INSTITUTE
OF GEOGRAPHY SLOVAK
ACADEMY OF SCIENCES

TEXTURE-BASED TECHNIQUES FOR VEGETATION CLASSIFICATION IN RIPARIAN ZONES USING HISTORICAL AERIAL ORTHOPHOTOS

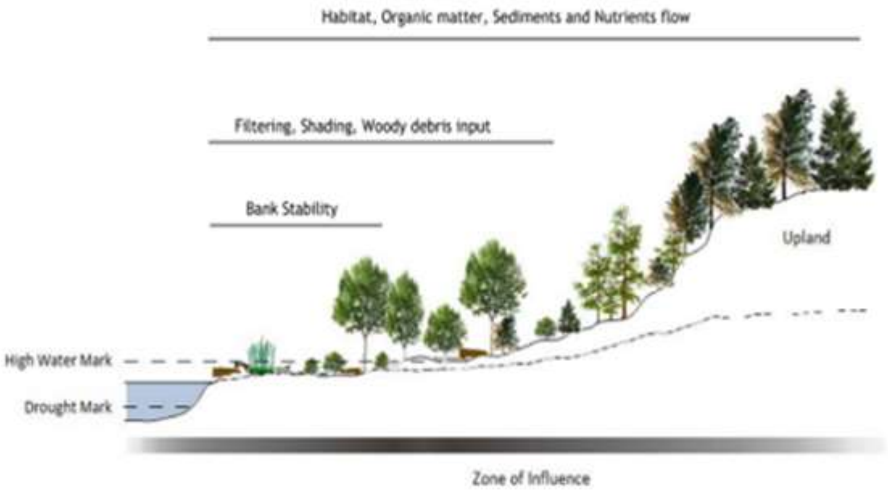
Afzali, Hamid, M.Sc.
Rusnák, Miloš, Mgr., PhD.
Opravil, Šimon, Mgr.

GEOKARTO 2024
5 - 6 SEPTEMBER

Introduction

➢ Riparian zone

- Biodiversity
- Water Resources
- Water Quality
- Soil Conservation
- Fauna and Flora
- Flood Hazards



Clerici Nicola,et al, 2011

➢ Historical Aerial Orthophotos

- Unique source of historical information
- Wide Area Coverage
- Cost-effective
- High Spatial Resolution

Study Area

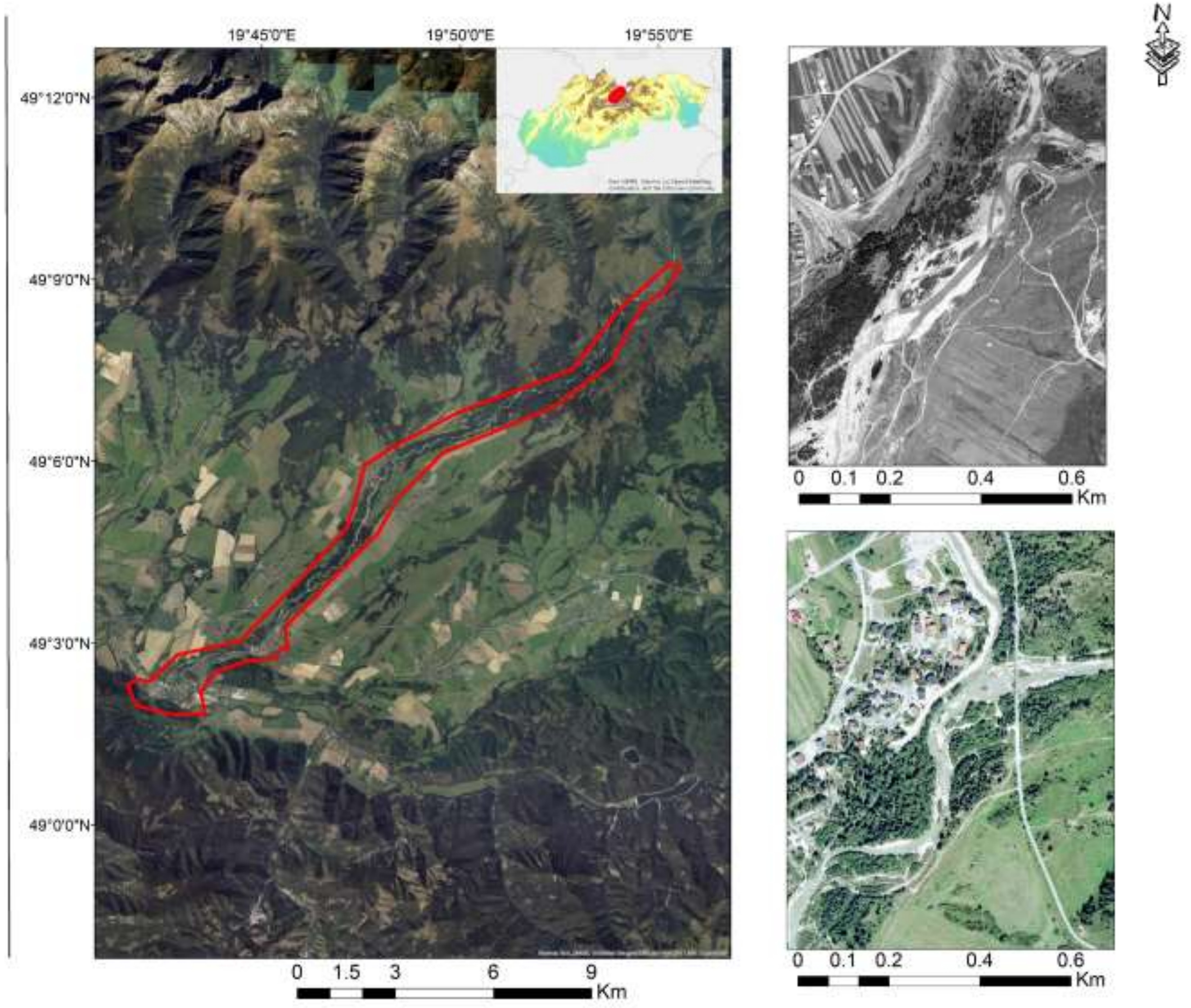
Bela River

- The largest right tributary of the upper Váh River
- 23.60Km long
- The catchment area amounts to 244 km2

Kidova et al, 2017

Dataset

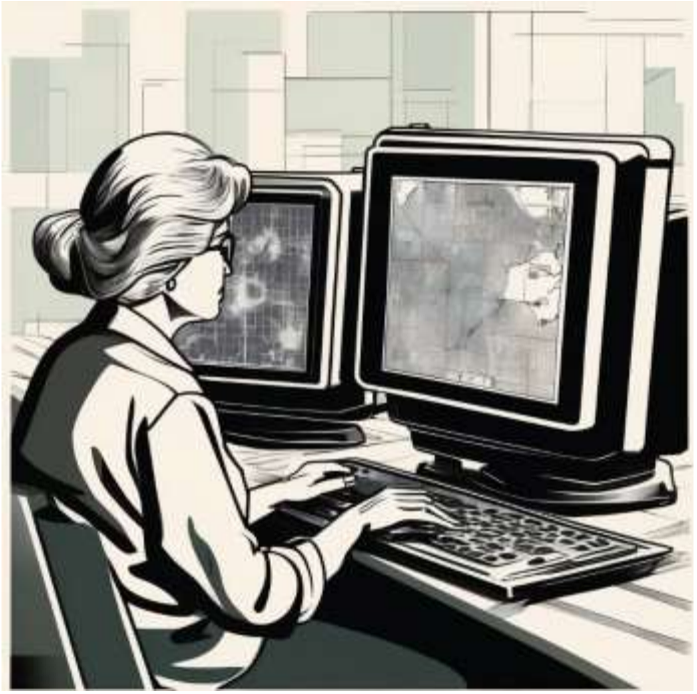
Gray scale Orthophotos	RGBN
• **1949** , 0.5 m	• **2002** , 0.5 m
• **1961** , 0.3 m	• **2006** , 0.4 m
• **1973** , 0.3 m	• **2009** , 0.25 m
• **1986** , 0.5 m	• **2012** , 0.25 m
• **1992** , 0.5 m	• **2015** , 0.2 m
	• **2018** , 0.2 m
	• **2022** , 0.2 m



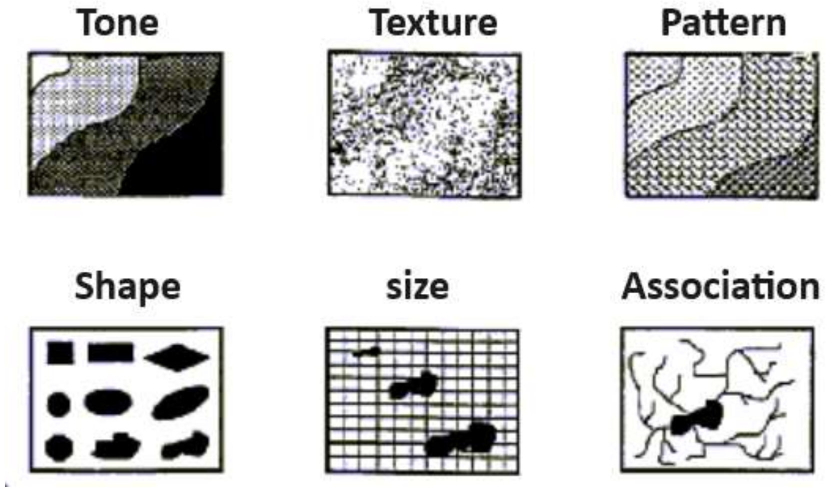
Historical Images

Visual Image Interpretation

- **Traditional use**



- **Government agencies**
- **Lawyers**
- **Land surveyor**
- **LULC Mapping and Cartography**
- **Military**
- **Archaeology**



Historical Images

Digital Image Processing

➤ **Modern use**

- **Gray-tone spatial dependencies**
- **Statistical properties of the intensity histogram**
- **Spatial relationship**

0	0	0	3	3	7	7
0	0	0	3	3	7	7
0	0	0	3	5	5	5
0	0	0	5	5	5	2
0	0	2	2	2	2	2
1	1	2	2	2	2	2
1	1	2	2	2	2	2

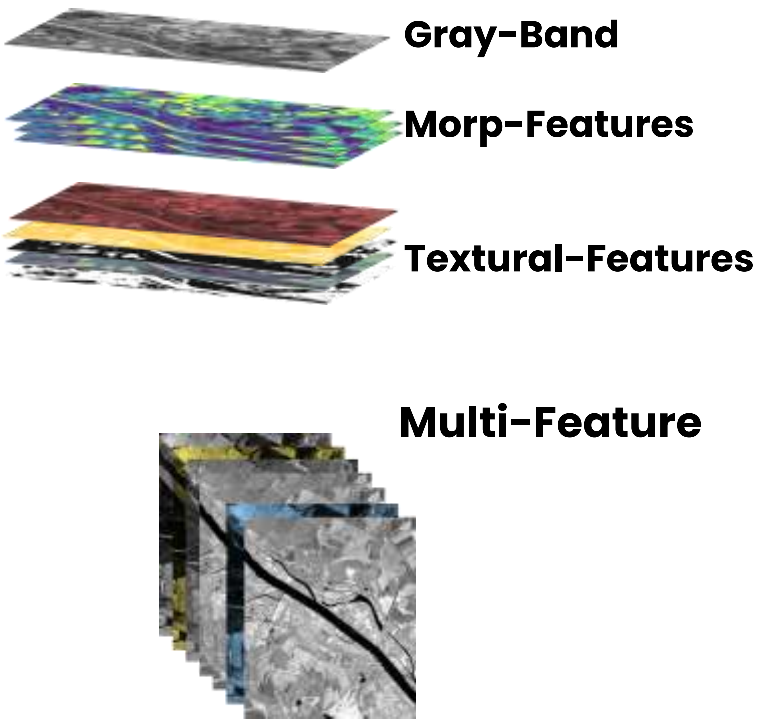
$$Correlation = \sum_i \sum_j \frac{p(i,j)\left[(i-\mu_i)\left(j-\mu_j\right)\right]}{\sigma_i \sigma_j}$$

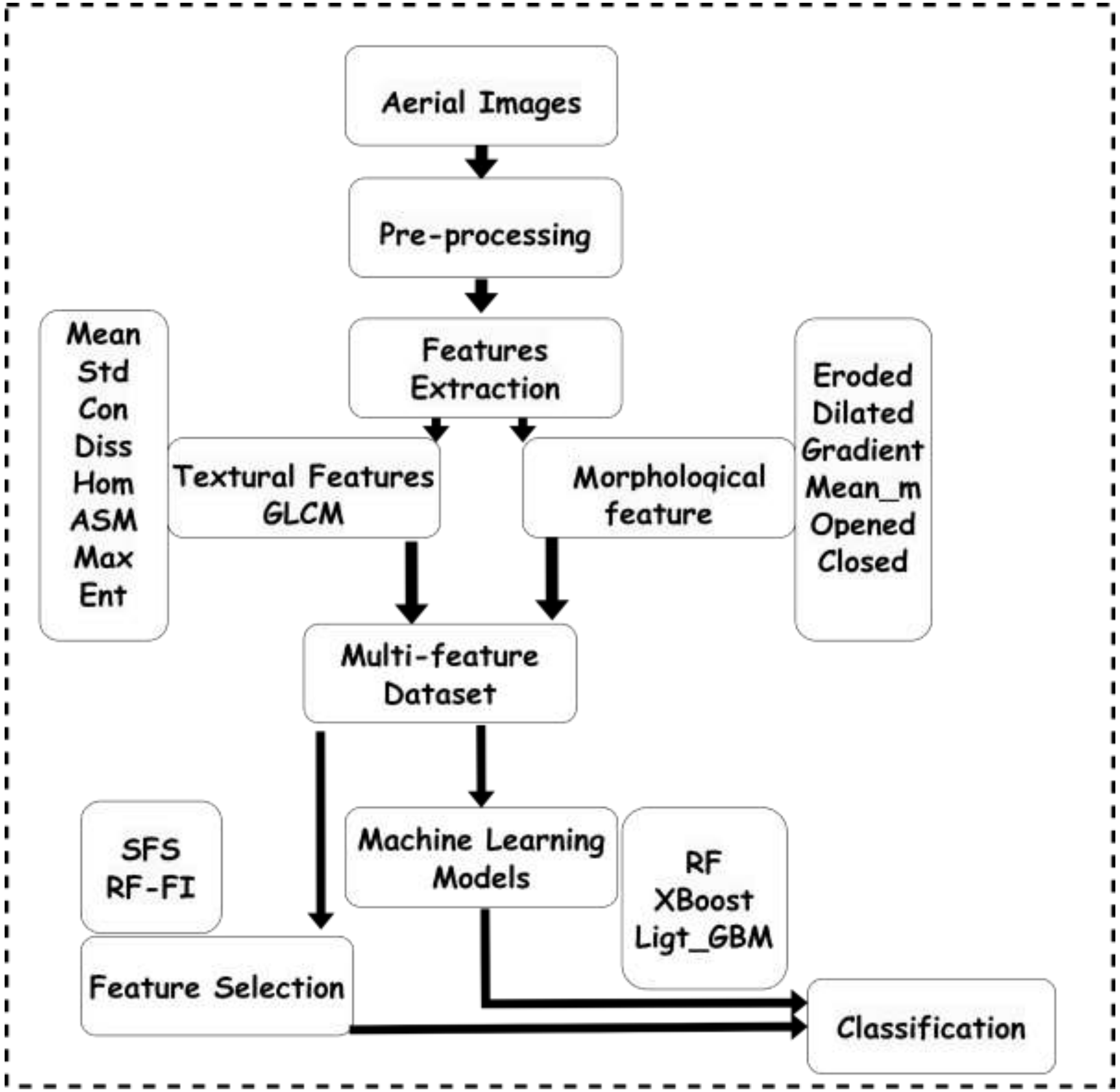
$$Energy = \sum_i \sum_j p(i,j)^2$$

$$Homogeneity = \sum_i \sum_j \frac{p(i,j)}{1+|i-j|}$$

$$Contrast = \sum_i \sum_j (i-j)^2 p(i,j)$$

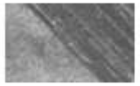
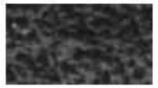
Methodology



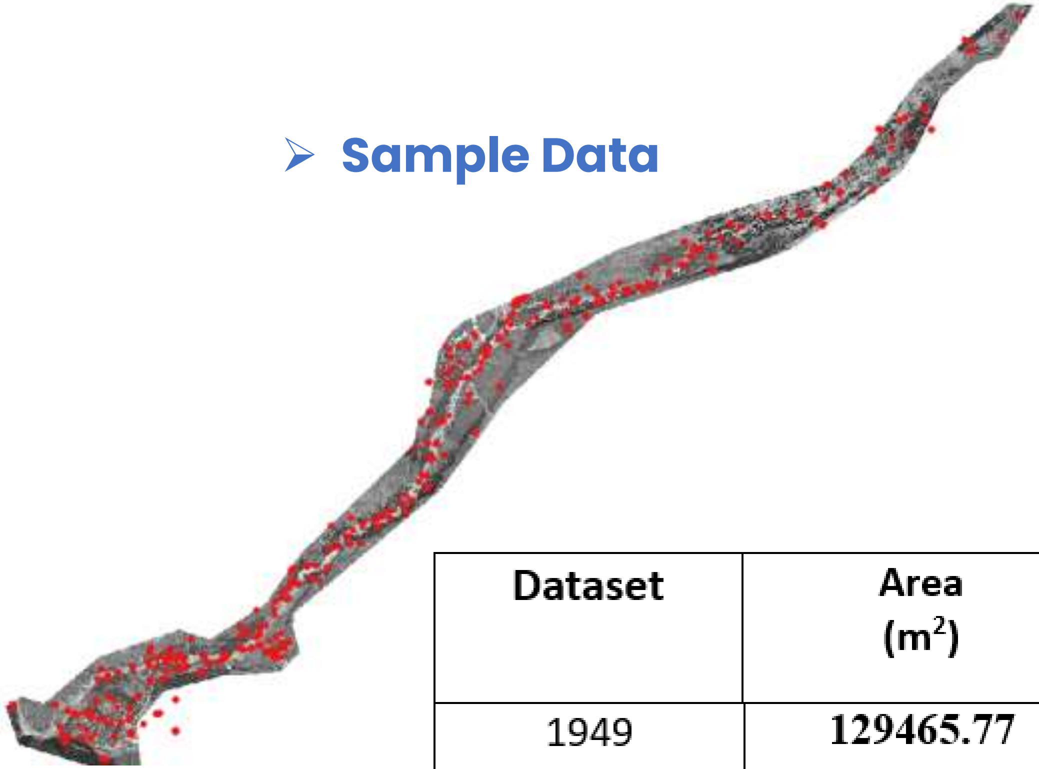


Results

➢ Class definition

Class Name	Definition	B&W	RGB
Artificial	Areas that changed by human activities, such as urban developments, roads		
Cropland	Regions used for agricultural purposes, including fields where crops , cultivated		
Water	Natural or artificial bodies of water		
Bare-land	Exposed land surfaces with no vegetation		
Grass-land	Areas dominated by grasses and other low vegetation		
Forest	Regions covered by dense tree		

➢ Sample Data



Dataset	Area (m^2)
1949	129465.77
1961	128831.95
1973	127969.25
1986	127095.02
1992	127944.51

Results

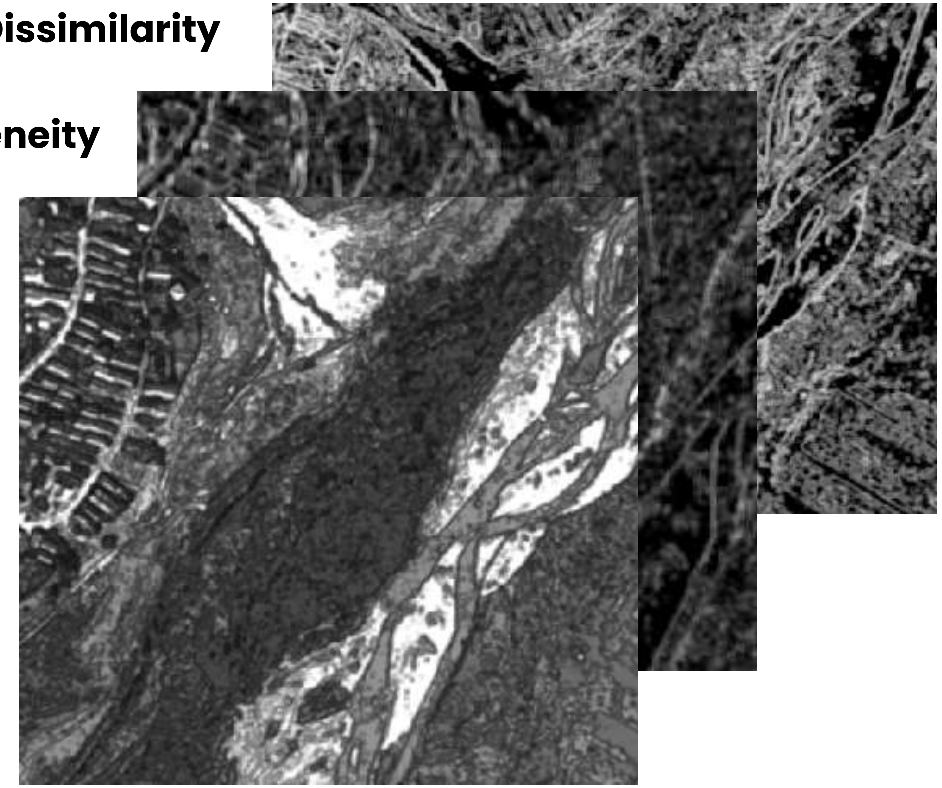
➢ Feature Extraction

GLCM-Feature

Dissimilarity

Homogeneity

Std

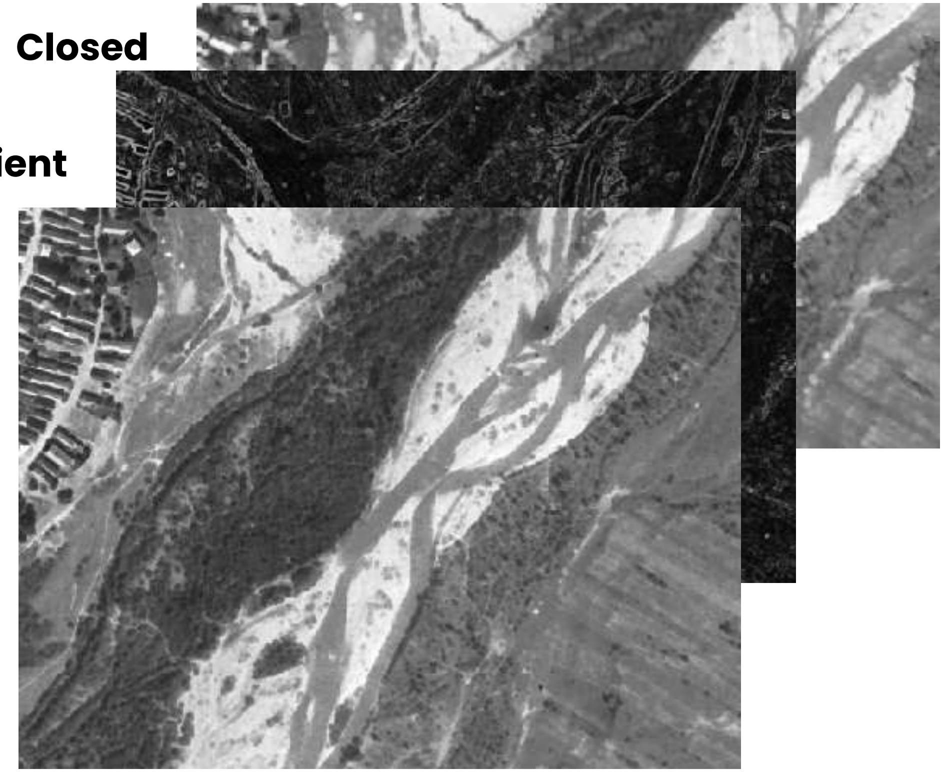


Morphological -Feature

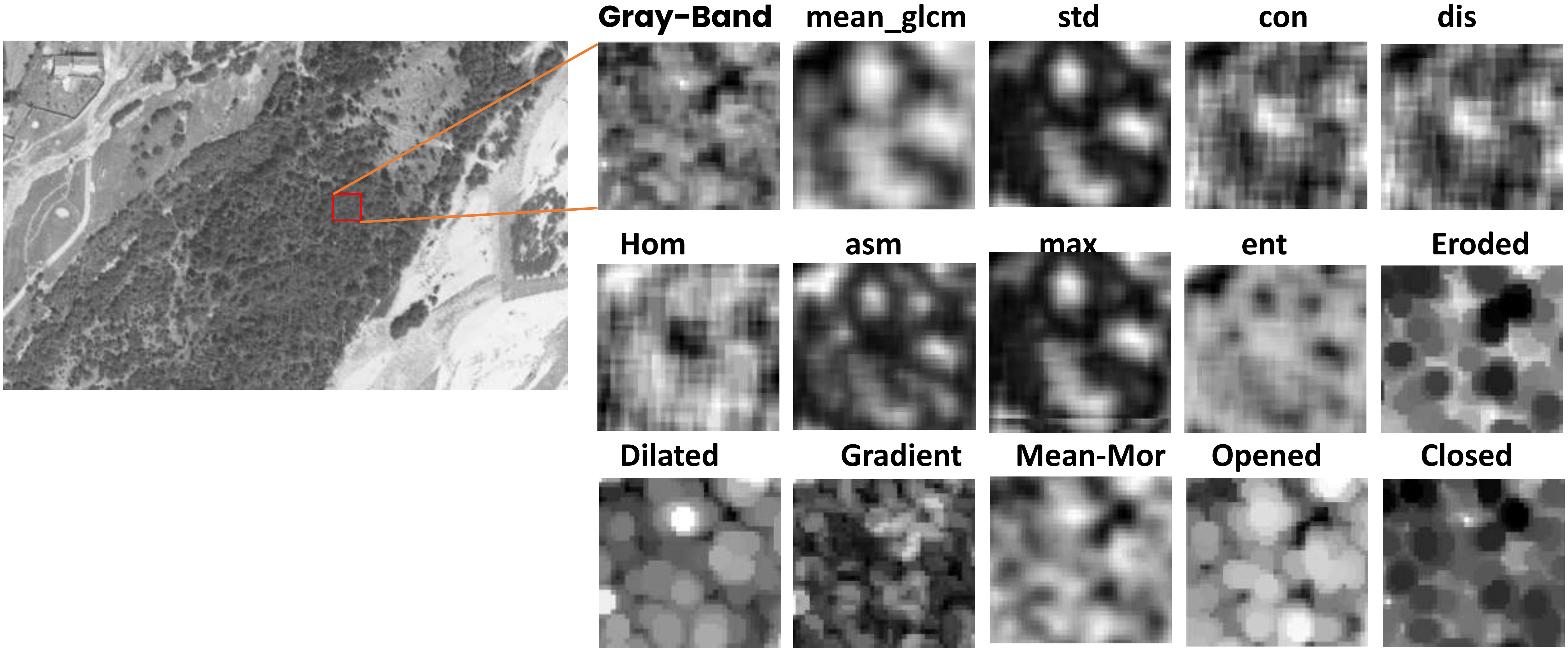
Closed

Gradient

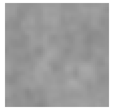
Opened



Results

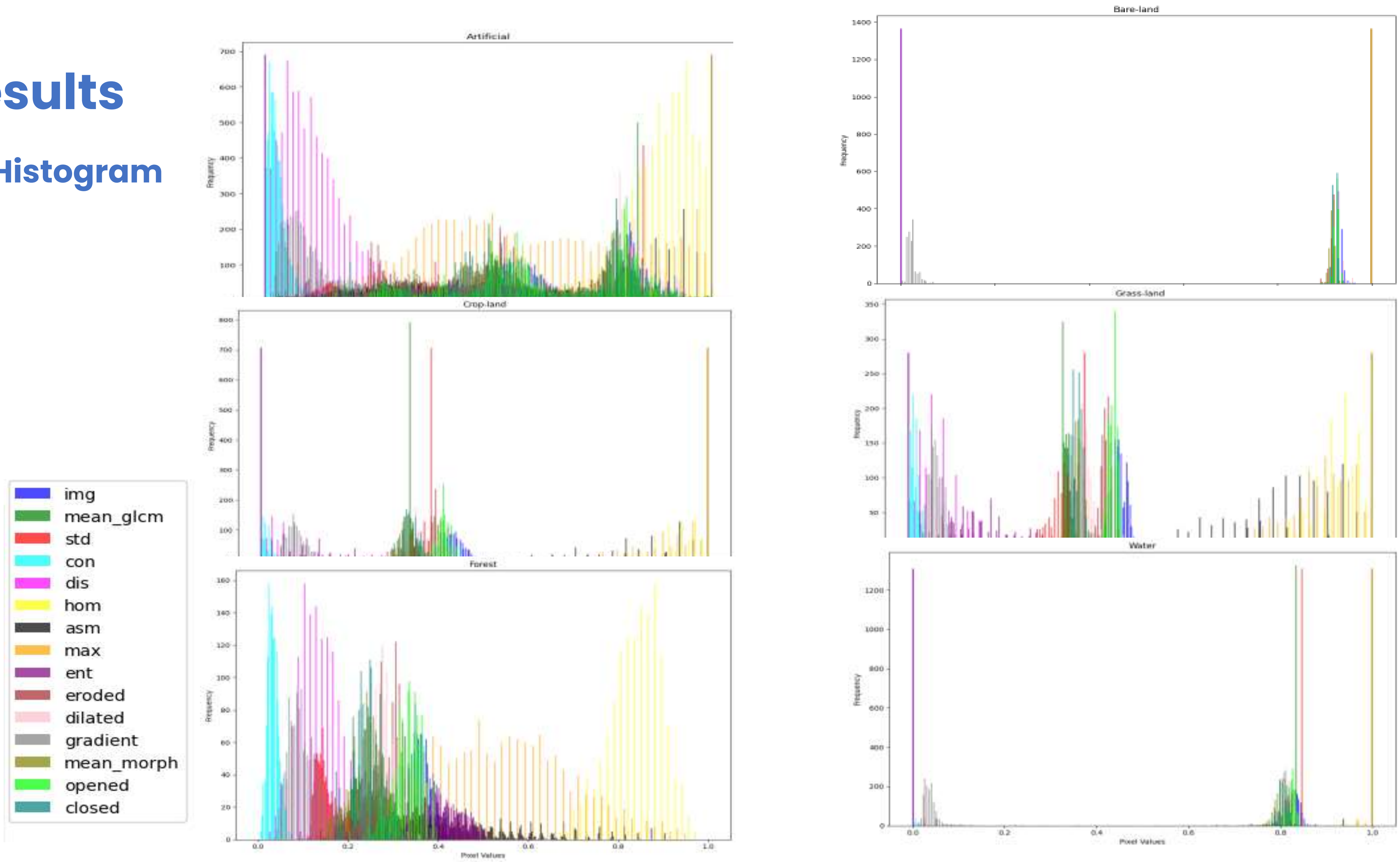

Gray-Band
mean_glcm
std
con
dis
Hom
asm
max
ent
Eroded
Dilated
Gradient
Mean-Mor
Opened
Closed

Results

GLCM (40*40 Pixel Object Level=256, Horizontal)		**Contrast**	**Dissimilarity**	**Homogeneity**	**Correlation**
Artificial		**3022.723**	26.333	0.145	0.585
Cropland		2896.085	15.82	0.225	-0.04
Water		3746.689	**49.411**	**0.027**	0.132
Bare-land		2925.122	13.847	**0.36**	**-0.023**
Grass-land		2427.002	19.905	0.078	0.057
Forest		2573.538	31.037	0.084	0.445

Results

➢ Histogram



Results

➢ Random Forest

```
rf_params = {
    'n_estimators': [  500],
    'criterion': ['entropy'],
    'max_depth': [None,  10],
    'min_samples_split': [ 5],
    'min_samples_leaf': [  4],


}
```

5. GridSearchCV

```
start_time = time.time()
grid_search = GridSearchCV(estimator=model, param_grid=rf_params, cv=5, scoring='accuracy', verbose=2, n_jobs=-1)
grid_search.fit(X_train_smote, y_train_smote)

end_time = time.time()
execution_time = end_time - start_time
print("Execution time:", execution_time / 60, "Minutes")
# Get the best model
best_model = grid_search.best_estimator_
```

```
Fitting 5 folds for each of 2 candidates, totalling 10 fits
Execution time: 31.626696328322094 Minutes
```

```
y_pred1 = best_model.predict(X_test)
accuracy = accuracy_score(y_test, y_pred1)
print(f'Best Model Accuracy: {accuracy}')

# Print best parameters
print('Best parameters found by grid search are:', grid_search.best_params_)
```

```
Best Model Accuracy: 0.9006016847172081
Best parameters found by grid search are: {'criterion': 'entropy', 'max_depth': None, 'min_samples_leaf': 4, 'min_samples_spl
it': 5, 'n_estimators': 500}
```

Results

Random Forest

➢ Model performance

Gray-Band

Execution time: 6.194768865903218 Minutes
Model Accuracy: 0.5260328920978741
Classification Report:

	precision	recall	f1-score	support
1	0.60	0.18	0.27	3112
2	0.30	0.36	0.33	1436
3	0.16	0.44	0.23	543
4	0.51	0.72	0.60	1541
5	0.38	0.41	0.39	1960
6	0.79	0.83	0.81	3114
7	0.98	1.00	0.99	759
accuracy			0.53	12465
macro avg	0.53	0.56	0.52	12465
weighted avg	0.57	0.53	0.51	12465

15-Band Dataset

Execution time: 72.08796071211496 Minutes
Model Accuracy: 0.9006016847172081
Classification Report:

	precision	recall	f1-score	support
1	0.92	0.88	0.90	3092
2	0.82	0.82	0.82	1487
3	0.78	0.90	0.84	501
4	0.90	0.92	0.91	1530
5	0.87	0.86	0.86	2032
6	0.95	0.95	0.95	3102
7	1.00	1.00	1.00	721
accuracy			0.90	12465
macro avg	0.89	0.90	0.90	12465
weighted avg	0.90	0.90	0.90	12465

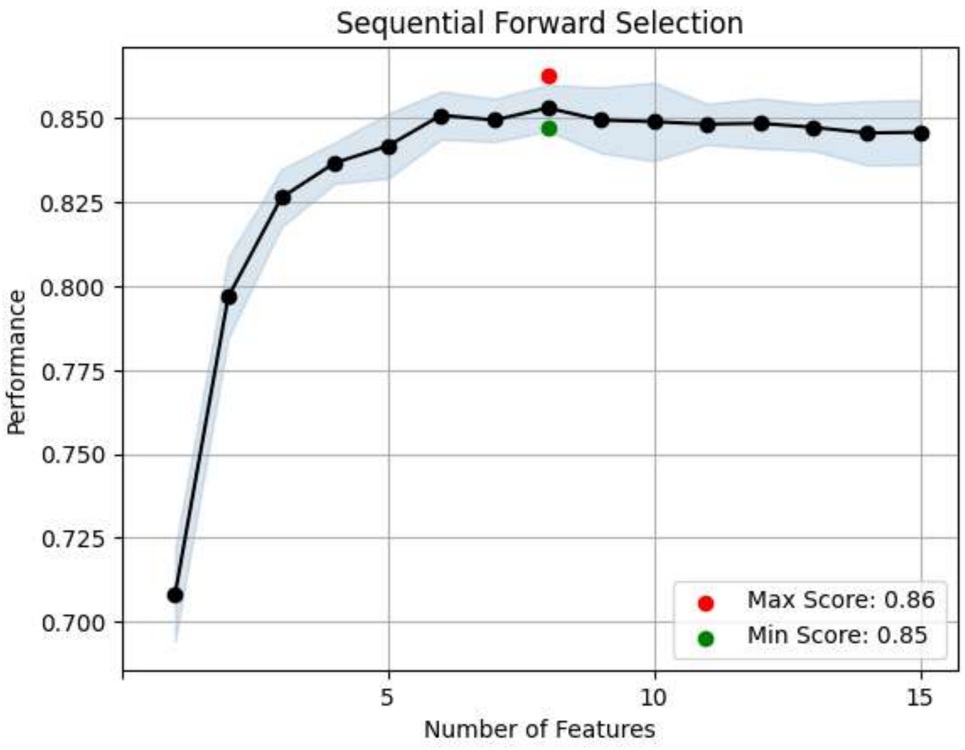
SFS (7 features)

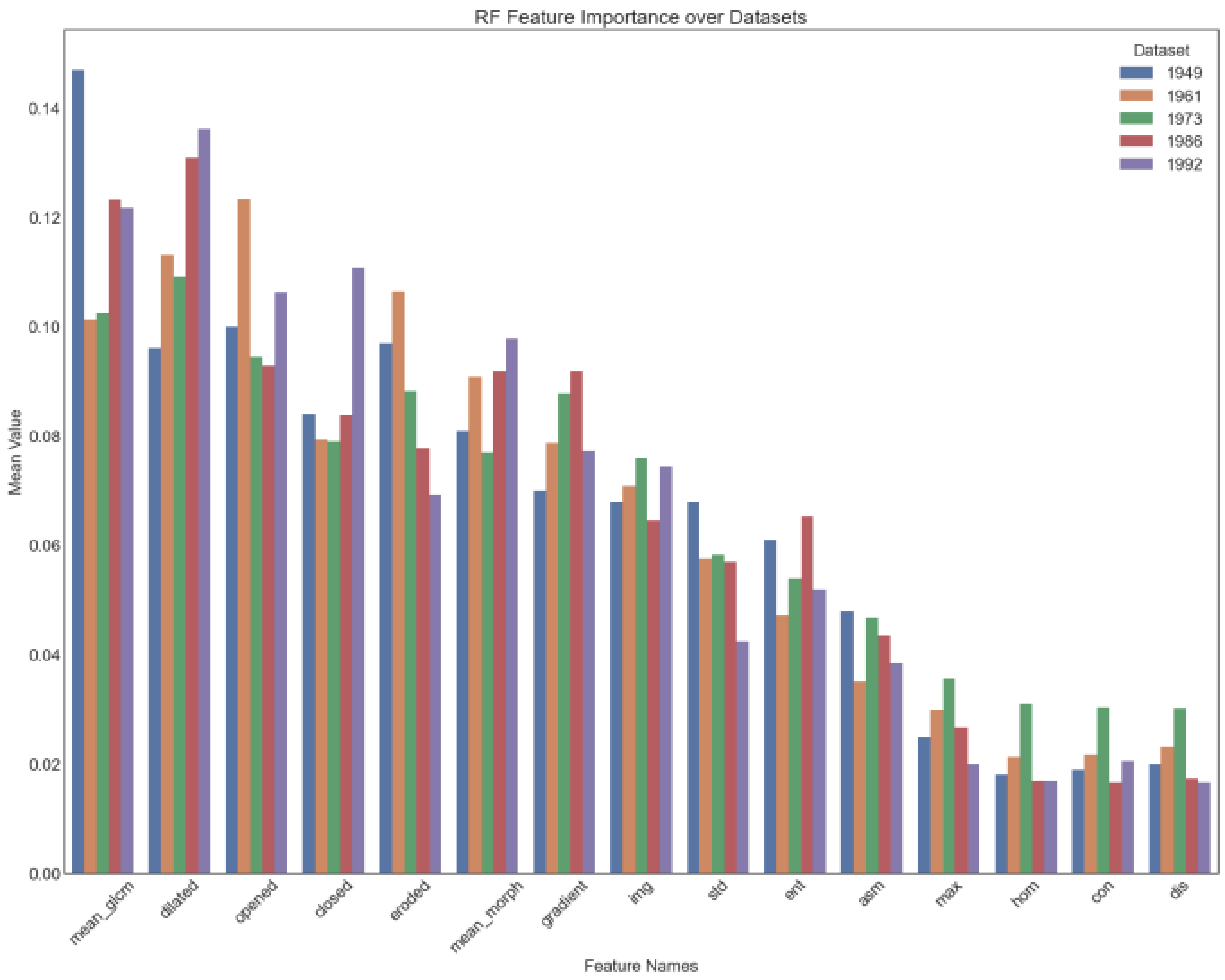
Execution time: 10.064293356736501 Minutes
Model Accuracy: 0.8835940633774568
Classification Report:

	precision	recall	f1-score	support
1	0.90	0.86	0.88	3119
2	0.78	0.81	0.79	1454
3	0.73	0.86	0.79	533
4	0.87	0.90	0.88	1425
5	0.85	0.84	0.85	2028
6	0.95	0.94	0.94	3148
7	1.00	1.00	1.00	758
accuracy			0.88	12465
macro avg	0.87	0.89	0.88	12465
weighted avg	0.89	0.88	0.88	12465

Results

➢ Feature Selection





Results

➢ Model Performance

Model Performances on 15-Band dataset

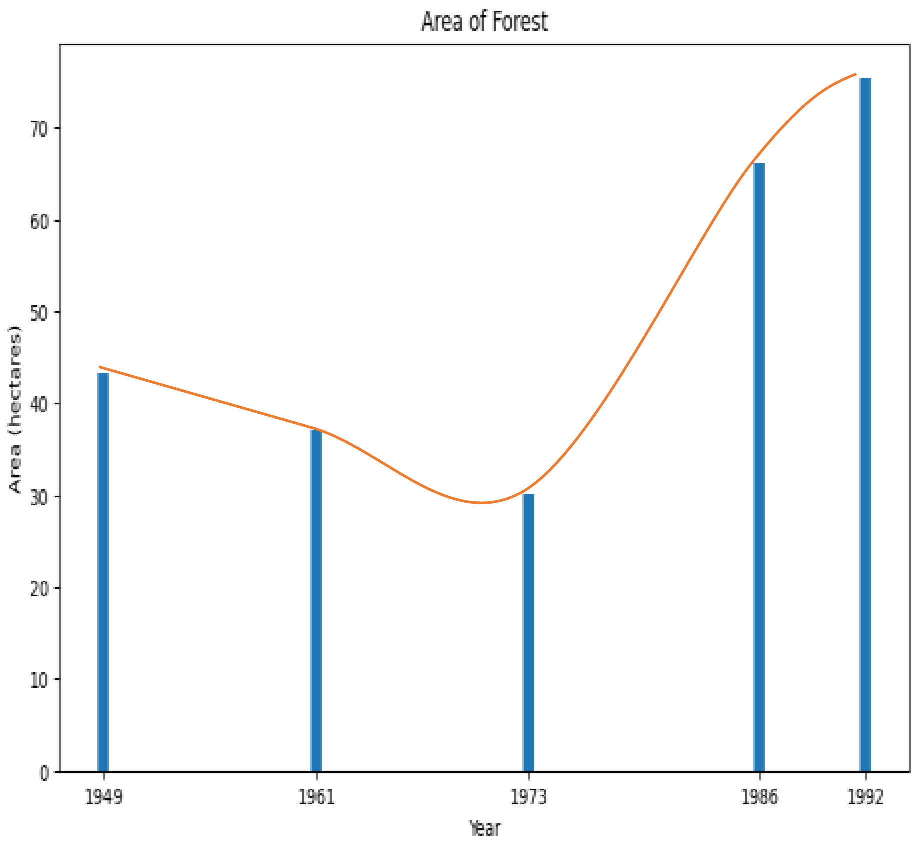
Classifier	Hyperparameters	Value	Model Performance		
			MA (%)	F1 (%)	AET(Minute)
RF	n_estimators criterion max_depth min_samples_split min_samples_leaf	500 entropy None 4 5	90.0	89.0	379
LightGBM	boosting_type learning_rate num_leaves n_estimators max_depth objective	gbdt 0.2 100 800 10 multiclass	78.0	76.0	425
XBoost	n_estimators max_depth gamma eta max_leaves subsample colsample_bytree	800 20 0.09 0.2 80 0.5 0.9	77.8	75.0	53

OS specification:

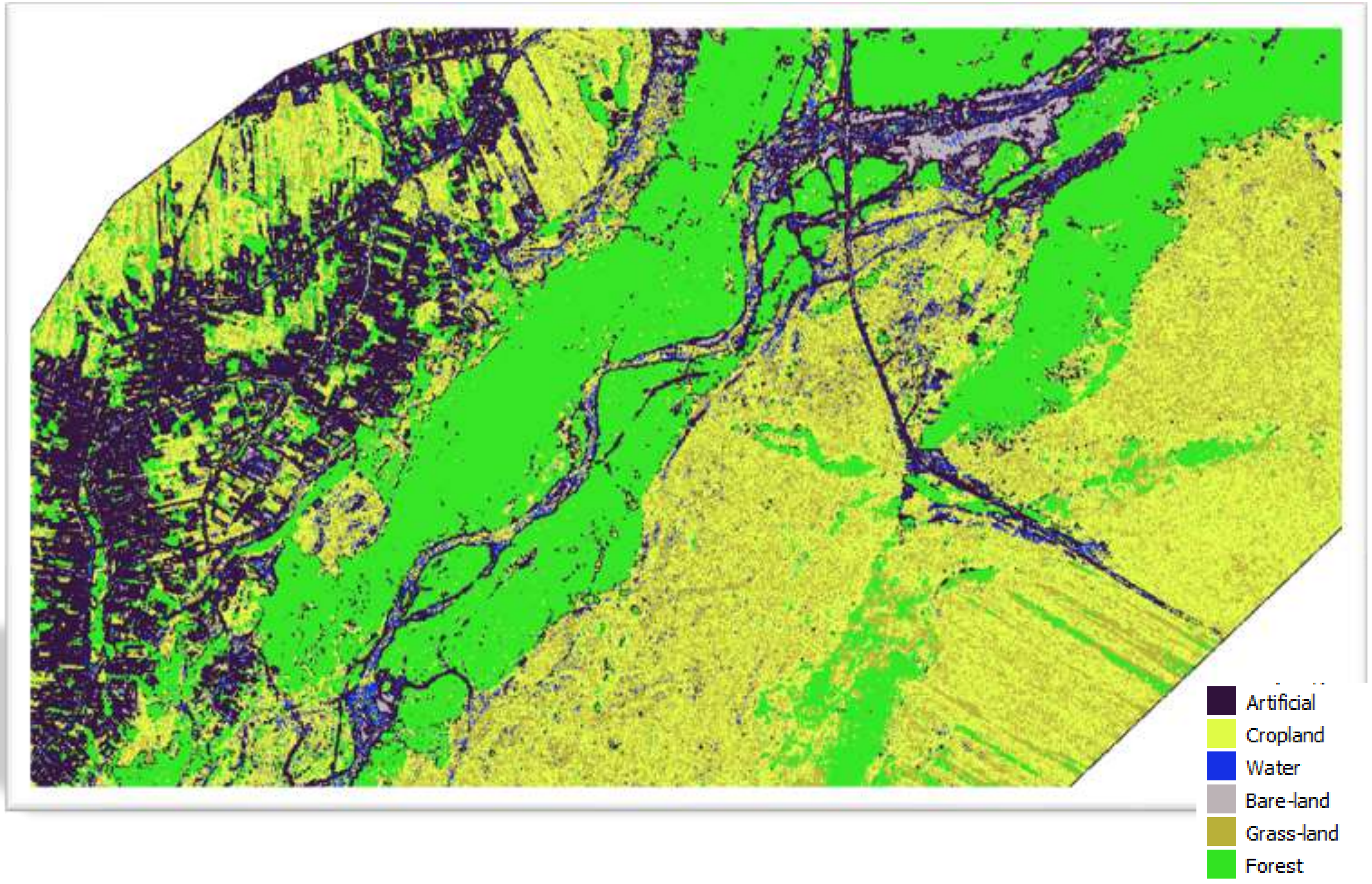
- 12th Gen Intel(R) Core(TM) i7-12700K 3.60 GHz
- 128 GM RAM

Results

➢ **RF Classification**



1992



Main Takeaways

- Incorporating GLCM and Geomorphological features significantly improves the classification accuracy
- Feature selection methods, including RF importance scores and SFS, are important in reducing data complexity and dimensionality
- We observed significant model accuracy with the RF compared to another models
- Despite advancements, issues such as shadow effects and radiometric differences in black-and-white orthophotos continue to pose challenges

INSTITUTE
OF GEOGRAPHY SLOVAK
ACADEMY OF SCIENCES

THANK YOU FOR YOUR ATTENTION

TEXTURE-BASED TECHNIQUES FOR VEGETATION CLASSIFICATION IN RIPARIAN ZONES USING HISTORICAL AERIAL ORTHOPHOTOS

github.com/hamidafzal/RF_aerialIMG

Afzali, Hamid, *hamid.afzali@savba.sk*

Rusnák, Miloš, Mgr., PhD.

Opravil, Šimon, Mgr.

GeoKARTO 2024

5 - 6 SEPTEMBER