Detection of Gas Flares Using Satellite Imagery

Alexander Troussov, Dmitry Botvich and Sergey Vinogradov
International laboratory for mathematical modelling of social networks,
RANEPA,
Moscow, Russia
e-mails: troussov@gmail.com, dbotvich@gmail.com, derbosebar@gmail.com

***Abstract*—During the extraction, transportation and processing of oil, associated petroleum gas is formed, which is usually disposed in flares. Monitoring these flares is an important environmental challenge. Objective instrumental methods for detecting gas flares and assessing the volumes of gas burnt on them are based on multi-spectral remote sensing of the nighttime Earth. To refine the list, a manual check is used, which includes a visual examination of the locations of the alleged gas flares on high-resolution daytime images. Automating this check reduces the cost of monitoring flares. The paper proposes a method for verifying the list of high-temperature anomalies, based on the classification of images of objects in daytime images. The classification is carried out using machine learning methods.**

Keywords—night lights; light pollution; viirs; image processing; machine learning; deep learning.

I. INTRODUCTION

During the extraction, transportation and processing of oil, associated petroleum gas is formed, which is usually disposed in flares. Monitoring these flares is an important environmental challenge (see materials of the World Bank's Global Gas Flaring Reduction Partnership). Objective instrumental methods for detecting gas flares and assessing the volumes of gas burnt on them are based on multi-spectral remote sensing of the nighttime Earth. The sensor Visible Infrared Imaging Radiometer Suite (VIIRS) is used by satellites (Suomi National Polar-orbiting Partnership (Suomi NPP) and NOAA-20 weather satellite) at nighttime to collect imagery and radiometric measurements of the land, atmosphere and oceans in the visible and infrared bands of the electromagnetic spectrum. These data make it possible to find high-temperature anomalies on the Earth's surface, and by the spectrum of radiation to distinguish gas flares from, for example, forest fires and greenhouses. The most famous algorithm for finding gas flares from Earth remote sensing data is the VIIRS Nightfire (VNF) algorithm [1], [2], [3].

The refinement of this algorithm for use on the territory of Russia is described in our paper [4]. However, the list of gas flares obtained using the VNF method is not completely correct. Among the main factors that make the presence of errors inevitable, we mention the low temporal resolution of satellite data (satellites do not often fly over objects, objects can be covered by clouds), and interference in observations that are difficult to eliminate (snow cover, polar nights). Additional work is required to compile a "final list" of gas flares from a "preliminary list" of gas flares. In the "final list" it is also desirable to classify the flares by type of enterprise (upstream, transportation, downstream processing).

In our paper [4], methods of manual correction of the list of gas flares for the territory of Russia are described based on the use of various additional sources and databases. For example, the presence of an object in the authoritative list CEDIGAS (The International Association For Natural Gas), is a confirmation of the eligibility of including the object in the "final list". If the coordinates of an object are not included in the boundaries of licensed areas for oil production, then in the conditions of Russia this object can be confidently excluded from objects of the upstream type. However, such manual corrections are laborious and time-consuming, and flare monitoring needs to be done on a continuous basis with a high temporal resolution. Automation of verification, even partial, reduces the cost of monitoring flares. If we consider the monitoring of gas flares in areas of armed conflict, where illegal oil production and processing may take place, then the automation of gas flare testing becomes uncontested.

In this paper, we propose a method for checking the list of high-temperature anomalies obtained on the basis of NTL data, based on the classification of images of objects in daytime images. The classification is carried out using machine learning methods.

The paper is organized as follows. In Section II, we describe the satellite data used in the paper. Our approach to the correction of gas flares list is outlined in Section III. In Section IV, we present the image transformation process, including image coding, image descriptor extraction, etc. In Section V, we present some preliminary results. Finally, Section VI concludes the paper.

II. SATELLITE IMAGE DATA

The original flare list used in this work is part of the list of associated petroleum gas flares in Russia obtained in the course of our work described in [4]. For each prospective gas flare, a high-resolution daytime image was selected that covered the location of the supposed flare. Some of these data are labeled - that is, it is known whether or not there is a gas flare a given point. In addition, during the training phase, we added a manually selected set of daytime images that do not contain gas flares, but there are production facilities that may visually resemble oil and gas production and processing facilities.

The VNF algorithm utilizes near-infrared and short-wave infrared data at night, gathered by the VIIRS on board satellites NPP and NOAA-20 [3]. In [4] we used a version of the VNF algorithm adapted to the conditions of Russia. The total number of gas flares detected in [4] is about two and a half thousand.

We used also daytime satellite imagery from the webserver TerraServer to compile a “final list” of gas flares based on a “tentative list” of gas flares. The resolution of the daytime images is about 1m per pixel.

III. CHECKING AND CORRECTION OF GAS FLARES LIST

First of all, we note that a wide variety of third-party external sources can also be used when considering whether a snapshot contains a gas flare. For example, in many countries the licensing of oil areas for oil production (for example, in the USA, Saudi Arabia, Russia, etc.) works quite effectively(and reliably. The lists of such areas (i.e., their coordinates) are publicly available and can be used for the evaluation if a particular area is actually the oil area.

Here, we consider a different approach based on the daytime satellite image classification. We are using machine learning techniques here to adjust the ”final list” of gas flares based on some ”tentative list” of gas flares. The input data for us are the geographic coordinates of the alleged gas flares, which are used to extract recent daily satellite data, and then these images are analyzed by our automatic classifier (called, ClassGasFlare), which determines the type of gas flare from the image (by specialization enterprises: production, transportation, upstream processing, downstream processing) or lack of flares in the picture.

Since manual correction of this data is quite laborious and time-consuming, automating the check of the list can significantly reduce this work and helps to close the entire verification technological cycle: from automatic acquisition of images from the data cloud to their classification according to the type of gas flare or the absence of gas flares. In those cases when the automatic classifier has a ”difficult” case, then only these ”difficult” cases can be left for manual additional verification, which significantly reduces the volume of manual work.

We train our classifier on a sufficiently large volume of historical images, where they are labeled as (flare-dobycha, flare-transport, flare-upstream, flare-downstream, flare-no) depending on the type of oil production or simply as (flare-yes, flare-no) to classify the presence or absence of oil production. Thus, formally, in terms of machine learning, we get the picture classification problem. In a first approximation, this is true, but there are some important aspects that significantly complicate this task. For example, aspects such as:

- images can cover partially our objects (where the gas flare is located),
- images can be ”naturally spoiled” (for example, by clouds that cover part of the surface in the image),
- images may have additional features such as snow cover, which also distorts the original.

All these factors significantly affect the quality of the classification and should be taken into account when developing a classifier.

IV. DATA FOR THE CLASSIFIERS TRAINING

This section describes the data and machine learning classifiers used for the training and evaluation of the satellite image classification. The image transformation techniques are also discussed in this section.

A. Training data sets

We use satellite imagery (from the TerraServer web site) for the period 2010-2017 (GasFlareData dataset). The resolution of images from this dataset is within 1 m per pixel. The images include both different types of gas flares and images where there are no gas flares. As a rule, pictures are presented for different moments in time (usually 3-10 variants), both in color and in black and white. Images also often contain ”clouds” and ”snow cover” as natural disturbances in our classification task. There are also in small numbers simply spoiled pictures. Figures 1, 2 and 3 show examples of images from the used dataset.

Figure 1: Examples of ”clean” upstream flare images.

Moreover, it would be more correct to consider our task as a task for the presence of certain types of scenes (scene detection), which in itself, as a rule, is much more complicated than the usual classification. We only note that we do not consider

Figure 2: Examples of ”clean” downstream flare images.

Figure 3: ”Unclean” examples of flare images: downstream flare, covered with clouds (on the left) and upstream flare, covered with snow (on the right).

the approach from the point of view of scene detection in this work.

B. Image extractors and descriptors: SIFT and SURF algorithms

This section provides an overview of the various experiments we use to evaluate the performance of the algorithm for image classification. The set of raw data used in our experiments is described here. The following describes methods for transforming images and methods for obtaining feature descriptors (including SIFT, SURF, etc.). Then, preliminary results of our work on evaluating the accuracy of the algorithm for classifying images with flares are presented.

Both of these methods have an important feature: they are relatively insensitive to both scaling and rotation of images. This is very important for our task, since our images can be with different scaling and resolution, and are oriented in different ways. The scale-invariant feature transform (SIFT) [5] [6] was the first algorithm with similar properties, and the speeded up robust features (SURF) algorithm [7] improves it overall: SURF is faster and more robust than SIFT. The result of SIFT and SURF is a certain finite-dimensional vector of descriptors, and the order of the elements itself is not important in this vector of descriptors, i.e., this means that in fact it is a set. An image descriptor vector is a kind of histogram of the visual elements it contains.

After using SIFT or SURF algorithms, one can use the corresponding descriptors of pictures instead of the pictures themselves, both at the stage of training and at the stage of using the classifier.

C. Image Data Set Augmentation

The quality and quantity of datasets is very important as this affects the accuracy of machine learning algorithms, as well as it is related to overfitting and underfitting problems. We increase the size of the training data by using augmentation techniques, e.g., cropping, rotating, shifting and scaling of original images. As SIFT and SURF are invariant with respect to rotating, shifting and scaling we do not use them with SIFT and SURF algorithms but we use cropping technique. On other hand, for deep learning models, in particular, for we use all of them: including rotating, shifting, scaling as well as cropping extensively. Overall, the image augmentation helps to improve the prediction accuracy of the model. In our experiments after applying the augmentation technique, the data set (about 11000 images) is increased approximately by factor 2 (about 22000 images).

D. Classifier training process

The encoded image feature vectors from each category of images are used in the classifier training process for the SVM and Random Forrest algorithms to create a predictive function of the model. Figure 4 demonstrates the steps of the image preparation used in our image classification approach.

It is based on a more generic framework for image data transformation discussed in [10].

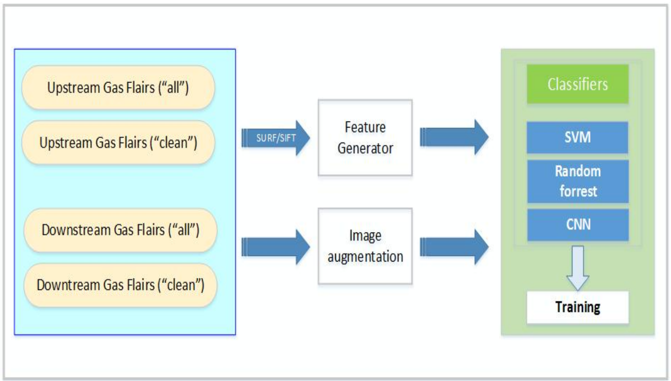

Figure 4: Stages of processing for gas flare images.

V. EXPERIMENTS WITH THE FLARE IMAGE CLASSIFICATION

In this section, we describe the experiments we use to evaluate the quality of classifiers for classifying gas flare images.

We are interested in the average accuracy measurement and the "the confusion matrix" for various experiments. For this we use different categories of images and their groups from the GasFlareData data set.

Note that the results below are *preliminary*, and the classifiers themselves are far from optimal. This is especially true for the results on the application of the deep learning (CNN classifier).

Experiment 1. Measuring classification accuracy from data ("all" or "clean") into categories (upstream and downstream) using SURF extractor for SVM [8], Random Forrest [9] and without using SURF in the case of the CNN classifier. Here a classifier (SVM, Random Forrest and CNN) is used to create a prediction model based on two image classes, including (1) "presence of a gas flare" and (2) "no gas flare". Each category of data sets (upstream and downstream) is considered separately here. In the training process, we use 70% of the entire set of images from each category for the training purposes. The rest of each image set (i.e., 30% of all images) are used during the forecast evaluation stage. Our measurements show that the achieved average forecast accuracy is about 75% (see TABLE I).

Experiment 2. Measuring classification accuracy from data ("all" or "clean") combined categories (upstream + downstream) using SURF extractor for SVM, Random Forrest and without using SURF for CNN. Here the SVM, Random Forrest and CNN classifier are used to generate a prediction model based on three image classes: ((1) "downstreanm gas flare presence", (2) "upstreanm gas flare presence" or (3) "gas flare absence"). During training we use 70% of the whole set of images (downstream + upstream) for data ("all" or "clean"). The rest of the images (i.e., 30% of all images) are included in the test data set and are used during the forecast evaluation stage. Preliminary experiments also demonstrate that the achieved average forecast accuracy for the three classes is about 70% (see TABLE II).

TABLE I. THE AVERAGE ACCURACY RESULTS FOR THE TWO CLASSES

Data	SVM	Random Forrest	CNN
Upstream "Clean"	0.81	0.79	0.75
Downstream "Clean"	0.78	0.78	0.74
Upstream "all"	0.79	0.76	0.75
Downstream "all"	0.77	0.76	0.74

TABLE II. THE AVERAGE ACCURACY RESULTS FOR THE THREE CLASSES

Data	SVM	Random Forrest	CNN
Upstream, "clean" + Downstream, "clean"	0.72	0.71	0.70
Upstream, "all + Downstream, "all"	0.70	0.70	0.68

VI. CONCLUSION AND FUTURE WORK

In this paper, a method for checking and correcting the list of high-temperature anomalies obtained on the basis of nighttime light data is proposed. The image classification of daytime satellite images is used for this purpose. It is carried out using machine learning methods.

Various machine learning methods and algorithms to classify the gas flare images from the GasFlareData set is applied. The up-to-date solutions that generally demonstrate the most acceptable classification quality in a wide variety of applications are used. The process of preparing images for classification, including special image encoding operations such as the well-known SIFT and SURF algorithms, is presented. For different variants of datasets: "clean" (including only those that do not contain clouds or snow) and "all" (including those that do not contain clouds and/or snow), the flare image classification is performed.

In experiments, the preliminary comparison of the classification quality for different machine learning methods (including SVM, random forest, Convolutional Neural Networks (CNN)) both with and without image encoding operations is carried out. Both the variants of the input data ("clean" or "all") are evaluated. The following average forecast accuracy is achieved in the experiments: about 75% for the two classes and about 70% for the three classes.

Future work concerns the improvement of the classification accuracy of the machine learning models used in the paper. We also plan to apply the other deep learning architectures to the problem of interest and compare them with the all other classifiers.

ACKNOWLEDGMENT

We thank Mikhail Zhizhin at the Colorado School of Mines, Golden, USA and Alexey Poyda at the National Research Centre "Kurchatov Institute", Moscow, Russia for fruitful discussions and collaboration in collecting samples of satellite images.

REFERENCES

[1] Earth Observation Group, Payne Institute, Colorado School of Mines, Golden, USA [Online]. Available from: https://payneinstitute.mines.edu/eog/viirs-nightfire-vnf/ [retrieved: March, 2021].

[2] C.Elvidge, M. Zhizhin, F-C. Hsu and K. Baugh, "VIIRS nightfire: Satellite pyrometry at night". Remote Sensing, 5, no. 9, 2013, pp.4423-4449.

[3] C. Elvidge, M. Zhizhin, K. Baugh, F-C. Hsu and T. Ghosh, "Methods for Global Survey of Natural Gas Flaring from Visible Infrared Imaging Radiometer Suite Data". Energies, 9(1), 14, 2016, pp.1-15.

[4] A. Matveev, A. Andreev, M. Zhizhin and A. Troussov, "Satellite Monitoring of Associated Gas Flaring Torches in Russia" (In Russian). SSRN Electronic Journal, DOI: 10.2139/ssrn.3362303, January 2019.

[5] D. G. Lowe, "Object recognition from local scale-invariant features". Proceedings of the International Conference on Computer Vision. 2. 1999, pp. 1150–1157.

[6] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints". International Journal of Computer Vision. 60 (2), 2004, pp.91–110.

[7] H. Bay, T. Tuytelaars and L. Van Gool, "Surf: Speeded up robust features". European conference on computer vision, 2006, pp.404-417.

[8] H. Drucker, C. Burges, L. Kaufman, A. J. Smola, and V. Vapnik, "Support Vector Regression Machines", in Advances in Neural Information Processing Systems 9, NIPS 1996, 1997, pp.155–161.

[9] L. Breiman, "Random Forests". Machine Learning, 45, pp.5-32, 2001.

[10] S. Loussaief and A. Abdelkrim, "Machine Learning framework for image classification". Advances in Science, Technology and Engineering Systems Journal, 3, 2018, pp.1-10.

 ISBN: 978-1-61208-850-1