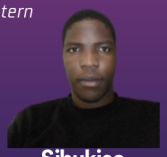
Improving Radio Frequency Modeling for Telecom Networks

Telnet Solutions

THE TEAM

Zamancwabe Makhathini
Data science intern



Sibukiso Nhlengethwa
Data science intern

Caroll Tshabane
Data science intern

Muwanwa Tshikovhi
Data science intern

Jan Motene
Data science intern

Nolwazi Mndebele
Data science intern

Nkadimeng Nkhubalale
Data science intern

Problem Statement

- High cost of clutter data for RFP in Wireless Communication
- Network Performance
- Clutter Impact
- Frequent Data Updates

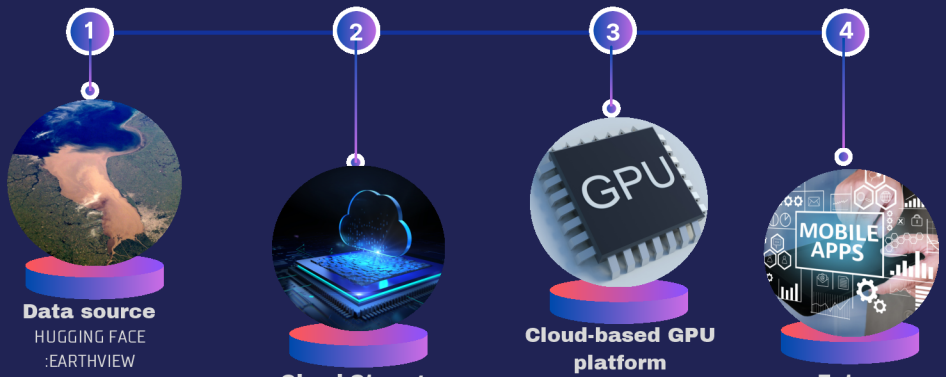
Aim

To develop a cost effective, scalable and accurate method for generating clutter data using satellite imagery and machine learning.

Objectives

- Create a segmentation model
- Design and implement a user-friendly interface or API for inputting areas of interest and retrieving clutter data

Data Pipeline Components



Data source
HUGGING FACE
:EARTHVIEW

Cloud Storage
AWS [Amazon S3 bucket and EC2 instance]

Cloud-based GPU platform
Google Collab[GPU]

Future Deployment
EC2 instance

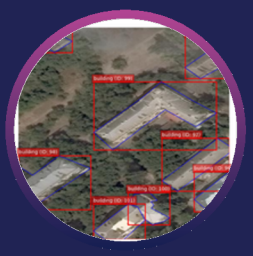
Data Quality Enhancement

Check Invalid Images

Automated checks to remove corrupted pictures

Remove Duplicates

Identify duplicate images



Label Checking

Checking for annotation consistency

Standardize Image Sizes

Resizing images

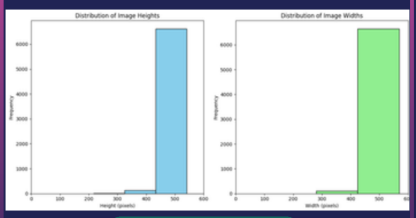
Data Quality Enhancement continues..

Total images

6,764

Total Annotations

89,113



Distribution

Min Height: 217, Max Height: 3, 456

Min Width: 280, Max Width: 4, 608

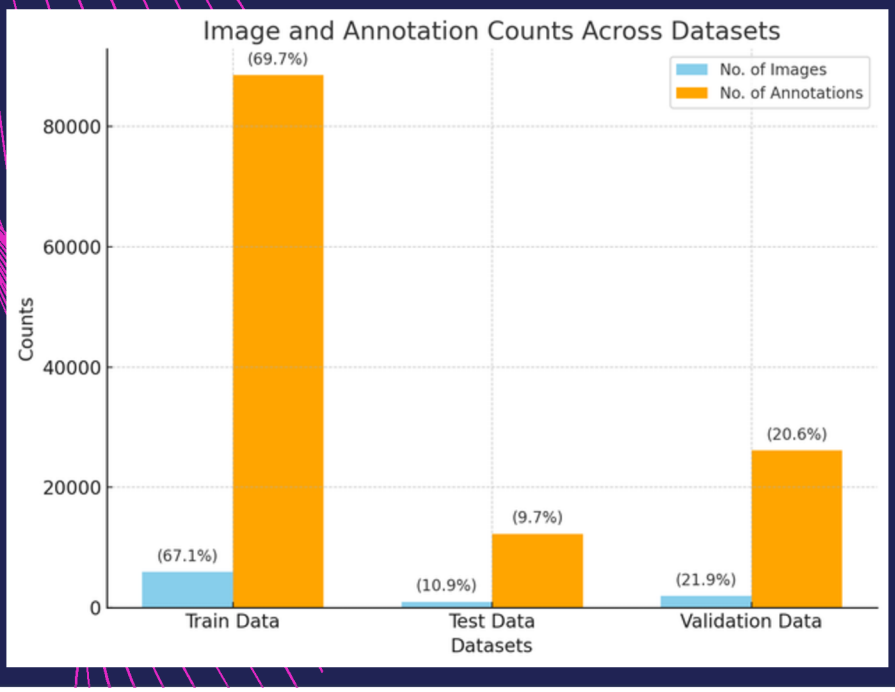
Images with annotations

5,932 out of 6,764

Exploratory Data ANALYSIS

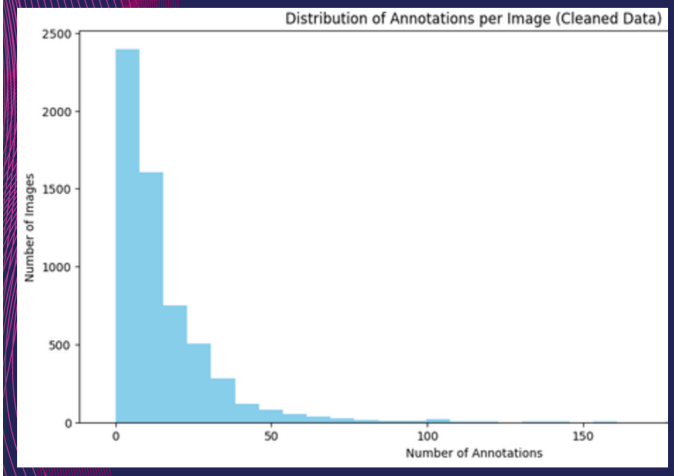
- Data understanding
- Feature insights
- Data validation

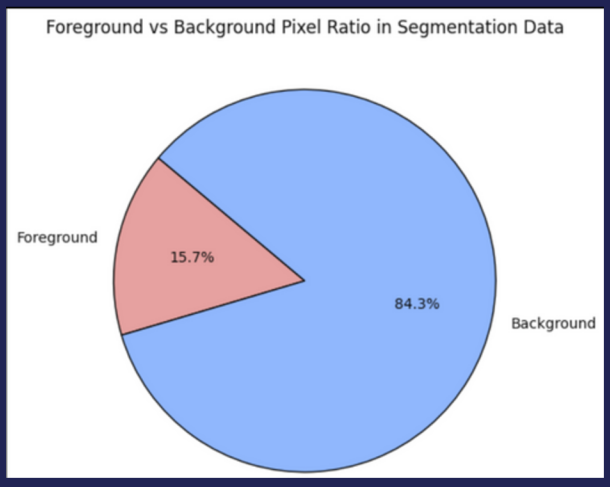
Image & Annotation Analysis



- Dataset distrribution
- Testing and validation balance
- Annotations importance

Distributions

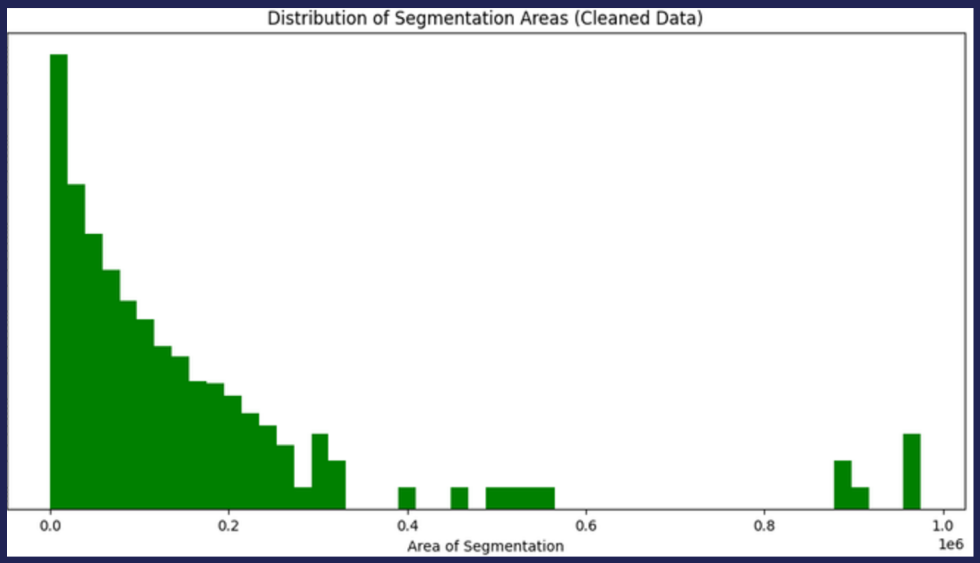




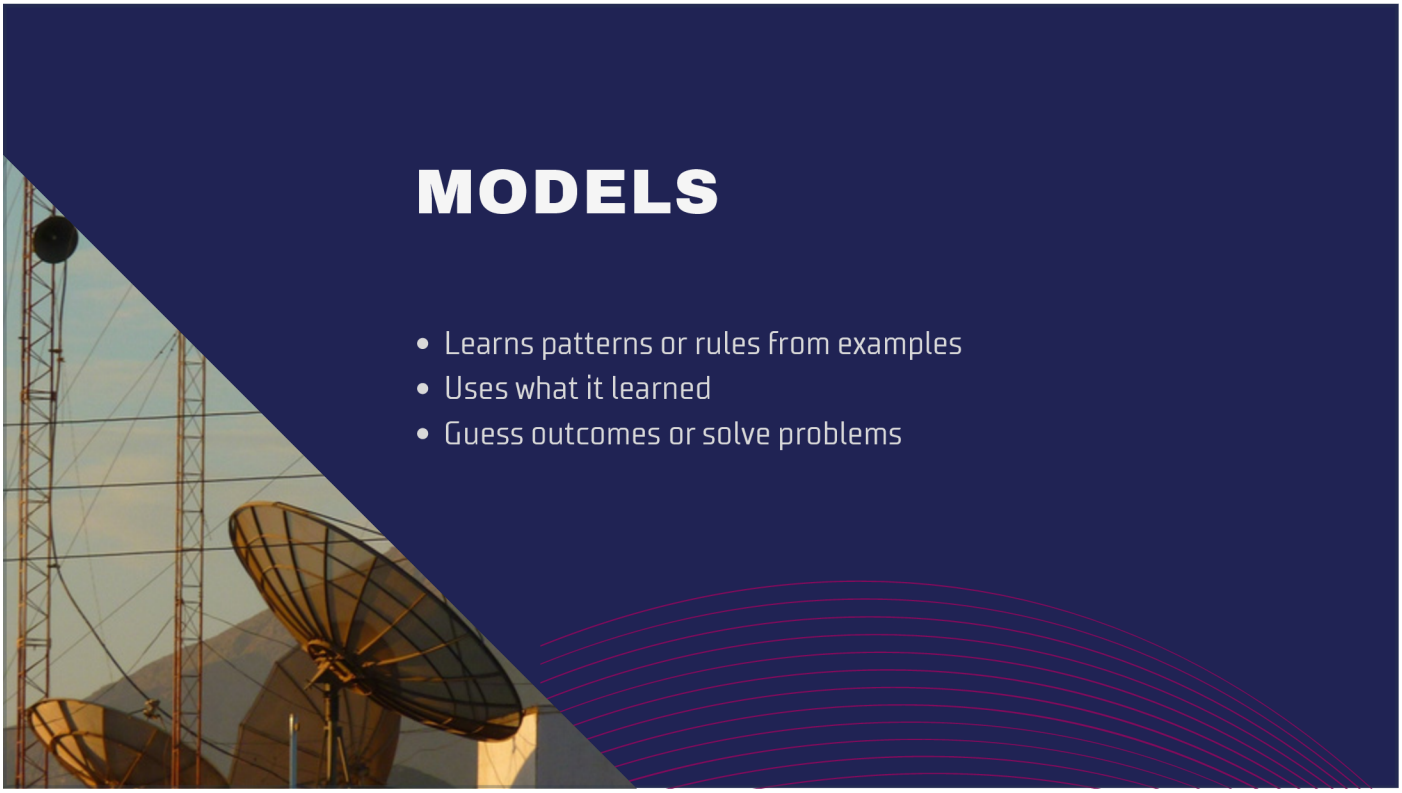
- Skewness to the Left
- Image Variety
- Impact on Modeling

- Foreground- buildings
- Background- trees, roads, vegetation, etc

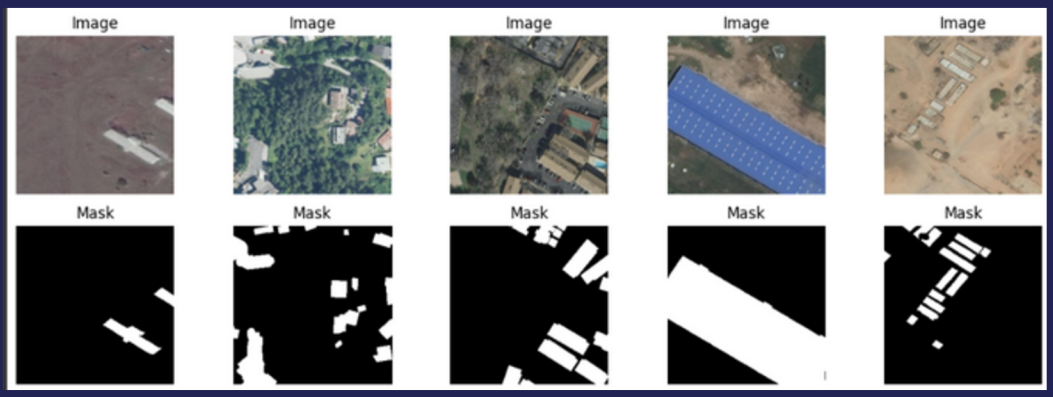
Distributions continue



- Heavily skewed towards smaller bounding box areas
- Many of the annotations are for small objects in the dataset.
- Few outliers of large bounding boxes


MODELS
• Learns patterns or rules from examples
• Uses what it learned
• Guess outcomes or solve problems

Preprocessing

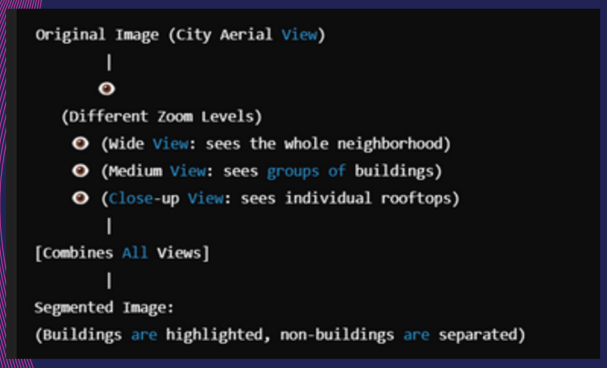


- BUILDINGS IN **WHITE** - 1
- NON-BUILDING AREAS AS **BLACK** - 0

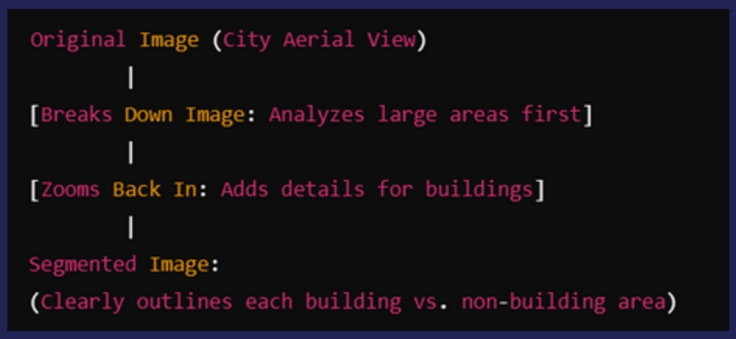
Preprocessing

1. **Batch Creation (Batch #1, #2, ...)**
   ↓
2. **Data Generator**
   ↓
3. **Preprocessing**
   - Resize, Normalize
     ↓
4. **Load to Model**
   ↓
5. **Forward & Backward Pass**
   - Loss Calculation & Optimization

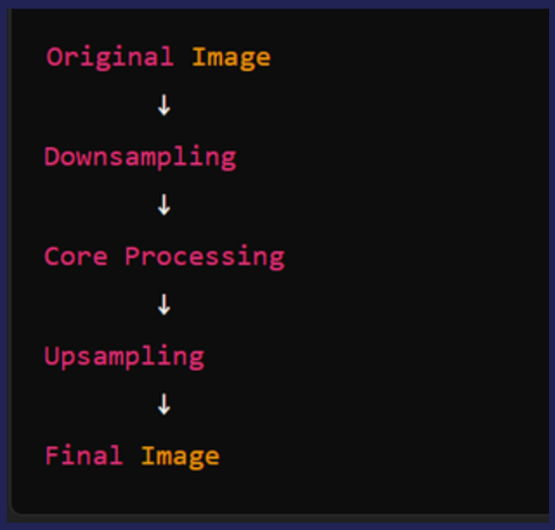
DEEPlabV3 model



U-net model



ResUNet model



- Great for imbalanced data
- Remembers the small, less common areas
- Gives them the attention they need

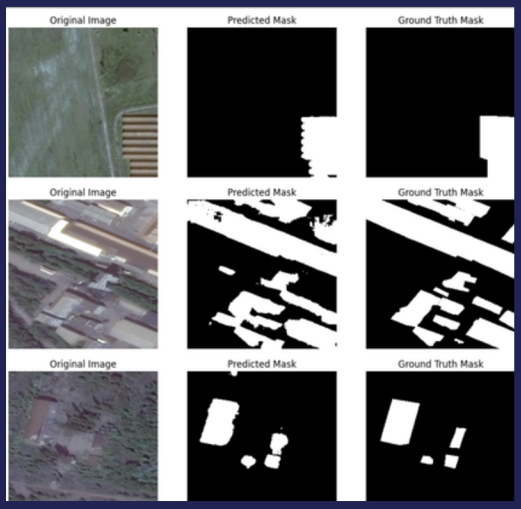
Important Terms

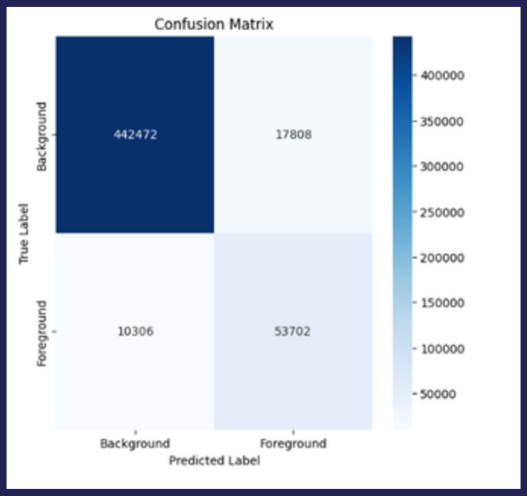
- **Accuracy - % of correct predictions**
- **Mean IOU - actual object/predicted object**
- **Dice Loss - difference between actual object and predicted object**
- **Ground Truth - original/actual mask**

Model perfomance comparison

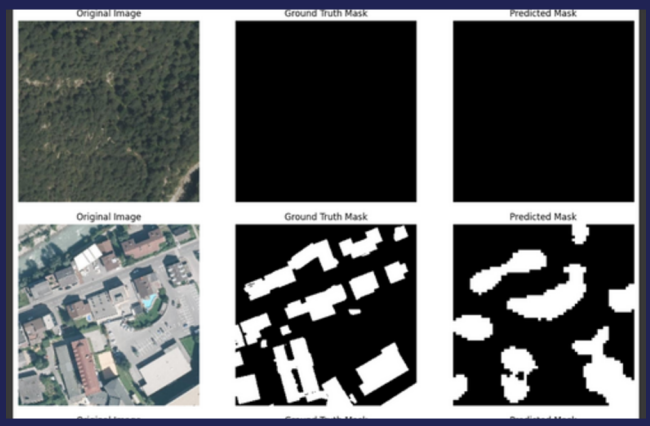
Aspect	ResUNet	DeepLabV3	U-Net
Loss	✅	❌	❌
Accuracy	✅	❌	✅
Mean IoU	✅	❌	❌
Predicted Mask Noise	✅	❌	❌
Building Boundaries	✅	❌	❌
Background (Correct)	✅	✅	❌
Background (Incorrect)	✅	❌	❌
Foreground (Correct)	✅	✅	❌
Foreground (Incorrect)	✅	❌	❌

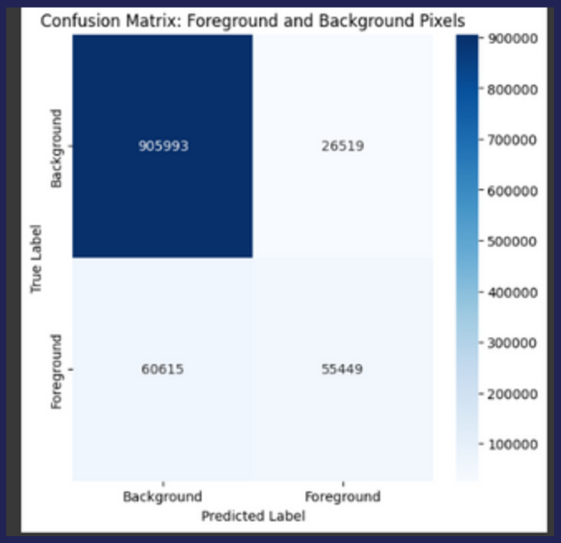
ResUNet model





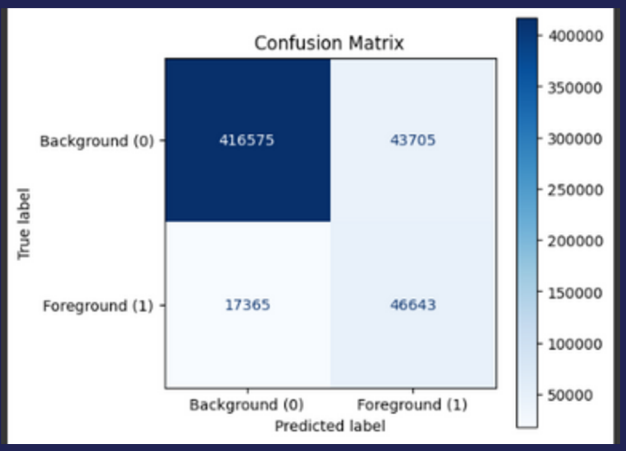
DeepLabv3


Original Image
Ground Truth Mask
Predicted Mask
Original Image
Ground Truth Mask
Predicted Mask


Confusion Matrix: Foreground and Background Pixels
True Label
Background
Foreground
905993
26519
60615
55449
Background
Foreground
Predicted Label
900000
800000
700000
600000
500000
400000
300000
200000
100000

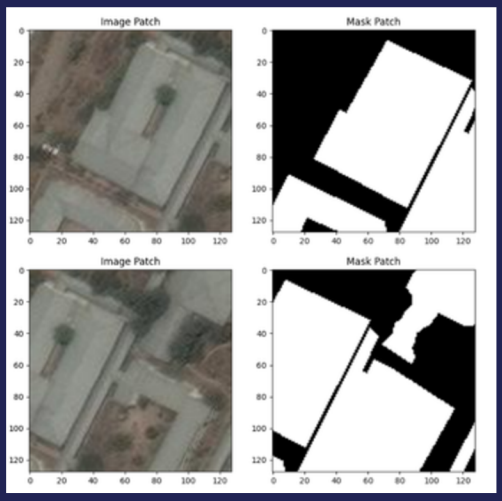
U-net Model


Input Image
True Mask
Predicted Mask


Confusion Matrix
Background (0)
Foreground (1)
True label
416575
43705
17365
46643
Background (0)
Foreground (1)
Predicted label
400000
350000
300000
250000
200000
150000
100000
50000

What's next?

1. Gradualy Adjust learning rate



- Screenshots in a large image
- Area of interest (buildings)
- Model becomes more familiar with building pixels
- Better predictions of buildings, with their different shapes and sizes

APP DEMO

Streamlit

Limitations

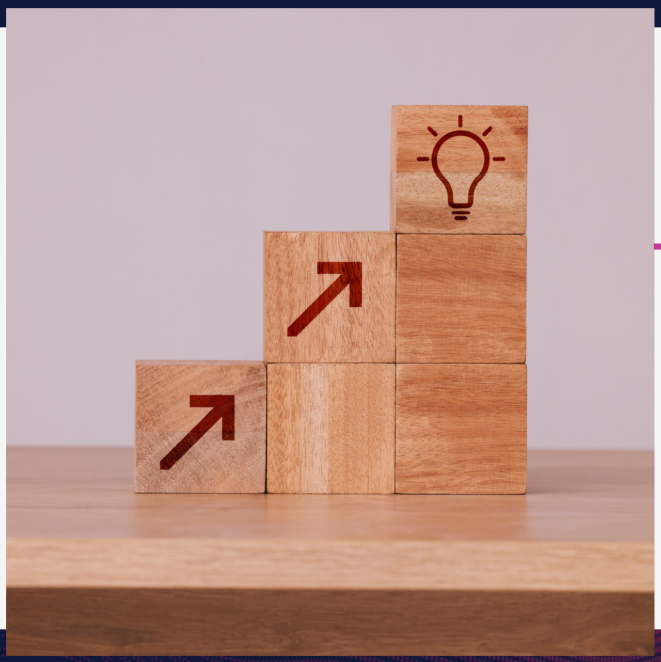
Data Quality

Class imbalance

Time & Computational constraints

Model limitation

- Incomplete or low-resolution data
- Overrepresentation of certain classes
- Models struggled with high-density areas



Recommendations & Conclusion

- Enhance Data Acquisition
- Ongoing optimization and updates needed
- Model Improvement
- Greater potential to improve telecom planning & performance


THANK
YOU