

Machine Learning and Multispectral Unmanned Aerial Vehicle Imagery Data for Agriculture.

Georgios Batsis

<https://github.com/gbatsis/VYSegML>

Image Segmentation

Image segmentation → Image is broken down into subgroups (Segments) → Assigning labels to pixels

Image segmentation – Deep Learning →

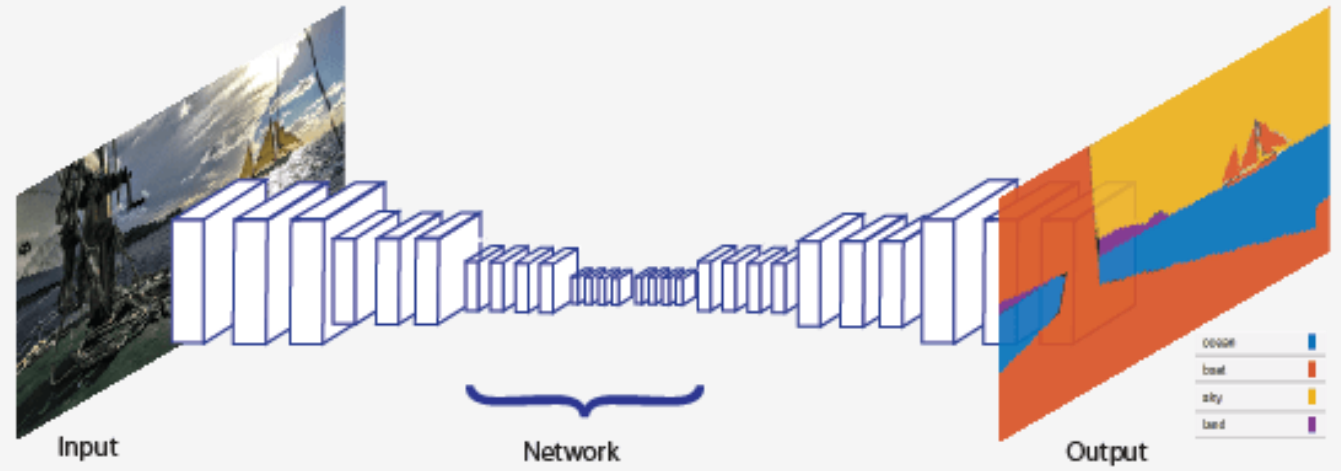
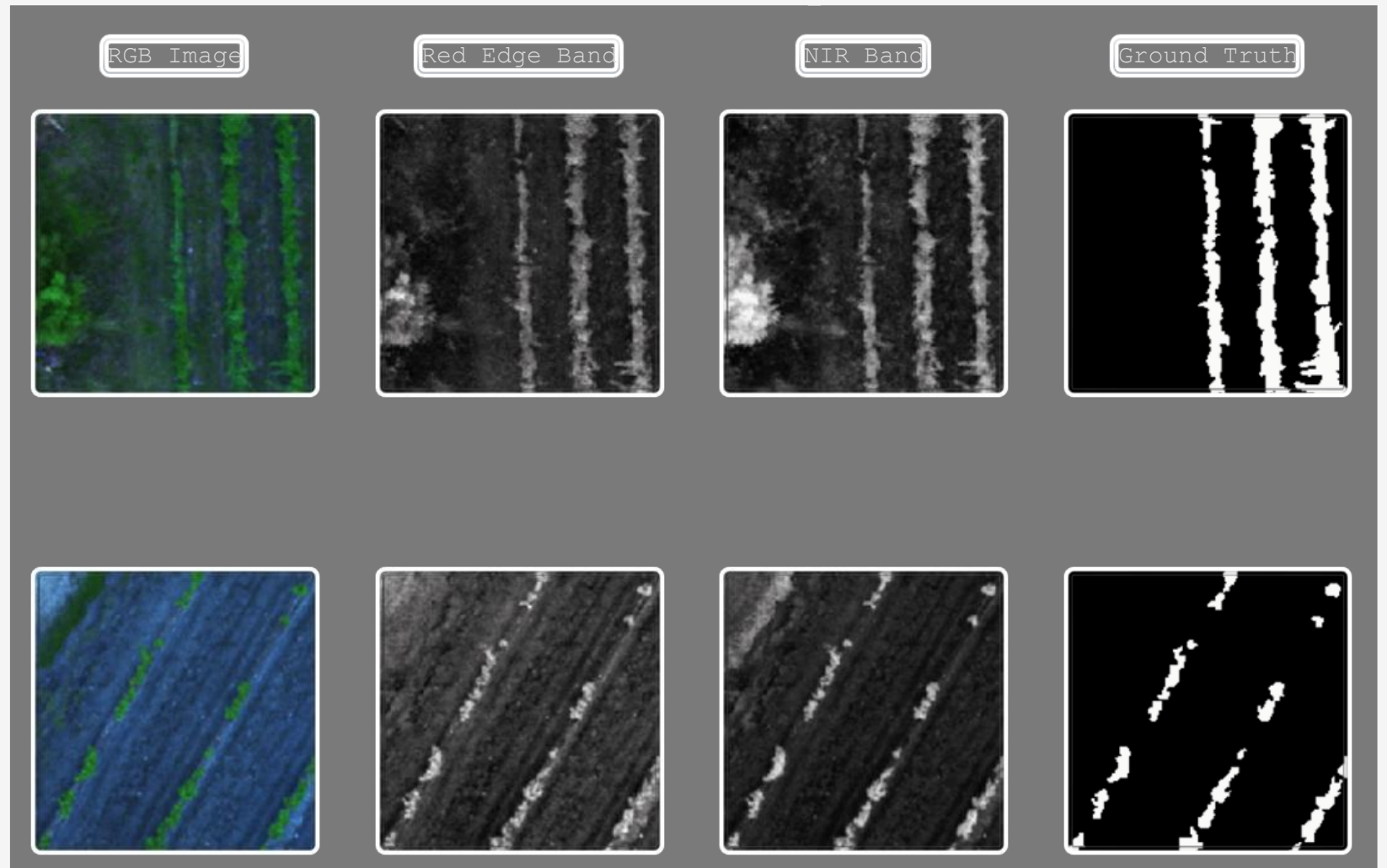
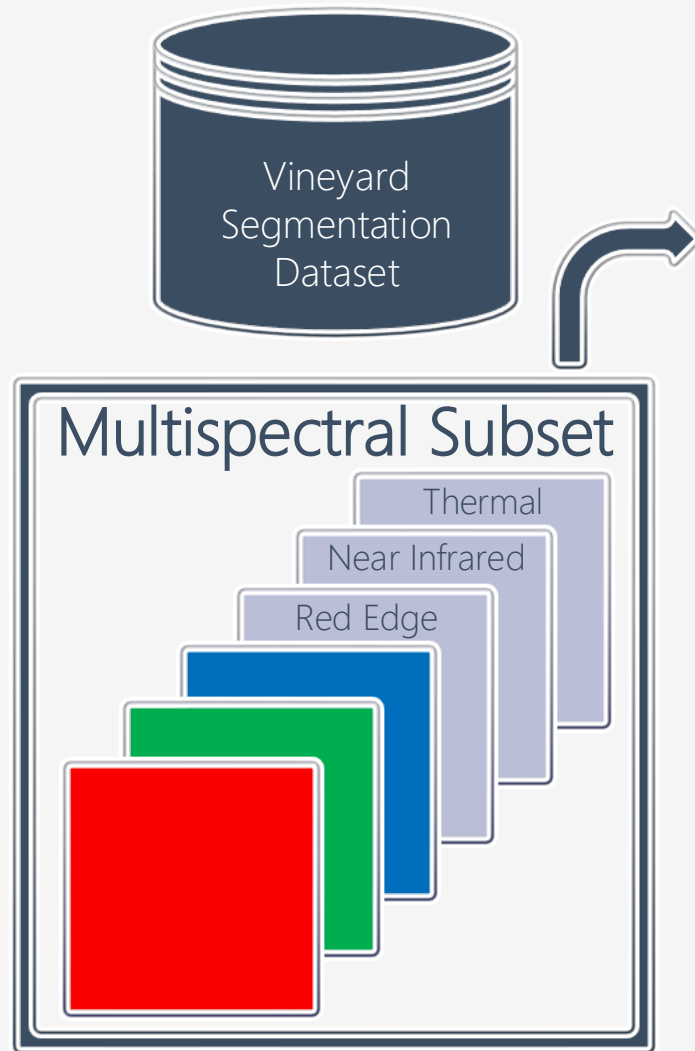
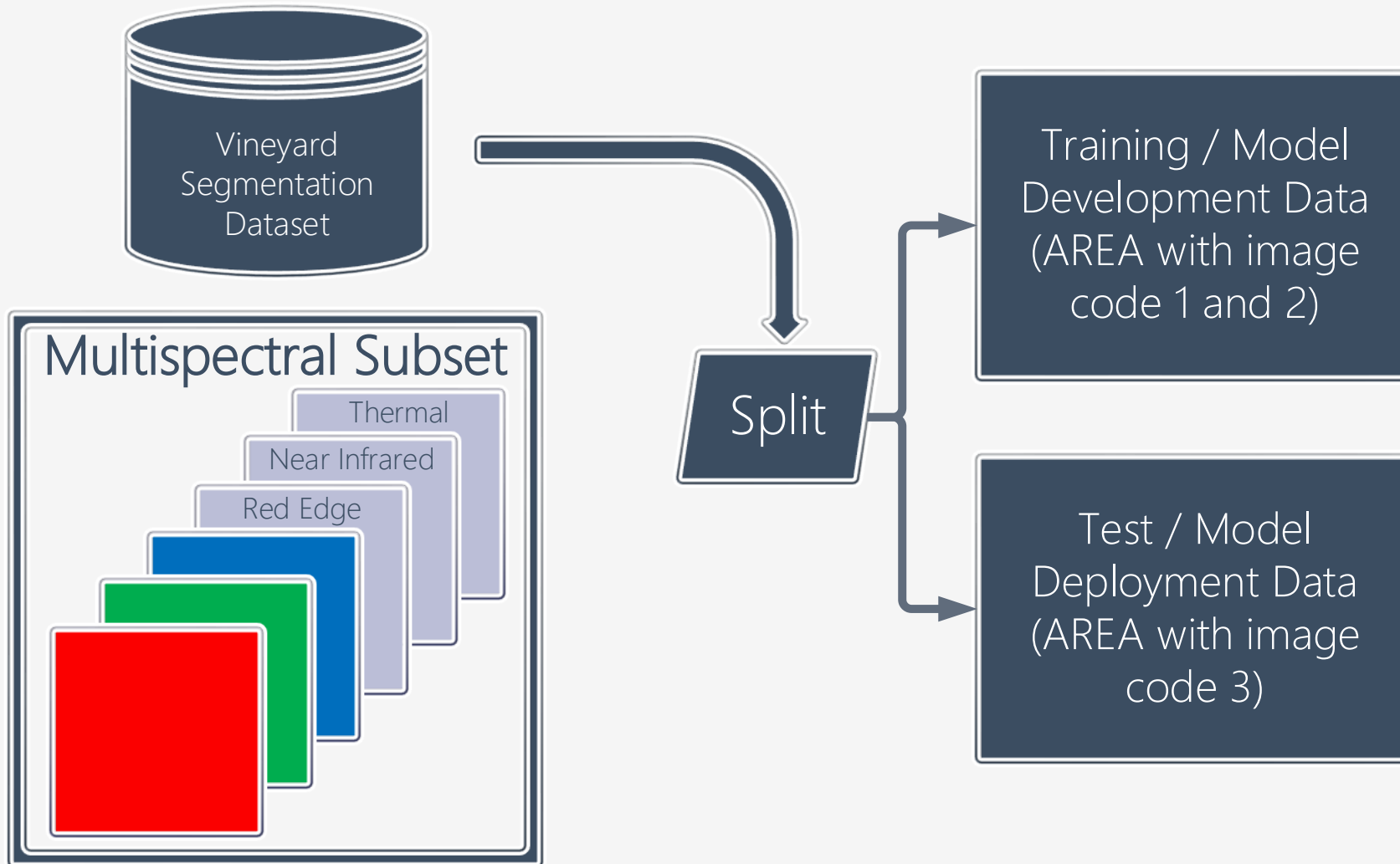


Image segmentation – Machine (Non Deep) Learning → Direct Pixel classification using informative Features.

Dataset



Dataset

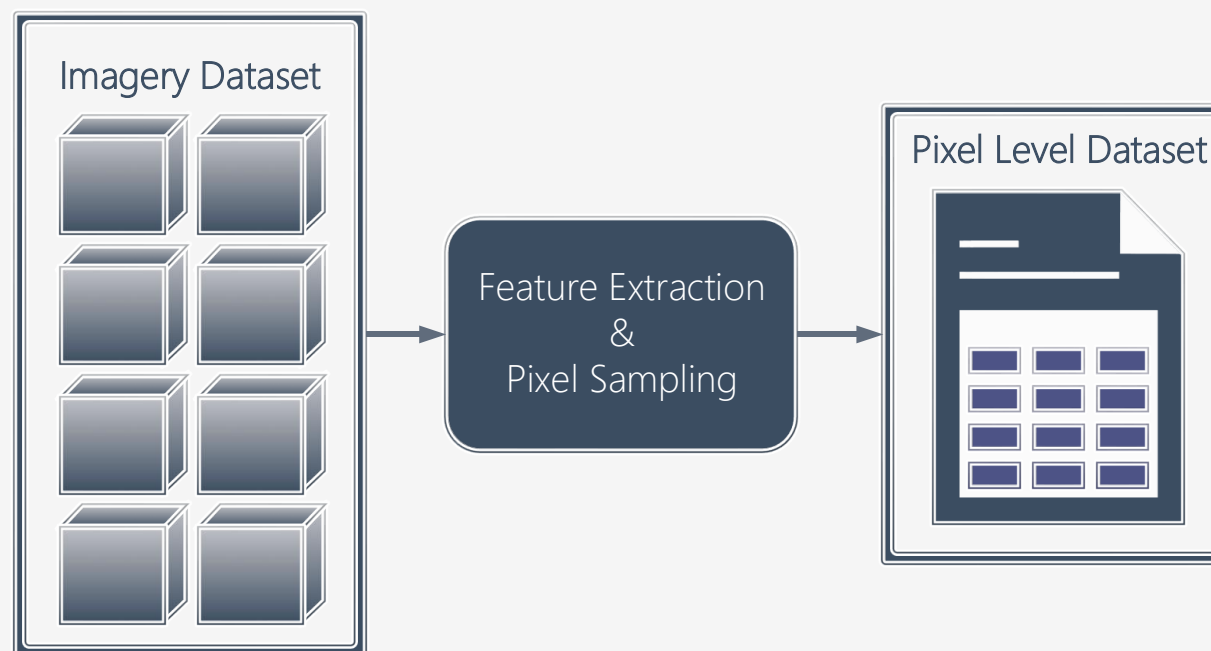


Changing Form of Dataset

Image segmentation – Machine (Non Deep) Learning → Direct Pixel classification using informative Features.

It is essential that we should change the form of Dataset to fit Machine Learning Classifiers and to perform Pixel Level Classification.

Extract a Pixel – Level Dataset from original Imagery Dataset.



Feature Extraction

Vegetation Indices

- A *Vegetation Index (VI)* is a spectral transformation metric for measuring the presence and state of vegetation.
- After its calculation, VI simulates a single channel image.

Red Edge Normalized Difference Vegetation index: $\frac{NIR - RedEdge}{NIR + RedEdge}$

Normalized Difference Vegetation index: $\frac{NIR - Red}{NIR + Red}$

Green Normalized Difference Vegetation index: $\frac{NIR - Green}{NIR + Green}$

Blue Normalized Difference Vegetation index: $\frac{NIR - Blue}{NIR + Blue}$

RedEdge-based Indices:

- $\frac{RedEdge - Red}{RedEdge + Red}$
- $\frac{RedEdge - Green}{RedEdge + Green}$
- $\frac{RedEdge - Blue}{RedEdge + Blue}$

Normalized Difference Green-Red index: $\frac{Green - Red}{Green + Red}$

Normalized Difference Green-Blue index: $\frac{Green - Blue}{Green + Blue}$

Feature Extraction

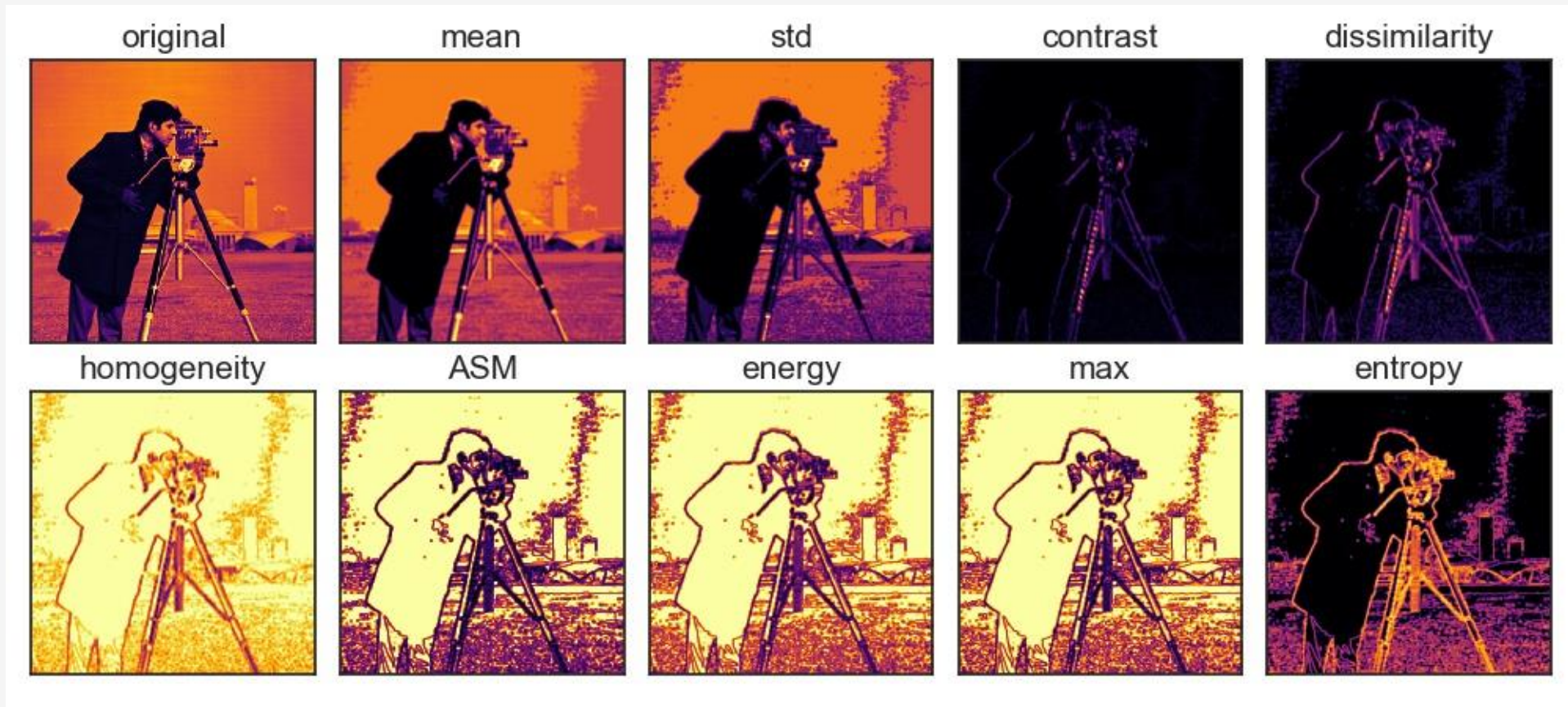
Texture Information

- **Gray Level Co-occurrence Matrix:** *Tabulation of how often pairs of pixel with specific values and in a specified spatial relationship occur in an image.*
- GLCM contains information about the relationship of intensity of a pixel and its neighborhood defined using a window.
- Texture Information are statistical measures extracted from the aforementioned matrix.
- Each measure simulates a single level image.

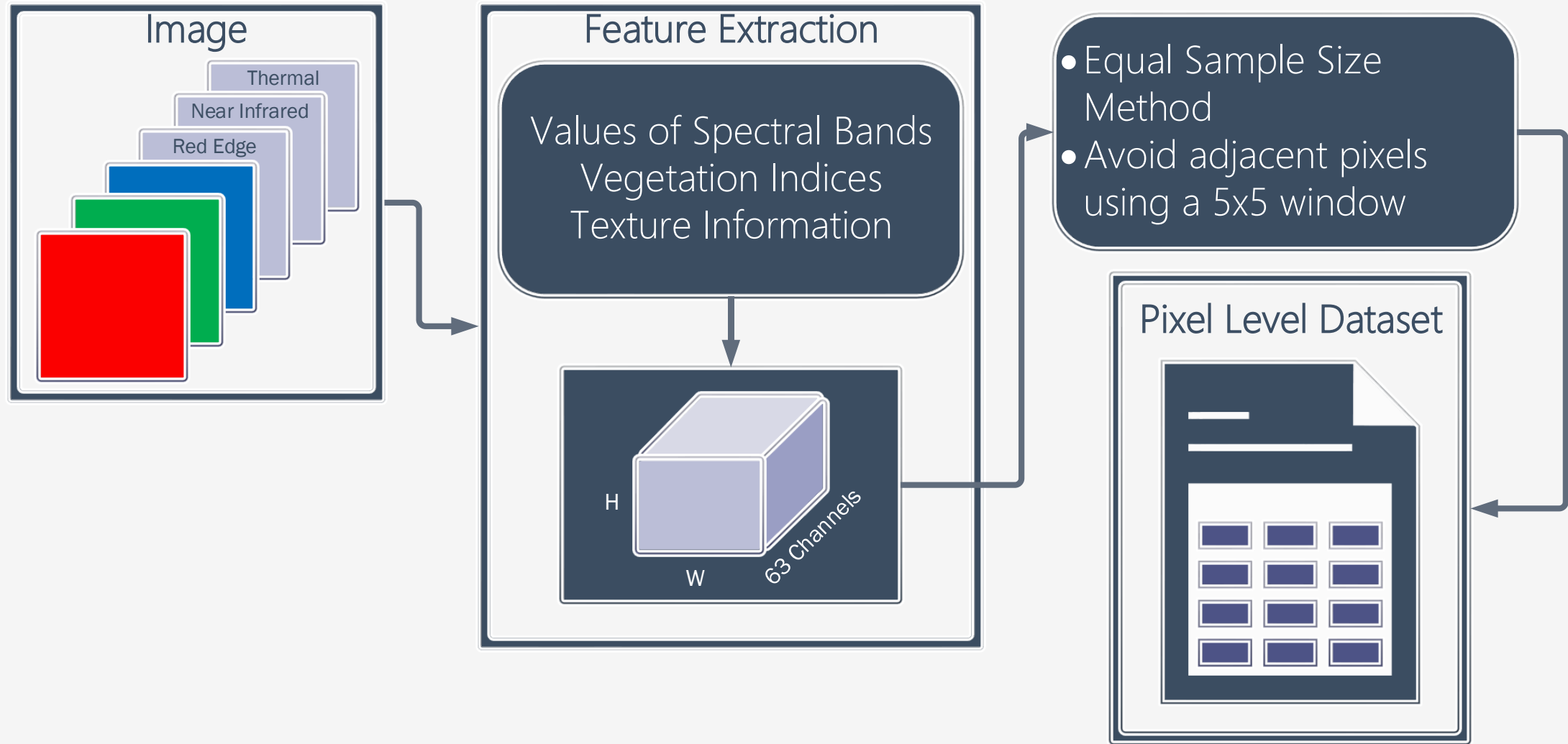
- Mean
- Standard Deviation
- Max
- Entropy
- Contrast = $\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} (i - j)^2$
- Homogeneity = $\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} \frac{P(i,j)}{1+(i-j)^2}$
- Angular second Moment (ASM) $\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i,j)^2$
- Energy $\sqrt{\sum_{i=0}^{N-1} \sum_{j=0}^{N-1} P(i,j)^2}$

Feature Extraction

Texture Information



Changing Form of Dataset using Feature Extraction & Sampling



Pixel Classification Dataset

	R	G	B	RE	NIR	TH	RENDVI	NDVI	GNDVI
0	37.0	99.0	39.0	85.0	96.0	254.0	0.060773	0.443609	1.000000
1	112.0	75.0	121.0	72.0	60.0	254.0	1.000000	1.000000	1.000000
2	56.0	84.0	64.0	150.0	136.0	254.0	1.000000	0.416667	0.236364
3	51.0	46.0	29.0	46.0	52.0	254.0	0.061224	0.009709	0.061224
4	61.0	121.0	40.0	117.0	124.0	254.0	0.029046	0.340541	0.012245

	BNDVI	...	T_std	T_contrast	T_homogeneity	T_ASM	T_energy
0	0.422222	...	173.627426	0.0	25.0	625.0	25.0
1	1.000000	...	173.627426	0.0	25.0	625.0	25.0
2	0.360000	...	173.627426	0.0	25.0	625.0	25.0
3	0.283951	...	173.627426	0.0	25.0	625.0	25.0
4	0.512195	...	173.627426	0.0	25.0	625.0	25.0

	T_max	T_entropy	label	imgName	mode
0	25.0	8.070781	0	img_10000000036	development
1	25.0	8.070781	1	img_20000000127	development
2	25.0	8.070781	1	img_30000000088	deploy
3	25.0	8.070781	0	img_30000000046	deploy
4	25.0	8.070781	0	img_10000000005	development

Comparison of Different Classifiers

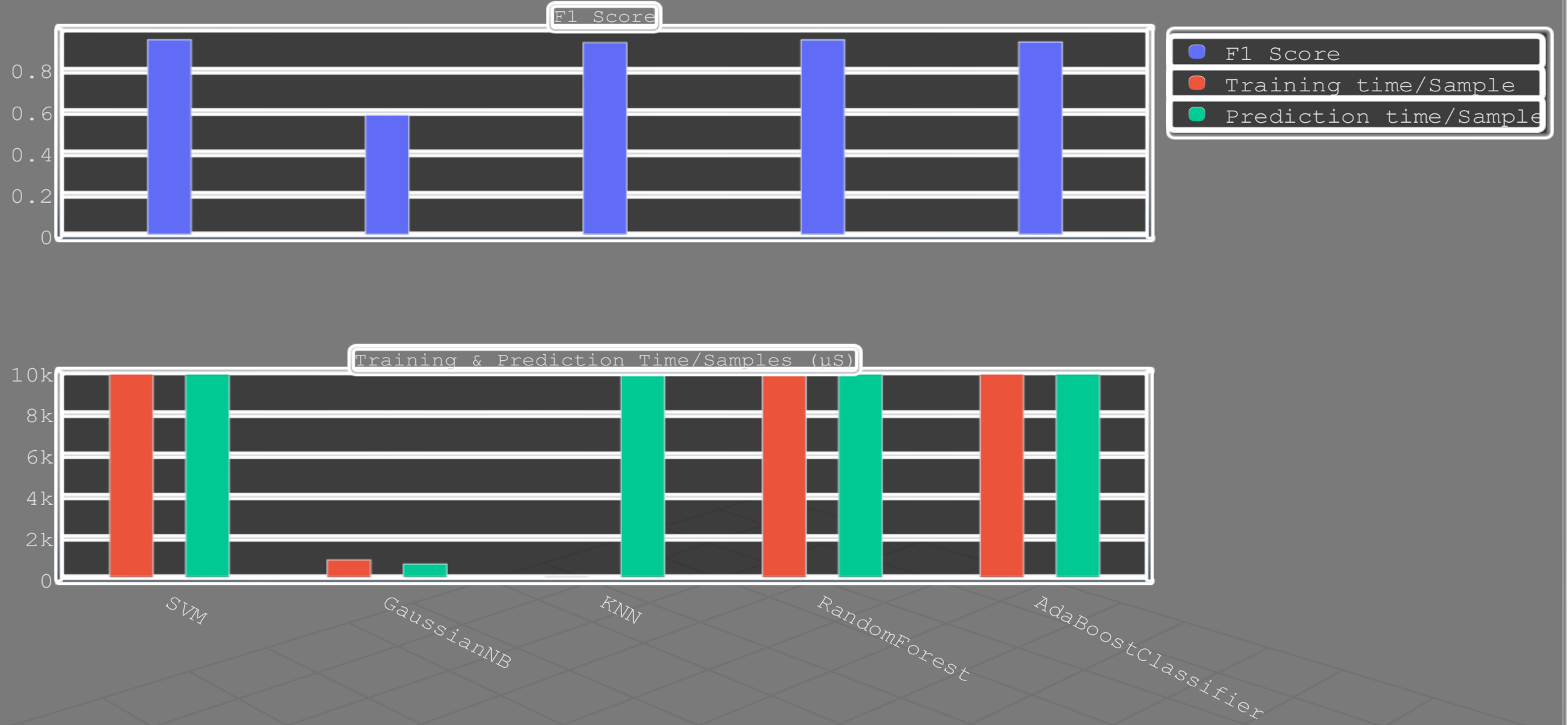
Comparison of baseline models in terms of F1 Score, training and prediction time.

Machine Learning Models: SVM-RBF, Gaussian Naïve Bayes, K-Nearest Neighbor, Random Forest, AdaBoost

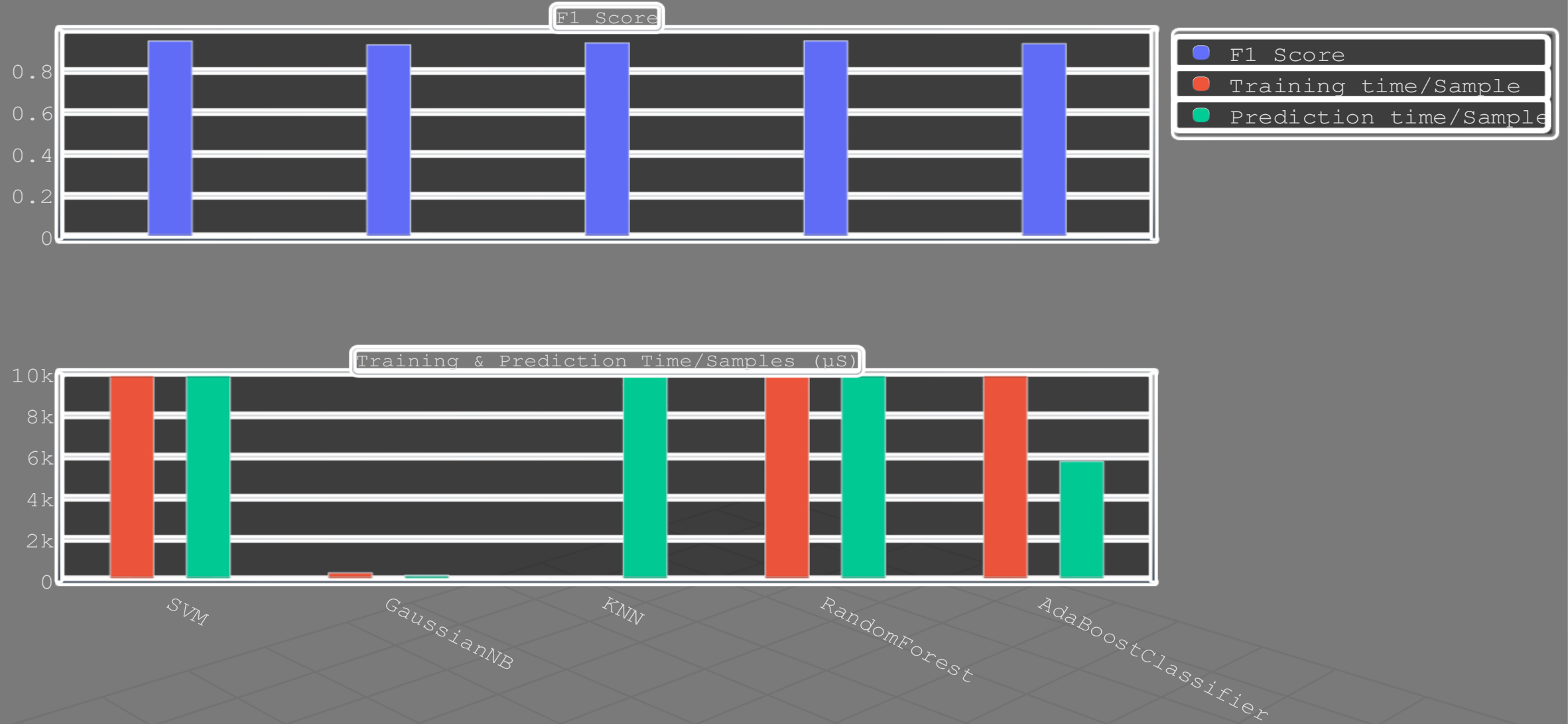
- *Comparison Using All Features*
- *Comparison Using Features Selected by Random Forest*
- *Comparison Using Features Selected by AdaBoost*
- *Comparison Using Features Selected by ANOVA statistical Test*

For this task, training dataset is used by performing a random split.

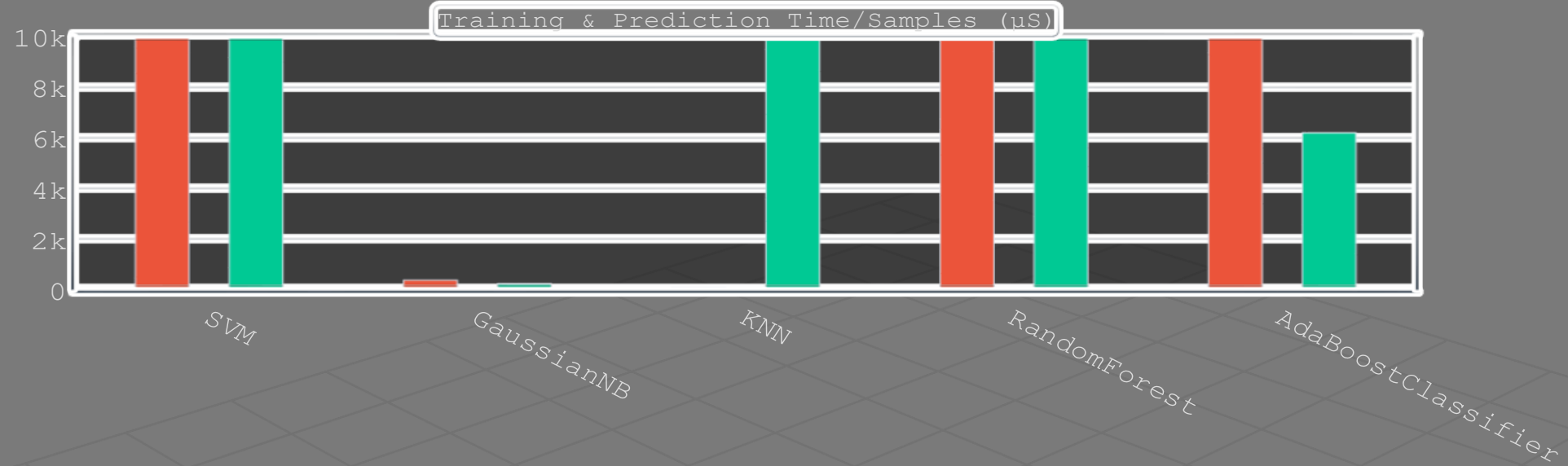
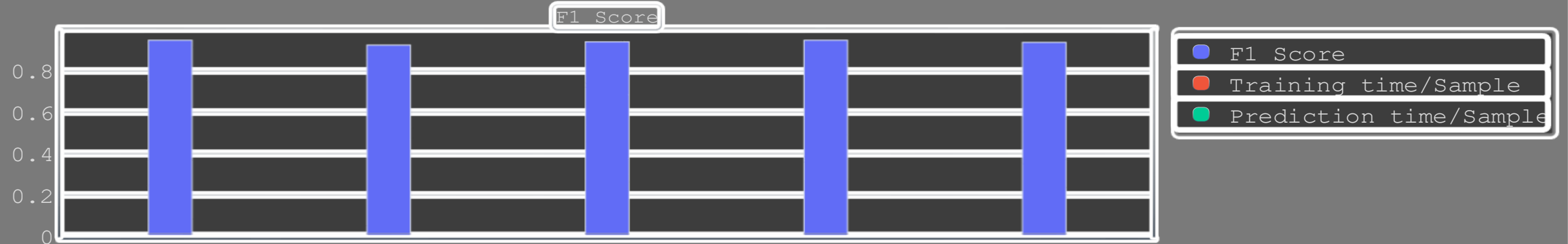
Model Comparison: All Features



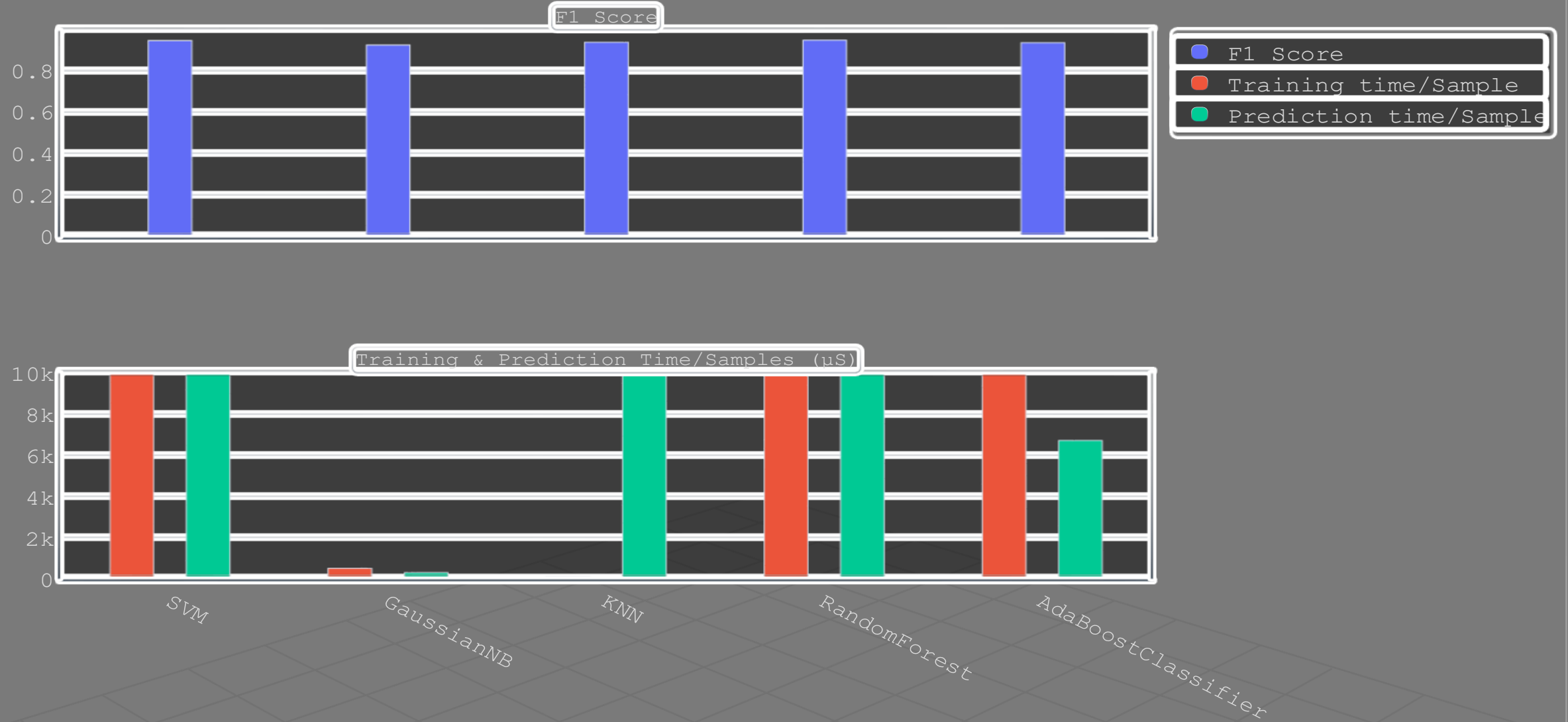
Model Comparison: Features selected by Random Forest



Model Comparison: Features selected by AdaBoost



Model Comparion: Features selected by ANOVA Method



Training of Selected Classifier

Selection: Gaussian Naïve Bayes combined with Features Selected By Random Forest.

How well a classifier generalizes ?

Which is the range of expected errors of the classifier ?

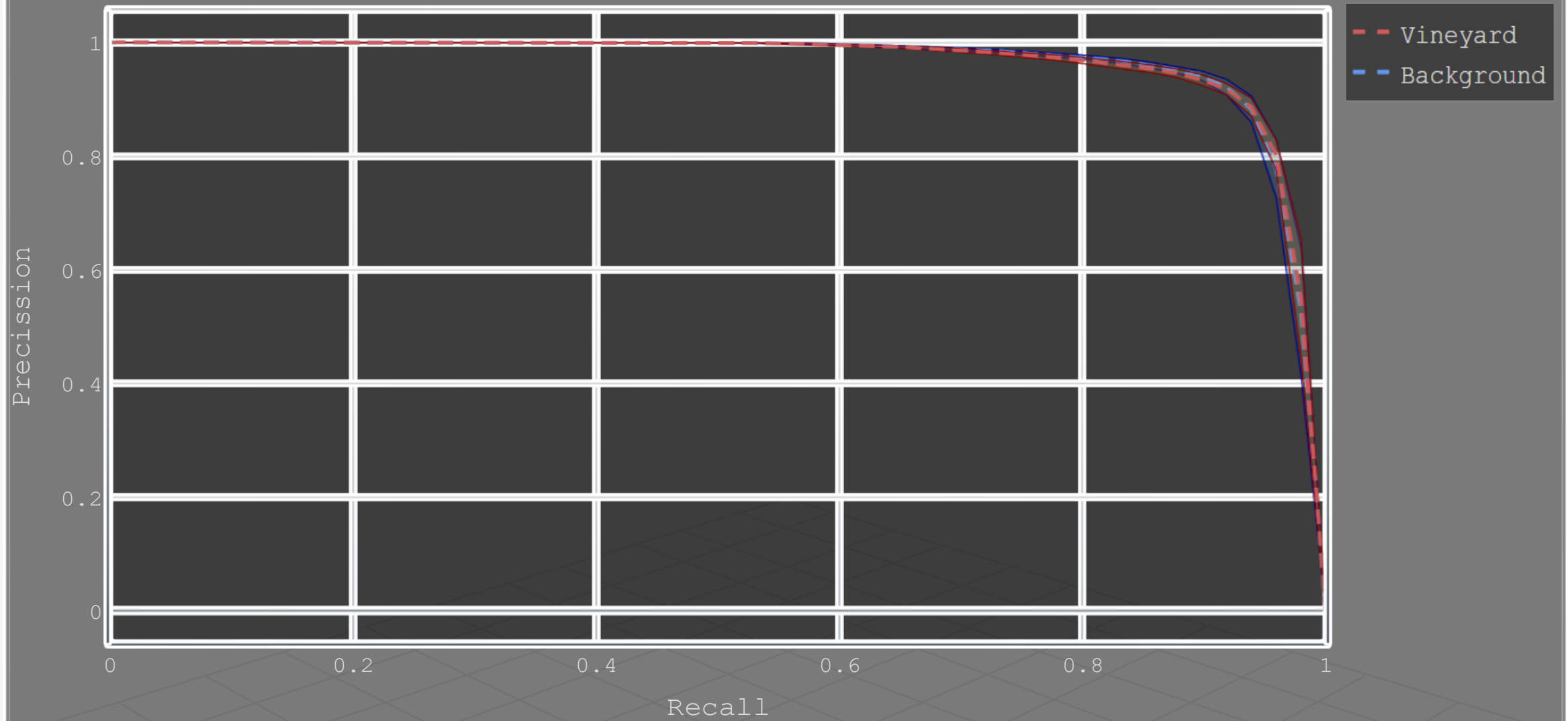
} **Cross Validation**

Mean F1: 0.918 | Sigma F1: 0.003

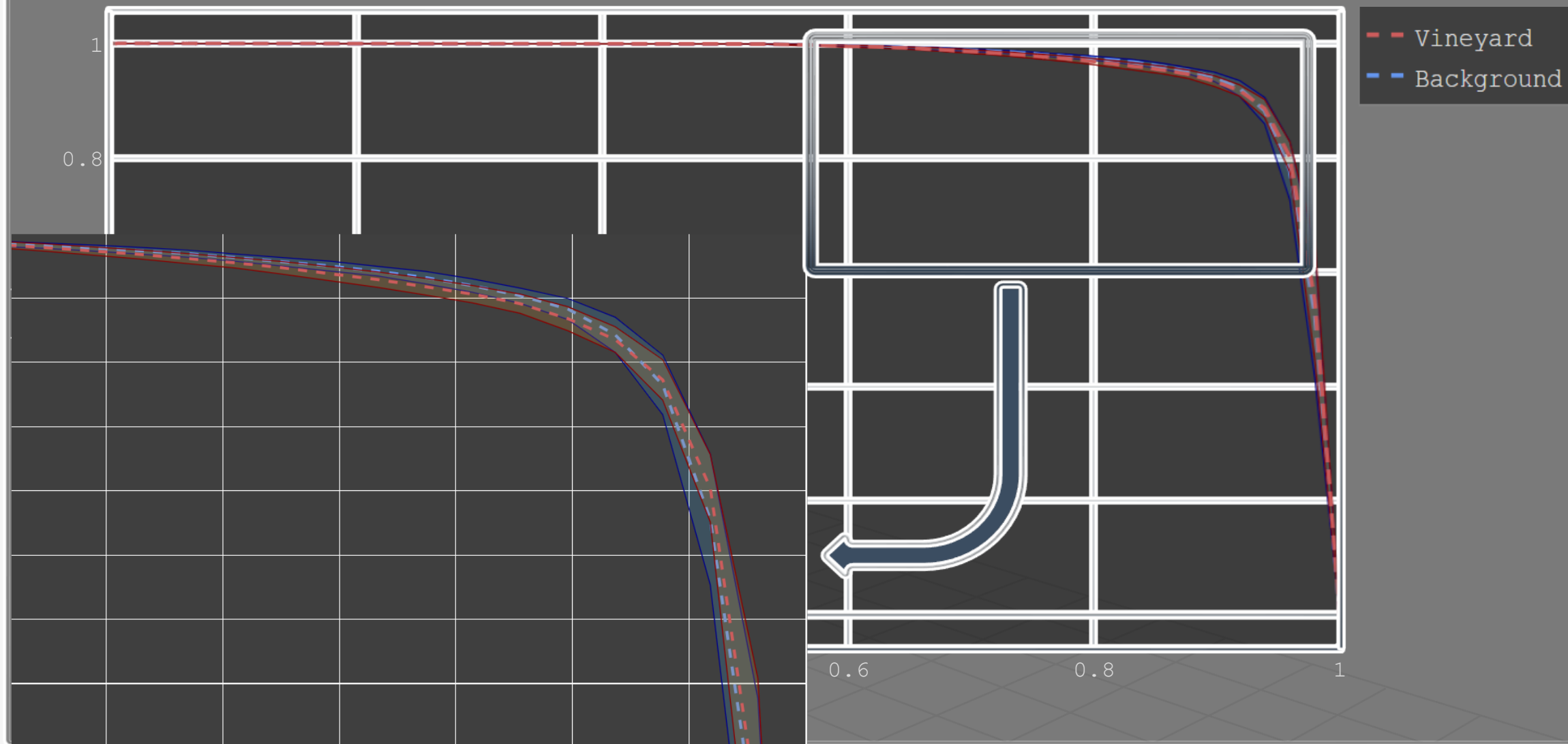
Finally, the entire dataset used to train the model for the last time.



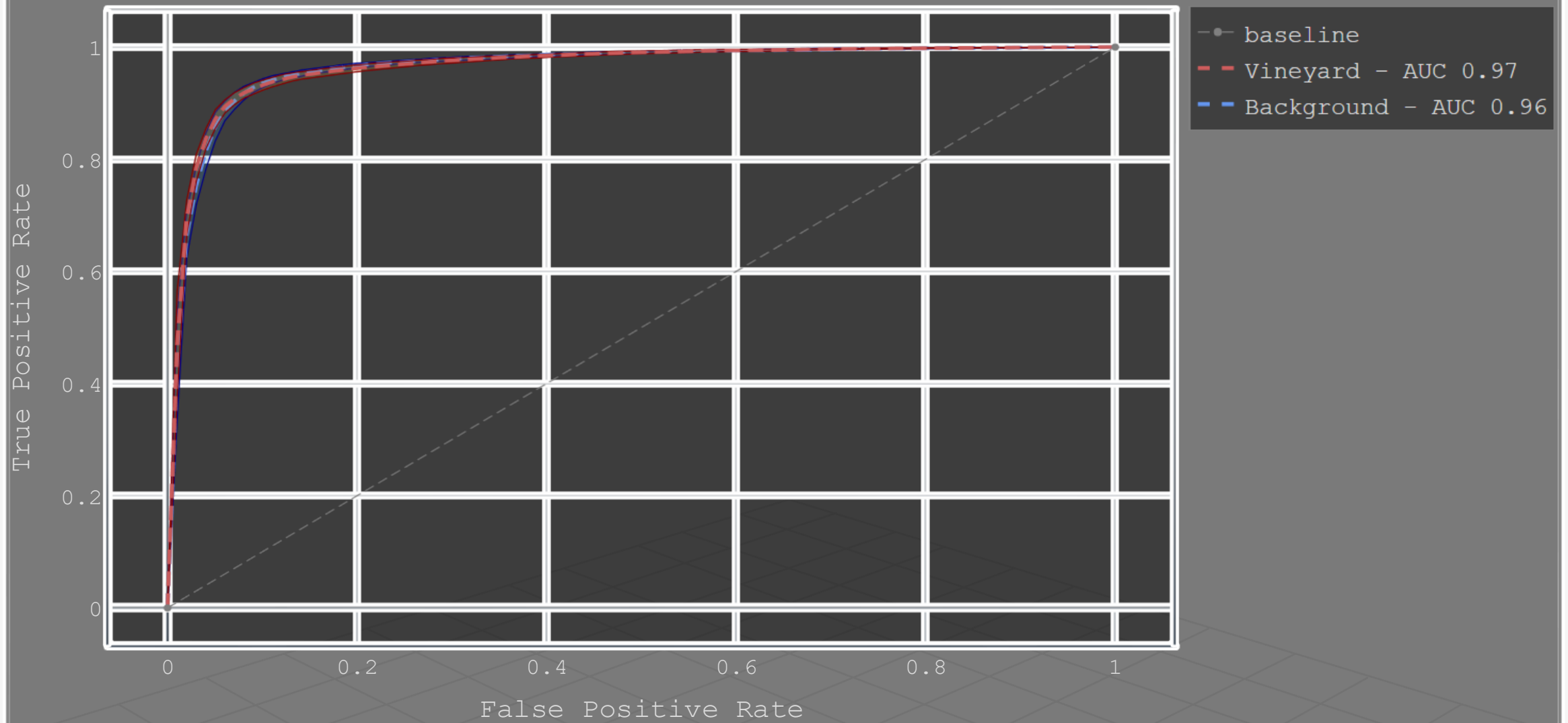
Cross Validation - Precision/Recall Curve



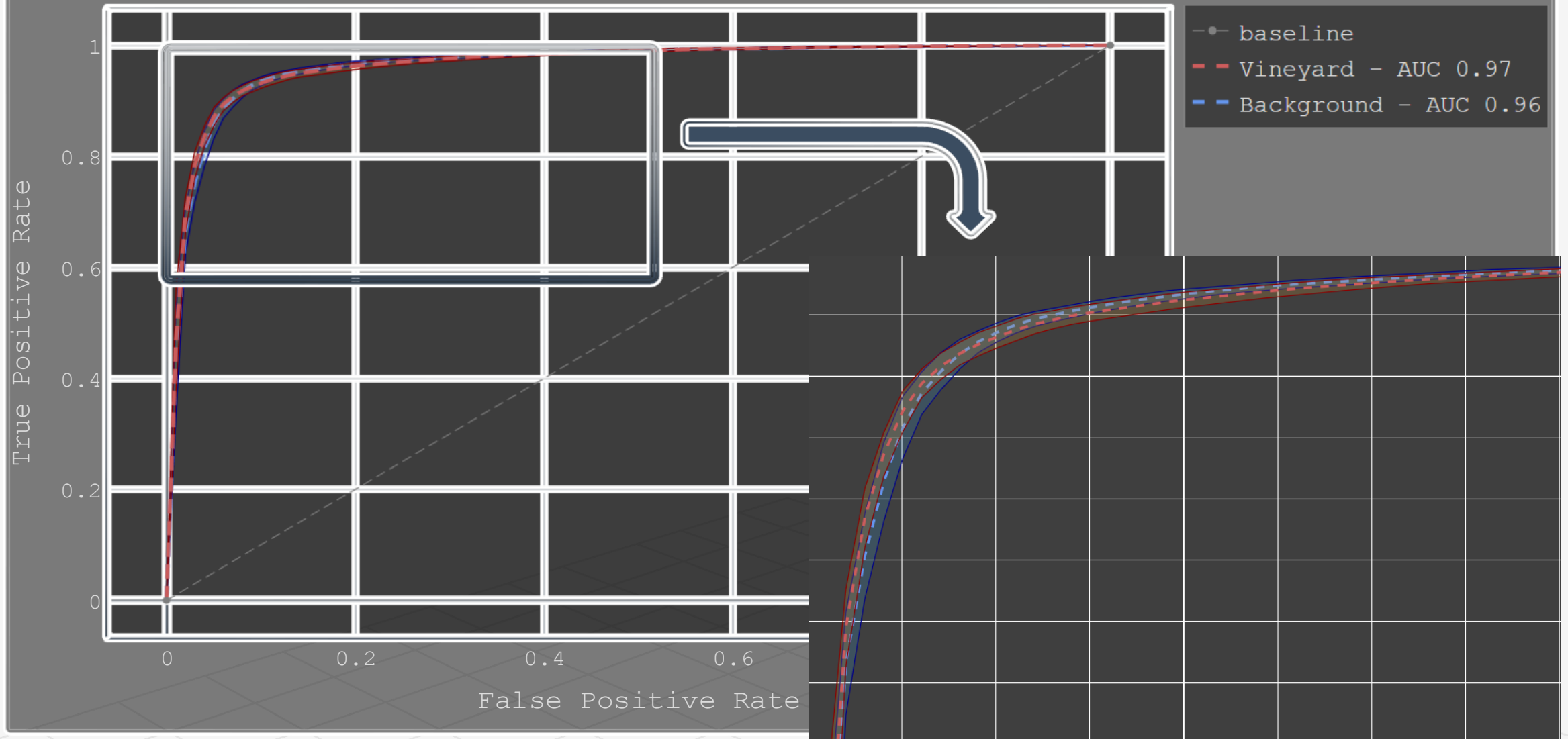
Cross Validation - Precision/Recall Curve



Cross Validation - ROC Curve



Cross Validation - ROC Curve



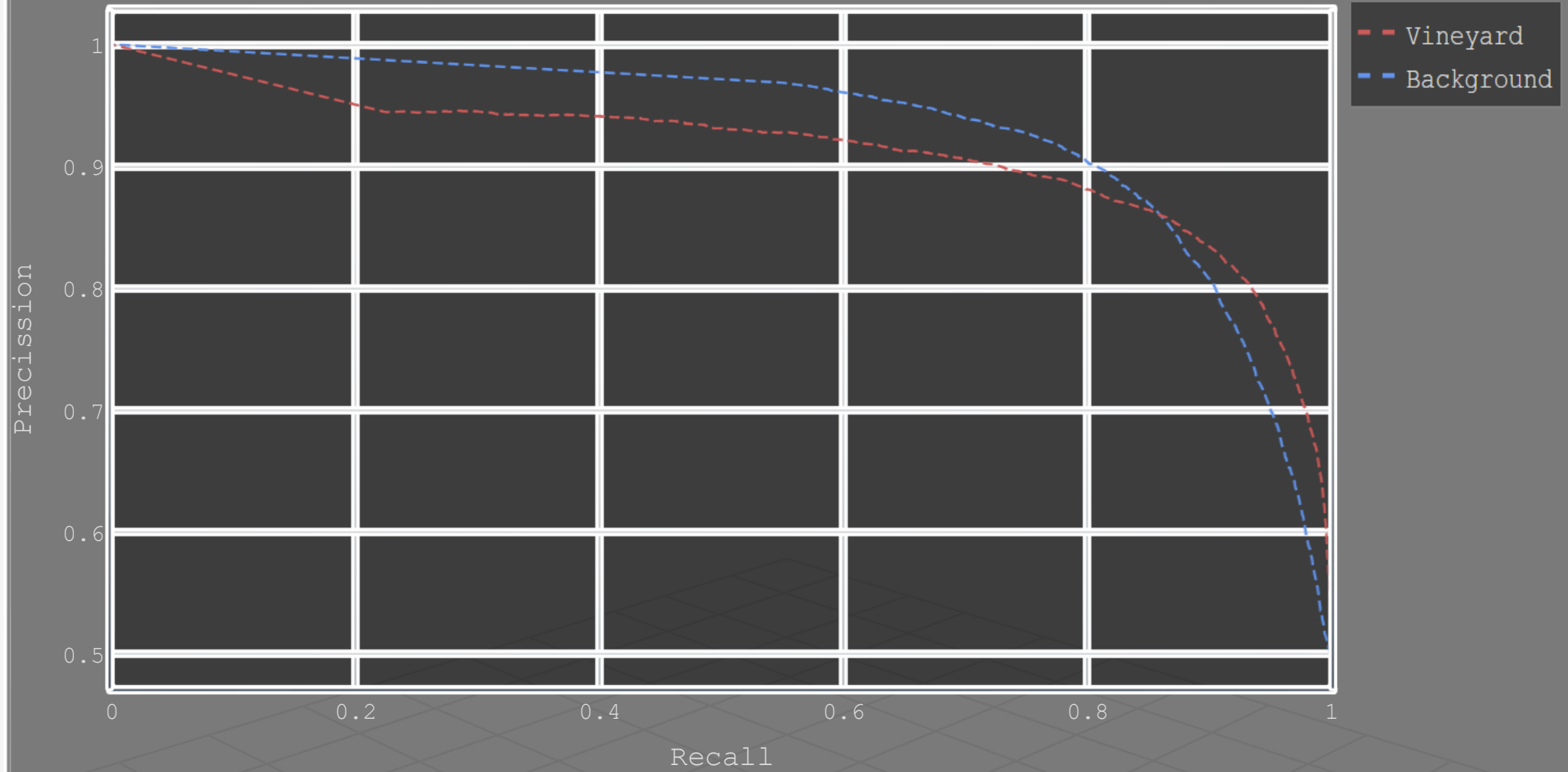
Using Classifier to predict unknown data (Area 3)

Evaluation of Test Dataset: Confusion Matrix

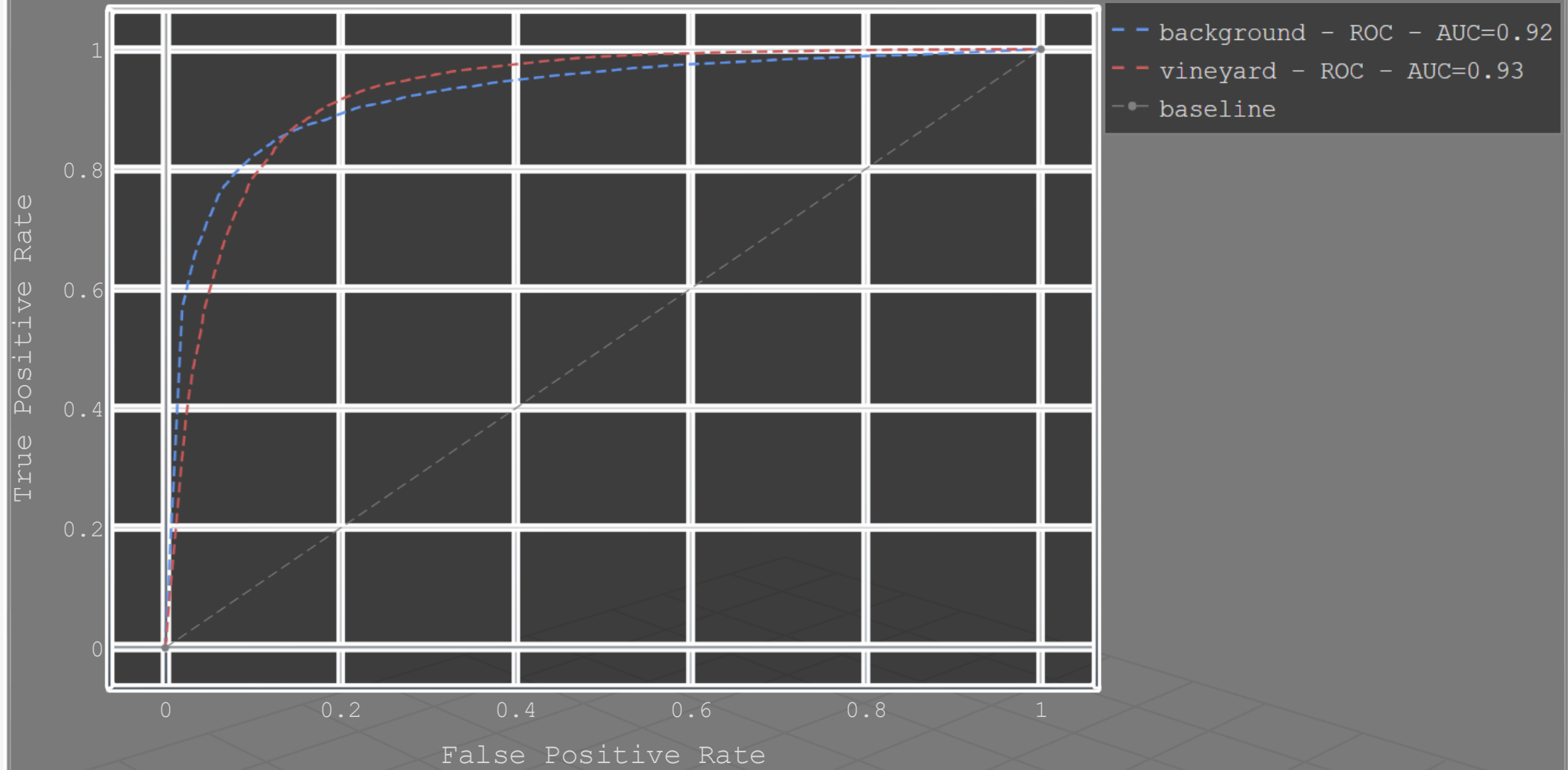
Actual	vineyard	1370	9217
	background	9755	1908
		background	vineyard
		Predicted	

Final Evaluation & Classification Report:				
	precision	recall	f1-score	support
0	0.84	0.88	0.86	11125
1	0.87	0.83	0.85	11125
accuracy				
macro avg	0.85	0.85	0.85	22250
weighted avg	0.85	0.85	0.85	22250

Evaluation of Test Dataset: Precision/Recall Curve



Evaluation of Test Dataset: ROC Curve



Predictions from Test Dataset

RGB Image

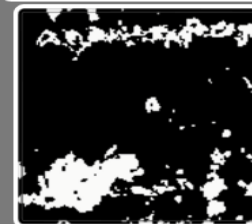
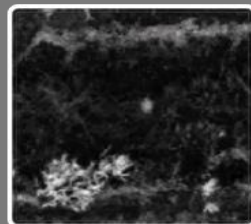
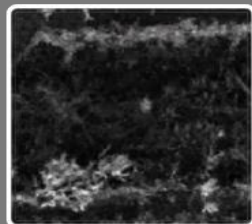
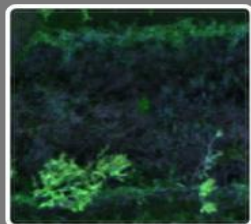
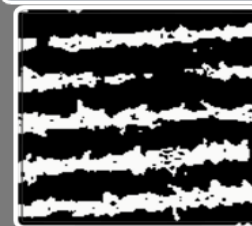
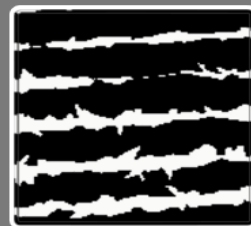
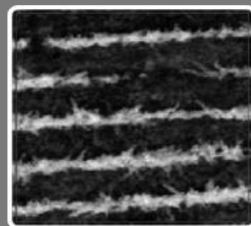
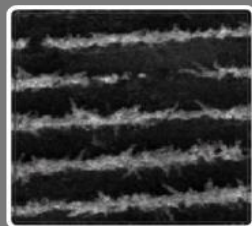
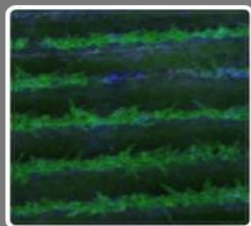
Red Edge Band

NIR Band

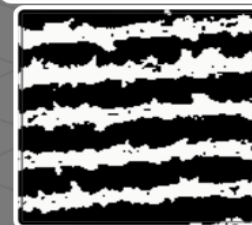
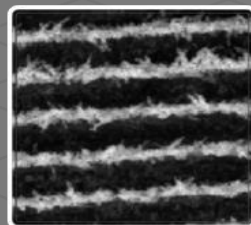
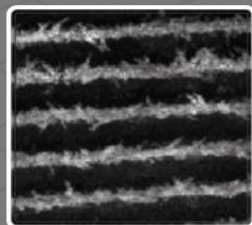
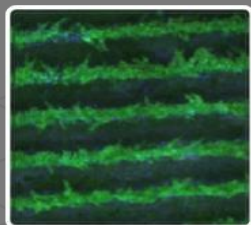
Ground Truth

Precision

Dice Score = 0.79



Dice Score = 0.43



Dice Score = 0.86

Thank you!

