Retraction Notice

Title of retracted article:	Real-Time Coal Emissions Monitoring System Using Geospatial Satellite Images and Deep Neural Networks
Author(s):	Aaryan Doshi
* Corresponding author:	Aaryan Doshi
Journal:	American Journal of Climate Change (AJCC)
Year:	2024
Volume:	13
Number:	4
Pages (from - to):	720-731
DOI (to PDF):	https://doi.org/10.4236/ajcc.2024.134033
Paper ID at SCIRP:	2361488
Article page:	https://www.scirp.org/journal/paperinformation?paperid=138003
Retraction date:	2024-12-20

Retraction initiative (multiple responses allowed; mark with **X**):
X All authors
☐ Some of the authors:
☐ Editor with hints from
☐ Journal owner (publisher)
☐ Institution:
☐ Reader:
☐ Other:
Date initiative is launched: 2024-12-20

Retraction type (multiple responses allowed):
☐ Unreliable findings
☐Lab error ☐Inconsistent data ☐Analytical error ☐Biased interpretation
☐Other:
☐ Irreproducible results
☐ Failure to disclose a major competing interest likely to influence interpretations or recommendations
☐ Unethical research

☐ Fraud
☐Data fabrication ☐Fake publication ☐Other:
☐ Plagiarism ☐ Self plagiarism ☐ Overlap ☐ Redundant publication *
☐ Copyright infringement ☐ Other legal concern:

☐ Editorial reasons
☐Handling error ☐Unreliable review(s) ☐Decision error ☐Other:

X Other:

Results of publication (only one response allowed):
☐ are still valid.
X were found to be overall invalid.

Author's conduct (only one response allowed):
☐ honest error
☐ academic misconduct
X none (not applicable in this case – e.g. in case of editorial reasons)

* Also called duplicate or repetitive publication. Definition: "Publishing or attempting to publish substantially the same work more than once."

Scientific Research Publishing

History
Expression of Concern:
☐ yes, date: yyyy-mm-dd
X no

Correction:
☐ yes, date: yyyy-mm-dd
☐ no

Comment:

The paper does not meet the standards of "American Journal of Climate Change (AJCC)".

This article has been retracted to straighten the academic record. In making this decision the Editorial Board follows COPE's Retraction Guidelines. Aim is to promote the circulation of scientific research by offering an ideal research publication platform with due consideration of internationally accepted standards on publication ethics. The Editorial Board would like to extend its sincere apologies for any inconvenience this retraction may have caused.

American Journal of Climate Change, 2024, 13, 720-731
https://www.scirp.org/journal/ajcc
ISSN Online: 2167-9509
ISSN Print: 2167-9495

Real-Time Coal Emissions Monitoring System Using Geospatial Satellite Images and Deep Neural Networks

Aaryan Doshi

Monta Vista High School, Cupertino, USA
Email: aaryandoshi@gmail.com

How to cite this paper: Doshi, A, (2024). Real-Time Coal Emissions Monitoring System Using Geospatial Satellite Images and Deep Neural Networks. *American Journal of Climate Change, 13,* 720-731.
https://doi.org/10.4236/ajcc.2024.134033

Received: September 5, 2024
Accepted: December 3, 2024
Published: December 6, 2024

Abstract

Global efforts to reduce emissions are hampered by incomplete and unreliable data from coal power plants regarding their release of greenhouse gases. This inability to acquire data, particularly relevant in regions lacking robust reporting mechanisms, hinders evidence-based policy-making and proper action against prominent emissions contributors. To address this issue, we propose using a transformer-based approach to determine the coal plant's operational status – entirely through satellite images. This method leverages the readily available global coverage of satellite imagery to provide independent, real-time monitoring of coal plant operations, even in areas where traditional data collection methods are challenging. The model learns to identify visual indicators, such as cooling tower plumes in unseen test images. Our transformer-based approach outperforms the best prior work with statistical significance, reaching an 80% accuracy on a held-out set of unseen coal plant satellite images. These results have far-reaching implications for climate monitoring and regulations. By providing more accurate and independently sourced information on coal plant operations worldwide, our approach can support more effective emissions tracking, inform climate negotiations, and introduce greater transparency into the process of climate regulation.

Keywords

Machine-Learning, Deep-Learning, Climate Change, Coal Emissions

1. Introduction & Background

The global urgency to mitigate climate change is intensifying as the effects of greenhouse gas emissions become clearer (Singh, 2021). Coal power plants, which remain among the most significant contributors to global carbon dioxide emissions, are at the forefront of scrutiny (Birol & Malpass, 2024). Our ability to reliably regulate and understand the impact of coal power plants on the climate compromise hinges on receiving comprehensive and reliable data on the operational status and emissions of these plants, particularly in regions where transparency and robust reporting mechanisms are lacking (Bauckloh et al., 2023; Datt et al., 2019). Unfortunately, this requirement of procuring reliable data has not been consistently met. The resulting data deficiency not only hampers accurate global emissions tracking but also undermines the development of evidence-based policies and international agreements crucial for curbing climate change (Rosenstock et al., 2017; Climate Works Australia, 2018).

Traditionally, monitoring coal plant operations has relied heavily on self-reported data, infrequent inspections, and ground-based measurements, all of which are prone to inaccuracies and inconsistencies. These methods fail to provide the real-time, granular data necessary for effective climate action (NSW EPA, 2018; Asia-Pacific Economic Cooperation, 2008). In many regions, especially in developing countries, the absence of stringent regulatory frameworks and monitoring infrastructure exacerbates these challenges, leading to significant gaps in emissions data (Li & Kroese, 2022; NGFS, 2022). Even in regions with more robust systems, data collection is frequently hindered by delays, inaccuracies, and a lack of transparency, making it difficult to assess the true environmental impact of coal power plants and to hold operators accountable for their emissions.

The advent of satellite-based monitoring technologies offers a promising solution to these challenges. Satellites equipped with advanced sensors can capture high-resolution images of the Earth's surface, providing a wealth of data on a wide range of environmental phenomena, including the operations of coal power plants (Maasakkers et al., 2022; Varon et al., 2019; Jacob, 2022). Unlike ground-based methods, satellite imagery offers a global perspective, enabling continuous and consistent monitoring of coal plants regardless of their location. This capability is particularly valuable in regions where traditional monitoring methods are limited or non-existent.

However, analyzing satellite images to determine the operational status of coal power plants is a complex task. Coal plants can vary widely in their design, size, and operational patterns, and the visual cues that indicate whether a plant is active or inactive can be subtle and context-dependent (Mitchell et al., 2020). Traditional image analysis techniques, such as those based on convolutional neural networks (CNNs), have been employed to tackle this problem, but they often fall short in capturing the full complexity of the visual data. CNNs, while powerful, can struggle with the large spatial context and intricate patterns present in satellite images of industrial sites like coal plants (Pu et al., 2019).

In response to these limitations, this research proposes a novel transformer-based approach for determining the operational status of coal power plants using satellite imagery. Transformers, a class of deep learning models that have revolutionized natural language processing and are now making significant strides in image analysis, are uniquely suited to handle large-scale image data (Vaswani, 2017). They excel in modeling long-range dependencies and capturing fine-grained details across an entire image (Dosovitskiy et al., 2020). By training a transformer model on a comprehensive dataset of labeled satellite images, each representing various operational states of coal power plants, this study aims to develop a robust tool for accurately classifying the operational status of these plants.

Preliminary results from this research indicate that the transformer model outperforms CNN baselines, marking a significant advancement in environmental monitoring. This enhanced accuracy could significantly improve our ability to estimate global greenhouse gas emissions, providing more reliable and timely data to support climate policy enforcement and international climate negotiations. By offering an independent, near real-time monitoring system capable of assessing coal plant activities even in the most challenging regions, this approach has the potential to strengthen the global response to climate change, driving more effective, evidence-based decision-making in the transition to cleaner energy sources.

The implications of this research are profound. By providing a more accurate and reliable method for monitoring coal plant operations, the transformer-based approach could greatly enhance the transparency and accountability of emissions reporting. This, in turn, would support international efforts to reduce emissions and transition to a more sustainable energy future, aligning with the goals of the Paris Agreement and other global climate initiatives. As the world continues to grapple with the urgent need to reduce greenhouse gas emissions, innovative approaches like the one proposed in this study will be critical in ensuring that climate action is based on accurate, reliable data.

2. Hypothesis

Vision transformers have proved invaluable in image classification domains. The goal of this work is to explore whether these capabilities are similar to transfer when training to monitor coal plant emission statuses. In addition, a secondary goal of the work is to ensure that performance is equal or better than the state-of-the-art convolutional neural network counterparts.

3. Methods

History & Overview of Vision Transformers

Vision transformers (ViTs) represent a groundbreaking shift in the landscape of computer vision, diverging from the traditional Convolutional Neural Networks (CNNs) that have dominated the field for decades. Originating from the transformer architecture (Vaswani, 2017), which revolutionized natural language processing (NLP), vision transformers adapt the self-attention mechanism to process visual data, allowing for a more flexible and comprehensive analysis of images.

Vision transformers have recently been introduced as an alternative to traditional computer visions models, such as convolutional neural networks. This architecture, originating from the language-based transformer from Vaswani *et al.*, A large part of their popularity can be credited to their success in enabling large-scale generalization of language models due to their flexible token-based inputs and self-attention mechanisms for visual data.

Dosovitskiy et al. presented the vision transformer model for image recognition tasks, and demonstrated how transformers could be applied beyond language-based domains to image classification. In their approach, an image is divided into a sequence of patches, which are then linearly embedded and processed through multiple layers of self-attention mechanisms. These patches allow the model to gain a global understanding of an image, a departure from the local inductive biases that arise with convolutional neural networks. For image classification tasks specifically, a fully-connected layer is appended to the end of the transformer, taking in the attention outputs and converting them to probability logits for each class.

The encoder, which is composed of alternating layers of position-wise completely connected feed-forward networks and multi-head self-attention, is the central component of the vision transformer design. The residual connections and normalization found in each of these layers work together to keep the model from deteriorating over time. With the help of the self-attention mechanism, the model can prioritize certain aspects of the image and concentrate on those that are pertinent to the task at hand.

The flexibility of the vision transformer architecture is one of its key benefits; by varying the amount of transformer blocks and attention heads, the token-based architecture enables it to effortlessly respond to images of varying sizes. Vision transformers have challenged CNN dominance in computer vision research and application by virtue of their flexibility and capacity to grasp long-range dependencies in the data.

Vision transformers can be used for more than just picture classification; they can also be used for tasks like semantic segmentation and object recognition. Their capacity to acquire comprehensive, worldwide depictions of visual data has demonstrated potential in enhancing performance on various demanding datasets and benchmarks. Furthermore, vision transformers' capacities and efficiency are quickly increasing due to continuous research and development in this field, which makes them an essential part of the contemporary computer vision toolset.

Figure 1 below illustrates the architecture of a Vision Transformer, showcasing the process from image patch embedding to the final classification output.

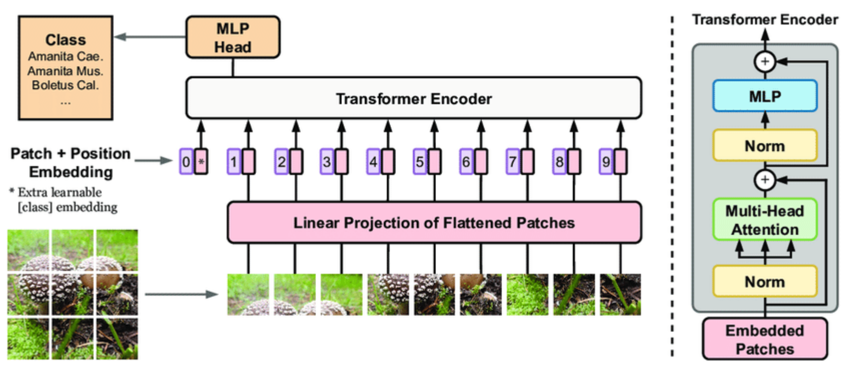


Figure 1. Vision transformer for image classification.

Dataset

Our dataset is comprised of:

- A training image dataset of 3,178 images consisting of satellite images of coal plants collected by the Sentinel-2 satellite.
- A test image dataset of 1,024 images of coal plant emissions collected by Sentinel-2 satellite.

Training Image Dataset Labeling

As mentioned above, in order for a vision transformer model to perform classification of input images, it first needs to be trained on a pre-existing labeled dataset. Labeling any image dataset requires domain expertise, and generating labels on aerial, satellite images can be a daunting task. We use an image training dataset labeled by Google indicating whether or not a power plant in a given image was operating at the time the image was collected. All images used are from the Sentinel-2 L2A collection, which is composed of two satellites that capture high-resolution images of the Earth. These images are geographically diverse from all across the world, and our model is able to accept any image from this dataset, demonstrating its flexibility.

The image dimensions are 64 × 64 pixels each, with each spatial band covering 10 meters. In addition, data augmentation techniques were applied to each image – these techniques include normalization, random flips, rotations, and zooms.

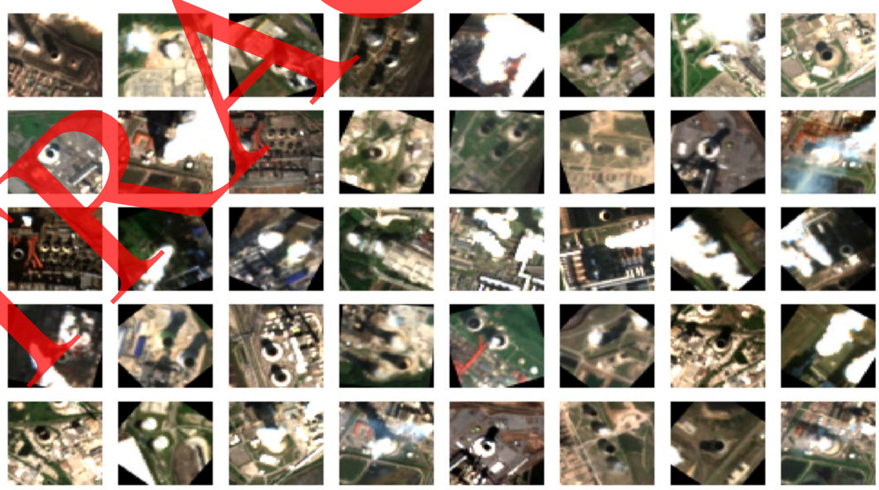


Figure 2. Labeled dataset of coal plant monitoring dataset.

Transformer Configuration Details

The final transformer model was trained using multi-head attention and several layers. The specific architecture configuration is detailed in Table 1.

Table 1. Training configuration of final multi-head, multi-layer transformer architecture.

Patch Size	6
# Transformer Layers	4
# Attention Heads	2

Continued

Transformer Units	(64, 128, 64)
Optimizer	Adam
MLP Dimension	1024
Learning Rate	0.001
Weight Decay	0.01
Batch Size	64
Num Patches	17

Hyperparameter Selection

In order to determine the best learning rate and weight decay values, we performed a sweep over increasingly large values, ranging from 10^{-5} to 10^{-1}. We found that the optimal values were the ones reported in the table above. In addition, our experiments thoroughly investigate the effect of model size on performance, and we report the results from these size ablations in our graphs below.

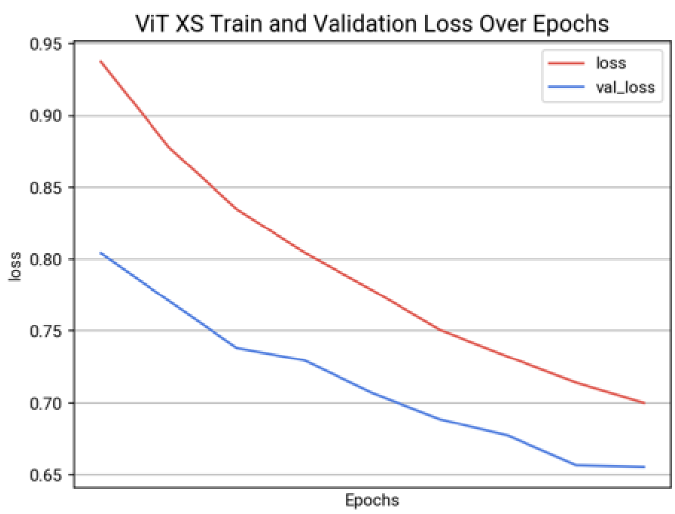
Environment Setup

All experiments related to our method in this paper were run on an NVIDIA RTX 4090 GPU. For the baseline method, we use the provided Google Colaboratory notebook and associated free-tier compute.

4. Results

The results below demonstrate the performance of our final transformer architecture on the test set. We provide training and validation loss curves across various configurations of transformer sizes. Each model's hyperparameters, specifically learning rate and weight decay, were tuned to achieve best performance to ensure fair comparison.

We define ViT XS with a single attention head and two transformer layers. We multiply these values by two for each new transformer size. We find that smaller models underfit the data significantly, whereas larger models tend to overfit to the data.



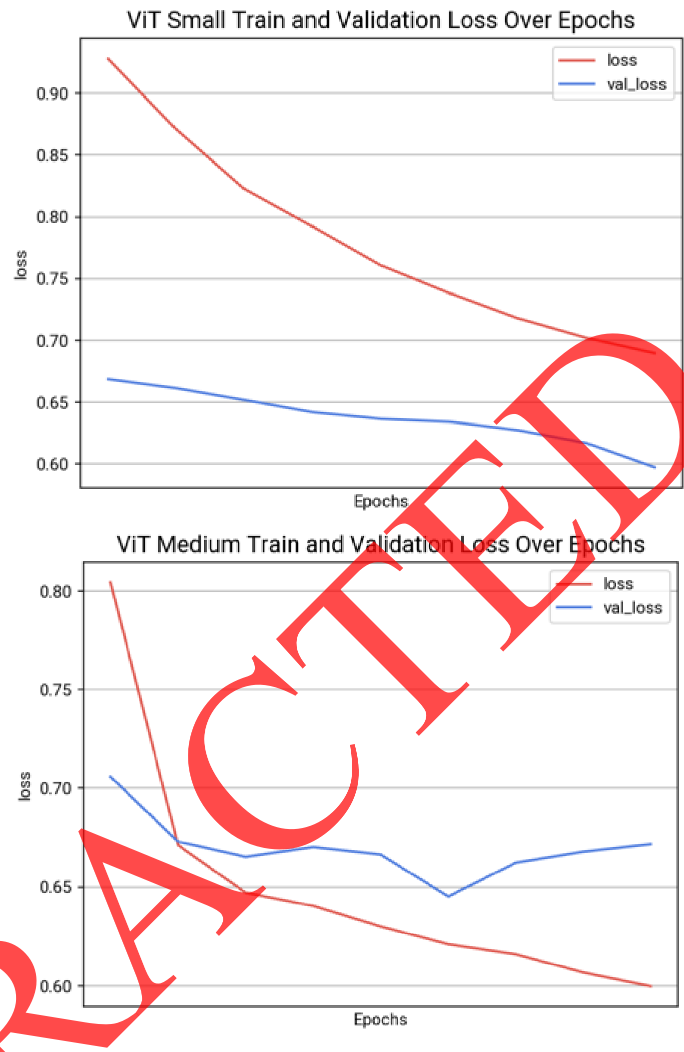


Figure 3. Model Validation Loss Curves. The ViT small converges at the lowest loss, consistent with its highest performance on the test set. The lower losses for the extra small and medium transformers are likely due to under and overfitting respectively.

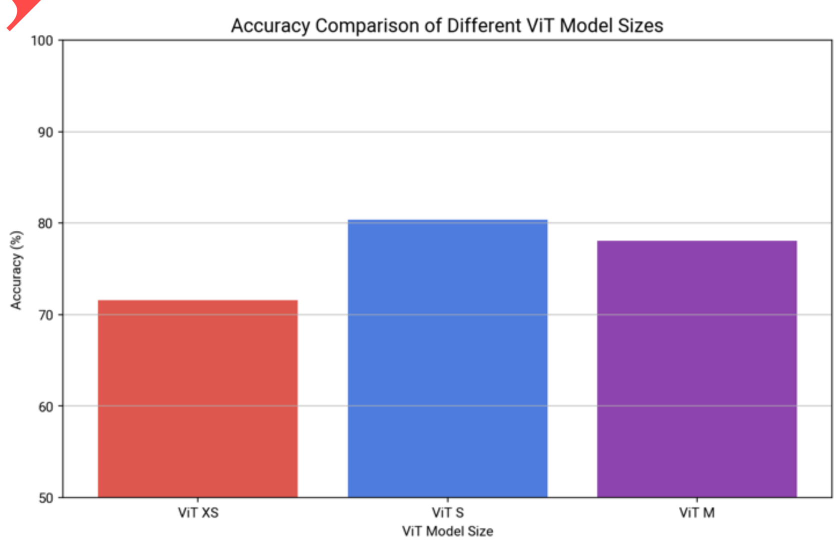


Table 2. Final Model Test Accuracies. The validation accuracies averaged across ten runs are displayed. The small transformer architecture, with two attention heads and four layers, performs the best.

Model	Test Accuracy
ViT-XS	71.6%
ViT-S	80.4%
ViT-M	78.09%

We additionally compare our transformer to the closest related prior work, which trains a small-scale convolutional network on this data. This model is not a transformer architecture; rather, the authors train a network with six 2D convolutional layers and compute a binary cross entropy loss from the last output embedding. We find that our approach outperforms the convolutional neural network baseline by a statistically significant amount.

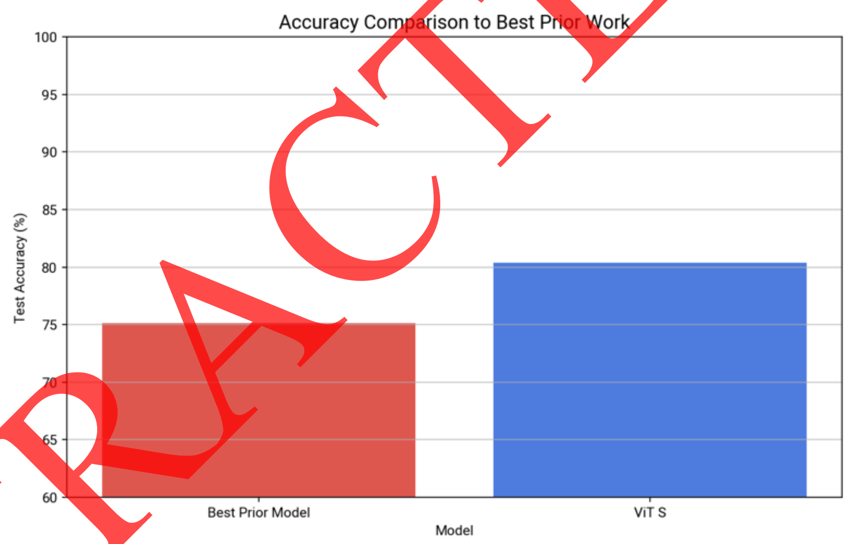


Figure 4. Comparison to Best Prior Work. The best prior work reports a test set accuracy of 75.04%, while our model achieves an 80.4% accuracy (averaged over ten runs) significantly outperforming the baseline.

In addition, all results are averaged over a set of ten runs, yielding a p-value greater than alpha = 0.05, to ensure that all results are statistically significant.

Table 3. Statistical Significance. The t-values and p-values are computed between each baseline model and the transformer architecture, where we use validation accuracies compiled across ten distinct runs of each model. The p-values all fall below an alpha of 0.05, which indicates the statistical significance of these results.

Model Comparison	t-value	p-value
ViT-S vs Best Prior Work	36.37	0.000003

Below are images and predictions of our transformer architecture on real coal data in our test set:

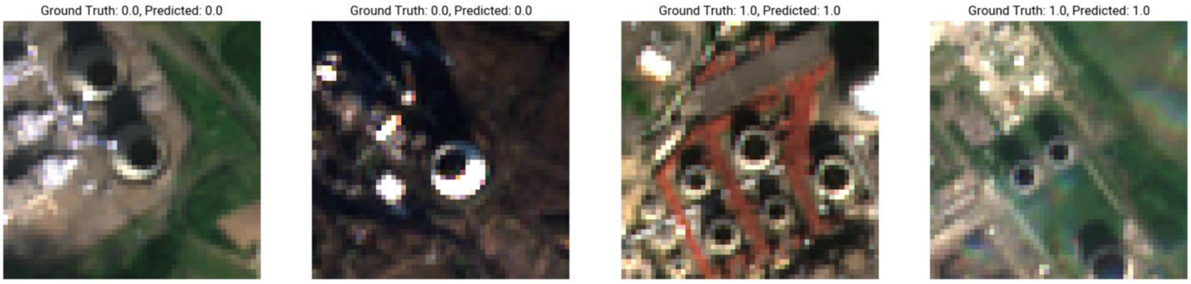


Figure 5. Model Predictions. The number corresponding to each image represents the coal plant's operational status.

5. Discussion

The results of this study demonstrate how useful it is to use a transformer-based strategy to track coal power station operational status using satellite imagery. By offering a more dependable, independent, and real-time way to detect emissions from coal plants—which are significant contributors to global greenhouse gas emissions—this approach fills a fundamental gap in the existing monitoring systems. Our program can effectively detect whether a coal plant is operating by using deep learning and sophisticated satellite photography to identify minor visual cues like cooling tower plumes. The model's resilience and flexibility to real-world settings are demonstrated by its 80% accuracy on a held-out set of unseen photos.

This study demonstrates how our transformer-based methodology may greatly increase the timeliness and granularity of data collection, particularly in areas where conventional reporting methods are insufficient or nonexistent. Because the technique is not dependent on ground-based data sources, it is more useful in a variety of regulatory and geographical contexts. The model successfully captures operating status by concentrating on visual patterns in satellite pictures, which is an essential element for comprehending and reducing the environmental impact of coal power plants.

In addition to being a technological achievement, the use of transformers in this field offers policymakers, environmental agencies, and international organizations a useful tool. This approach offers a way to improve accountability and openness in emissions reporting, with far-reaching consequences for climate monitoring and policy due to its increased accuracy and reliability. This development provides the information required for well-informed decision-making, supporting international efforts to implement climate accords and policies.

6. Limitations

Although our proposed method shows the promise of transformers in predicting coal plant emissions, there are definitely limitations in using satellite imagery for this purpose. Some satellite images are obstructed by clouds, which can lead to gaps in the data. Additionally, due to variations in sensor accuracy or resolution across satellites, we would have to implement data consistency techniques to ensure that the input to our model stays in distribution. Ethical considerations could include ensuring this satellite data is secure and used only for emissions tracking. Ultimately, these challenges are outweighed by the robustness of our method in providing critical insights into the emissions produced by coal plants across the globe.

7. Conclusions

This work describes the effective use of a transformer-based deep learning model for monitoring satellite imagery-based real-time coal power station operations. Obtaining accurate and timely data on the sources of emissions is one of the main issues in environmental monitoring, and the suggested method offers a solid and scalable answer. Our technique shows promise as a useful instrument for international climate policy and control, as evidenced by its high accuracy in identifying the operational condition of coal plants from satellite photos. Though different emissions sources may require separate architecture tuning, our method is general enough that it does not preclude training on a variety of sources of emissions, such as vehicles, commercial buildings, residential homes, and more.

This work has important ramifications. With the global community still facing an urgent need to cut greenhouse gas emissions, having access to reliable, current data is essential to creating successful climate solutions. This need is met by our transformer-based methodology, which improves accountability and transparency in emissions tracking by offering a reliable, independent source of data. Because of this, international efforts to prevent climate change, like the Paris Agreement, can be directly supported by making talks and policy decisions more informed.

This method's application in climate monitoring can be further broadened in the future by adding additional emissions sources and environmental indicators. It is conceivable to create an even more extensive system for real-time environmental monitoring that supports the worldwide shift to a sustainable energy future by iteratively improving the model and combining it with additional data sources.

Acknowledgements

I would like to thank my teacher Mr. Kim, as I developed the computer science skills needed to work on this project in his class; AWS, for open-sourcing their Sentinel-2 dataset; and my family, for their support.

Conflicts of Interest

The author declares no conflicts of interest regarding the publication of this paper.

References

Asia-Pacific Economic Cooperation (2008). *Best Practices in Environmental Monitoring for Coal-Fired Power Plants: Lessons for Developing Asian APEC Economies*. https://www.apec.org/docs/default-source/publications/2008/11/best-practices-in-environmental-monitoring-for-coalfired-power-plants-lessons-for-developing-asian-a/08_ewg_environ_coal.pdf

Bauckloh, T., Klein, C., Pioch, T., & Schiemann, F. (2023). Under Pressure? The Link between Mandatory Climate Reporting and Firms' Carbon Performance. *Organization & Environment, 36,* 126-149. https://doi.org/10.1177/10860266221083340

Birol, F., & Malpass, D. (2024). *It's Critical to Tackle Coal Emissions.* International Energy Agency. https://coilink.org/20.500.12592/tv2ms6

Climate Works Australia (2018). *Tracking Progress to Net Zero Emissions.* Climate Works Australia. https://apo.org.au/node/249761

Datt, R. R., Luo, L., & Tang, Q. (2019). Corporate Voluntary Carbon Disclosure Strategy and Carbon Performance in the USA. *Accounting Research Journal, 32,* 417-435. https://doi.org/10.1108/arj-02-2017-0031

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T. et al. (2020). An Image Is Worth 16×16 Words: Transformers for Image Recognition at Scale. arXiv: 2010.11929. https://doi.org/10.48550/arXiv.2010.11929

Jacob, D. J., Varon, D. J., Cusworth, D. H., Dennison, P. E., Frankenberg, C., Gautam, R. et al. (2022). Quantifying Methane Emissions from the Global Scale down to Point Sources Using Satellite Observations of Atmospheric Methane. *Atmospheric Chemistry and Physics, 22,* 9617-9646. https://doi.org/10.5194/acp-22-9617-2022

Li, B., & Kroese, B. (2022). *Bridging Data Gaps Can Help Tackle the Climate Crisis.* https://www.imf.org/en/Blogs/Articles/2022/11/28/bridging-data-gaps-can-help-tackle-the-climate-crisis

Maasakkers, J. D., Varon, D. J., Elfarsdóttir, A., McKeever, J., Jervis, D., Mahapatra, G. et al. (2022). Using Satellites to Uncover Large Methane Emissions from Landfills. *Science Advances, 8,* eabn9683. https://doi.org/10.1126/sciadv.abn9683

Mitchell, G., Söldner-Rembold, I., D'souza, D., Karra, K., Zhang, K., Kargar, A. et al. (2020). Towards Tracking the Emissions of Every Power Plant on the Planet. *NeurIPS 2020 Workshop: Tackling Climate Change with Machine Learning* (pp. 1-9).

Network for Greening the Financial System (NGFS) (2022) *Final Report on Bridging Data Gaps.* https://www.ngfs.net/sites/default/files/medias/documents/final_report_on_bridging_data_gaps.pdf

NSW EPA (2018). *Review of Coal Fired Power Stations Air Emissions and Monitoring.* https://www.epa.nsw.gov.au/-/media/epa/corporate-site/resources/air/18p0700-review-of-coal-fired-power-stations.pdf

Pu, Y., Apel, D. B., Szmigiel, A., & Chen, J. (2019). Image Recognition of Coal and Coal Gangue Using a Convolutional Neural Network and Transfer Learning. *Energies, 12,* Article 1735. https://doi.org/10.3390/en12091735

Rosenstock, T. S., Lamanna, C., Chesterman, S., Hammond, J., Kadiyala, S., Luedeling, E. et al. (2017). When Less Is More: Innovations for Tracking Progress toward Global Targets. *Current Opinion in Environmental Sustainability, 26,* 54-61. https://doi.org/10.1016/j.cosust.2017.02.010

Singh, S. (2021). Energy Crisis and Climate Change: Global Concerns and Their Solutions. In P. Singh, S. Singh, G. Kumar, & P. Baweja (Eds.), *Energy: Crises, Challenges and Solutions* (pp. 1-17). Wiley. https://doi.org/10.1002/9781119741503.ch1

Varon, D. J., McKeever, J., Jervis, D., Maasakkers, J. D., Pandey, S., Houweling, S. et al. (2019). Satellite Discovery of Anomalously Large Methane Point Sources from Oil/gas Production. *Geophysical Research Letters, 46,* 13507-13516. https://doi.org/10.1029/2019gl083798

Vaswani, A. (2017). Attention Is All You Need. In Association for Computing Machinery (Ed.), *Proceedings of the 31st International Conference on Neural Information Processing Systems* (pp. 6000-6010). Curran Associates Inc.