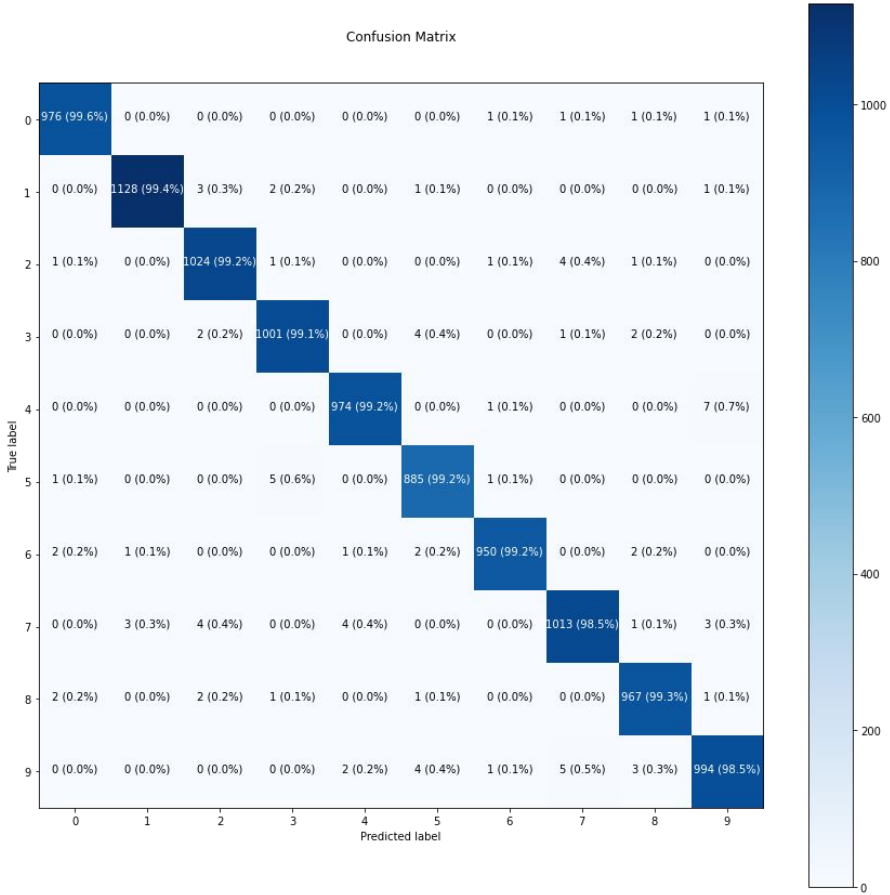




Fully Connected Network

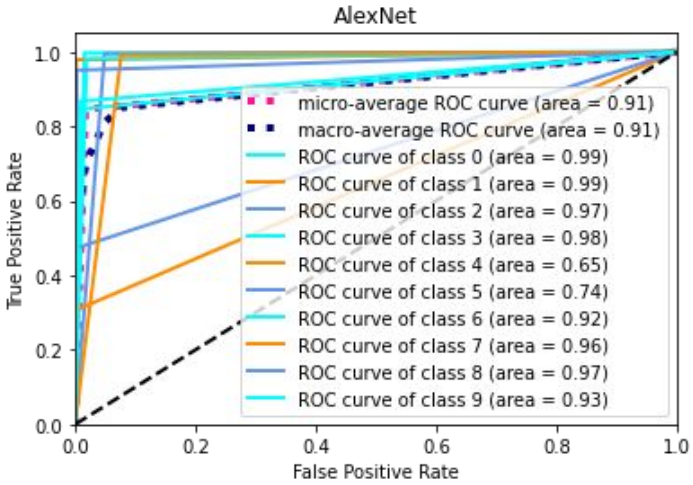
	precision	recall	f1-score	support
0	0.99	1.00	0.99	980
1	1.00	0.99	1.00	1135
2	0.99	0.99	0.99	1032
3	0.99	0.99	0.99	1010
4	0.99	0.99	0.99	982
5	0.99	0.99	0.99	892
6	0.99	0.99	0.99	958
7	0.99	0.99	0.99	1028
8	0.99	0.99	0.99	974
9	0.99	0.99	0.99	1009
accuracy			0.99	10000

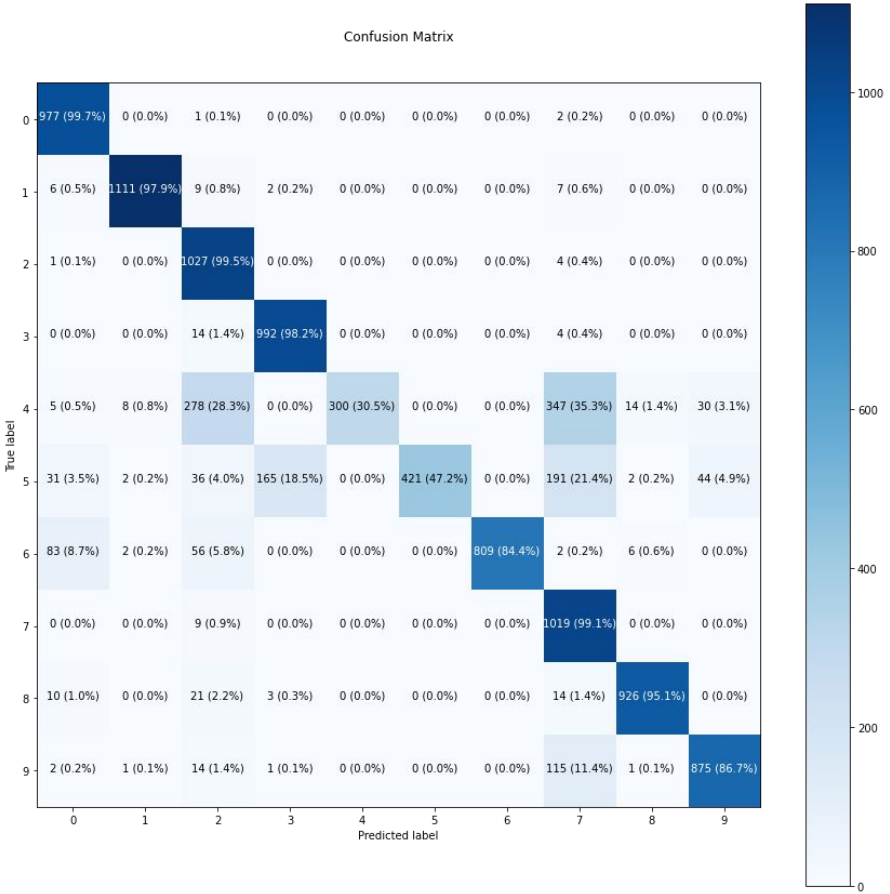




ConvNet

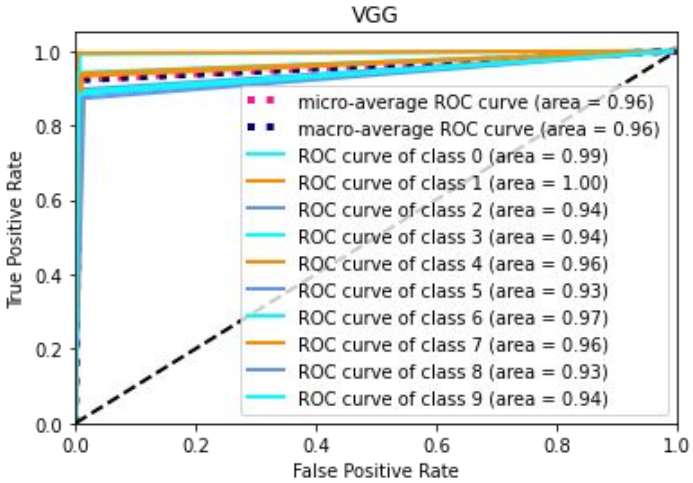
	precision	recall	f1-score	support
0	0.88	1.00	0.93	980
1	0.99	0.98	0.98	1135
2	0.70	1.00	0.82	1032
3	0.85	0.98	0.91	1010
4	1.00	0.31	0.47	982
5	1.00	0.47	0.64	892
6	1.00	0.84	0.92	958
7	0.60	0.99	0.75	1028
8	0.98	0.95	0.96	974
9	0.92	0.87	0.89	1009
accuracy			0.85	10000

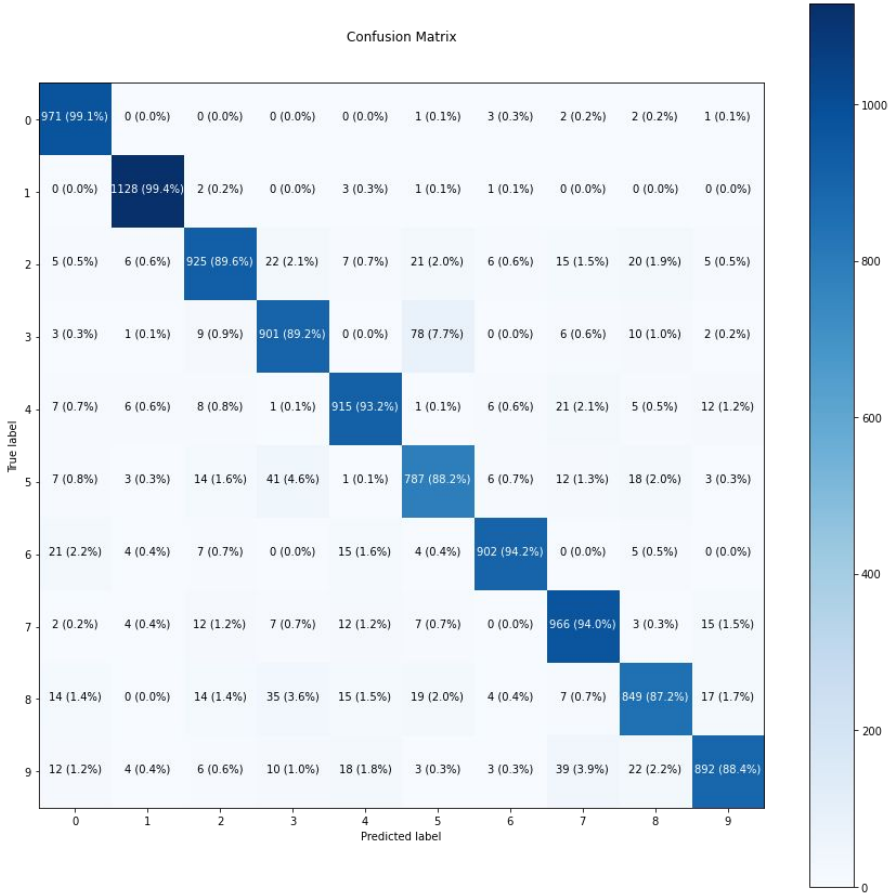




AlexNet

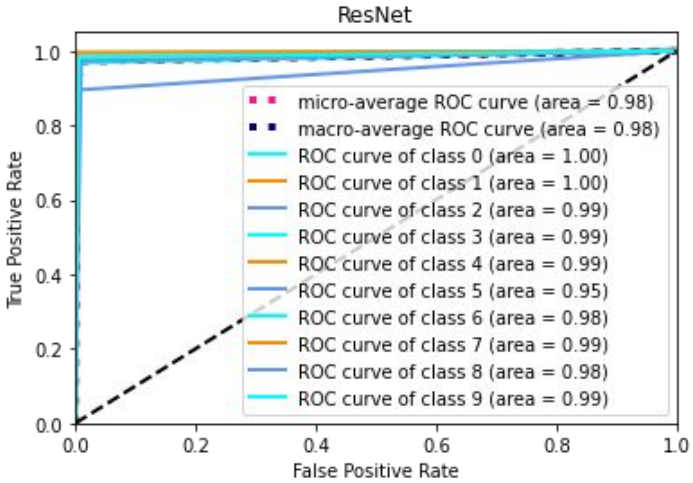
	precision	recall	f1-score	support
0	0.93	0.99	0.96	980
1	0.98	0.99	0.98	1135
2	0.93	0.90	0.91	1032
3	0.89	0.89	0.89	1010
4	0.93	0.93	0.93	982
5	0.85	0.88	0.87	892
6	0.97	0.94	0.96	958
7	0.90	0.94	0.92	1028
8	0.91	0.87	0.89	974
9	0.94	0.88	0.91	1009
accuracy			0.92	10000

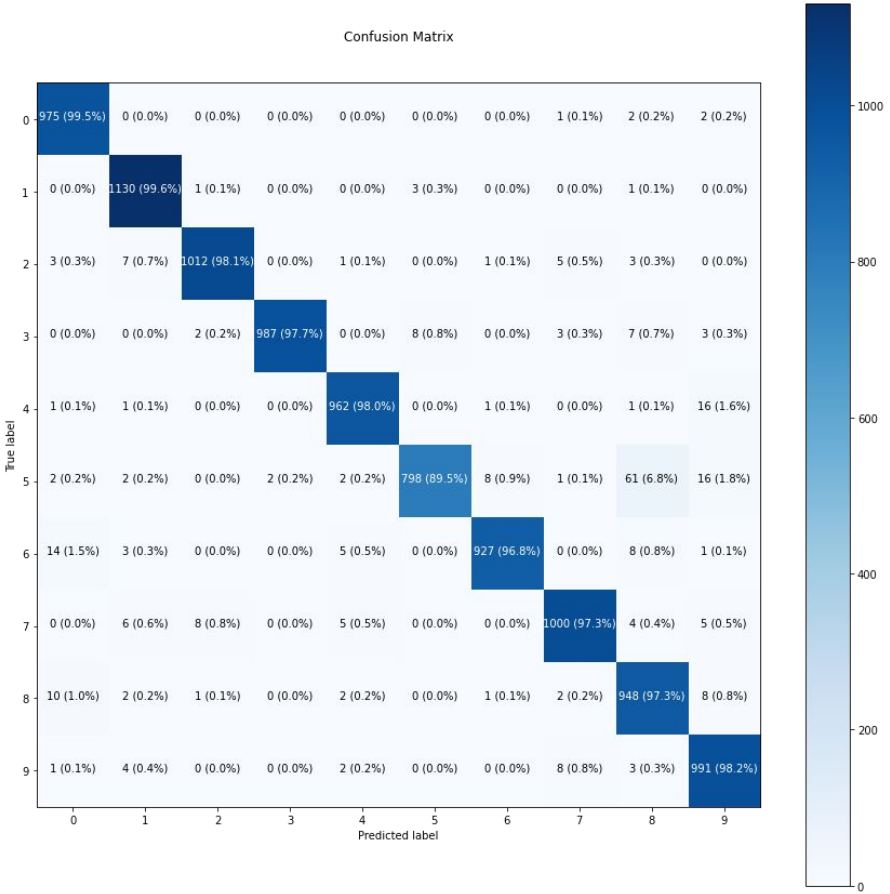




VGG

	precision	recall	f1-score	support
0	0.97	0.99	0.98	980
1	0.98	1.00	0.99	1135
2	0.99	0.98	0.98	1032
3	1.00	0.98	0.99	1010
4	0.98	0.98	0.98	982
5	0.99	0.89	0.94	892
6	0.99	0.97	0.98	958
7	0.98	0.97	0.98	1028
8	0.91	0.97	0.94	974
9	0.95	0.98	0.97	1009
accuracy			0.97	10000





ResNet

Conclusions

- ConvNet performed the best with an accuracy of 99.12%
- Models considered to be state-of-the-art performed poorly on MNIST
- Dataset was relatively simpler for such models.
- The size of the dataset was also small to train models with more number of layers.

Key Learnings