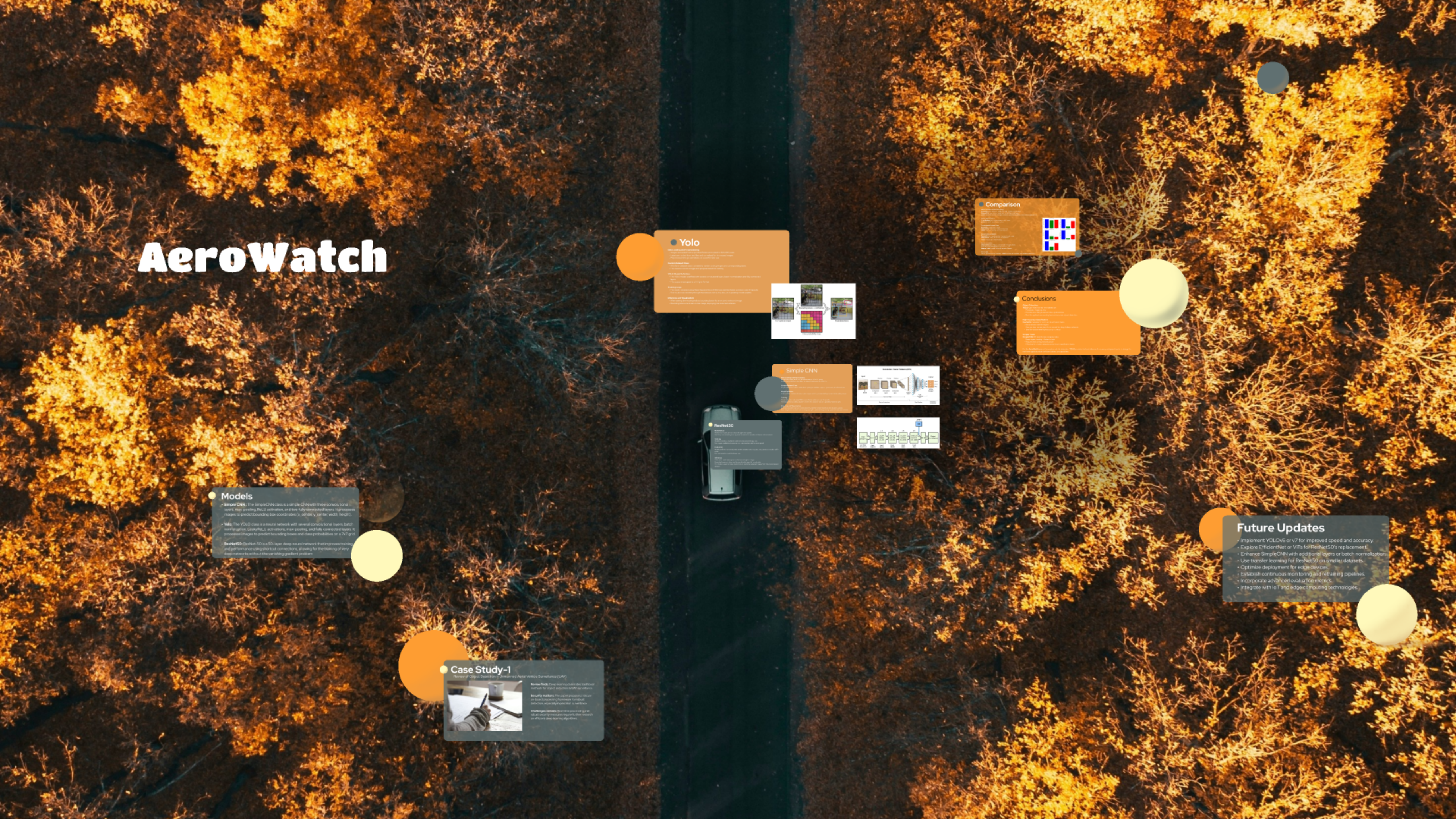

AeroWatch
Yolo
Comparison
Conclusions
Simple CNN
ResNet50
Models
- Simple CNN : The SimpleCNN class is a simple CNN with three convolutional layers, max-pooling, ReLU activation, and two fully connected layers. It processes images to predict bounding box coordinates (x_center, y_center, width, height).
- Yolo: The YOLO class is a neural network with several convolutional layers, batch normalization, LeakyReLU activations, max pooling, and fully connected layers. It processes images to predict bounding boxes and class probabilities on a 7x7 grid.
- ResNet50: ResNet-50 is a 50-layer deep neural network that improves training and performance using shortcut connections, allowing for the training of very deep networks without the vanishing gradient problem.
Case Study-1
Review of Object Detection in Unmanned Aerial Vehicle Surveillance (UAV)
Future Updates
• Implement YOLOv5 or v7 for improved speed and accuracy.
• Explore EfficientNet or ViTs for ResNet50's replacement.
• Enhance SimpleCNN with additional layers or batch normalization.
• Use transfer learning for ResNet50 on smaller datasets.
• Optimize deployment for edge devices.
• Establish continuous monitoring and retraining pipelines.
• Incorporate advanced evaluation metrics.
• Integrate with IoT and edge computing technologies.

Case Study-1

Review of Object Detection in Unmanned Aerial Vehicle Surveillance (UAV)

Review finds: Deep learning dominates traditional methods for object detection in UAV surveillance.

Security matters: The paper proposes a secure on-board processing framework for robust detection, especially in precision surveillance.

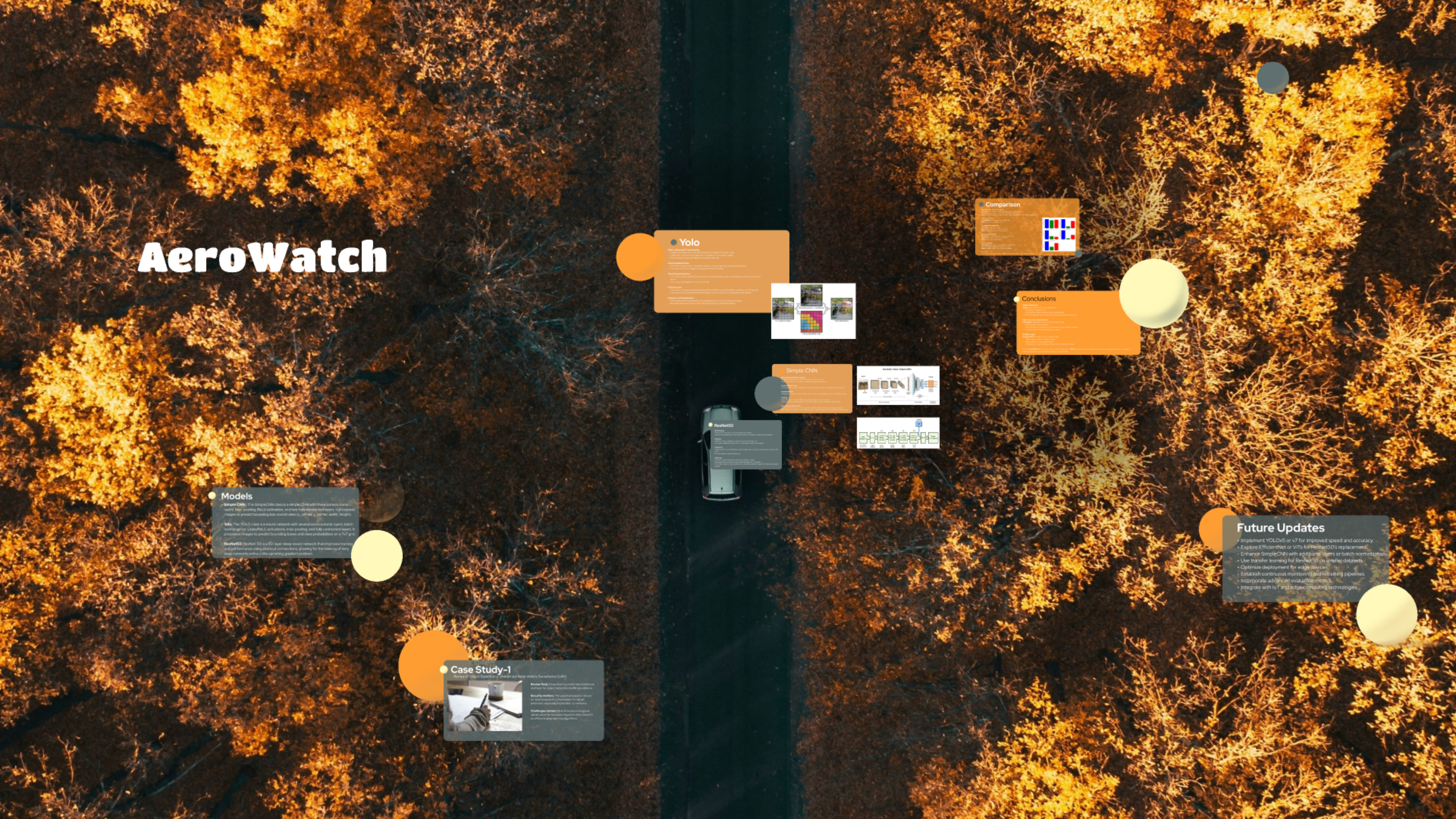
Challenges remain: Real-time processing and robust security measures require further research on efficient deep learning algorithms.

Models

- **Simple CNN :** The SimpleCNN class is a simple CNN with three convolutional layers, max-pooling, ReLU activation, and two fully connected layers. It processes images to predict bounding box coordinates (x_center, y_center, width, height).

- **Yolo:** The YOLO class is a neural network with several convolutional layers, batch normalization, LeakyReLU activations, max-pooling, and fully connected layers. It processes images to predict bounding boxes and class probabilities on a 7x7 grid.

- **ResNet50:** ResNet-50 is a 50-layer deep neural network that improves training and performance using shortcut connections, allowing for the training of very deep networks without the vanishing gradient problem.


AeroWatch
Yolo
Comparison
Conclusions
Simple CNN
ResNet50
Models
- Simple CNN: The SimpleCNN class is a simple CNN with three convolutional layers, max-pooling, ReLU activation, and two fully connected layers. It processes images to predict bounding box coordinates (x_center, y_center, width, height).
- Yolo: The YOLO class is a neural network with several convolutional layers, batch normalization, LeakyReLU activations, max pooling, and fully connected layers. It processes images to predict bounding boxes and class probabilities on a 7x7 grid.
- ResNet50: ResNet-50 is a 50-layer deep neural network that improves training and performance using shortcut connections, allowing for the training of very deep networks without the vanishing gradient problem.
Future Updates
• Implement YOLOv5 or v7 for improved speed and accuracy.
• Explore EfficientNet or ViTs for ResNet50's replacement.
• Enhance SimpleCNN with additional layers or batch normalization.
• Use transfer learning for ResNet50 on smaller datasets.
• Optimize deployment for edge devices.
• Establish continuous monitoring and retraining pipelines.
• Incorporate advanced evaluation metrics.
• Integrate with IoT and edge computing technologies.
Case Study-1
Review of Object Detection in Unmanned Aerial Vehicle Surveillance (UAV)

Yolo

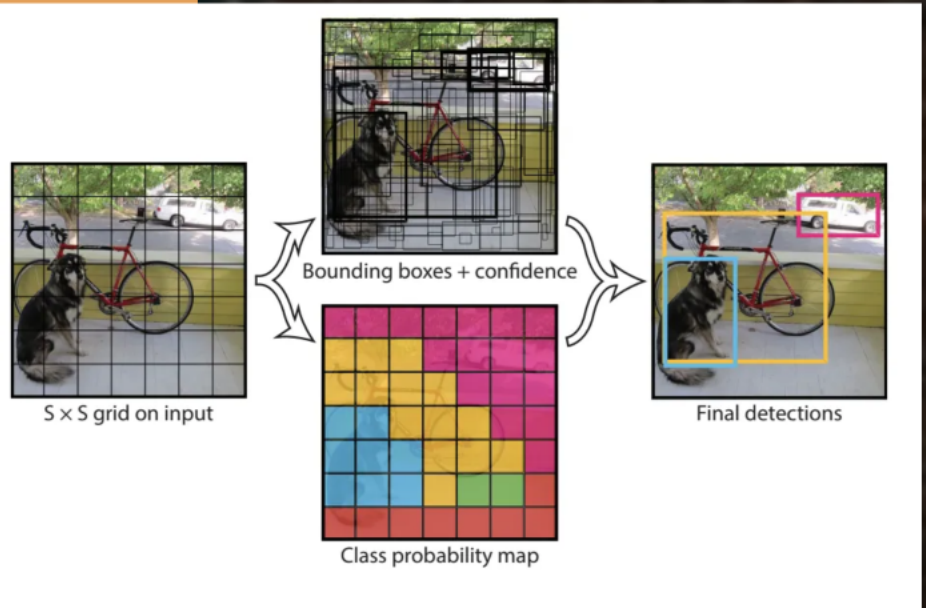
Data Loading and Preprocessing:
- Images are loaded from a specified folder and resized to 320x320 pixels.
- Labels are loaded from text files and normalized for the resized images.
- Preprocessed images and labels are saved for later use.

Custom Dataset Class:
- A PyTorch Dataset class is created to handle loading images and corresponding labels.
- The class transforms images and prepares labels for training.

YOLO Model Definition:
- The YOLO model is defined with several convolutional layers, batch normalization, and fully connected layers.
- The output is reshaped to a 7x7 grid format.

Training Loop:
- The model is trained using Mean Squared Error (MSE) loss and the Adam optimizer over 20 epochs.
- Training involves iterating through the dataset, computing loss, and updating model weights.

Inference and Visualization:
- After training, the model predicts bounding boxes for a randomly selected image.
- Bounding boxes are drawn on the image, displaying the detected vehicles.



Simple CNN

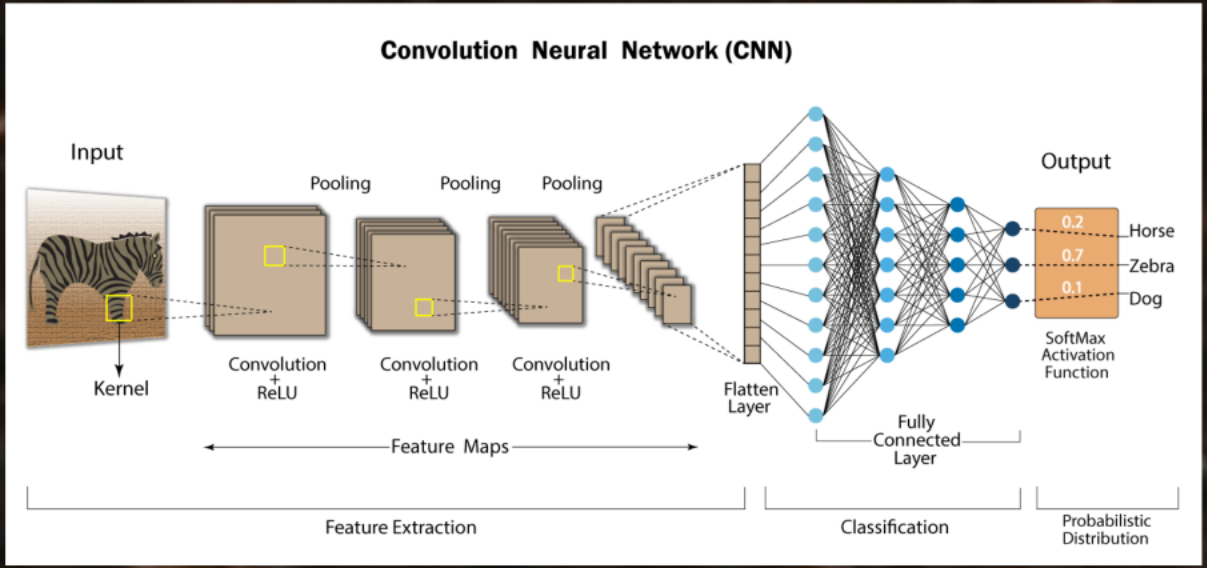
Data Loading and Preprocessing:
Images are loaded from a folder and resized to 320x320 pixels.
Labels are loaded from text files, normalized, and saved as a CSV file.

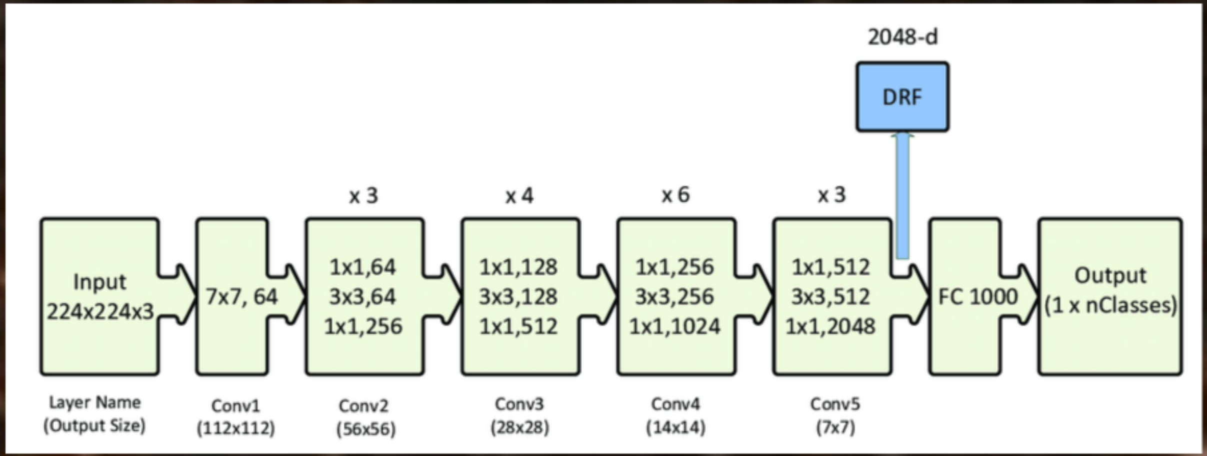
Custom Dataset Class:
A PyTorch Dataset class handles loading images and labels, applying necessary transformations.

Model Definition:
A SimpleCNN model with convolutional layers and fully connected layers predicts bounding boxes.

Training:
The model is trained using MSE loss and Adam optimizer over 20 epochs.
Training involves padding targets to match the output shape and updating model weights.

Inference and Visualization:
A preprocessed image is selected randomly, and the trained model predicts bounding boxes.
Bounding boxes are drawn on the image using PIL, and the image is displayed with detected vehicles.





dataset.

ion, recall, and F1-

ResNet50

Model Setup:
ResNet-50 model pre-trained on ImageNet is loaded.
Final fully connected layer is replaced to match the number of classes in the dataset.

Training:
Model is trained using Adam optimizer and cross-entropy loss.
Training and validation losses are calculated and printed for each epoch.

Evaluation:
Model performance is evaluated on the validation set using accuracy, precision, recall, and F1-score.
Trained model is saved for future use.

Inference:
A function is defined to predict the class of a given image.
Bounding boxes are drawn on the predicted images for visualization.
The model is used to make predictions on randomly selected images from the preprocessed dataset.

Comparison

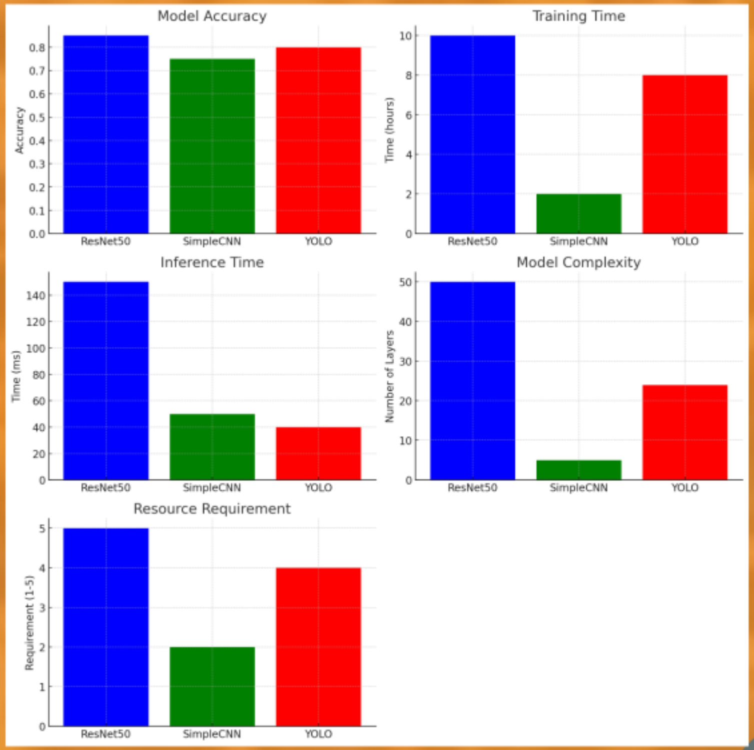
Model Architecture and Complexity
ResNet50: Deep, 50 layers, suitable for high-accuracy classification.
SimpleCNN: Less complex, faster training, ideal for smaller tasks.
YOLO: Designed for real-time object detection, provides bounding boxes and class probabilities.

Performance Metrics
Classification: Accuracy, precision, recall, F1 score.
Detection: mAP, IoU, detection speed.

Training and Inference Time
ResNet50: Longer due to depth.
SimpleCNN: Faster, lower computational cost.
YOLO: Fast, optimized for real-time detection.

Resource Requirements
ResNet50: High, needs significant computational power.
SimpleCNN: Lower, can use standard hardware.
YOLO: Requires GPU for efficiency.

Recommendation
Object Detection: YOLO is best for real-time applications.
High-Accuracy Classification: ResNet50 is ideal.
Simpler Tasks: SimpleCNN is efficient and effective.



Conclusion
For vehicle detection in AeroWatch, **YOLO** provides the best balance of accuracy and speed.

Conclusions

Object Detection:
YOLO: Optimized for real-time detection

- Processes images quickly
- Provides bounding boxes and class probabilities
- Best for applications needing fast and accurate object detection

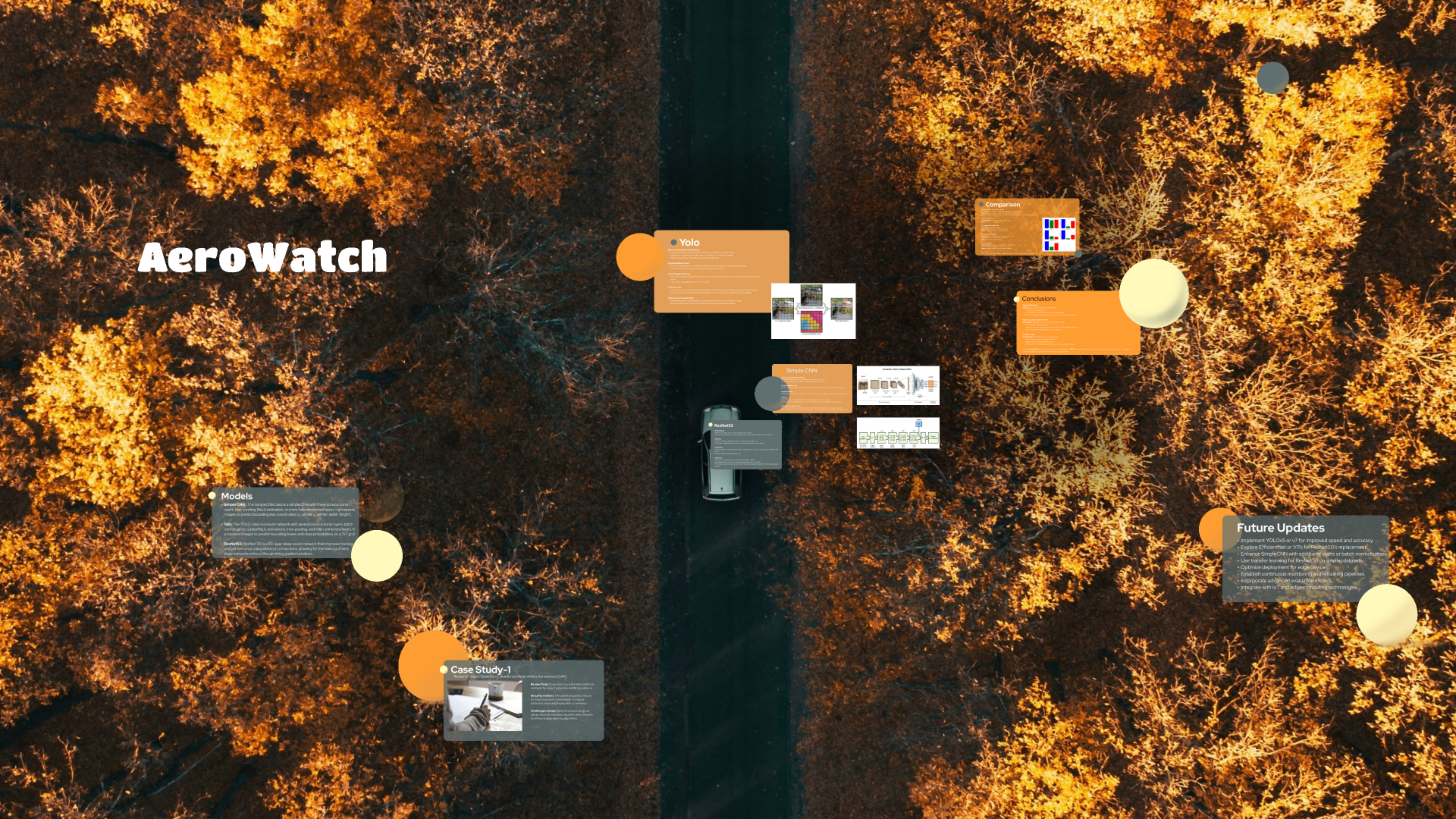
High-Accuracy Classification:
ResNet50: Suitable for complex classification tasks

- Deep network with 50 layers
- Uses residual connections for improved training of deep networks
- Ideal for tasks where high accuracy is critical

Simpler Tasks:
SimpleCNN: Efficient for less complex tasks

- Fewer layers, making it faster to train
- Requires less computational power
- Effective for smaller datasets and simpler classification tasks

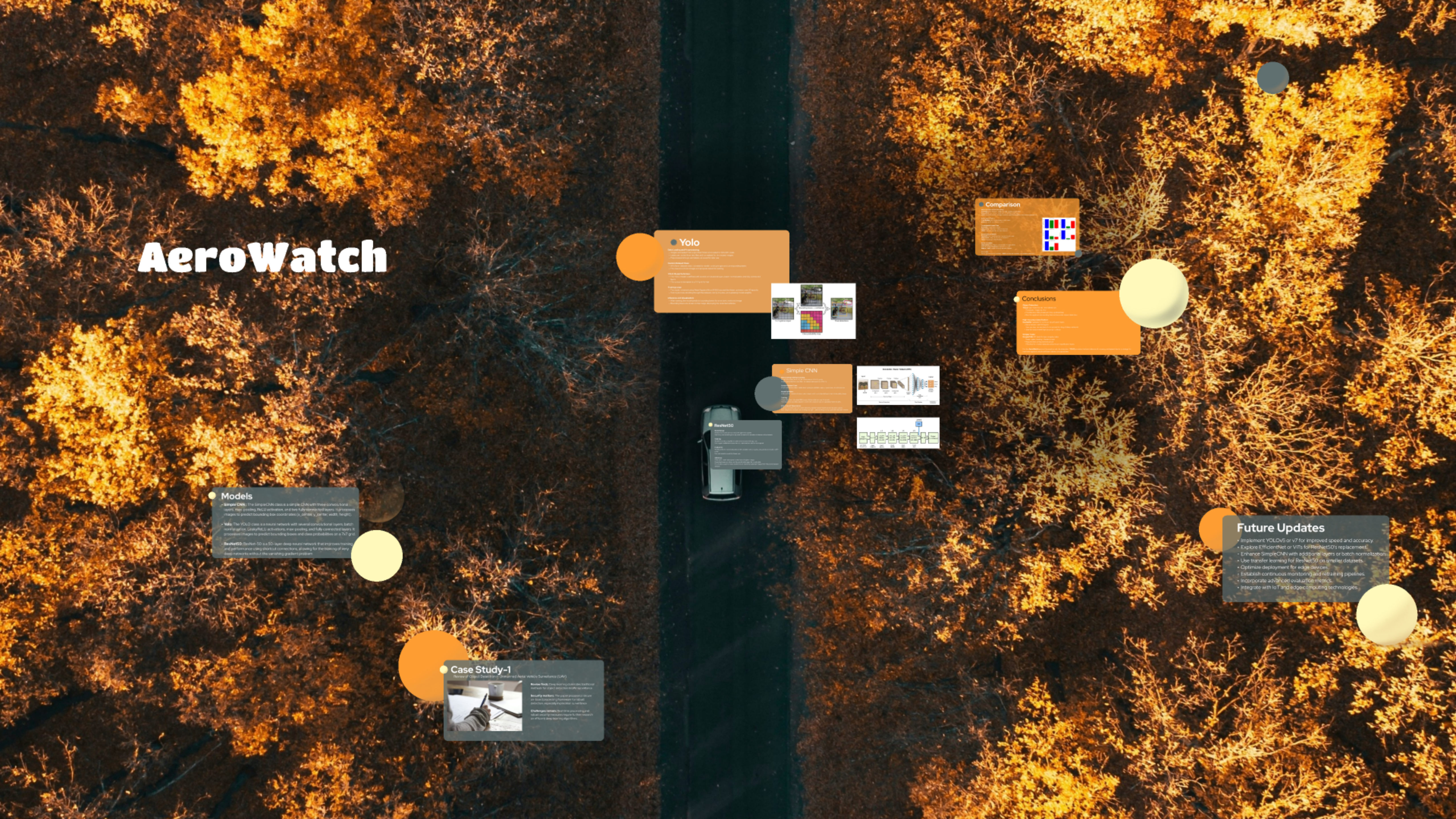
For the **AeroWatch** project focused on vehicle detection, **YOLO** provides the best balance of accuracy and speed due to its design for real-time applications and efficient detection capabilities.



AeroWatch
Yolo
Comparison
Conclusions
Simple CNN
ResNet50
Models
- Simple CNN: The SimpleCNN class is a simple CNN with three convolutional layers, max-pooling, ReLU activation, and two fully connected layers. It processes images to predict bounding box coordinates (x_center, y_center, width, height).
- Yolo: The YOLO class is a neural network with several convolutional layers, batch normalization, LeakyReLU activations, max pooling, and fully connected layers. It processes images to predict bounding boxes and class probabilities on a 7x7 grid.
- ResNet50: ResNet-50 is a 50-layer deep neural network that improves training and performance using shortcut connections, allowing for the training of very deep networks without the vanishing gradient problem.
Case Study-1
Future Updates
• Implement YOLOv5 or v7 for improved speed and accuracy.
• Explore EfficientNet or ViTs for ResNet50's replacement.
• Enhance SimpleCNN with additional layers or batch normalization.
• Use transfer learning for ResNet50 on smaller datasets.
• Optimize deployment for edge devices.
• Establish continuous monitoring and retraining pipelines.
• Incorporate advanced evaluation metrics.
• Integrate with IoT and edge computing technologies.

Future Updates

- Implement YOLOv5 or v7 for improved speed and accuracy.
- Explore EfficientNet or ViTs for ResNet50's replacement.
- Enhance SimpleCNN with additional layers or batch normalization.
- Use transfer learning for ResNet50 on smaller datasets.
- Optimize deployment for edge devices.
- Establish continuous monitoring and retraining pipelines.
- Incorporate advanced evaluation metrics.
- Integrate with IoT and edge computing technologies.



AeroWatch
Yolo
Comparison
Conclusions
Simple CNN
ResNet50
Models
- Simple CNN: The SimpleCNN class is a simple CNN with three convolutional layers, max-pooling, ReLU activation, and two fully connected layers. It processes images to predict bounding box coordinates (x_center, y_center, width, height).
- Yolo: The YOLO class is a neural network with several convolutional layers, batch normalization, LeakyReLU activations, max pooling, and fully connected layers. It processes images to predict bounding boxes and class probabilities on a 7x7 grid.
- ResNet50: ResNet-50 is a 50-layer deep neural network that improves training and performance using shortcut connections, allowing for the training of very deep networks without the vanishing gradient problem.
Case Study-1
Future Updates
• Implement YOLOv5 or v7 for improved speed and accuracy.
• Explore EfficientNet or ViTs for ResNet50's replacement.
• Enhance SimpleCNN with additional layers or batch normalization.
• Use transfer learning for ResNet50 on smaller datasets.
• Optimize deployment for edge devices.
• Establish continuous monitoring and retraining pipelines.
• Incorporate advanced evaluation metrics.
• Integrate with IoT and edge computing technologies.