



HAL
open science

Machine Learning and Feature Extraction for Industrial Smoke Plumes Detection from Sentinel-2 Images

Florentin Poucin, Elyes Ouerghi, Simon Lajouanie, Hugo de Almeida Rodrigues, Gabriele Facciolo, Carlo de Franchis, Charles Hessel

► To cite this version:

Florentin Poucin, Elyes Ouerghi, Simon Lajouanie, Hugo de Almeida Rodrigues, Gabriele Facciolo, et al.. Machine Learning and Feature Extraction for Industrial Smoke Plumes Detection from Sentinel-2 Images. 2023 IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2023), Jul 2023, Pasadena, CA, United States. pp.6113-6116, 10.1109/IGARSS52108.2023.10282662. hal-04497731

HAL Id: hal-04497731

<https://hal.science/hal-04497731v1>

Submitted on 10 Mar 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MACHINE LEARNING AND FEATURE EXTRACTION FOR INDUSTRIAL SMOKE PLUMES DETECTION FROM SENTINEL-2 IMAGES

Florentin Poucin¹ Elyes Ouerghi² Simon Lajouanie¹ Hugo de Almeida Rodrigues¹
Gabriele Facciolo² Carlo de Franchis^{1,2} Charles Hessel^{1,2}

¹Kayrros SAS

²Université Paris-Saclay, CNRS, ENS Paris-Saclay, Centre Borelli, France

ABSTRACT

The detection of smoke plumes by satellite imagery is a comprehensive research topic that can be used to better monitor activity and emissions from the energy and industrial sectors. In this study, we propose a machine learning methodology based on the extraction of relevant features from Sentinel-2 images to perform industrial smoke plume detection. This computer vision problem is modeled as an image classification task based on the presence or absence of plumes for previously identified sources. A dataset of nearly 17,000 hand-labeled images of smoke plumes for activity classification has been compiled to train and evaluate our detection models. The final Gradient Boosting model only uses the 3 RGB bands of Sentinel-2 and after a post-processing step reaches an accuracy of 95%.

Index Terms— Industrial plumes, Sentinel-2 images, Machine learning, Feature extraction, Pattern recognition

1. INTRODUCTION

Monitoring the activity of industrial and power generation plants is a major issue for estimating greenhouse gas emissions from these sectors. The observation over time of industrial smoke plumes emitted by these plants is a good indicator of the state of their activity. These plumes can be of different sizes and compositions depending on the plant type, but many of them can be identified on medium-resolution satellite images. For this purpose, optical images from the Sentinel-2 constellation provide a good trade-off between spatial resolution and revisit frequency. The problem addressed in this study is the classification of satellite images based on the presence or absence of industrial smoke plumes. To tackle this problem, we propose a machine-learning method based on the extraction of relevant features.

The smoke plume detection of our study is carried out on cooling towers of different sites around the world, such as power or industrial plants. The prior identification of plume-emitting sources makes the computer vision task easier but greatly increases the performance of the detection models. The different plume sources are then identified beforehand.

They are of various sizes and shapes. This diversity is an important stake making the robustness of our method necessary. In order to train and evaluate our models, we built a dataset of nearly 17,000 images of smoke plume sources labeled by hand for activity classification.

The main idea of the approach is to extract the most valuable features from Sentinel-2 images in order to detect the presence of smoke plumes. This detection is achieved by utilizing reference classification models, such as gradient boosting or ResNet. The value of the presented method lies in the pre-processing of the images. First, we apply a process aimed at eliminating the image background, which tends to highlight possible plumes. Then we reduce the number of parameters by grouping information into relevant features. For this purpose, various parameters are incorporated into the method, including source positions, seasonality, satellite orbits, and image context.

2. RELATED WORK

The detection of industrial smoke plumes has been extensively studied in the scientific literature. Hsu *et al.* introduce Project RISE [1], which specifically focuses on recognizing and identifying industrial smoke emissions. To accurately characterize these plumes, various papers have investigated the physical properties of industrial smoke. Mommert *et al.* [2, 3] and Calassou *et al.* [3] have contributed to this area of research, studying the characteristics and behavior of industrial smoke plumes. In addition to industrial smoke, Calassou *et al.* [3] also propose a comprehensive characterization of aerosol plumes. Furthermore, Ehret *et al.* [4] have explored the detection of methane plumes. These studies provide valuable insights and methodologies for effectively detecting and characterizing different types of plumes in various industrial and environmental contexts.

Satellite data are a rich source of information for detecting smoke plumes. The European Space Agency (ESA) provides detailed documentation [5] allowing the exploitation of Sentinel-2 satellite images [2, 4, 6]. The usability of satellite images often depends on weather conditions. For this pur-

pose, ground visibility is studied in [7, 8], and cloud detectors are developed in [9, 10].

To perform image classification, a wide range of models is available, ranging from logistic regression to the latest image transformers. In this study, we chose to utilize a gradient boosting method, XGBoost [11], along with a well-known convolutional neural network, ResNet [12]. The main focus of this study lies in the pre-processing of the data; therefore, exhaustive testing of different classifiers was not conducted. However, the relevance of the method should extend to any other model thus enhancing its performance.

3. METHOD

Our smoke plume detection method is based on the computation of features to train a classification model. The exact positions of the towers that can emit plumes are considered to be known. This allows us to build, for each tower, a time series of observations. This time series takes into account the seasonality to ensure that the different observations of the towers are sufficiently similar: the time difference between the first and last observation of the time series should not exceed 3 months. Satellite images are also grouped by orbit in time series to which a global registration [13] is applied. The background of the observed scene is then computed as the average of this time series. We then compute the difference between the image to be tested and the average of the time series to remove this background.

We then extract square crops of several sizes centered around each tower. The idea behind these centered crops is to retrieve information from a potential plume signal at different scale levels around the source. This is illustrated by Figure 1. A very small patch to get concentrated information on whether something is coming out of the tower, a medium-sized patch to get information from the environment close to the source, and a much larger patch to assess the overall context of the image. We use crops of 4, 10, and 40 pixels of side. With a Sentinel-2 resolution of 10 meters, this corresponds to patches of 40, 100, and 400 meters of side. Using crops of different sizes, especially small ones, ensures that the plume is emitted from the observed tower and not from another nearby tower. These crops allow computing features in the proximity of the towers, such as the mean and the standard deviation of the pixels of the image to be tested and the residual of the image with the time series. We also use the mean ratio of the red band over the blue band of the Sentinel-2 image.

These features are then used to train a Gradient Boosting model [11]. The classification of activities is based on whether or not a tower is emitting a smoke plume, but for more than half of the images, the presence of clouds makes this inference impossible. These images are nevertheless kept in the dataset in order to make the model capable of differentiating between atmospheric conditions and industrial activity signals. Therefore, the model is trained to classify the data

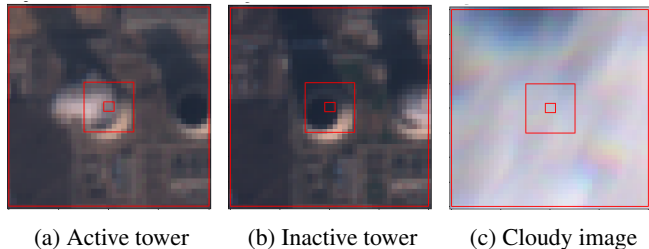


Fig. 1: Patches of different sizes (red contours), centered on a large cooling tower. We employ patches of varying sizes to capture the image context at different scales.

into 3 classes: *Active tower*, *Inactive tower*, and *Cloudy*. The dataset is separated into a training set and a test set by making sure that the areas seen in the training set are never present in the test set, even at a different date. This ensures that the model will have a stable performance in all types of areas.

To limit the errors of the model we add a posteriori a fourth class: *Unclear*. This class allows us to isolate the observations where we cannot trust the prediction of the model. We determine if an observation is *Unclear* by calculating the entropy of the distribution that the model associates with the observation. It should be noted that the *Unclear* class does not mean that the model has made an error in its prediction but rather that even if the prediction is correct we cannot trust it.

4. EXPERIMENTS AND RESULTS

The model thus constructed is tested on a manually labeled plume dataset. This dataset contains nearly 17,000 images, labeled as active, inactive, or cloudy, and separated into training, validation, and test sets. The method gives an accuracy of 95% on the test set. In particular, the model without the addition of the *Unclear* class a posteriori gives an accuracy of 92%. The confusion matrix of the model after the addition of the *Unclear* class is presented in Table 1. The other main strength of the method is background subtraction. As can be seen in Figure 2, the number of potential false positives is drastically reduced by switching from the RGB image to the image without a background. The accuracy scores obtained by our method are quite similar to those obtained in [2]. However, it should be noted that the task and the dataset are not the same. In our case, we limit ourselves to classification and assume the position of the plume sources is known. This allows us to have a more interpretable and lighter model than convolutional networks, and still detect a wide variety of plumes even if they are small. In addition, we note a strong seasonality in the detectability of smoke plumes whose physical properties depend greatly on temperature and humidity.

Although the performance of the method is quantitatively strong, several improvements to the model have been iden-

	Predicted Active	Predicted Inactive	Predicted Cloudy
Label Active	99	26	5
Label Inactive	4	746	3
Label Cloudy	8	10	1116

Table 1: Confusion matrix obtained by applying the model on the test set. Those data correspond to the results after the addition of the Unclear class which means that the data labeled as Unclear have been removed.

tified in this study. The introduction of the *Unclear* class, based on the entropy of the model’s class predictions, makes it possible to identify the cases of limitation of the method. Figure 3 illustrates the inference by the model of an *Unclear* state, in which illumination makes it difficult to differentiate between smoke plumes and illuminated roofs. Furthermore, comparing Figures 2 and 3, images of the same plant were taken from two different orbits. The tops of the towers from which the smoke comes out are not exactly in the center of the patches following the orbit. As mentioned before, this difference is taken into account for the calculation of the time series and the background subtraction. However, for robustness reasons, the relative position of the top of each source with respect to the Sentinel-2 orbit is not provided as input to the model. Incorporating this information would probably improve the accuracy of the model.

We also compare the results of our method to the results obtained with a deep neural network. We make several comparisons. First of all, we compare our method with the one presented in [2]. For this we use the architecture proposed by [2] and we train it on our dataset. This allows us to test the viability of this model on a different set of plumes. The other comparison is to use a classical deep learning architecture such as a Resnet and to perform the classification using the features extracted in our method. The more complex architecture of a network and its significantly larger number of weights make image reduction using patches less important. We directly use as input the largest crops centered on the plume sources. The network should be able to extract by itself the information that we would have extracted with smaller crops. We use crops of size 40 which are the largest we have used with the gradient boosting. The input of the Resnet is then an image of size 40×40 with 6 channels. The first 3 channels are those of the RGB image and the next 3 are the residual of the time series. The results of the experiments proposed here are summarized in Table 2. The methods using a Resnet are indicated by the label Resnet + TS to indicate that we use the time series to obtain the features used as input.

We can see in Table 2 that our method using gradient boosting gives the best global accuracy with a score of 95.3. It is closely followed by the Resnet152 with a score of 95.2. Despite these very close scores in terms of global accuracy,

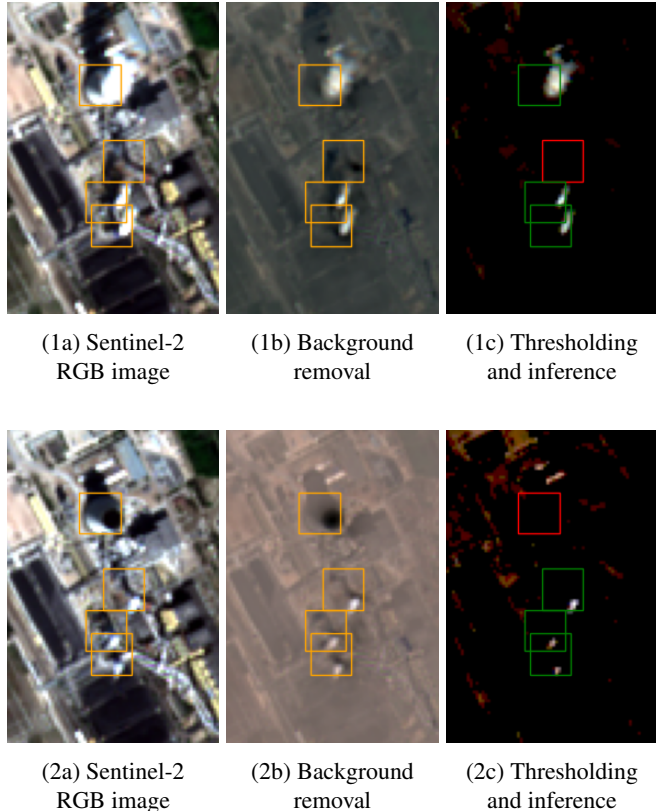


Fig. 2: Image processing for feature extraction. Plume sources are initially identified on Sentinel-2 RGB images (orange). Background removal is obtained after time series image registration [13] and subtraction of a reference image obtained by temporal averaging. Features are calculated for each source from centered crops of different sizes, from which the model infers whether the source is active (green) or inactive (red). The thresholding shown in this figure is applied only for visualization.

the results of the two methods are significantly different when we look at the results by class. The gradient boosting outperforms all other methods on the inactive label while the Resnet152 far outperforms the other methods on the active label. It is this latter label that has the greatest variance between the different methods. The difference between Resnet152 and Gradient boosting is much greater on the active label than on the inactive label. However, in the dataset, we have much more images with the inactive label than with the active label which explains why we obtain a better global accuracy with the gradient boosting. In the cloudy class, all methods have very close results although the Resnet152 gives slightly better results.

Concerning the method of Mommert *et al.* [2], we can note that it is the one that obtains the worst score in all the classes. Its results on the inactive and cloudy classes are just slightly below the others. But, the difference is especially

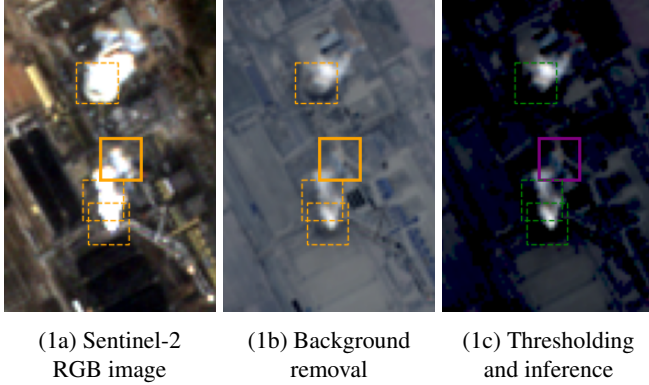


Fig. 3: Inference of an *Unclear* state of activity. In this example, the model prediction for the middle source (purple box in the right image) is considered not reliable enough. All four sources are predicted to be active by the model but this one has a distribution entropy equal to 0.63 compared to less than 0.03 for the other three. The prediction of this tower is then put to *Unclear*.

	Acc.	Inactive	Active	Cloudy
Gradient Boost.	95.3	99.1	76.2	98.4
Mommert <i>et al.</i>	89.5	94.7	45.9	97.4
Resnet50+TS	93.9	97.3	67.6	98.3
Resnet152+TS	95.2	95.9	81.1	98.6

Table 2: Comparison of classification results between Gradient Boosting and deep learning methods. The scores displayed for the Active, Inactive, and Cloudy columns correspond to the accuracy for each class in percentage. The column Acc. corresponds to the global accuracy. Those data correspond to the results after the addition of the *Unclear* class which means that the data labeled as *Unclear* have been removed

significant in the active class where it has only 45.9% of accuracy. These results are explained by the fact that in the dataset proposed by [2], the plumes are significantly larger than in our dataset. Thus, the approach of [2] is less suitable for small plumes. As for the choice to be made between Gradient Boosting and Resnet, it depends on the metric one wishes to maximize. If one prefers to maximize the recall on the inactive areas then gradient boosting will be better, while to maximize the recall on plume detection one will prefer Resnet.

5. REFERENCES

- [1] Y.C. Hsu et al., “Project rise: recognizing industrial smoke emissions,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2021, vol. 35, pp. 14813–14821.
- [2] M. Mommert et al., “Characterization of industrial smoke plumes from remote sensing data,” *arXiv preprint arXiv:2011.11344*, 2020.
- [3] G. Calassou et al., “Aerosol plumes characterization by hyperspectral images coupled with sentinel-2 products,” *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 43, pp. 791–797, 2020.
- [4] T. Ehret et al., “Global tracking and quantification of oil and gas methane emissions from recurrent sentinel-2 imagery,” *Environmental Science & Technology*, vol. 56, no. 14, pp. 10517–10529, 2022.
- [5] European Space Agency, “Sentinel-2 msi user guide,” 2015, <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi>.
- [6] G. Calassou et al., “Industrial plume properties retrieved by optimal estimation using combined hyperspectral and sentinel-2 data,” *Remote Sensing*, vol. 13, no. 10, pp. 1865, 2021.
- [7] R. G. von Gioi et al., “Ground visibility in satellite optical time series based on a contrario local image matching,” *Image Processing On Line*, vol. 11, pp. 212–233, 2021.
- [8] T. Dagobert et al., “Temporal repetition detector for time series of spectrally limited satellite imagers,” *Image Processing On Line*, vol. 10, pp. 62–77, 2020.
- [9] T. Dagobert et al., “Cloud detection by luminance and inter-band parallax analysis for pushbroom satellite imagers,” *Image Processing On Line*, vol. 10, pp. 167–190, 2020.
- [10] R. G. von Gioi et al., “Cloud detection by inter-band parallax and a-contrario validation,” *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. 3, pp. 39–46, 2022.
- [11] T. Chen and C. Guestrin, “Xgboost: A scalable tree boosting system,” in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, 2016, pp. 785–794.
- [12] K. He et al., “Deep residual learning for image recognition,” in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770–778.
- [13] C. Hessel et al., “A global registration method for satellite image series,” in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*. IEEE, 2021, pp. 3121–3124.