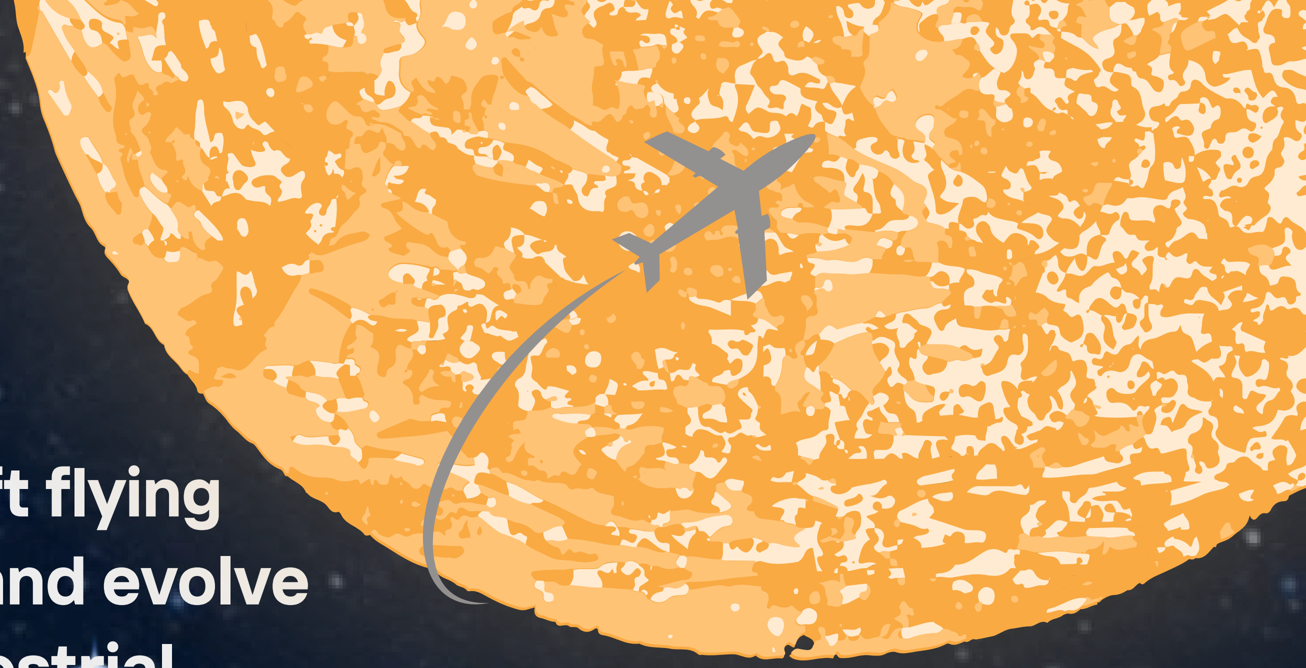
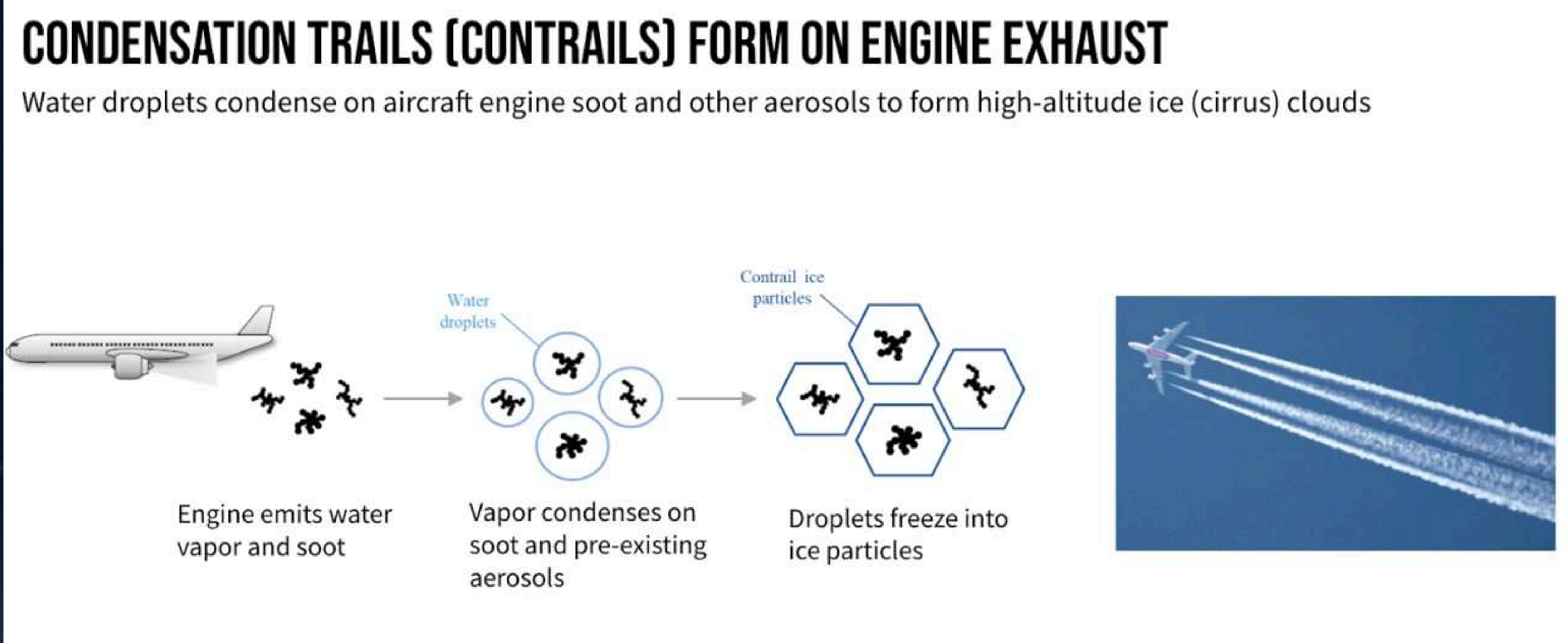
IDENTIFYING AND MITIGATING AVIATION-INDUCED CONTRAILS USING MACHINE LEARNING AND BIG DATA ANALYTICS

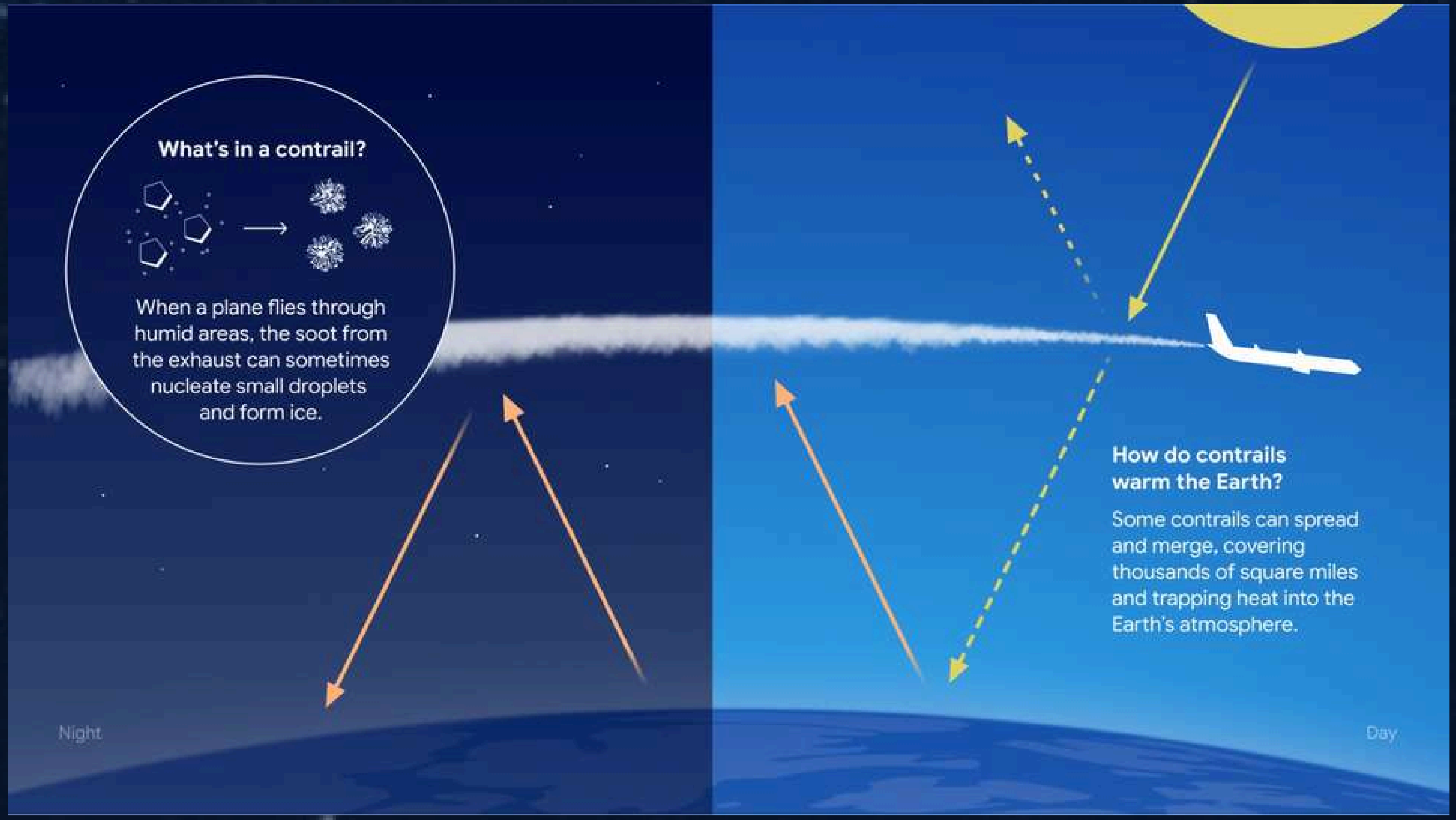
1. Introduction



Project Background:

- Contrails, or condensation trails, are line-shaped clouds formed by aircraft flying through humid regions of the atmosphere. These formations can persist and evolve into cirrus clouds, contributing to the greenhouse effect by trapping terrestrial radiation. Recent studies suggest that contrails account for a substantial portion of aviation's climate impact, with some estimates attributing up to 35% of aviation-induced global warming to contrail formations.

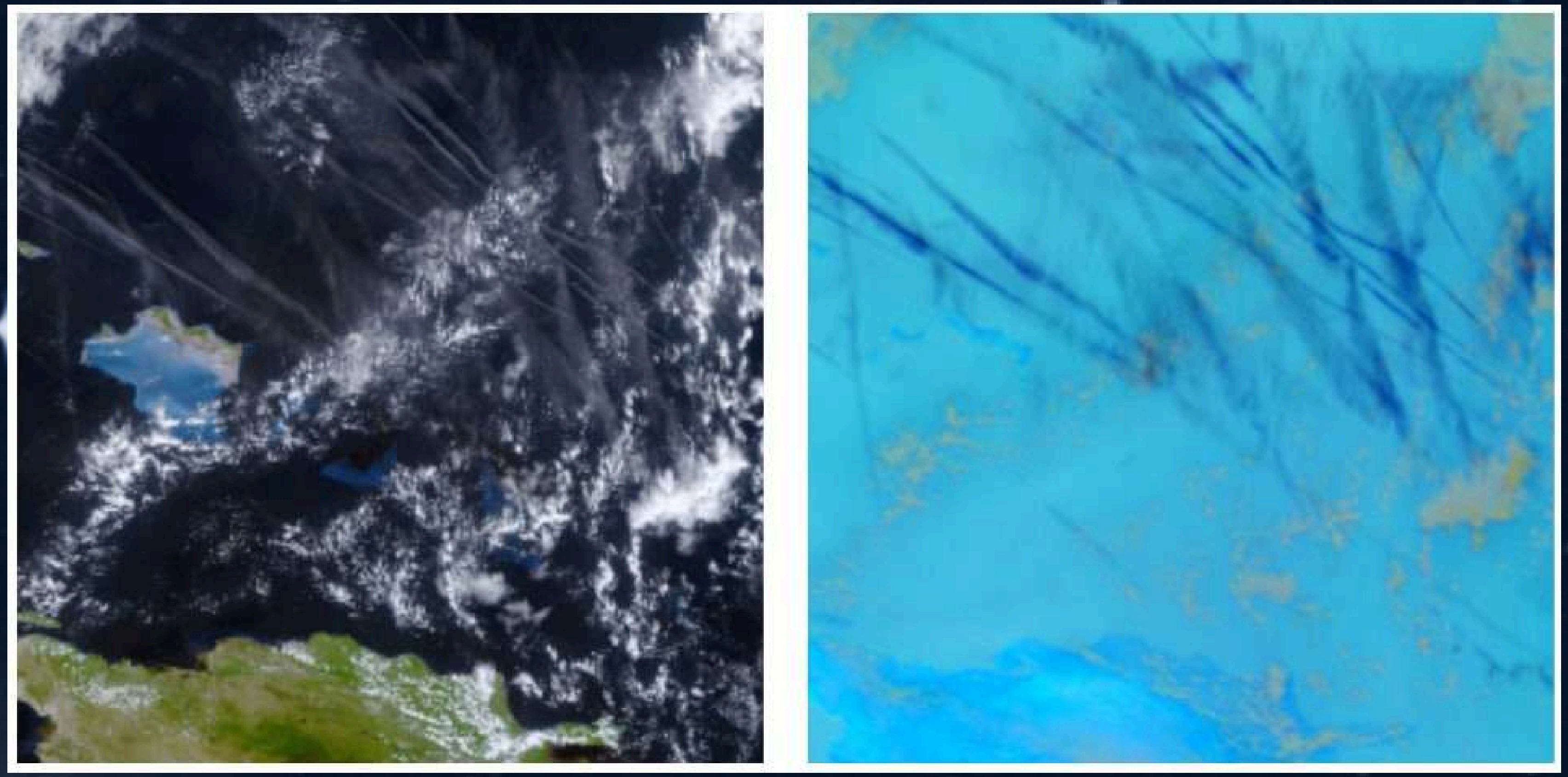




1. Introduction

Project Background:

- To address this environmental challenge, the Google Research initiative has launched a competition focused on developing machine learning models capable of detecting contrails in satellite imagery. The objective is to enable airlines and aviation authorities to adjust flight operations, thereby minimizing contrail formation and its associated climate impact. This paper presents our approach to the competition, detailing the data analysis, model development, and potential implications for sustainable aviation practices



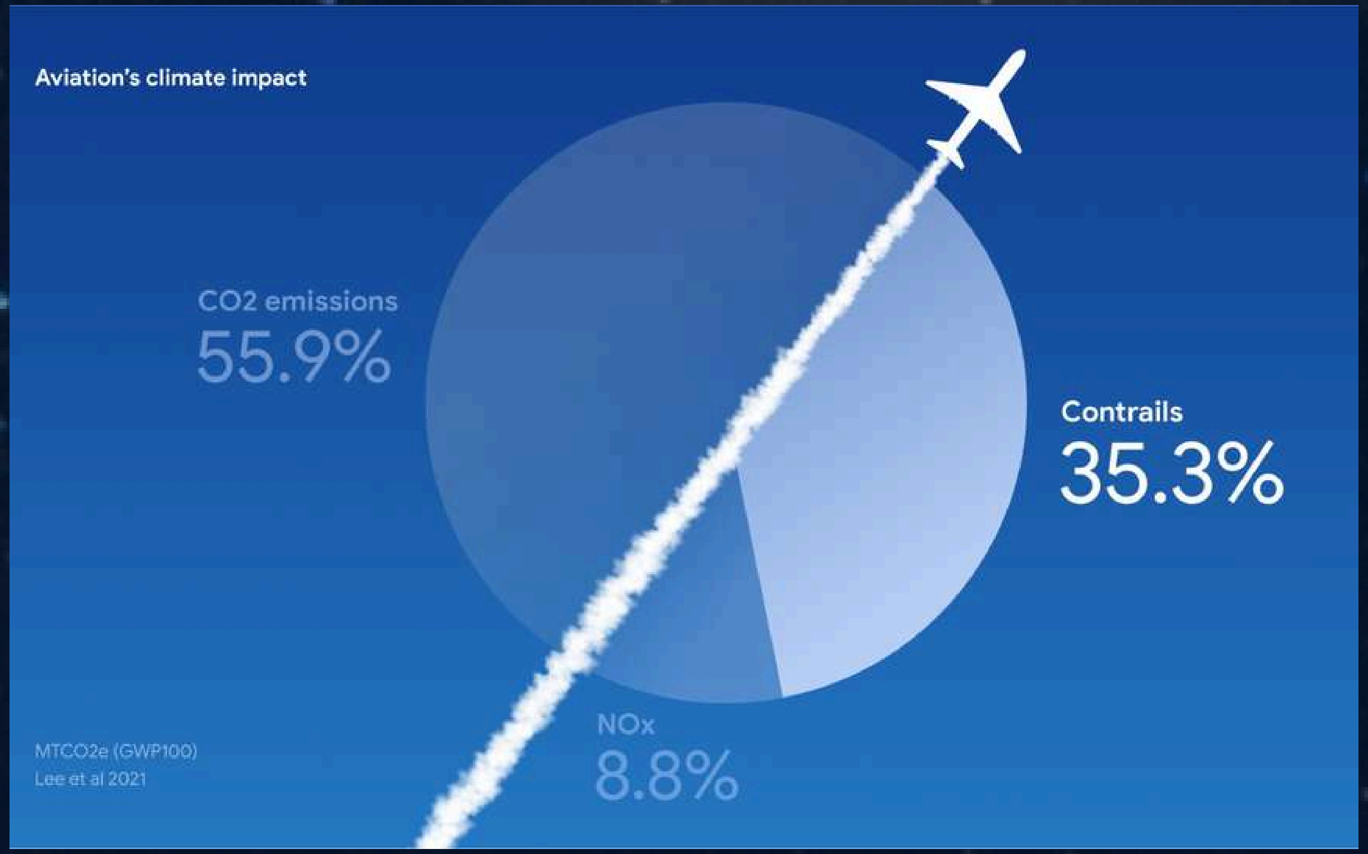
1. Introduction

Problem Statement:

Contrails, short for condensation trails, are artificial clouds formed by the water vapor emitted from aircraft engines. Under certain atmospheric conditions, these trails can persist and contribute to radiative forcing, exacerbating global warming. The challenge is to identify contrails accurately and develop strategies to mitigate their climate impact.

Proposed Solution:

The proposed solution involves leveraging machine learning and satellite data to detect and classify contrails effectively. By using high-resolution imagery and environmental data, the goal is to identify the specific flights or flight patterns contributing significantly to persistent contrails. This information can guide airline operational changes, such as route optimization, to minimize contrail formation.



1. Introduction

Role of Big Data:

Big data plays a critical role in solving the contrail problem. Large-scale satellite imagery, flight path data, meteorological information, and other datasets are essential for training and deploying machine learning models. These models can analyze vast amounts of data to identify contrails with high accuracy, making it feasible to scale the solution globally.

2. Literature Survey

Title	Authors	Year	Summary
Deep Semantic Contrails Segmentation of GOES-16 Satellite Images: A Hyperparameter Exploration	Gabriel Jarry, Philippe Very, Amine Heffar, Valentin Tordjman-Levavasseur	2024	Explores hyperparameter optimization for deep semantic segmentation models aimed at contrail detection in GOES-16 satellite imagery.
Contrail Altitude Estimation Using GOES-16 ABI Data and Deep Learning	Vincent R. Meijer, Sebastian D. Eastham, Ian A. Waitz, and Steven R. H. Barrett	2024	Develops a deep learning algorithm to estimate contrail altitudes based on GOES-16 ABI infrared imagery, crucial for understanding contrail formation and persistence in climate models.
Exploring Models and Band Selection for Improved Contrail Detection in Geostationary Satellite Imagery	Irfan Darmawan, Mochamad Al-Husaini	2024	Utilizes deep learning models and band selection techniques to enhance contrail detection in geostationary satellite imagery, focusing on Landsat-8 data.
Optimizing Contrail Detection: A Deep Learning Approach with EfficientNet-b4 Encoding	Qunwei Lin, Qian Leng, Zhicheng Ding, Chao Yan, Xiaonan Xu	2024	Presents a deep-learning approach utilizing EfficientNet-b4 for feature extraction, along with misalignment correction and soft labeling techniques to improve contrail detection.
The Application of a Convolutional Neural Network for the Detection of Contrails in Satellite Imagery	Jay P. Hoffman, Timothy F. Rahmes, Anthony J. Wimmers, Wayne F. Feltz	2023	Presents a CNN-based approach for contrail detection in satellite imagery, emphasizing the significance of monitoring contrails for climate change.
Combining UPerNet and ConvNeXt for Contrail Identification to Reduce Global Warming	Zhenkuan Wang	2023	Focuses on aircraft contrail detection in satellite images, developing a novel preprocessing technique for NOAA GOES-16 satellite images and integrating segmentation models for enhanced contrail identification.
Performance Evaluation of Deep Segmentation Models for Contrail Detection	Akshat Bhandari, Sriya Rallabandi, Sanchit Singhal, Aditya Kasliwal, Pratinav Seth	2022	Benchmarks several segmentation models to detect contrails in low-orbit satellite imagery, addressing the challenge of limited labeled data and providing an open benchmark for contrail segmentation.

Research Gap

Despite advancements in satellite imaging and machine learning, there is a lack of comprehensive datasets linking contrail properties to specific atmospheric and flight conditions. Additionally, there is limited research on real-time operational strategies for contrail avoidance without significantly increasing fuel consumption or flight time. Addressing these gaps is crucial for developing a practical and scalable solution.

ISRO'S Existing Proposal

- ISRO's Focus: The National Remote Sensing Centre's proposal aims to study the impact of atmospheric constituents on the Indian Summer Monsoon, focusing on emissions, air quality, aerosols, clouds, and their interactions with the monsoon.
- Key Areas: The proposal emphasizes the importance of understanding how pollutants affect monsoon behavior and the broader implications for climate change and air quality.

Existing Proposal Identification Info

- Proposal No: RES-NRSC-2024-002
- Title: Impact of the Indian Summer Monsoon by Atmospheric Constituents
- Area of Research: Remote Sensing & GIS, Earth Observations
- Co-PI (Focal Point) from ISRO Centre/Unit: Ms. K.L. Kanchana
- ISRO/DOS Centre/Unit: National Remote Sensing Centre, Hyderabad

Relevance to ISRO's Objectives:

- Environmental Monitoring: Your project aligns with ISRO's focus on using satellite data for environmental monitoring and management.
- Sustainable Aviation: Contributes to ISRO's goal of promoting sustainable practices in various sectors, including aviation.
- Data Integration: Utilizes data from ISRO satellites (e.g., MODIS, Sentinel) and integrates satellite and ground-based observations, supporting ISRO's mission to enhance data utilization for environmental stewardship.
- Climate Research: Supports ISRO's broader objectives in climate research and understanding atmospheric phenomena, which is crucial for future climate modeling and prediction.

3. Primary Objectives

1. To identify and classify contrails using satellite imagery and machine learning.
2. To establish correlations between atmospheric conditions, flight patterns, and contrail persistence.
3. To develop actionable insights for airlines to minimize contrail-induced radiative forcing.
4. To evaluate the impact of contrail mitigation strategies on overall carbon emissions and climate change.
5. To create a framework for real-time contrail monitoring and mitigation on a global scale.

4. Methodology

Data Collection:

- **Sources:** The dataset used is the Carvana Image Masking Challenge dataset from Kaggle, which includes high-resolution images of cars along with their corresponding segmentation masks.
- **Pre-Processing:**
  - Resizing: Images are resized to a consistent size to ensure uniform input dimensions for the model.
  - Normalization: Pixel values are normalized to a range of [0, 1] to facilitate better convergence during training.
  - Data Augmentation: Techniques such as random rotations, flips, and brightness adjustments are applied to increase the diversity of the training data and improve the model's generalization.
-

4. Methodology

Machine Learning Model:

- U-Net architecture, a convolutional neural network designed for semantic segmentation tasks. U-Net is known for its encoder-decoder structure, which captures context and enables precise localization. The implementation is based on the PyTorch framework, leveraging its dynamic computation graph and GPU acceleration capabilities.

- Model Selection: The U-Net model is chosen due to its effectiveness in image segmentation tasks, particularly in medical imaging and other domains requiring precise localization. Its symmetric architecture with skip connections allows for efficient feature extraction and reconstruction, making it suitable for the Carvana Image Masking Challenge dataset.

4. Methodology

Training and Testing:

- Loss Function: The binary cross-entropy loss is used, which is appropriate for binary segmentation tasks.
- Optimizer: The Adam optimizer is employed for its adaptive learning rate properties, aiding in faster convergence.
- Batch Size: A batch size of 16 is used, balancing memory constraints and training stability.
- Epochs: The model is trained for 50 epochs, with early stopping implemented to prevent overfitting.
- Validation: A validation set is used to monitor the model's performance during training and adjust hyperparameters accordingly.

4. Methodology

Evaluation Metrics:

- The primary evaluation metric used in the notebook is the Dice coefficient, a measure of overlap between the predicted segmentation mask and the ground truth mask. A higher Dice coefficient indicates better performance of the segmentation model.

5. Results

```
Train : 100% ████████████████████ 642/642 [04:54<00:00, 2.34it/s]

Valid: 100% ████████████████████ 58/58 [00:16<00:00, 5.18it/s]

  Train Loss: 0.0108 | Val Loss: 0.0124 | Val Dice: 0.6181
  Learning rate: 7.491063393793129e-06
  EPOCH: 29

Train : 100% ████████████████████ 642/642 [04:54<00:00, 2.34it/s]

Valid: 100% ████████████████████ 58/58 [00:14<00:00, 6.56it/s]

  Train Loss: 0.0108 | Val Loss: 0.0123 | Val Dice: 0.6196
  Learning rate: 1.8015407097782265e-08
  Total Training Time: 9333.491 seconds
```

5. Results

In [28]:

```
# Finding the Best Threshold
bdice = -1
bi = None
for i in tqdm(np.arange(0, 1.01, 0.01)):
    val_dice = dice_coef(ground_truths, predictions, i)
    if val_dice > bdice:
        bdice = val_dice
        bi = i
```

```
100% 101/101 [01:27<00:00, 1.16it/s]
```

In [29]:

```
print(f'Best Threshold: {bi}')
print(f'Best Validation Dice Score: {bdice}')
```

```
Best Threshold: 0.01
Best Validation Dice Score: 0.6216237299559233
```

6. references

1. Jarry, G., Very, P., Heffar, A., & Tordjman-Levavasseur, V. (2024). Deep semantic contrails segmentation of GOES-16 satellite images: A hyperparameter exploration. Retrieved from https://www.sesarju.eu/sites/default/files/documents/sid/2024/papers/SIDs_2024_paper_028%20final.pdf
2. Author(s) Not Specified. (2024). Contrail altitude estimation using GOES-16 ABI data and deep learning. Retrieved from https://amt.copernicus.org/articles/17/6145/2024/
3. Darmawan, I., & Al-Husaini, M. (2024). Exploring models and band selection for improved contrail detection in geostationary satellite imagery. Journal of Advances in Information Technology, 15(6), 682-690. Retrieved from https://www.jait.us/articles/2024/JAIT-V15N6-682.pdf
4. Lin, Q., Leng, Q., Ding, Z., Yan, C., & Xu, X. (2024). Optimizing contrail detection: A deep learning approach with EfficientNet-b4 encoding. arXiv preprint arXiv:2404.14441. Retrieved from https://arxiv.org/abs/2404.14441
5. Hoffman, J. P., Rahmes, T. F., Wimmers, A. J., & Feltz, W. F. (2023). The application of a convolutional neural network for the detection of contrails in satellite imagery. Remote Sensing, 15(11), 2854. Retrieved from https://www.mdpi.com/2072-4292/15/11/2854
6. Wang, Z. (2023). Combining UPerNet and ConvNeXt for contrail identification to reduce global warming. arXiv preprint arXiv:2310.04808. Retrieved from https://arxiv.org/abs/2310.04808
7. Bhandari, A., Rallabandi, S., Singhal, S., Kasliwal, A., & Seth, P. (2022). Performance evaluation of deep segmentation models for contrail detection. arXiv preprint arXiv:2211.14851. Retrieved from https://arxiv.org/abs/2211.14851