Semantic-Search: A Knowledge-Driven Classification Method for Plant Diseases

Presenter: Kiran Kumar Kethineni

Kiran Kumar Kethineni[1], Saraju P. Mohanty[2], E. Kougianos[3]

University of North Texas, USA.[1,2,3], USA

Email: kirankumar.kethineni@unt.edu[1], saraju.mohanty@unt.edu[2], elias.kougianos@unt.edu[3]

Agriculture: The Foundation of Life

- Agriculture is the foundation of the food system.
- Agriculture is a major contributor to the global economy.
- The global human population is projected to reach 9.7 billion by 2050 and 10.9 billion by 2100.
- Ensured Food security and food safety.

Source of Food

Source of Livelihood

Source of Employment

Source of Income

Contribution to Trade

Evolution of Smart Agriculture

Traditional Agriculture

- Manual labor, experience-based decision-making, simple tools.

Green Revolution

- Synthetic fertilizers, pesticides, and high-yield crop varieties .

Precision Agriculture

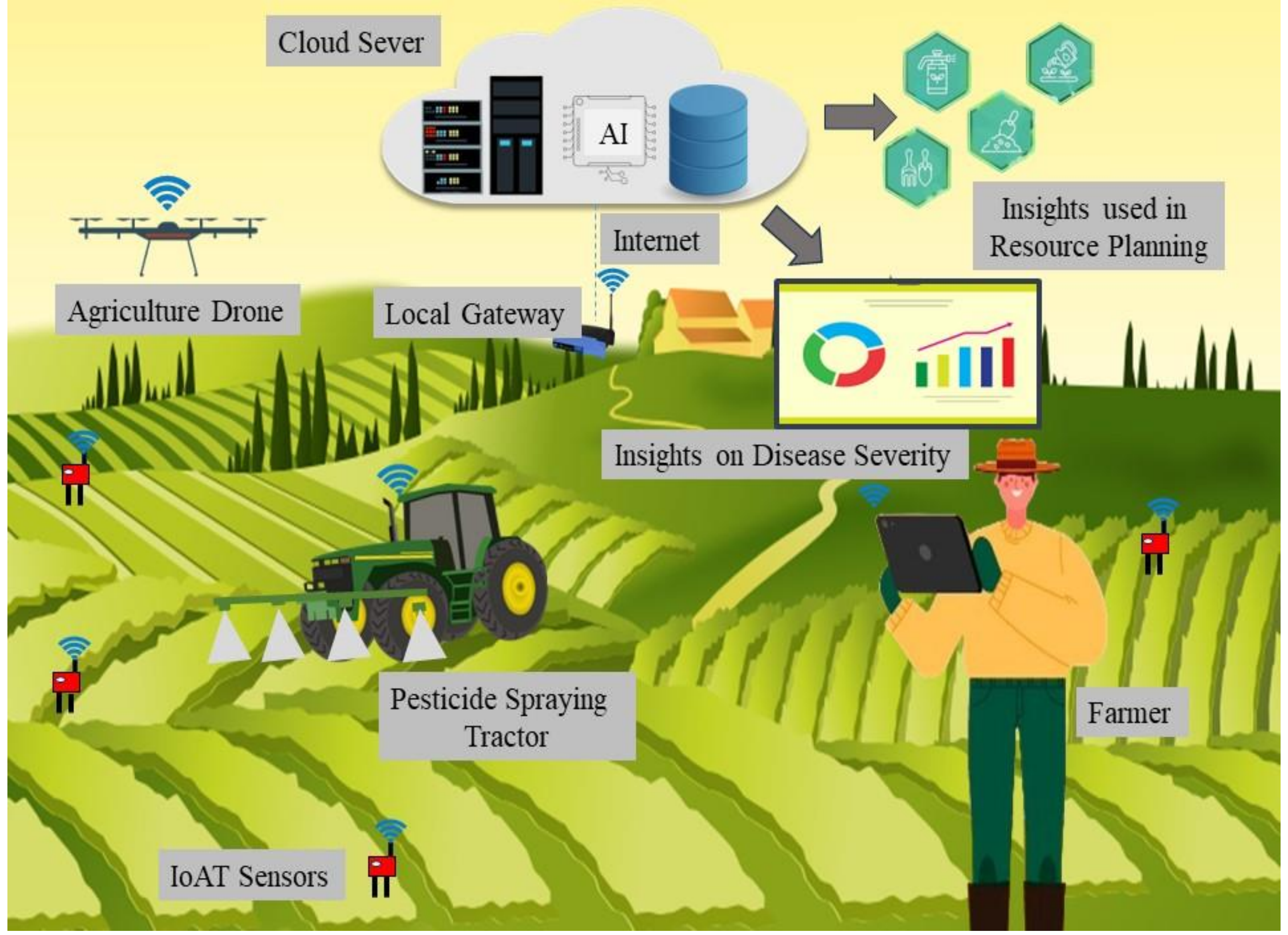
- sensors, and data analytics for automation.

Smart Agriculture

- IoT, machine learning, and big data analytics.

Agricultural Cyber-Physical Systems (ACPS)

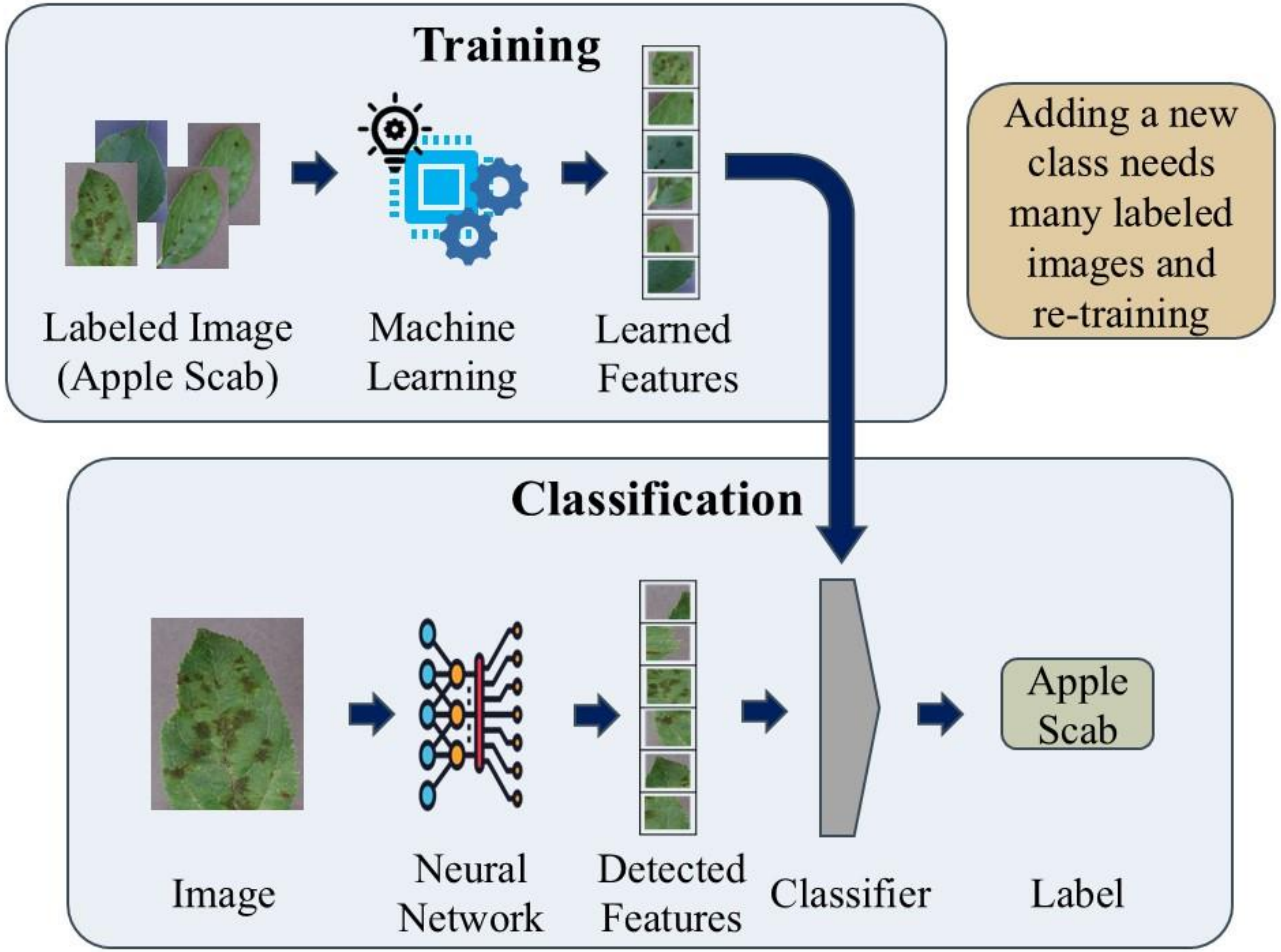
- ACPS: A system that integrates physical entities, sensors, and digital technologies for real-time monitoring and control of agricultural processes.
- Plant diseases lead to economic losses, need timely intervention.
- ACPS facilitates early disease detection using computer vision.
- Continuous monitoring via ACPS allows for effective damage control.



Disease management ACPS

How Computer Vision Works?

Data Driven

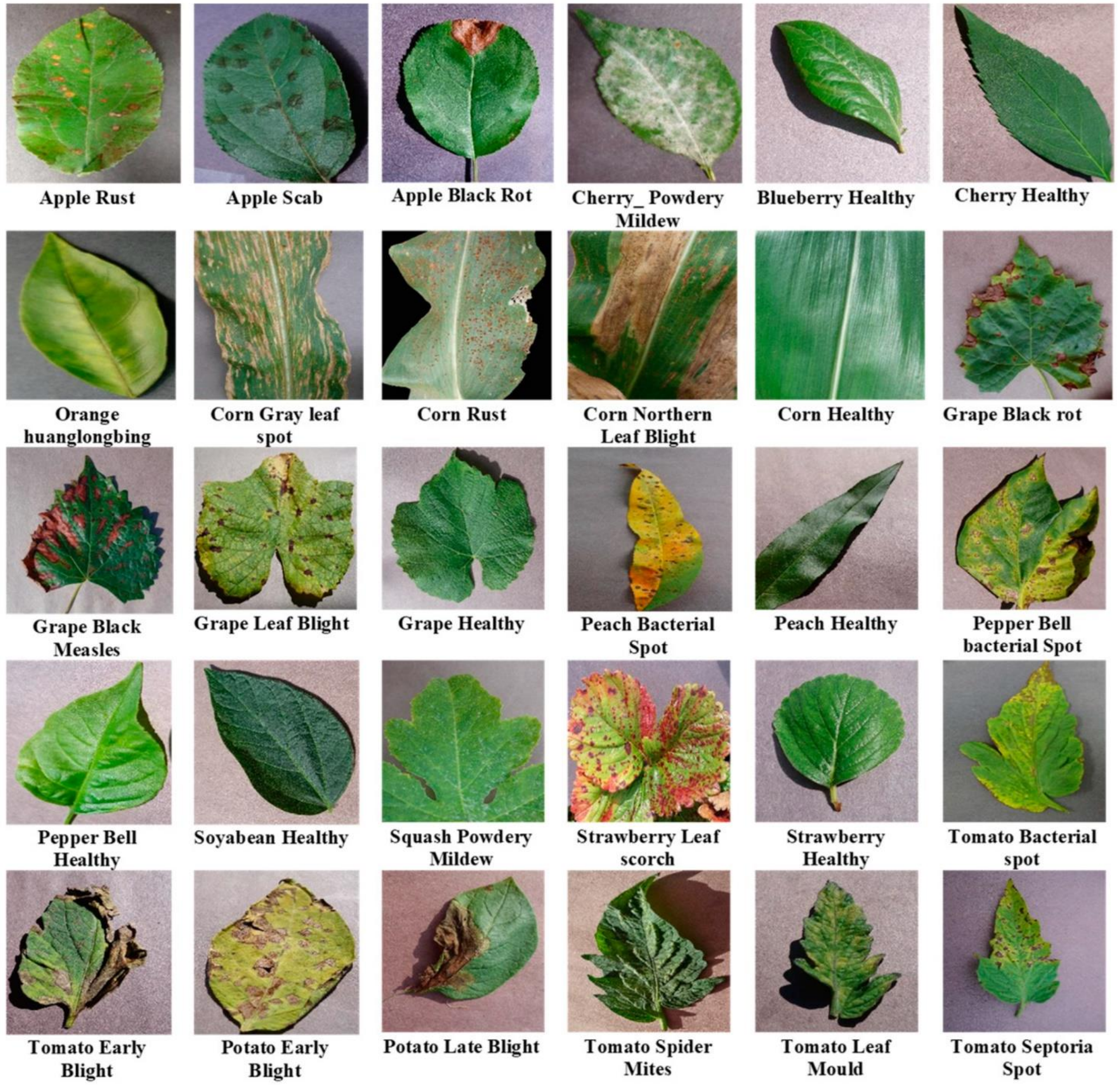


The Problem

Adding a new class to the classifier needs computationally intensive retraining.

Developing a model that can detect all the diseases needs a huge number of labeled images.

There are large number of plant types and various diseases in each of them.

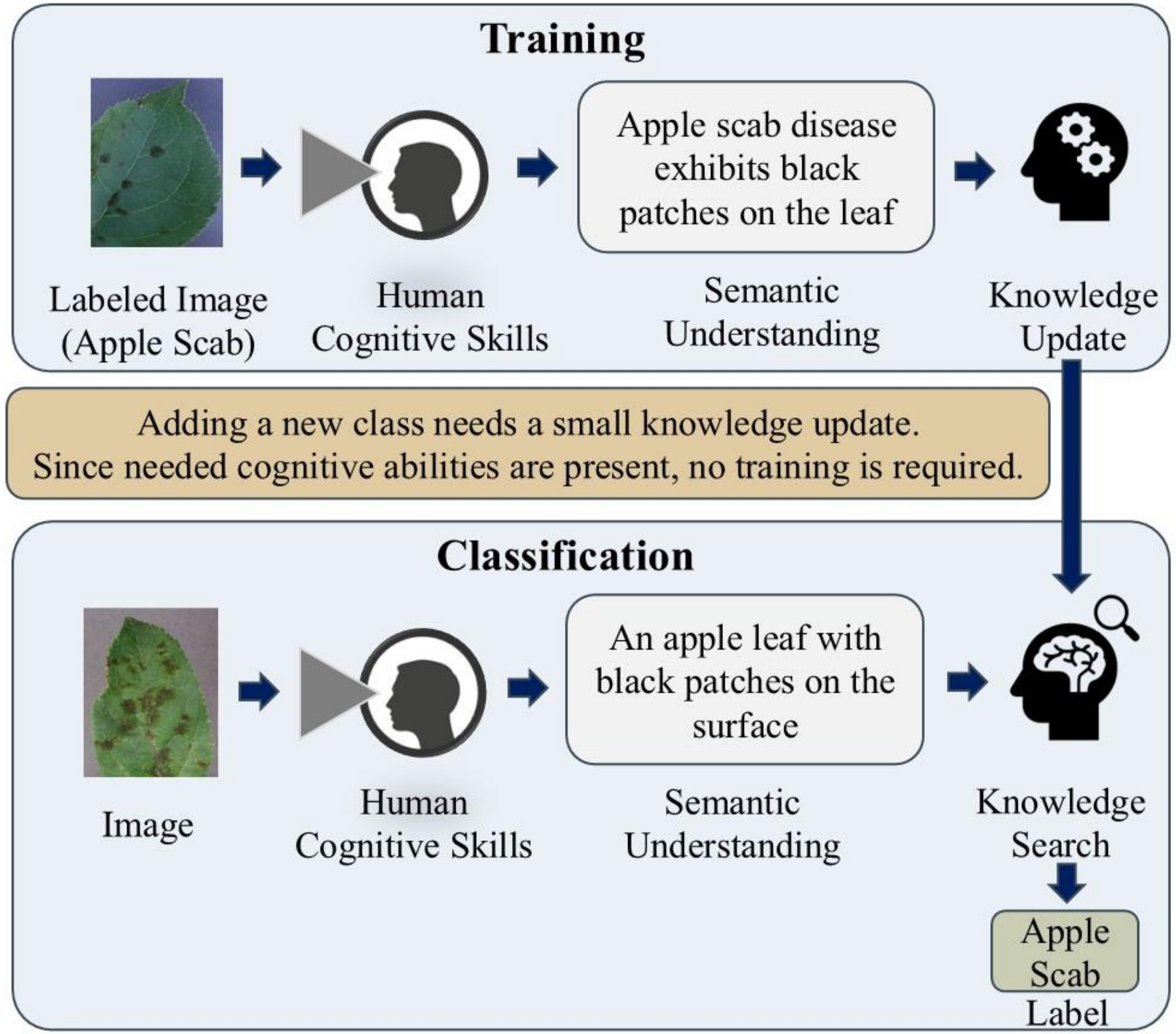


Different types of plant diseases

Alternative to Data Driven Methods?

Knowledge Driven!

- Ex: Human Vision



Related Works

Work	Year	Factors considered	Remark
Park, et al.	2004	Texture features classified by neural network	Lacks semantic understanding
Agrawal, et al.	2011	Color histograms are classified by a SVM	Lacks semantic understanding
Vailaya, et al.	2001	Low level features are used hierarchically to classify image	Adding new class needs retraining of Bayesian networks
Yang, et al.	2007	Bag of visual words	Lacks semantic understanding
Su, et al.	2012	Bag of visual words and semantic attributes	Adding a new class needs re-training

Related Works

Work	Year	Factors considered	Remark
Marino, et al. and Menglong, et al.	2017 2019	Searching knowledge map with objects detected in the image	Implementing and traversing knowledge maps is complex
Jearanaiwongkul, et al.	2018	Ontology based classification using farmer’s findings	Not fully automated, farmers findings are used as inputs
Semantic-Search (Current Paper)	**2024**	**Semantic understanding with knowledge base search**	**Has semantic understanding and does not need retraining**

Proposed Solution

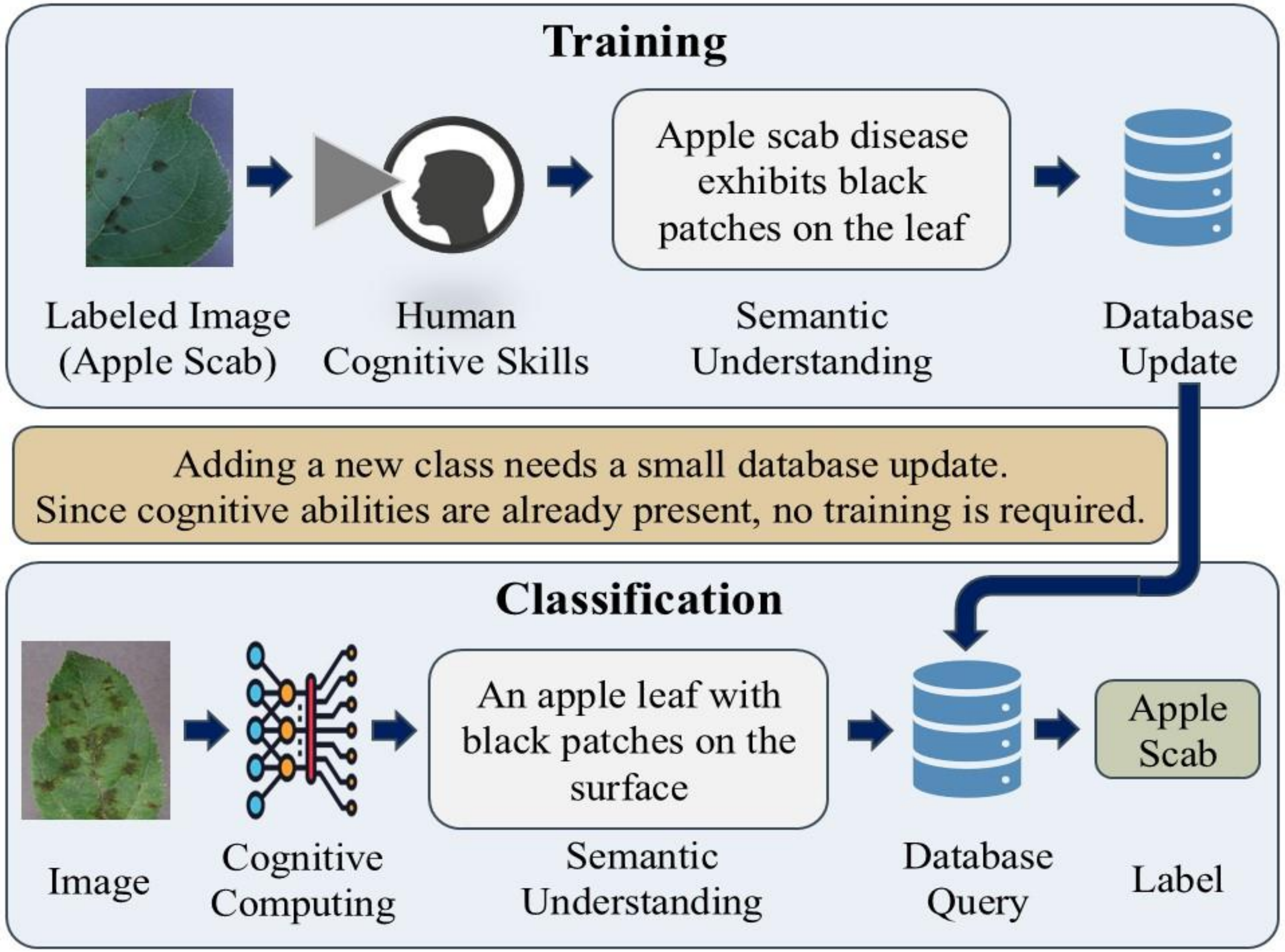
- Each disease in a plant is just a combination of a few visual features like patterns, shapes, textures, and colors.
  - For example, Apple Scab and Grape Leaf Blight exhibit brown to black spots on the leaf.
- By understanding the disease semantically and searching the knowledge base with the information, most of the diseases can be classified.


Apple Scab

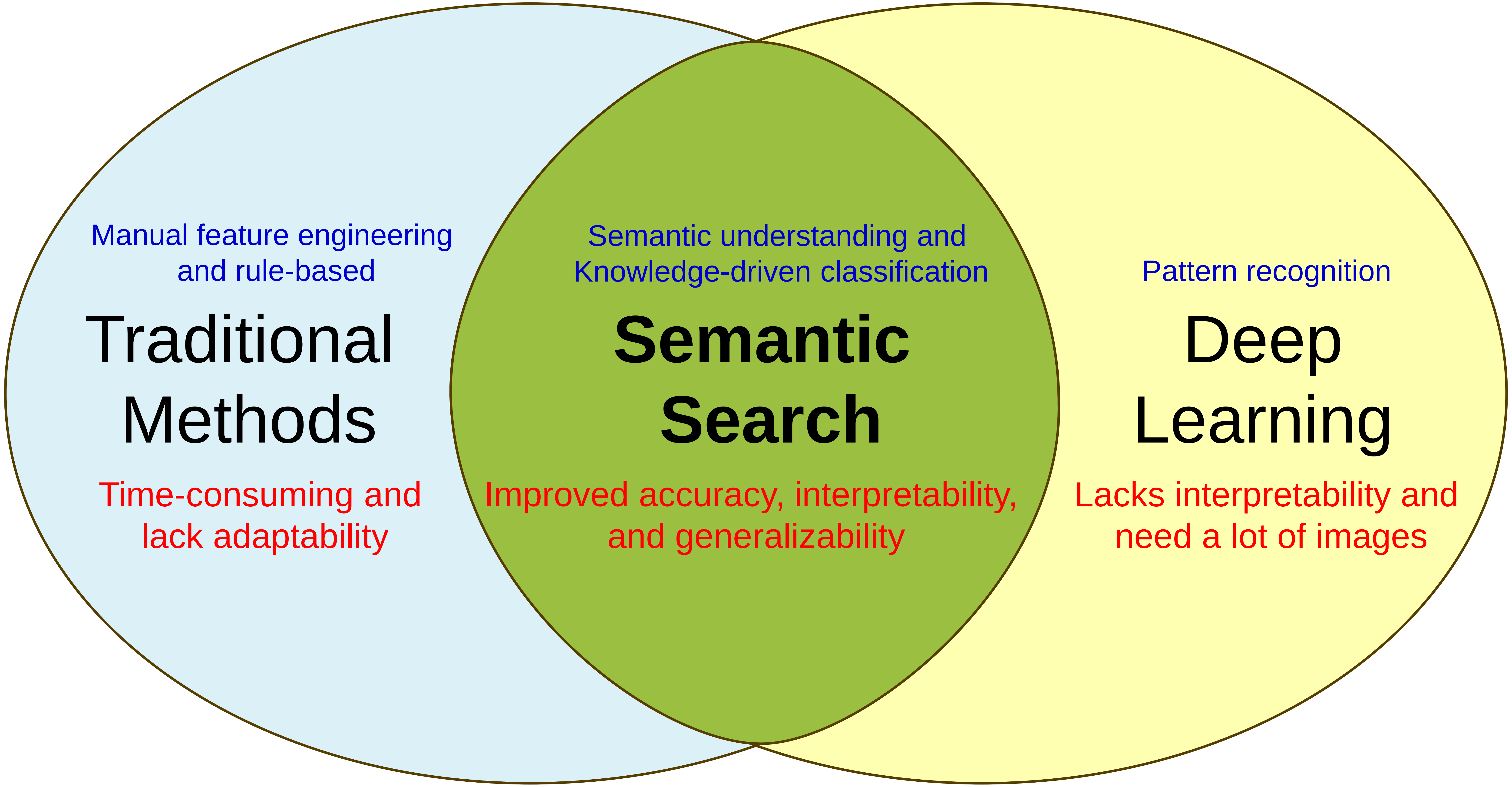

Grape Leaf Blight

Overview of Proposed "Semantic-Search"

- The Neural Networks are trained to detect Disease Semantics.
- An expert creates a database for all plant diseases, with their semantic features.
- By querying this database, the disease can be classified.



Comparison to State-of-the-Art



Novel Contributions

Semantic Understanding

- Method focuses on analyzing patterns, and objects in the diseased area.

Knowledge-Driven Approach

- The classification proposed is driven by a knowledge Database.

Interpretability and Explainability

- Presents the description of the disease explaining the classification decision, not a block box.

Generalizability

- Proposed Neural Networks have the cognitive ability to detect semantics, and can be used for any plant disease.

Proposed Feature Engineering

- 20 different plant diseases were examined to semantically describe each disease and hand-pick the semantics.

Semantics	Instances
Shape	Spot (Spots, Lesions, Patches), Flecks, Curls, Stripes
Color	Yellow, Purple, Orange, Black, Brown, White, Red
Texture	Powdery, Mosaic, Velvety

Overview of Semantics engineered

Proposed Feature Engineering

Plant	Disease	Symptom	Semantics
Apple	Black rot	Flecks or lesions which are brown in the center and purple at margin	Objects: Flecks, Lesions Colors: Brown, Purple
Apple	Powdery mildew	White velvety patches on the underside of leaves	Texture: Velvety Color is redundant
Grape	Leaf blight	Small, brown-black spots	Objects: Spots Colors: Brown, Black
Tomato	Black mold	Appearance of black or brown lesions	Objects: Lesions Colors: Black, Brown
Tomato	Mosaic virus	Infected leaves exhibit dark green mosaic	Texture: Mosaic Color is redundant
Tomato	Blight	Yellow chlorotic lesions	Objects: Lesions Colors: Yellow

A brief overview of a few disease semantics.

Proposed Entity Relationship

Diseases

Id	Plant	Disease
1	Grape	Leaf blight
2	Tomato	Black mold
3	Apple	Powdery mildew

Disease colors

Disease Id	Color Id
1	1
1	2
2	1
2	2
3	Null

Disease Shapes

Disease Id	Color Id
1	1
2	1
2	2
3	Null

Disease Textures

Disease Id	Color Id
1	Null
2	Null
3	2

Texture

Id	Value
1	Velvet
2	Powder
3	Mosaic

Shape

Id	Value
1	Spot
2	Fleck
3	Curl
4	Stripe

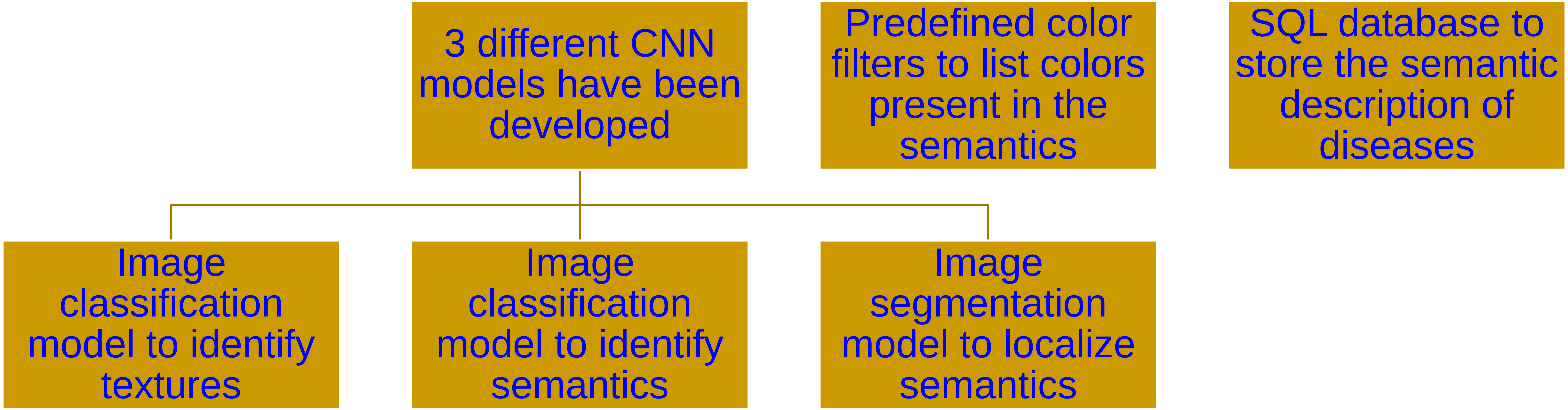
Color

Id	Value	Id	Value
1	Black	5	Orange
2	Brown	6	Purple
3	Yellow	7	White
4	Red		

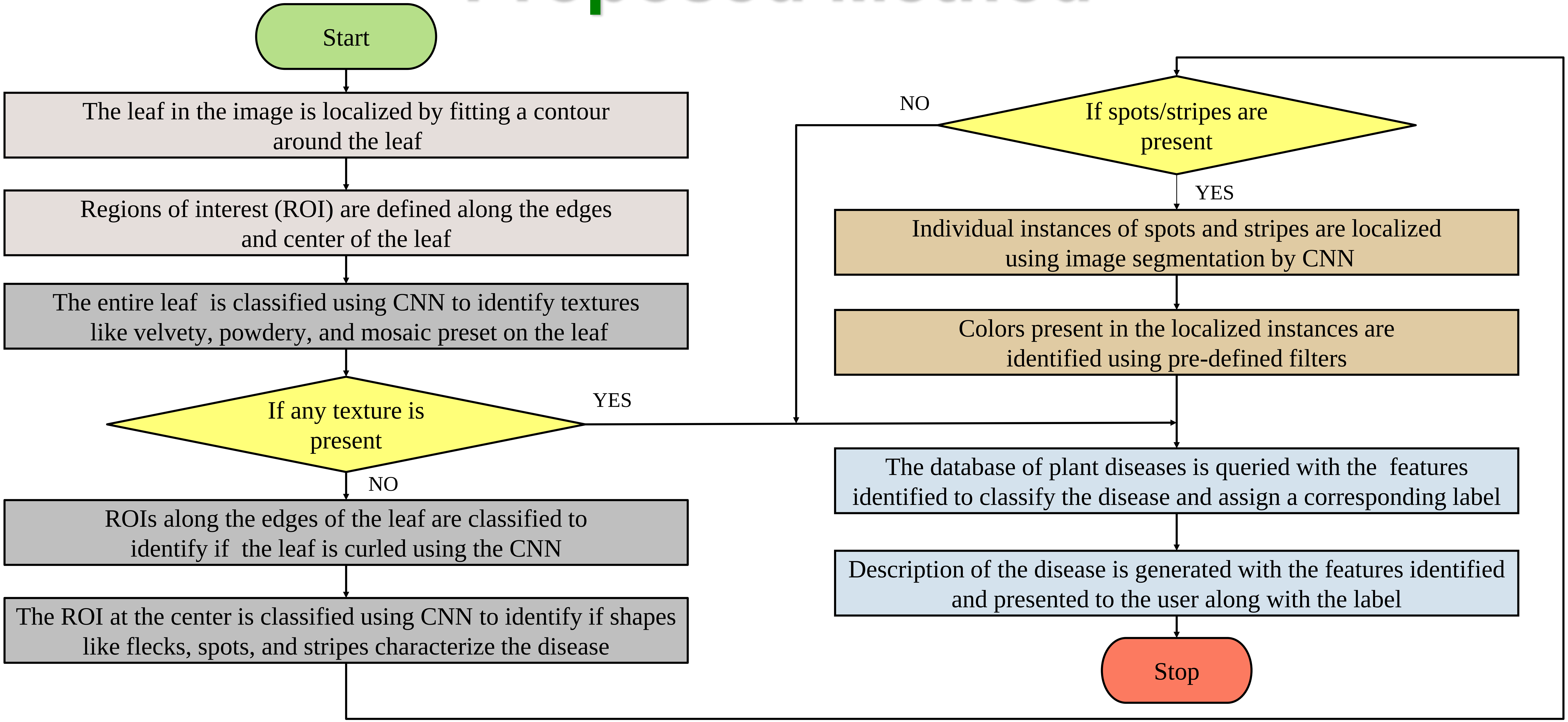
Entity-Relationship diagram of the database.

Proposed Method

- The proposed Semantic-Search has the following components working in sequence.



Proposed Method

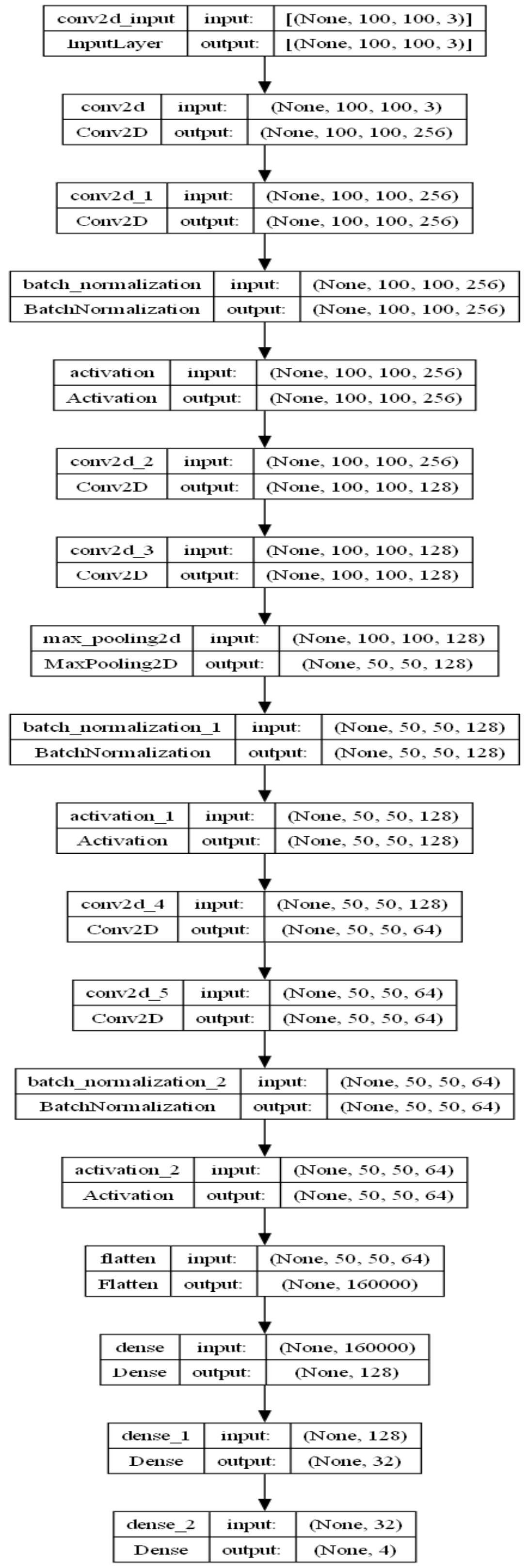


Working of Semantic-Search.

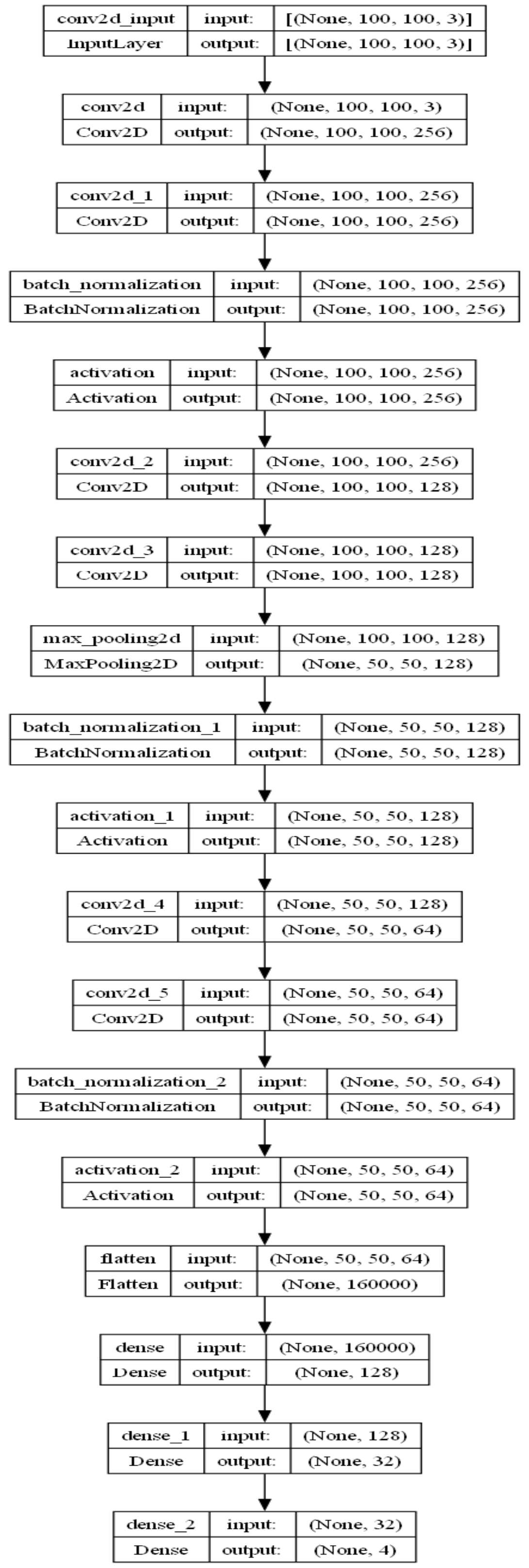
Implementation

- The proposed "Semantic-Search" has been experimentally validated on the **PlantVillage** database with **2000** images of **20** different diseases from **5** types of plants.
- The solution was developed in Python using Tensorflow and Keras libraries.
- CNNs were used for image detection, and classification.
- SQLite was used to create and manipulate SQL Server
- Proposed method achieved an accuracy of **94%**.

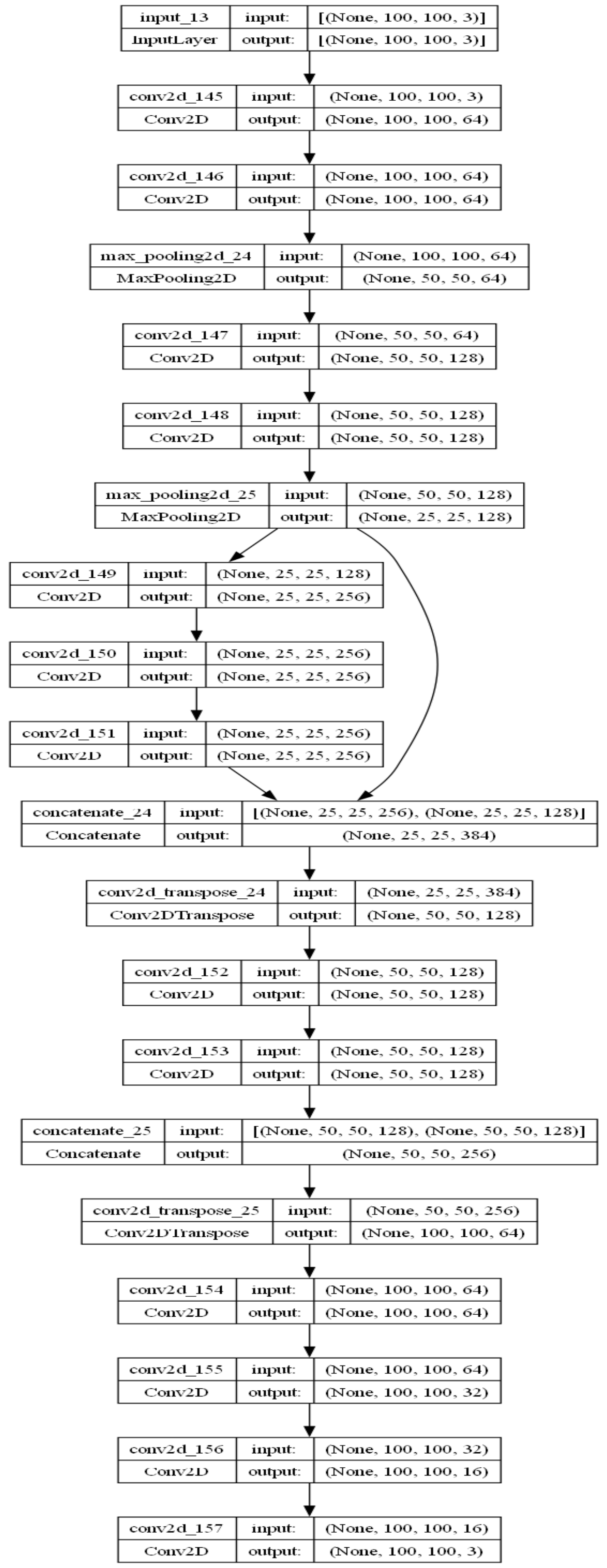
Implementation



Texture detection



Object detection



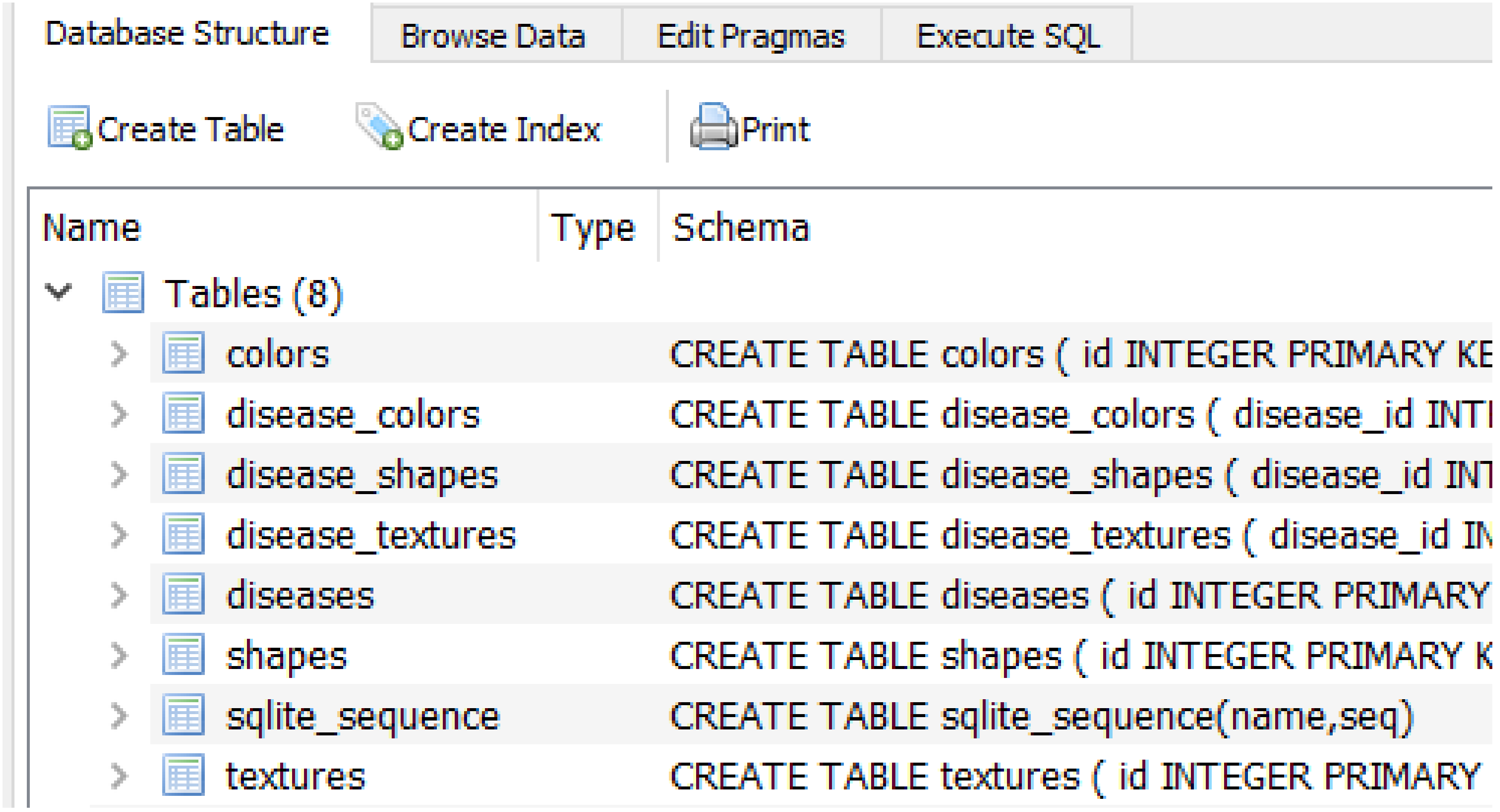
Object Localization

Implementation

```python
# Define the lower and upper bounds for each color in HSV color space
color_ranges = {
    'orange':(np.array([ 25, 80,80]), np.array([29, 255, 255])),
    'yellow':(np.array([ 39, 80,80]), np.array([46, 255, 255])),
    'brown': (np.array([ 10, 80,80]), np.array([24, 255, 255])),
    'black': (np.array([  0,  0,  0]), np.array([255, 80,  80])),
    'purple':(np.array([187, 80,80]), np.array([201,255, 255])),
    'white': (np.array([  0,220,  0]), np.array([255,255,  20])),
    'red':   (np.array([  0, 80,80]), np.array([9,255, 255])),
    'green': (np.array([ 50, 80, 80]), np.array([110,255, 255])),

}
```

Color Filters

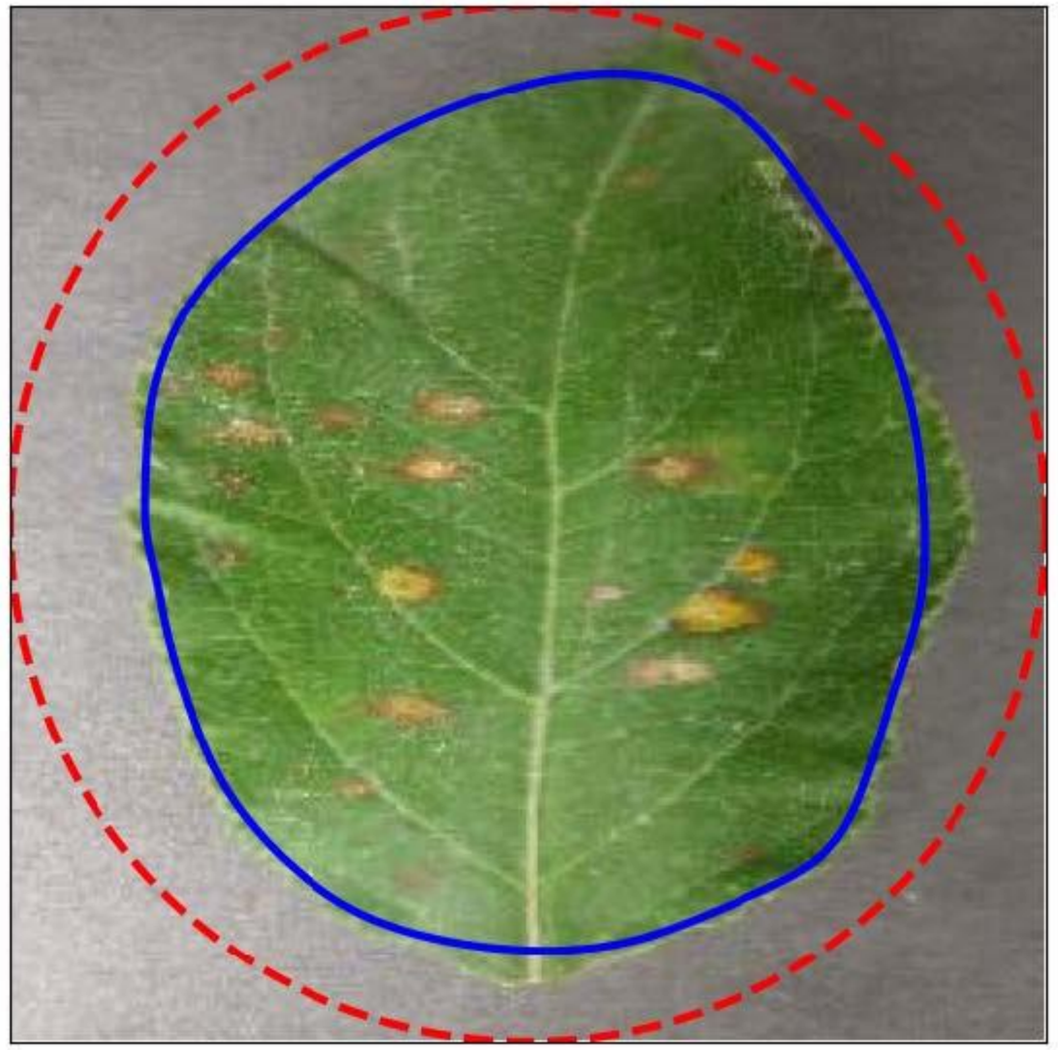


SQL Database

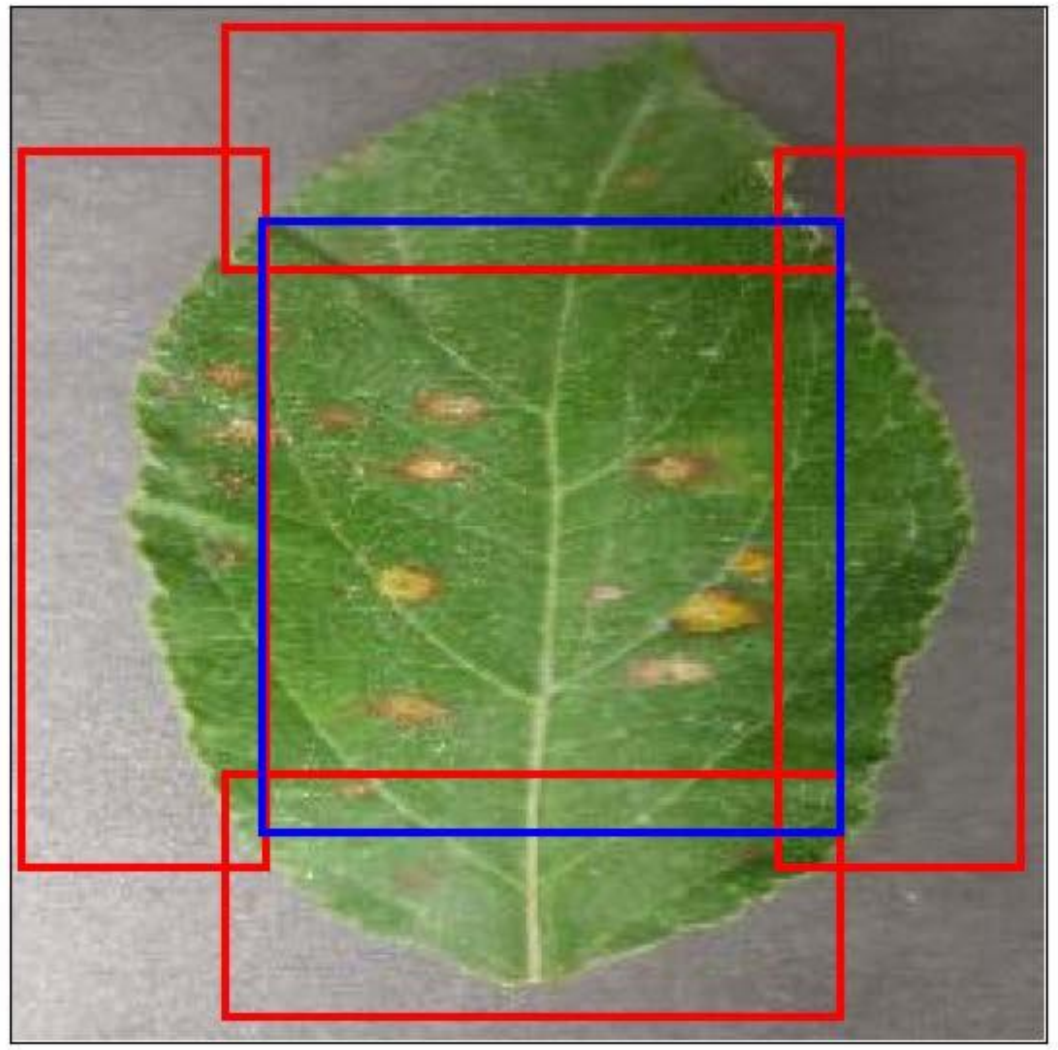
Results: Instance 1

Apple leaf with **Ceder Rust** disease

Input Image



Leaf localized by contour fitting



RoI defined

Results: Instance 1

Apple leaf with **Ceder Rust** disease

```
(1, 100, 100, 3)
Shape of images array: (1, 100, 100, 3)
1/1 [==============================] - 3s 3s/step
Teaxture of leaf is: Others
```

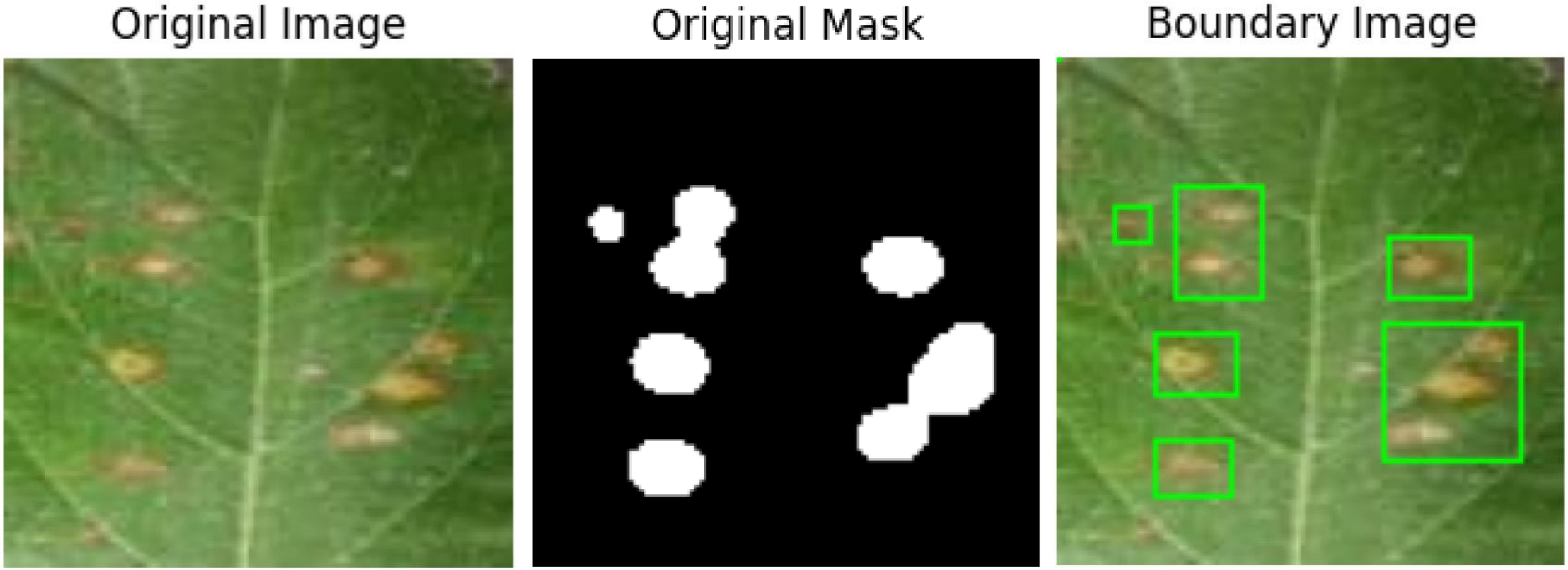
Texture prediction

```
(5, 100, 100, 3)
Shape of images array: (5, 100, 100, 3)
1/1 [==============================] - 2s 2s/step
Objects in RoI defined in leaf: Spots
Objects in RoI defined in leaf: None
Objects in RoI defined in leaf: None
Objects in RoI defined in leaf: Spots
Objects in RoI defined in leaf: None
The leaf has spots
```

Object prediction

Results: Instance 1

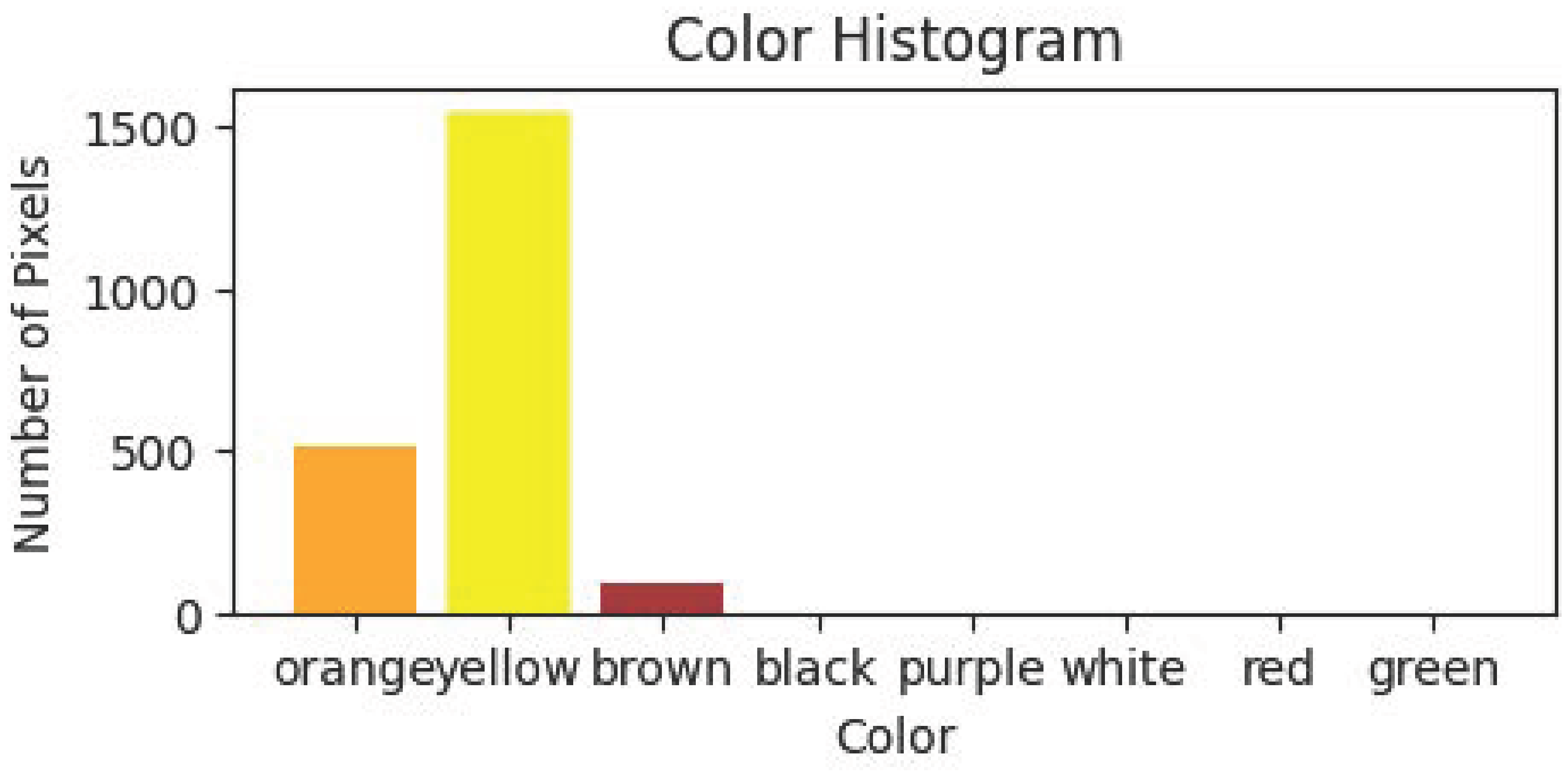
Apple leaf with **Ceder Rust** disease



Object localization by segmentation

Results: Instance 1

Apple leaf with **Ceder Rust** disease



Colors identified

SQL 1

```sql
SELECT d.disease_name
FROM diseases d
JOIN disease_shapes ds ON d.id = ds.disease_id
JOIN shapes s ON ds.shape_id = s.id AND s.value = 'spots'
JOIN disease_colors dc ON d.id = dc.disease_id
JOIN colors c ON dc.color_id = c.id AND c.value IN ('yellow', 'brown', 'orange')
WHERE d.plant_name = 'apple'
GROUP BY d.disease_name
HAVING COUNT(DISTINCT s.id) = 1
AND COUNT(DISTINCT c.id) = 3;
```

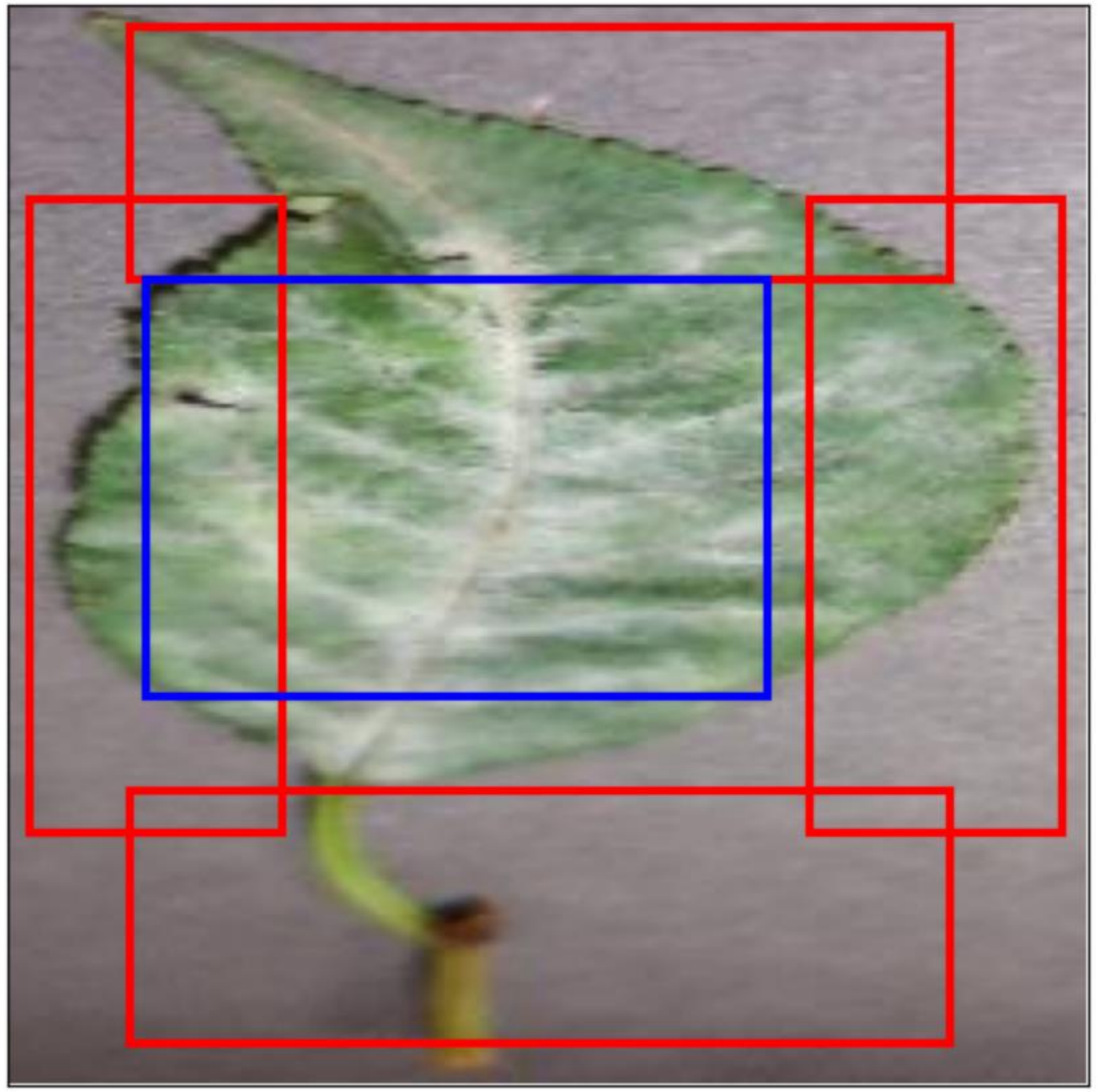
	disease_name
1	cedar rust

⇨ Label predicted by **Semantic-Search**

Query performed with identified semantics

Results: Instance 2

Cherry leaf with **Powdery Mildew** disease



Leaf localized, RoI defined

```
(1, 100, 100, 3)
Shape of images array: (1, 100, 100, 3)
1/1 [==============================] - 0s 93ms/step
Teaxture of leaf is: Powdery
```

Texture prediction

SQL 1

```sql
SELECT d.disease_name
FROM diseases d
JOIN disease_textures dt ON d.id = dt.disease_id
JOIN textures t ON dt.texture_id = t.id
WHERE d.plant_name = 'cherry' AND t.value = 'powder';
```

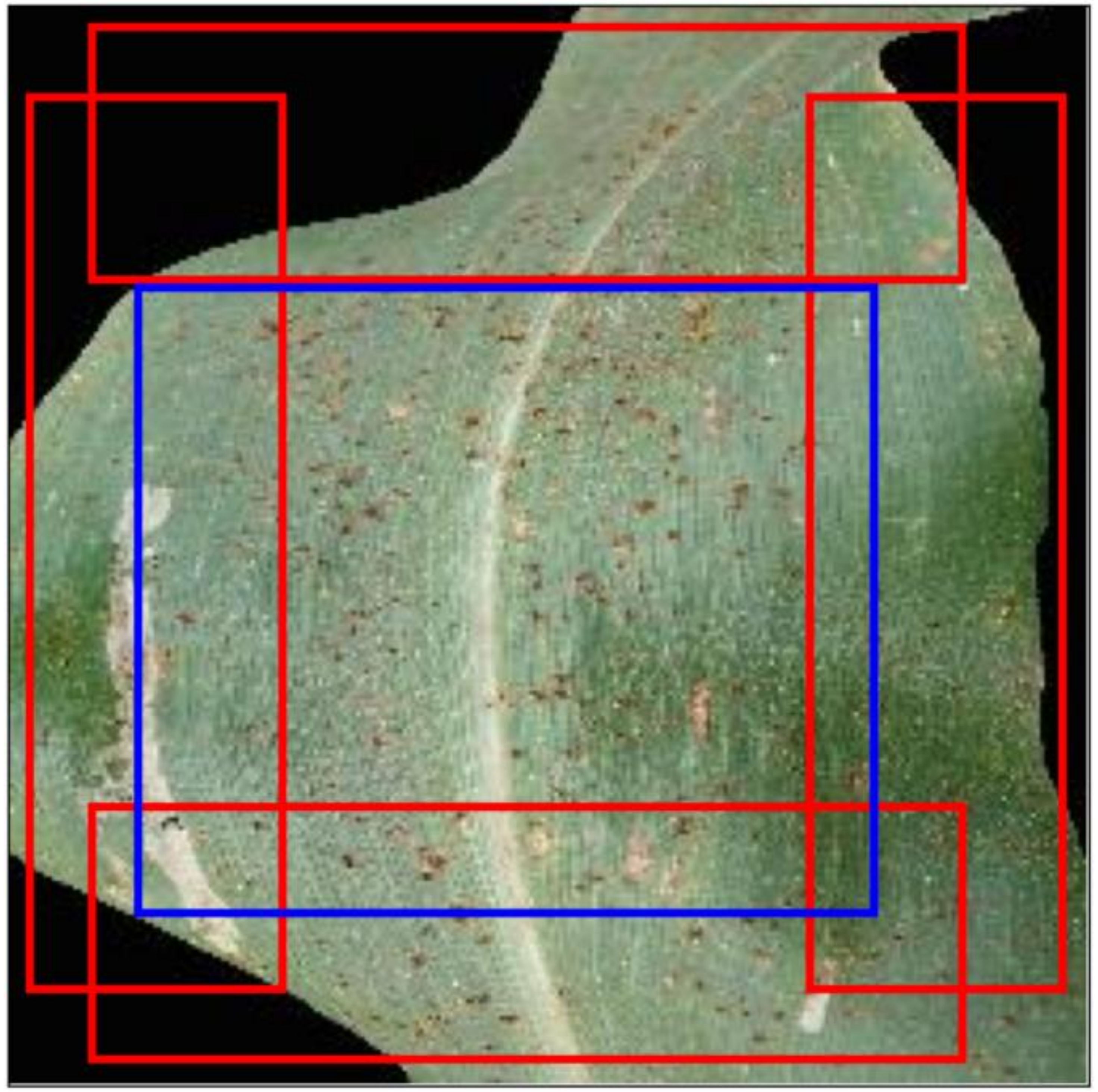
	disease_name
1	powdery mildew

⇨ Label predicted by **Semantic-Search**

Query performed with identified semantics

Results: Instance 3

Corn leaf with **Common Rust** disease



Leaf localized, RoI defined

```
(1, 100, 100, 3)
Shape of images array: (1, 100, 100, 3)
1/1 [==============================] - 0s 96ms/step
Teaxture of leaf is: Others
```

Texture prediction

Results: Instance 3

Corn leaf with **Common Rust** disease

```
(5, 100, 100, 3)
Shape of images array: (5, 100, 100, 3)
1/1 [==============================] - 0s 94ms/step
Objects in RoI defined in leaf: Flecks
Objects in RoI defined in leaf: Stripes
Objects in RoI defined in leaf: Spots
Objects in RoI defined in leaf: Flecks
Objects in RoI defined in leaf: Spots
The leaf has Flecks
```

Object prediction

SQL 1

```sql
SELECT d.disease_name
FROM diseases d
JOIN disease_shapes ds ON d.id = ds.disease_id
JOIN shapes s ON ds.shape_id = s.id
WHERE d.plant_name = 'corn' AND s.value = 'flecks';
```

	disease_name
1	common rust

Label predicted by **Semantic-Search**

Query performed with identified semantics

Conclusion and Future Work

- The process of defining ROI in the leaf tries to find the largest rectangle possible inside the contour.
- In cases of narrow and curved leaves, the ROI defined could become smaller and may not include the diseased regions
- The classification method used is static and does not change with the semantics detected.
- The usage of attention mechanisms and transformers to develop dynamic methods specific to the semantics present in the diseased can be explored in future works.

Thank You !!