A MULTIMODAL APPROACH FOR EVENT DETECTION: STUDY OF UK LOCKDOWNS IN THE YEAR 2020

Joëlle Hanna, Linus Scheibenreif, Michael Mommert and Damian Borth

AIML Lab, School of Computer Science, University of St.Gallen
firstname.lastname@unisg.ch

ABSTRACT

Satellites allow spatially precise monitoring of the Earth, but provide only limited information on events of societal impact. Subjective societal impact, however, may be quantified at a high frequency by monitoring social media data. In this work, we propose a multi-modal data fusion framework to accurately identify periods of COVID-19-related lockdown in the United Kingdom using satellite observations (NO_2 measurements from Sentinel-5P) and social media (textual content of tweets from Twitter) data. We show that the data fusion of the two modalities improves the event detection accuracy on a national level and for large cities such as London.

Index Terms— remote sensing, NLP, data fusion

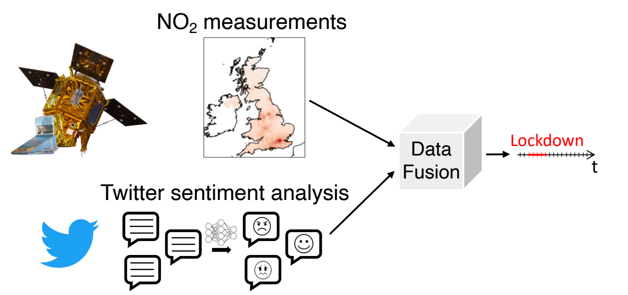

Fig. 1: Illustration of our data fusion approach, in which we combine remote sensing (NO_2 column density from Sentinel-5P) and social media (textual content of tweets from Twitter).

1. INTRODUCTION

Events with societal or environmental impact may be harder to detect with quantitative measures than others, such as natural disasters. While satellite observations provide an objective view of the Earth and its environment, subjective societal impact is much harder to quantify. Social networks provide stream of data that might be useful for this purpose.

Satellite data have the advantage of high spatial resolution (information on their exact geolocation), but unfortunately come with low temporal coverage. As for data collected on social networks, they present little or no detailed spatial information, but are abundant and have a high temporal frequency. For event detection, the usefulness of each of these modalities alone varies, but by merging their spatial and temporal resolutions, we can obtain a stronger prediction and more detailed information on the event.

In this work, we aim to automatically identify COVID-19 lockdown periods in the UK in a multimodal data fusion approach from remote sensing and social media data. This work serves as a feasibility study for automated detection of event with complex implication for people and the environment.

Our hypothesis is that during times of lockdown, human mobility and industrial activities are reduced, improving air quality, while at the same time negative sentiments rise in the population. By quantifying the concentration of Nitrogen Dioxide (NO_2), which has a short lifetime in the atmosphere and is primarily emitted from anthropogenic sources, we can use it as a suitable proxy for current industrial activity: we expect the values of NO_2 concentrations during periods of lockdown to decrease compared to the same period the previous year. At the same time, during lockdowns, we expect people to be more anxious, annoyed and pessimistic than usual. As a proxy for the sentiment among the population, we use textual content of tweets from Twitter. We therefore expect to see a higher than usual number of negative tweets during these periods. While we expect each modality to provide some information on on-going lockdowns, these signals might be hard to identify due to other events that may also affect people's sentiments. By employing data fusion, we expect to not only strengthen the prediction, but also obtain detailed information of the locations of lockdowns (spatial information from remote sensing) and their precise timeline (temporal information from social media sentiment). For the basis of this study, we chose data from the UK from 2020, as it is a English-speaking country that imposed major lockdowns on a national level.

Our contribution is two-fold: (1) we propose a transformer-based deep learning model to predict the sentiment of tweets in a multi-label setup using two independent datasets, and (2) we present a unique multimodal data fusion method that makes use of textual and remotely sensed information in order to boost the event detection performance.

2. RELATED WORK

Regarding the impact of restrictions due to COVID-19 on air pollution, [1] observed a decrease in the concentration of NO_2 and PM2.5 affected by the evolution of traffic during confinements. The same conclusion was drawn by [2], where they found Lisbon and Madrid to have historic low of NO_2 levels over the past 2 years for most of April 2020. In parallel, many works have studied sentiment analysis on COVID-19-related tweets, such as [3], who performed binary sentiment classification (positive/negative) and [4] who noticed a drop in the number of positive tweets during periods of lockdowns.

To the best of our knowledge, this is the first study effectively taking advantage of both textual social media and remote sensing information for this purpose.

3. DATASETS

3.1. Nitrogen Dioxide Column Densities

We use tropospheric column density measurements of Nitrogen Dioxide from the Sentinel-5P satellite of the European Space Agency's Copernicus program [5]. Sentinel-5P is a polar orbiting satellite carrying the TROPOMI instrument for differential optical absorption spectroscopy. This allows Sentinel-5P to measure the concentration of trace-gases like NO_2, Ozone or Methane across the Earth's atmosphere. With a swath width of ∼2600km and a resolution of 7×3.5km, the satellite provides trace-gas measurements with daily global coverage (weather permitting). We obtain 1,362 Sentinel-5P NO_2 Level-2 products over the UK in the 2019-2020 timespan. The data is pre-processed by mapping all measurements onto a common rectangular grid with resolution 0.05×0.05°(∼5×5km) and averaging overlapping measurements to daily frequency. Additionally, all measurements with a quality value below 75 are dropped, following ESA recommendations.

3.2. Textual Data from Twitter

We train and validate our sentiment analysis model on the textual content from 5000 labelled COVID-19-related English tweets [6]. Each tweet has one or more of the following 11 labels: *Optimistic*, *Thankful*, *Empathetic*, *Pessimistic*, *Anxious*, *Sad*, *Annoyed*, *Denial*, *Surprise*, *Official report*, *Joking*. We split the data in two datasets: 80% for training and 20% for validation.

We predict sentiment based on this model and a large-scale COVID-19 Twitter dataset taken from [7]. It contains millions of tweets from around the world written in different languages, but without detailed location information. We only use the textual content of tweets and their timestamps. We filter the dataset by country, keeping only tweets from the UK written in English and randomly sampling 11,000 tweets per month for the year 2020.

4. METHODS

Our approach consists of the following steps:

1. Retrieval and aggregation of NO_2 column density data for the UK during the years 2019 and 2020.
2. Sentiment analysis of COVID-19-related English tweets from the UK during the year 2020.
3. Data fusion of both modalities to detect the exact timeline of UK lockdowns in 2020.

4.1. Nitrogen Dioxide Column Densities

We focus on the country as a whole, as well as two specific regions of the UK with different levels of industrial activity and population density: London as an example of a heavily industrialized and densely populated region, and the Scottish Highlands as an example of a rural region. For each of these locations, we calculate a moving average over $N = 21$ days in order to smooth the otherwise noisy data. We are interested in identifying the periods of 2020 when the concentration of NO_2 is lower than at the same days in 2019. In order to quantify a statistical confidence on the concentration drop during a given period, we calculate for each such period $\Delta t = t_2 - t_1$ a SNR value defined as follows:

$$\overline{\Delta \mathrm{NO_2}} = (A_{2019} - A_{2020})|_{t_1}^{t_2} / \Delta t$$

$$\bar{\sigma} = \sqrt{\left(\frac{\sigma_{2019}}{\sqrt{N-1}}\right)^2 + \left(\frac{\sigma_{2020}}{\sqrt{N-1}}\right)^2}$$

$$\mathrm{SNR} = \overline{\Delta \mathrm{NO_2}} / \bar{\sigma}$$

where for the year x, A_x is the area under the curve and σ_x is the standard deviation over the N days rolling window.
We will only consider periods where $\mathrm{SNR} > 2.0$, for which NO_2 concentrations are significantly lower in 2020 compared to 2019 as lockdown candidates.

4.2. Twitter Sentiment Analysis

In order to classify the unlabelled tweets, we use a pre-trained transformer BERT [8] model, uncased, with 12 hidden layers and a hidden size of 768, as a feature extractor for the tweets. We append to the transformer model a multi-layer bi-directional Gated Recurrent Unit (GRU) to efficiently capture the semantics between different words. We freeze the transformer's layers and only train the GRUs using a binary cross entropy loss, for multilabel classification. To evaluate the model, we use the macro f1-score, a measure that treats all classes equally, and the micro f1-score that aggregates the contributions of all classes before averaging them. Our model reaches a macro f1-score of 0.46 and a micro f1-score of 0.52.

We use our trained model to predict the sentiment of tweets from the UK during 2020. We then calculate the distribution of the different sentiments on a daily basis. To follow

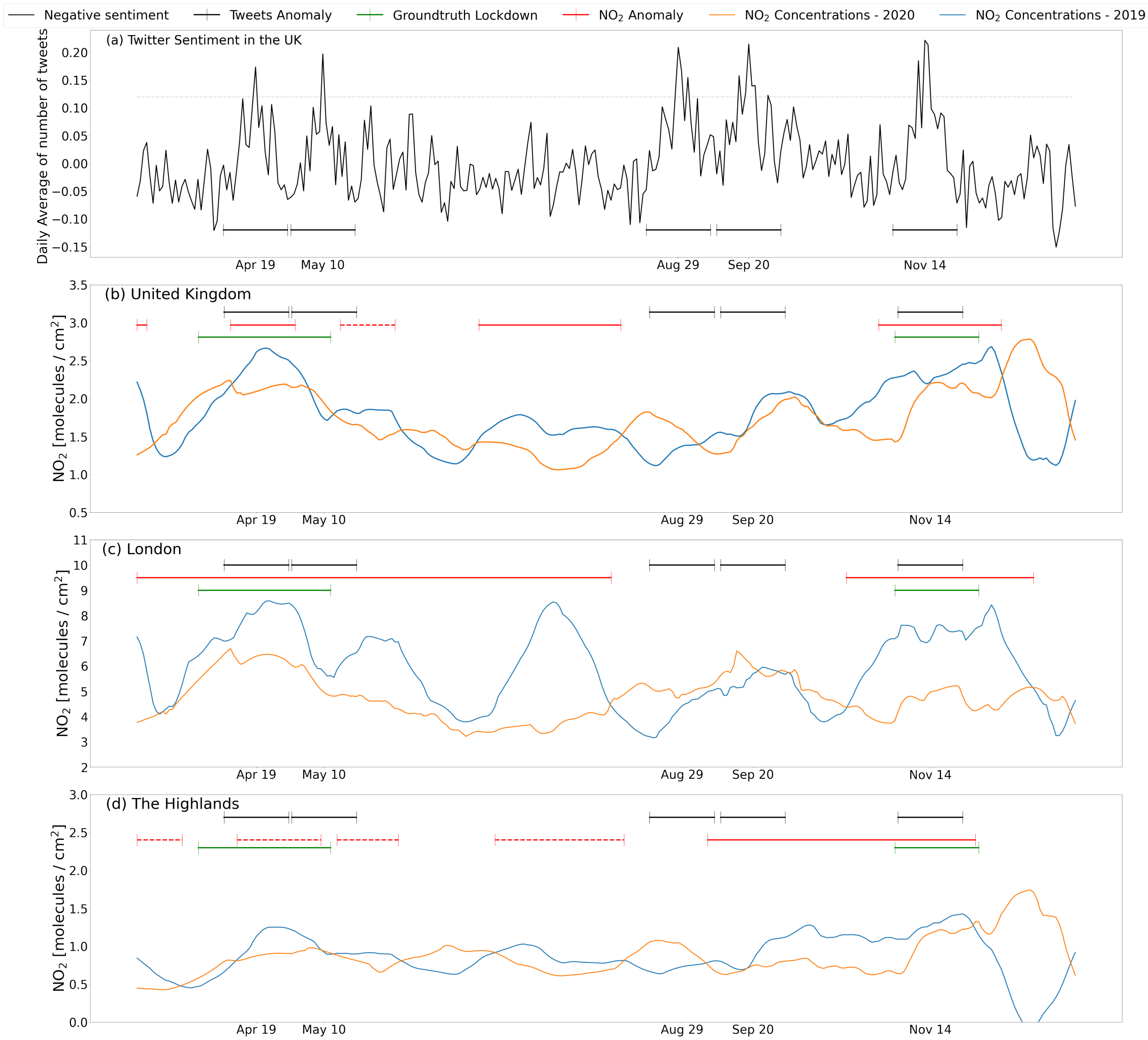

Fig. 2: Visualisation of our results. **(a)**: the average daily number of negative tweets for the entire UK; black bars indicate periods of extreme negative sentiments during with negative sentiments exceeds 3 times the MAD (see Section 4.2). **(b, c, d)**: column densities of NO_2 in 2019 and 2020, the red bars indicating periods when the 2020 curve is lower than that of 2019, dashed red bars representing periods where $\mathrm{SNR} < 2$, and green bars indicating ground-truth lockdown periods on a national level, in London and in the Highlands, respectively.

sentiment evolution in the UK during the pandemic, we treat each sentiment as a time series and use the median absolute deviation (MAD) to detect unusual amounts of negative sentiment (we consider labels annoyed/anxious/pessimistic as negative sentiments for this part), indicating emotional stress in the population. We define a value x_i as an outlier if $|x_i - \tilde{x}|/\mathrm{MAD} \geq T$, where $\tilde{x}$ is the median of values over the year and T is a chosen threshold, which we set to 3. We consider such outliers as lockdown candidates.

5. EXPERIMENTS AND RESULTS

To illustrate our method, we first display in Figure 2 (a) the time series of Twitter sentiment in the UK, while panels (b), (c) and (d) show localized NO_2 concentrations for the United Kingdom (at a national level), London and The Scottish Highlands, respectively. We only consider periods in which the 2020 NO_2 deficiency has a $\mathrm{SNR} > 2$, based on the definition in Section 4.1. From [9] we only consider "StayHomeOrder" measures for the year 2020 (March 24 – May 9, and November 5 – December 1), which we adopt as ground-truth periods of lockdowns.

We consider NO_2 deficiencies in 2020 and periods of negative sentiment as indicators for lockdowns. While Twitter sentiment is measured on a national level, only, NO_2 column densities are aggregated in this work on a national scale, for London and the Highlands. We expect the effects of lockdowns on our modalities to be strongest in London (densely populated and highly industrialized) and weakest in the Highlands (extremely rural area). This is based on the expected impact of lockdowns on industrial activities and traffic, thus affecting NO_2 concentration levels. Based on the Twitter sentiment analysis, some negative anomalies coincide with actual lockdowns, but not all. Based on the NO_2 analysis, there is some overlap in the lockdown periods for the UK on a national level, for London and for the Highlands. By combining the two modalities, we are able to detect the two periods of lockdown nationally and London, and the second period of lockdown in the Highlands.

In order to quantify the overlap between the detected lockdowns and the actual ones, we compute the intersection over union (IoU), by dividing the number of days correctly identified as lockdown by the union of the number of days detected as lockdown with the groundtruth. For each selected location, we compute the IoU for each source alone, and for their fusion (see Table 1).

Table 1: IoU of predicted and actual periods of lockdown

	Twitter Source	NO_2 Source	Both Sources
UK	0.54	0.33	0.57
London	0.54	0.36	0.75
The Highlands	0.54	0.23	0.37

6. DISCUSSION

Table 1 shows that the combination of the two modalities works well for the UK at a national level and for London. For the Highlands, the Twitter source appears to be more accurate than the NO_2 concentrations and the two sources combined. The modalities have different performances on the detection precision (IoU) because, as expected, the population density and therefore the traffic varies between selected regions, which is an important factor that highly affects the concentration of NO_2. Note that the Twitter signal is the same for all regions since we do not have detailed information on the exact location of tweets.

7. CONCLUSION

In this study, we propose a multimodal data fusion framework that combines two independent sources (textual content of tweets and NO_2 column densities) to detect actual periods of lockdown in the UK, taking advantage of spatial information from NO_2 concentrations, and temporal information from tweets. We focus on the country as a whole, as well as two specific regions: densely populated London and the more rural Highlands. This method is able to boost the prediction of periods of lockdown on a national level and for London specifically, but is not as effective for the Highlands.

8. REFERENCES

[1] Jan P. Musiał Patryk T. Grzybowski, Krzysztof M. Markowicz, "Impact of restriction due to covid-19 on air pollution in poland in spring 2020," 2021.

[2] Mariana Ávila Pedro Silva, Márcia Gonçalves and Emanuel Castanho, "Air quality improvement during covid-19 pandemic: Study of land and maritime pollution," 2021.

[3] Arunava Kumar Chakraborty, Sourav Das, and Anup Kumar Kolya, "Sentiment analysis of covid-19 tweets using evolutionary classification-based lstm model," *ArXiv*, vol. abs/2106.06910, 2021.

[4] Ishaani Priyadarshini et al., "A study on the sentiments and psychology of twitter users during covid-19 lockdown period," *Multimedia Tools and Applications*, pp. 1 – 23, 2021.

[5] JP Veefkind et al., "Tropomi on the esa sentinel-5 precursor: A gmes mission for global observations of the atmospheric composition for climate, air quality and ozone layer applications," *Remote sensing of environment*, vol. 120, pp. 70–83, 2012.

[6] "Sentiment analysis of covid-19 related tweets," *url: https://www.kaggle.com/c/sentiment-analysis-of-covid-19-related-tweets*.

[7] Juan M. Banda et al., "A large-scale COVID-19 Twitter chatter dataset for open scientific research - an international collaboration," Nov. 2021.

[8] Jacob Devlin et al., "Bert: Pre-training of deep bidirectional transformers for language understanding," *ArXiv*, vol. abs/1810.04805, 2019.

[9] European Centre for Disease Prevention and Control, "Data on country response measures to covid-19," 2021.