Universität St.Gallen

Code: github

Multitask Learning for Estimating Power Plant Greenhouse Gas Emissions from Satellite Imagery

Joëlle Hanna, Michael Mommert, Linus Scheibenreif, Damian Borth
AIML Lab, School of Computer Science, University of St.Gallen

Tackling Climate Change with Machine Learning workshop at NeurIPS 2021
December 14, 2021

"From insight to impact"

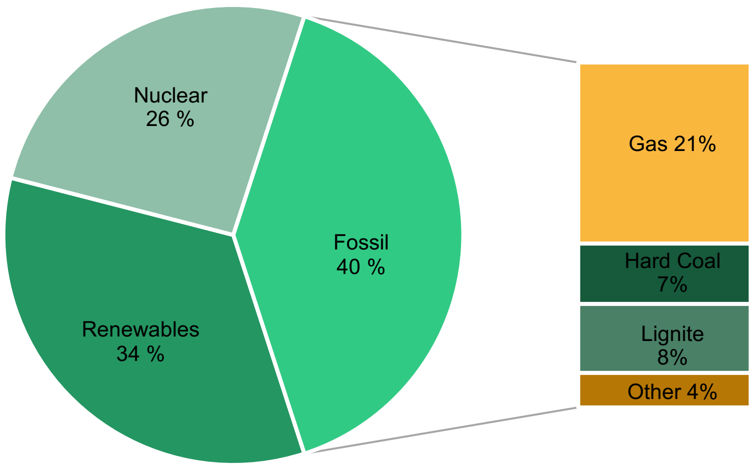
Universität St.Gallen

Introduction

Greenhouse gas emissions (mainly CO_2 and CH_4) are a major driver of climate change, and therefore, they represent the cause of all its adverse effects, namely extreme weather events, wildfires, droughts, sea level rise, etc.

Most CO_2 emissions come from the combustion of fossil fuels (hard coal, lignite, gas, oil, peat, etc.) for instance for the generation of power, on which we will focus in this study.

Share of Fossil-Fuels in Europe's Power Generation



https://energypost.eu/europes-energy-revolution-marches-one-third-power-supply-now-comes-renewables/

Motivations

To control the amount of greenhouse gas emissions, and therefore to mitigate the effect of climate change, emission trading systems have been implemented that aim to provide economic incentives for reducing the emission of pollutants.

For this to work properly, accurate and detailed emission reports are needed. These can be obtained using, for example, dedicated measuring devices.

What about regions where these reports are missing / not required?

Universität St.Gallen

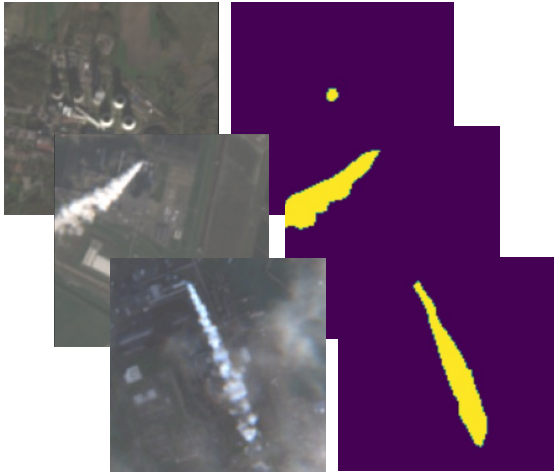
Contributions

The main objective of our work is to estimate the amount of emitted CO_2 for fossil fuel power stations. Our contribution is threefold:

I. We compile a satellite image data set of fossil fuel active power plants in Europe together with concurrent power generation rates and weather information and annotate their plumes in segmentation maps.

II. We propose a multitask learning approach able to simultaneously segment plumes, predict the type of fired fuel, as well as the power generation rate.

III. We estimate CO_2 emission rates using the predicted power generation rates and a derived emission factor.



Data

Our dataset is constituted of 1,639 Sentinel-2 satellite observations of 146 different fossil fuel power plants with their segmentation label. We supplement our processed satellite image data with concurrent weather[1] and power generation rates[2,3]. Our sample includes power plants that use 4 different types of fuel[2,3]: hard coal (41%), gas (29%), lignite (29%), and peat (≤1 %).

Example data

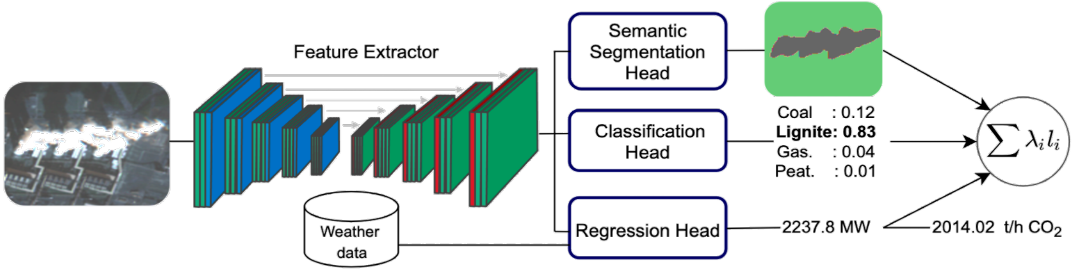
True color image	Segmentation map	Power generation	Temperature	Humidity	Wind	Fuel
		2050 MW	290 K	53.2 %	(-1.5, 2.47) m/s	Lignite
		500 MW	274 K	84.1 %	(5.4, 1.57) m/s	Hard Coal

[1] K. Kanellopoulos et al. Jrc open power plants database (jrc-ppdb-open), December 2019. https://doi.org/10.5281/zenodo.3574566.
[2] H. Hersbach et al. The era5 global reanalysis. Quarterly Journal of the Royal Meteorological Society, 146:1999–2049, 2020
[3] Entsoe. Actual generation per generation unit table, 2021. https://transparency.entsoe.eu/generation/r2/actualGenerationPerGenerationUnit/show.

Our Approach

We propose an end-to-end multitask deep learning approach able to simultaneously predict three different tasks:

semantic segmentation of plumes, classification of the type of fuel and regression of the power generation.



(1) We use a U-Net[1] as our feature extractor. It outputs a unique representation that will be used by the three specific tasks.

(2) Each task requiring a specific loss, we use a weighted sum of them for the full training.

(3) We estimate CO_2 emissions by multiplying the predicted power generation by a fuel-dependent emission factor.

[1] Ronneberger, O., Fischer, P., and Brox, T. U-Net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, May 2015

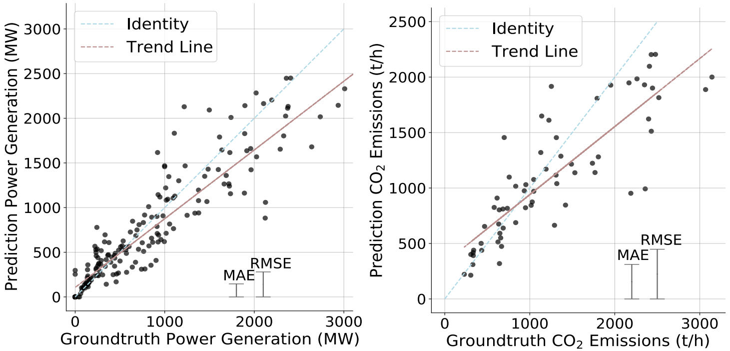
Universität St.Gallen

Results

We note that the multi-task model outperforms the single-task models trained on the same data: we observe a relative improvement of 5% on the segmentation IoU, 39 % on the regression MAE, 26 % on the regression R^2, and 10% on the classification accuracy

Test set multitask vs single task baseline performances

Loss	Task Weights (λ_i)			Seg. IoU	Reg MAE / R^2	Cls Acc.
	Seg.	Reg.	Cls.			
Seg. only	1	0	0	0.640	-	-
Reg. only	0	1	0		225 / 0.66	-
Cls. only	0	0	1	-	-	0.775
Reg + Seg	0.4	0.6	0	0.643	145 / 0.81	-
Reg + Cls	0	0.6	0.4	-	151 / 0.81	0.779
All tasks	0.15	0.7	0.15	**0.668**	**139 / 0.83**	**0.853**

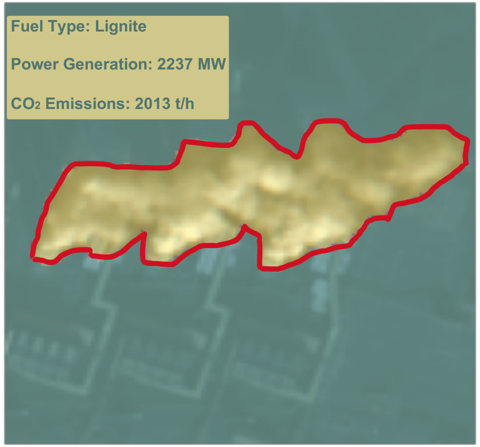


Our multitask approach enables us to predict power generation rates within 139 MW (MAE) or 19 % (MAPE) and estimate CO_2 emission rates within 311 t/h (MAE) or 34 % (MAPE) for our test set power plants.

Universität St.Gallen

Conclusions

- We propose a deep **multitask framework to predict power generation rates** with high confidence from single Sentinel-2 satellite images.
- Our framework is trained on a combination of **three tasks** and experiments confirmed that auxiliary tasks can indeed boost the network performance.
- Our model is able to predict power generation rates from individual images with **R^2=0.83** or within **139 MW (MAE)** and CO_2 emission rates within **311 t/h (MAE)**.
- The results from this work are able to supplement emission information in countries that do not require the reporting of such information



Responsible Author:

Joëlle Hanna

joelle.hanna@unisg.ch