A comprehensive report on the experiments performed on the old dataset

Experiment – 1

Running the previously generated Deep Learning model on the new dataset after creating HSV, LAB versions of the images and extracting the statistical features (mean, standard deviation, min, max and variance).

Architecture used:

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	[(None, 45)]	0
dense (Dense)	(None, 40)	1840
dense_1 (Dense)	(None, 20)	820
dropout (Dropout)	(None, 20)	0
dense_2 (Dense)	(None, 3)	63

Output (Non quantized):

Epochs: ***150***

```
Test loss:  0.31195950508117676
Test accuracy:  0.8999999761581421
```

Output (After Quantization Aware Training):

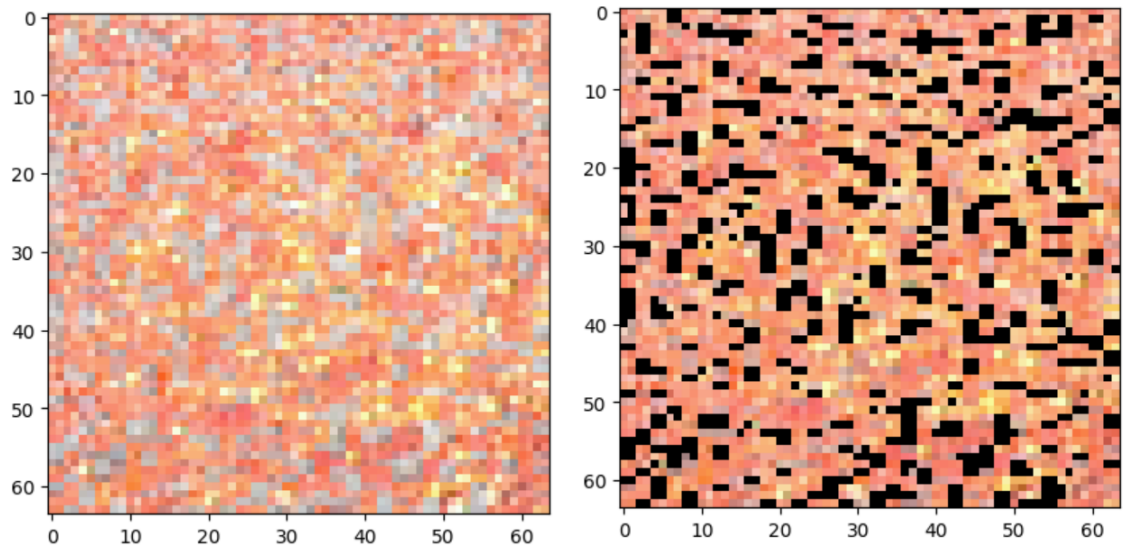
Epochs: ***150*** *(Normal Training) +* ***1*** *(Quantization Aware Training)*

```
Test loss:  0.3207307457923889
Test accuracy:  0.8999999761581421
```

Experiment – 2

Running the previously generated Deep Learning model using the old images after background subtraction

Architecture used:



Output (Non quantized):

Epochs: ***150***

Accuracy: ***93.33%***

Ouput (After Quantization Aware Training):

Epoch: ***150*** *(Normal Training) +* ***2*** *(Quantization Aware Training)*

Accuracy: ***93.33%***

Experiment – 3

Running the previously generated Deep Learning model using the new phase 2 images after background subtraction

```
_________________________________________________________________
 Layer (type)                 Output Shape              Param #
=================================================================
 input_1 (InputLayer)         [(None, 45)]              0

 dense (Dense)                (None, 40)                1840

 dense_1 (Dense)              (None, 20)                820

 dropout (Dropout)            (None, 20)                0

 dense_2 (Dense)              (None, 3)                 63

=================================================================
Total params: 2,723
Trainable params: 2,723
Non-trainable params: 0
_________________________________________________________________
```

Output (Non quantized):

Epochs: ***150***

```
Loss:  0.3499535322189331
Accuracy:  0.8888888955116272
```

Ouput (After Quantization Aware Training):

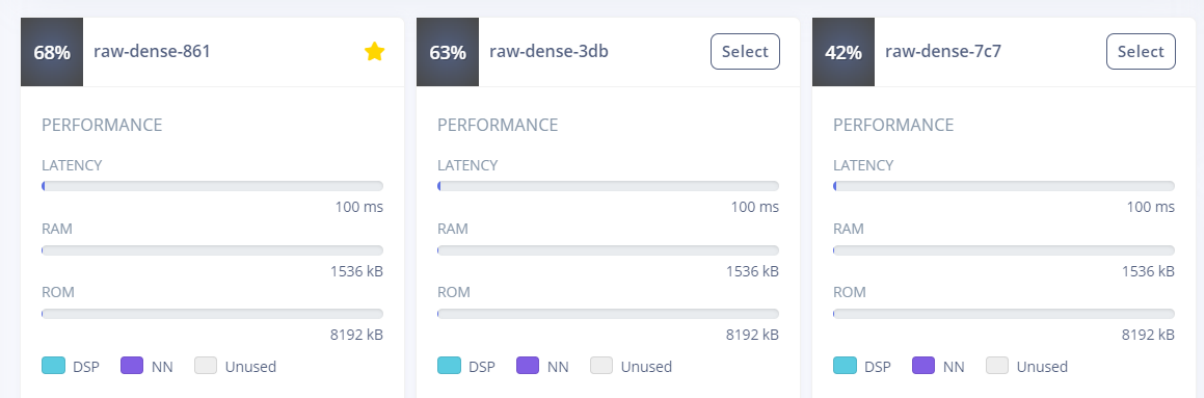
Epoch: ***150*** *(Normal Training) +* ***2*** *(Quantization Aware Training)*

```
Loss:  0.33639606833457947
Accuracy:  0.8777777552604675
```

Experiment – 4

Running the Eon Tuner on the newly phase 2 images after extracting the statistical features

Top 3 model:



The architecture of the top model:

Confusion matrix

	0	1	2
0	**100%**	0%	0%
1	0%	42.9%	**57.1%**
2	0%	24.1%	**75.9%**
F1 SCORE	1.00	0.51	0.66

Data exploration graph

Inference: Day_1 images are easily separated from the rest of the images by the model while both Day_2 and Day_3 images are not very well distinguished

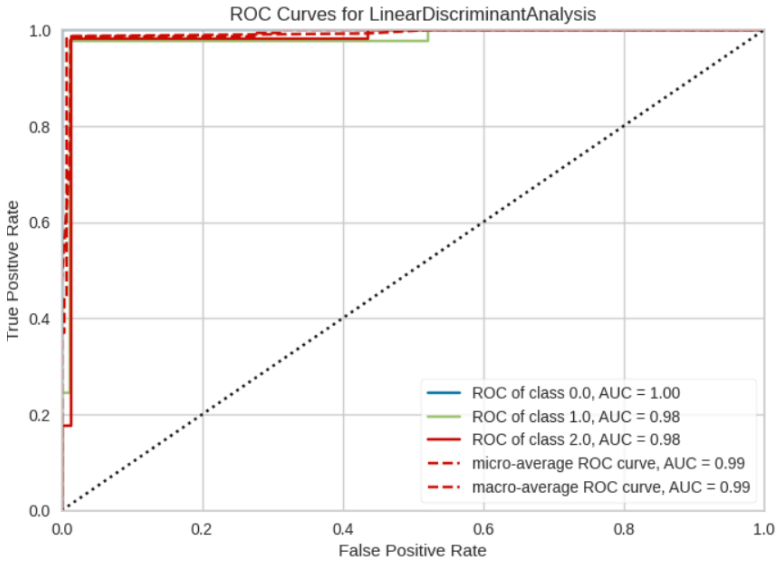
Experiment – 5

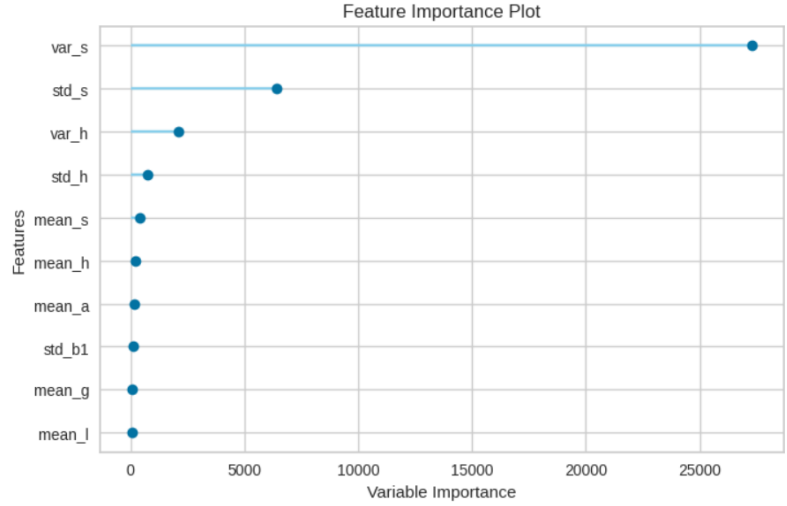
Various AutoML libraries which use only ML algorithms (NO Deep Learning) were experimented with using the new dataset and the results of each are listed below

1. ***PyCaret***
   This library uses one of these models as the final selection
   1. scikit-learn
   2. XGBoost
   3. LightGBM
   4. CatBoost

5. spaCy
6. Optuna
7. Hyperopt
8. Ray

	Model	Accuracy	AUC	Recall	Prec.	F1	Kappa	MCC	TT (Sec)
lda	Linear Discriminant Analysis	0.9650	0.9816	0.9650	0.9678	0.9648	0.9463	0.9479	
lr	Logistic Regression	0.9618	0.9939	0.9618	0.9645	0.9616	0.9414	0.9431	
qda	Quadratic Discriminant Analysis	0.9586	0.9681	0.9586	0.9628	0.9581	0.9360	0.9388	
et	Extra Trees Classifier	0.9521	0.9821	0.9521	0.9562	0.9516	0.9263	0.9290	
ridge	Ridge Classifier	0.9431	0.0000	0.9431	0.9454	0.9429	0.9125	0.9139	
xgboost	Extreme Gradient Boosting	0.9428	0.9835	0.9428	0.9475	0.9424	0.9124	0.9154	
lightgbm	Light Gradient Boosting Machine	0.9427	0.9858	0.9427	0.9450	0.9419	0.9120	0.9140	
rf	Random Forest Classifier	0.9426	0.9842	0.9426	0.9466	0.9423	0.9119	0.9145	
gbc	Gradient Boosting Classifier	0.9237	0.9772	0.9237	0.9262	0.9234	0.8828	0.8845	
dt	Decision Tree Classifier	0.8822	0.9064	0.8822	0.8890	0.8820	0.8200	0.8239	
knn	K Neighbors Classifier	0.8087	0.9190	0.8087	0.8102	0.8076	0.7057	0.7077	
svm	SVM - Linear Kernel	0.6775	0.0000	0.6775	0.6602	0.6319	0.5150	0.5699	
ada	Ada Boost Classifier	0.6657	0.7737	0.6657	0.4784	0.5458	0.4653	0.5869	
dummy	Dummy Classifier	0.4173	0.5000	0.4173	0.1743	0.2459	0.0000	0.0000	
nb	Naive Bayes	0.3952	0.6640	0.3952	0.3378	0.3106	0.1483	0.2087	





2. ***TPOT***

For a sample of 5 generations, the below shows the cross-validation score

```
Generation 1 - Current best internal CV score: 0.9940298507462686

Generation 2 - Current best internal CV score: 0.9940298507462686

Generation 3 - Current best internal CV score: 0.9940298507462686

Generation 4 - Current best internal CV score: 0.9940298507462686

Generation 5 - Current best internal CV score: 0.9940298507462686

Best pipeline: MLPClassifier(PCA(input_matrix, iterated_power=3, svd_solver=randomized), alpha=0.01, learning_rate_init=0.001)
0.9823008849557522
```

3. ***FLAML***

This library gives support to the following machine learning models
['rf', 'lgbm', 'catboost', 'xgboost', 'extra_tree', 'xgb_limitdepth', 'lrl1']

```
[flaml.automl.logger: 04-26 02:24:06] {2619} INFO - retrain xgb_limitdepth for 0.9s
[flaml.automl.logger: 04-26 02:24:06] {2622} INFO - retrained model: XGBClassifier(base_score=None, booster=None, callbacks=[],
              colsample_bylevel=0.4814471959023239, colsample_bynode=None,
              colsample_bytree=0.6050207253592859, early_stopping_rounds=None,
              enable_categorical=False, eval_metric=None, feature_types=None,
              gamma=None, gpu_id=None, grow_policy=None, importance_type=None,
              interaction_constraints=None, learning_rate=0.07962498837600937,
              max_bin=None, max_cat_threshold=None, max_cat_to_onehot=None,
              max_delta_step=None, max_depth=2, max_leaves=None,
              min_child_weight=0.0068282719220722, missing=nan,
              monotone_constraints=None, n_estimators=464, n_jobs=-1,
              num_parallel_tree=None, objective='multi:softprob',
              predictor=None, ...)
[flaml.automl.logger: 04-26 02:24:06] {1930} INFO - fit succeeded
[flaml.automl.logger: 04-26 02:24:06] {1931} INFO - Time taken to find the best model: 192.90792417526245
```