9th International Workshop on Climate Informatics
October 02-04, 2019
Hosted by École Normale Supérieure, Paris, France

U-NET FOR LEARNING AND INFERENCE OF DENSE REPRESENTATION OF MULTIPLE AIR POLLUTANTS FROM SATELLITE IMAGERY

Jacquelyn Shelton[1], Przemyslaw Polewski[2], Wei Yao[1]

***Abstract*—Air pollution is an important topic on countless fronts and is an active area of research. The goal of this work is to provide a machine learning model for learning and inference of pollution concentrations and air quality measures, namely Particulate Matter 2.5, $NO3$, Nitrate Pollution, and $NH4$, Atmospheric Ammonium, with high granularity by using easily obtainable satellite imagery data. In order to achieve this, we propose the fully convolutional network U-net that, unlike previous work, can predict these pollutant values at a pixel-level high-resolution instead of being able only to predict a single value for an entire geographical region. We demonstrate that this approach can reconstruct the considered pollutant concentrations on ground-truth data and can predict the concentrations and their spatial structure reasonably well, even for data that the network has temporally not yet seen. Finally, we illustrate that the model's pollutant predictions can offer valuable insights into the current *COVID-19* pandemic.**

I. INTRODUCTION

In 2020 UNEP, UN-Habitat and IQAir announced the release of the largest air quality data platform to date. While this inititive's announcement focuses on the health aspects of pollution the value of this data set are not to be understated. In 2019 with the outbreak of SARS-CoV-2, commonly referred to as *COVID-19* and the discovery of SARS-CoV-2 RNA on air pollution the value in this data and the potential to not just measure local but also gain insight into the movement of air pollution became of interest. While SARS-CoV-2 is primarily related to $PM_{2.5}$ and PM_{10}, Particulate Matter pollution of 2.5 and 10 micrometers and smaller in diameter respectively, other interests, health, environmental and agricultural, are also interested in other pollution and air quality measures, in particular $NO3$, Nitrate Pollution, and $NH4$, Atmospheric Ammonium.

Corresponding author: J Shelton, jacquelyn.ann.shelton@gmail.com, [1]Hong Kong Polytechnic University, [2]TomTom Location Technology Germany GmbH

Traditionally, $PM_{2.5}$ concentration data can be obtained from ground sensors and measurement stations. However, the spatial resolution of these measurements is greatly limited by the sparsity of sensor networks. Thus, detailed pollution maps have been developed that integrate heterogeneous data sources (see e.g. [2]). Namely, they proposed a $PM_{2.5}$ estimation of fusing aerosol optical depth information computed from satellite imagery with a chemical transport model and regional ground-based observations. With this, they published a database of estimated monthly pollutant concentrations over several countries. Although this historical data helped establish this causality relationship, up-to-date/live pollution information is still needed to monitor pollutants such as $PM_{2.5}$, NO_3 , and NH_4 .

Recently, there has been a growing interest in retrieving spatio-temporal air quality products based on high resolution optical satellite missions such as *Landsat 8* [3], utilizing top-of-atmosphere image reflectances as a basis for directly predicting pollution parameters [4], [5].

The primary goal of this work is to provide a machine learning model capable of estimating the pollutants' $PM_{2.5}$, NO_3 , and NH_4 concentrations with high granularity comparable to the data gathered by [2] (details in Sec.III) using plentiful, publicly available satellite imagery. Unlike previous work, the fully convolutional neural network U-net we employ can predict these pollutant values at a pixel-level high-resolution instead of predicting only a single value for an entire region.

The goal of the present work is to model air pollution concentrations $PM_{2.5}$ at higher resolution than has been done previously using readily available satellite imagery. The paper is organized as follows: Sec. II introduces the proposed U-net model, Sec. III presents the data, Sec. IV describes the experiments and results, and finally Sec. V provides a summary of the work and outlook.

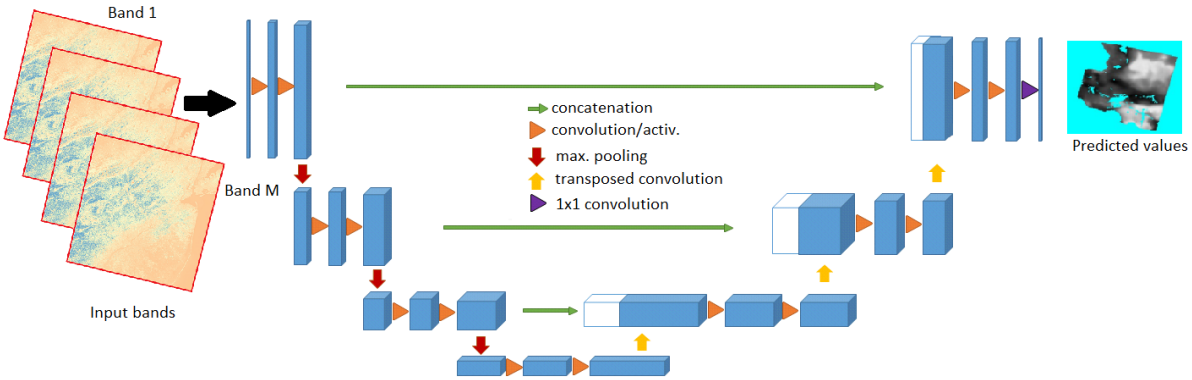

Fig. 1. Network structure for predicting pollutant ($PM_{2.5}$, NO_3, NH_4) concentrations from multispectral imagery. The network follows the classic U-Net [1] architecture.

II. U-NET MODEL FOR POLLUTANT AGENTS

We build upon the well known U-net architecture [1] to predict dense (per-pixel) pollutant concentrations from multispectral satellite imagery. Since its introduction in 2015, the U-net has been successfully applied to various semantic segmentation tasks, E.g. in medical imaging [6] and astronomy [7]. As a fully convolutional network (and hence devoid of fully connected layers), the U-net can be applied to an image of arbitrary size in a sliding window/tile based fashion. The U-net consists of two symmetrical parts. The encoder path downsamples the original image into meaningful features by means of convolutional filters and pooling operations, whereas the upsampling path aims at decoding these features into a full-sized output map using transposed convolution operations, driven by an appropriate loss function. Moreover, upsampling layers are augmented with feature maps from the downsampling path at the corresponding resolution, to provide more context information. The original U-net was meant for classification and featured a softmax layer after the top-level upsampling layer's output, resulting in per-class posterior probabilities meant to be optimized with using a cross-entropy type objective with discrete ground-truth labels. However, the U-net has also been used directly for dense regression, by removing the softmax layer and optimizing the squared difference between the upsampled output and a continuous target variable [8], [9]. In this work, we adopt a similar approach. Let $S = \{(x_i, y_i)\}, 1 \leq i \leq N$ denote a set of predictor variable vectors x and matching continuous response variables y (i.e. the target pollutant concentration), while $f(x|\theta)$ indicates the prediction function (i.e. the U-net) parameterized by θ. The optimization objective can be formulated as:

$$\underset{\theta}{\operatorname{argmin}} \sum_i w_i |y_i - f(x_i|\theta)|^2 \tag{1}$$

It is well known that least squares regression estimates the conditional mean of the response variable given the predictors, and is therefore sensitive to outliers. To alleviate this, we eliminate outliers from the training set by discarding pairs (x_i, y_i) such that y_i is below the 0.01-th or above the 0.99-th quantile of its distribution. This is achieved by setting the weights w_i corresponding to such pixels within the training images to 0. Otherwise, the respective weights are set to unity.

III. DATA

In order to address this problem, we use satellite imagery as a source of predictor variables, and approximate pollutant concentrations published by [2] for the time period of 2000-2017 in the role of ground-truth.

Satellite data. This study was based on Landsat 8 satellite imagery published by the United States Geological Survey [3]. Landsat 8 is the latest installment in a series of Earth observation missions dating back to 1972 (Landsat 1), which provide planet-wide imagery from multispectral, spaceborne optical sensors. Landsat uses its own coordinate system [10], which assigns coordinates to an image frame based on nominal satellite orbital tracks and the frame's latitudinal center line, resulting in a 2-tuple of (path,row) coordinates. The ground coverage cycle length is 16 days, which means that new data for an image frame is published approximately twice per month. Landsat 8 images contain a total of 11 spectral bands, ranging in wavelengths from 0.435 um (band 1 - coastal/aerosol) to 12.51 um (band 11 - thermal infrared). The spatial resolution (i.e. ground

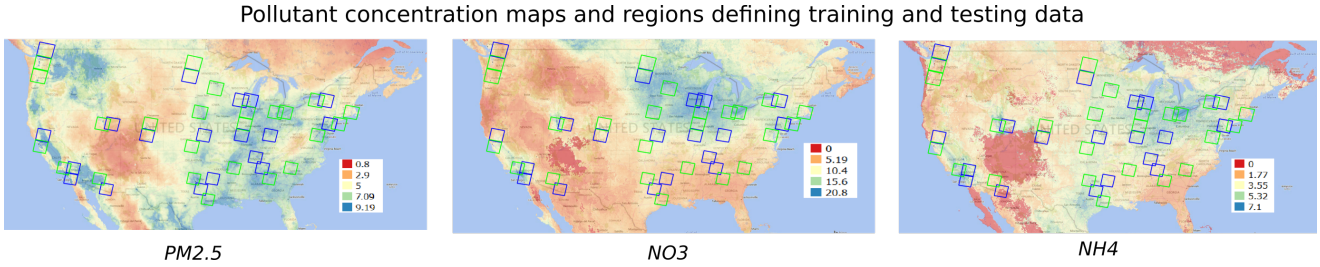

Fig. 2. Locations of US cities chosen to reasonably represent the diverse US geography for our study. Rectangles indicate the bounding boxes of the Landsat images' locations. Green boxes define regions used for training and validation, whereas blue boxes represent regions used exclusively for testing. The background color maps represents average concentrations, in micrograms per cubic meter, of $PM_{2.5}$, NO_3, and NH_4 in 2017. The corresponding distributions of the testing data for these pollutants are shown in Fig. 5

sampling distance per pixel) of the images ranges from 15 m (panchromatic band 8) through 30 m (bands 1-7, 9), up to 100 m (bands 10-11). We used the USGS Earth Explorer [11] as well as the Landsat dataset mirror on Amazon AWS [12] to download the imagery.

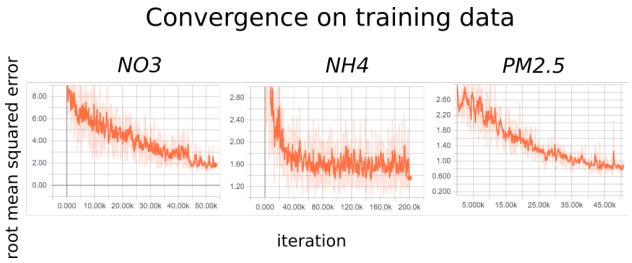

Fig. 3. Convergence of training process measured by root mean squared error loss function on training set respectively for NO_3, NH_4, and $PM_{2.5}$.

Pollution data. We downloaded monthly total $PM_{2.5}$,NO_3, and NH_4 concentration maps for North America from the FTP server of Dalhousie University, Canada, published by Aaron van Donkelaar and collaborators [13]. These maps are made available at a resolution of 0.01 degrees per pixel and feature a standard WGS84 coordinate reference system. The pollutant particulate concentration was obtained via a hybrid approach which simulates geophysical relationships between aerosol optical depth (estimated from multi-source satellite imagery) and $PM_{2.5}$ concentration based on the GEOS-Chem chemical transport model, and fuses the estimates with ground measurements obtained from the US Environmental Protection Agency's Air Quality System as well as Canada's National Air Pollution Surveillance [2]. Moreover, estimates of various component pollutants contributing to the total $PM_{2.5}$ are derived from simulated relative composition. The authors report good long term agreement of predicted pollutant agent concentrations with ground-truth validation sites (R^2 coefficients of 0.59-0.90). We considered only the period after the Landsat 8 mission became active, resulting in a time interval from March 2013 to December 2017.

Preprocessing. The Landsat imagery was first reprojected to the WGS84 coordinate system, and downsampled to the ground-truth resolution of 0.01 degrees (the panchromatic band was dropped). To match the temporal resolution of the ground-truth, we computed per-band average images grouped by month of acquisition. Next, we derived a per-pixel mask of regions within the image covered by cirrus clouds and hence not suitable for analysis. This was done on the basis of the Landsat band 9 (1.360-1.390 um), which was designed to detect high-altitude clouds. We used the estimated cloud cover percentage [14], available as part of image metadata, to remove the pixels locations having reflectance values in band 9 within the corresponding top quantile. Finally, we combined the cloud cover mask with the data availability mask from the ground-truth $PM_{2.5}$ maps, representing other features that do not have available data E.g. missing data over large bodies of water. Also, pixels corresponding to the top and bottom 1% of ground-truth values were masked out as outliers (see Fig. 5). All input Landsat imagery bands were normalized to the interval [0;1] individually per band.

IV. NUMERICAL EXPERIMENTS

We selected an initial number $N = 114$ Landsat images, spanning March 2013 to December 2017, of 24 major cities from representative regions of the United States (see Fig. 2, green boxes) and preprocessed them as described in Sec. III. Next, we defined our ground-truth data as these images paired with corresponding pollutant images from which we use an 80:20% random split to create training/testing set of 91:23 images. A U-net model based on adapting the implementation by [7] was trained separately for each pollutant ($PM_{2.5}$, NO_3, NH_4) using an ADAM optimized over 500 epochs with

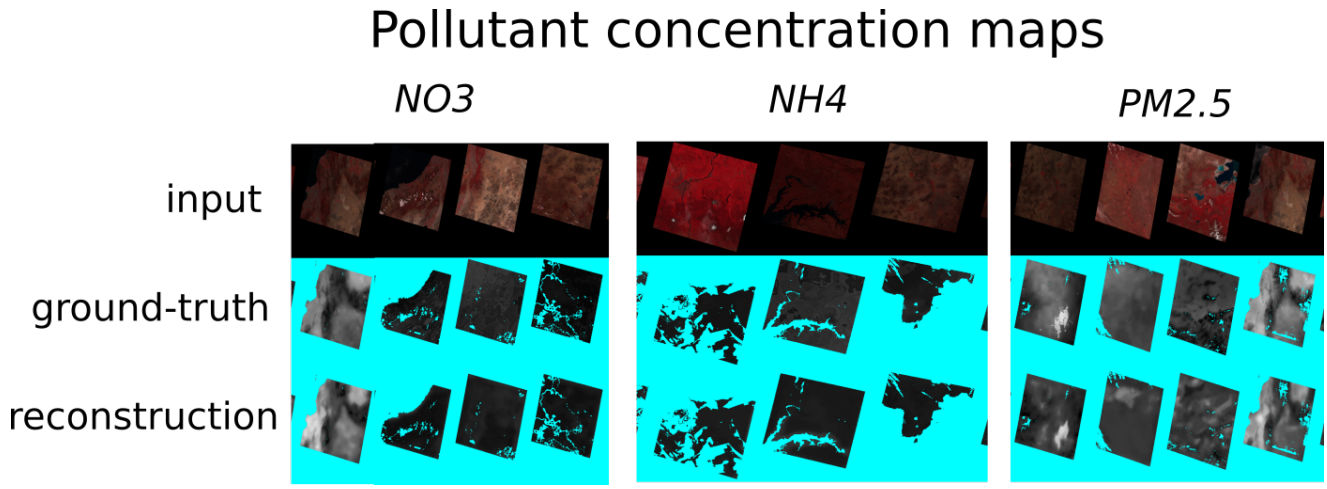

Fig. 4. Examples of pollutant concentrations predicted by the U-Net for new temporal data at locations used during training, respectively for NO_3, NH_4, and $PM_{2.5}$.

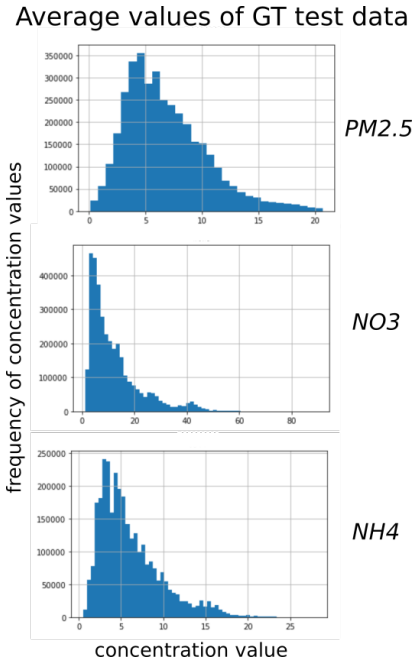

Fig. 5. Distributions of the average pollutant concentration values of ground-truth training data. These are calculated from values within the green squares in Fig. 2 for $PM_{2.5}$, NO_3, and NH_4 in 2017.

minibatch size of 15 and 100 internal iterations, with the following parameterization: dropout ratio was 0.5, learning rate 0.00005.

1: Sanity check with ground-truth.

At the end of training, the networks converged to root mean square error values of 1.24, 2.10, and 1.64 $\mu g/m^3$ respectively for $PM_{2.5}$, NO_3, and NH_4 (see Fig. 3). The corresponsing mean absolute errors of the predictions were 0.81, 1.44, and 1.14 $\mu g/m^3$. Considering the inter-quantile ranges between the 0.01-th and 0.99-th quantiles for the ground truth distribution of the three pollutants (see Fig.5), the mean absolute deviations constituted 4%, 3%, and 7% of the respective pollutant concentration ranges.

2: Generalizability to temporally novel data.

To assess the ability of our model to generalize in the temporal domain, we utilized 23 previously unseen images that overlapped with the training set spatially but not temporally. The test set contained images from 15 of the 24 cities. The mean absolute deviations for predictions on this test set were 2.23, 4.75, and 2.24 $\mu g/m^3$ respectively for $PM_{2.5}$, NO_3, and NH_4. This corresponded to 11.3%, 11.1%, and 14.1% of the inter-quantile distance. See Fig. 4 for an example of predicted pollutant concentration maps on previously unseen data.

3: Cities with similar pollution profile.

We selected 24 new images at 20 additional locations within the United States, showing visually similar distributions of groundtruth pollutant concentrations within the Landsat image frames to the original training cities, in order to evaluate the performance of predicting the concentrations at locations unseen during training. The obtained mean absolute error was 2.46, 4.07, 2.27 $\mu g/m^3$ respectively for $PM_{2.5}$, NO_3, and NH_4, yielding percentages of the inter-quantile ranges of 12.4%, 9.5%, and 14.2%. This shows that the network was able to generalize equally well in the spatial and temporal domains.

4: Predicting pollutant $PM_{2.5}$ for *COVID-19*.

Following the outbreak of *COVID-19*, there has been a great deal of research into its understanding. Importants results [15], [16], [17] suggest a two factor vulnerability to SARS-CoV-2 caused by increased particulate matter in the air: On one hand it increases the liklihood of having a more sever reaction to infection, and the other it serves a transmission vector. Furthermore, detailed pollution maps have been developed that integrate heterogeneous data sources (see e.g. [2]) which were

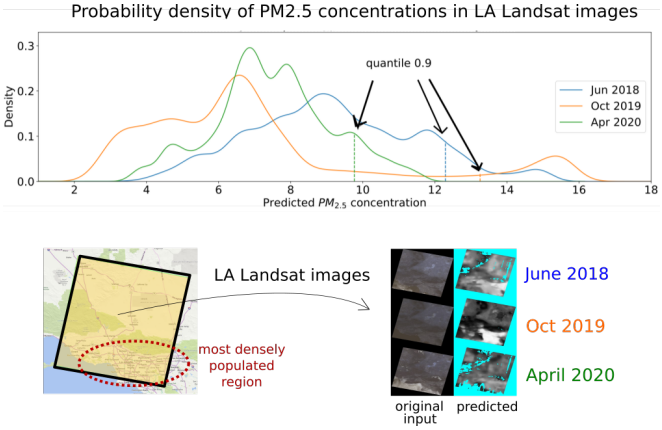

Fig. 6. Density of predicted $PM_{2.5}$ concentrations in Los Angeles for time periods before and during the COVID-19 outbreak (top) and corresponding Landsat images with their predicted $PM_{2.5}$ pollution maps (bottom). As expected, we see a significant reduction in the 0.9-quantile of the predicted $PM_{2.5}$ pollution distributions between April 2020 (immediately after the lock-down began) and the previous years. The predicted $PM_{2.5}$ maps match this trend, showing relatively less pollution in 2020 (darker pixels). Interestingly, yet unsurprisingly, the most densely populated regions of LA (encircled in red on the map) are predicted to have higher $PM_{2.5}$ concentrations for all of the years, as shown in the predicted $PM_{2.5}$ pollution maps.

used by [16] directly linking $PM_{2.5}$ concentrations to *COVID-19* mortality rates. Inferring air pollutants with reasonable accuracy, which we have demonstrated in Experiments 1-3 allows us to – essentially for free – highlight more dangerous regions wherein one is more likely to contract the virus.

The goal of the fourth experiment is to verify that our model can predict expected $PM_{2.5}$ concentration trends at time points before and after the government mandated lock-down in March 2020 intended to hinder the spread of SARS-CoV-2 (which consequently e.g. drastically reduced industrial emissions). We applied the U-net learned in Exp. 3 to satellite images of Los Angeles, CA from 2018, 2019, and early 2020. The results, shown in Fig. 6, illustrate that the U-net was able to learn $PM_{2.5}$ concentrations consistent with world events at the time. Namely, the pollution is inferred to be notably higher before the lock-down than after the lock-down, shown by the considerable shift of the predicted $PM_{2.5}$ distribution's 0.9 quantile from 13.2 in October 2019 to 9.7 in April 2020. Additionally, our approach was able to successfully isolate regions in LA that are the most densely populated by predicting higher pollutant $PM_{2.5}$ values. This suggests that our approach can generalize to new data and reliably predict pollutant $PM_{2.5}$ concentrations and their spatial structure, which informs on the presence and lethality of SARS-CoV-2.

V. DISCUSSION

The capability of measuring, monitoring, and predicting pollutant concentrations and their behavior is of crucial importance. Currently, the database established by [2] containing estimates of monthly pollutant concentrations over several countries is among the best sources for this information.

The goal of the present work was to learn and infer air pollutants, in particular $PM_{2.5}$, particulate matter 2.5, NO_3, Nitrate Pollution, and NH_4, Atmospheric Ammonium, at higher temporal and spatial resolution than has been done previously and to do so using readily available satellite imagery. For this, we proposed the use of a fully convolutional neural network known as a U-net that can, unlike previous work, predict these pollutant values at a pixel-level high-resolution instead of only predicting a single value for an entire region.

Additionally, we have demonstrated an important practical application of our approach. SARS-CoV-2 lethality has been concretely linked to the concentration of pollutant particulate matter, where a slight increase in $PM_{2.5}$ can drastically enhance the morbidity rate. Reliable predictions of the spatial structure and concentrations of $PM_{2.5}$ can help identify regions where SARS-CoV-2 may be particularly transmissible and lethal can provide critical advice to implementable public health strategies.

In future work, we plan to include other types of pollutants in training the network. Furthermore, we intend to explore simultaneously learning multiple pollutant concentrations using a form of mult-task learning [18].

ACKNOWLEDGMENTS

We would like to thank USGS for providing sattelite imagery and Aaron Van Donkelaar, Randall Martin, and team for providing their data estimates of PM2.5 exposure. Additionally, we thank Drew Lipman for his inspirational discussions.

REFERENCES

[1] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional Networks for Biomedical Image Segmentation," *CoRR*, vol. abs/1505.04597, 2015.

[2] A. van Donkelaar, R. V. Martin, C. Li, and R. T. Burnett, "Regional Estimates of Chemical Composition of Fine Particulate Matter Using a Combined Geoscience-Statistical Method with Information from Satellites, Models, and Monitors," *Environmental Science & Technology*, vol. 53, no. 5, pp. 2595–2611, 2019.

[3] "Landsat—Earth observation satellites," tech. rep., Reston, VA, 2016. Report.

[4] Q. Yang, Q. Yuan, L. Yue, T. Li, H. Shen, and L. Zhang, "Mapping pm2.5 concentration at a sub-km level resolution: A dual-scale retrieval approach," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 165, pp. 140 – 151, 2020.

[5] H. Shen, T. Li, Q. Yuan, and L. Zhang, "Estimating regional ground-level pm2.5 directly from satellite top-of-atmosphere reflectance using deep belief networks," *Journal of Geophysical Research: Atmospheres*, vol. 123, no. 24, pp. 13,875–13,886, 2018.

[6] H. Dong, G. Yang, F. Liu, Y. Mo, and Y. Guo, "Automatic Brain Tumor Detection and Segmentation Using U-Net Based Fully Convolutional Networks," in *Medical Image Understanding and Analysis* (M. Valdés Hernández and V. González-Castro, eds.), (Cham), pp. 506–517, Springer International Publishing, 2017.

[7] J. Akeret, C. Chang, A. Lucchi, and A. Refregier, "Radio frequency interference mitigation using deep convolutional neural networks," *Astronomy and Computing*, vol. 18, pp. 35–39, 2017.

[8] W. Yao, Z. Zeng, C. Lian, and H. Tang, "Pixel-wise regression using U-Net and its application on pansharpening," *Neurocomputing*, vol. 312, pp. 364 – 371, 2018.

[9] C. Payer, D. Štern, H. Bischof, and M. Urschler, "Regressing Heatmaps for Multiple Landmark Localization using CNNs," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016* (S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, eds.), (Cham), pp. 230–238, Springer International Publishing, 2016.

[10] "Landsat World Reference System Description." https://landsat.gsfc.nasa.gov/the-worldwide-reference-system/. Accessed: 2020-06-06.

[11] "USGS Earth Explorer." https://earthexplorer.usgs.gov/. Accessed: 2020-06-06.

[12] "Amazon AWS Landsat dataset." https://registry.opendata.aws/landsat-8/. Accessed: 2020-06-06.

[13] "PM2.5 concentrations published by Aaron van Donkelaar." ftp://stetson.phys.dal.ca/Aaron/V4NA02.MAPLE/ASCII/Monthly/PM25/. Accessed: 2020-06-06.

[14] S. Foga, P. Scaramuzza, S. Guo, Z. Zhu, R. Dilley, T. Beckman, G. Schmidt, J. Dwyer, M. Hughes, and B. Laue, "Cloud detection algorithm comparison and validation for operational landsat data products," *Remote Sensing of Environment*, vol. 194, pp. 379–390, 2017.

[15] E. Conticini, B. Frediani, and D. Caro, "Can atmospheric pollution be considered a co-factor in extremely high level of SARS-CoV-2 lethality in Northern Italy?," *Environ. Pollut.*, vol. 261, p. 114465, Jun 2020.

[16] X. Wu, R. C. Nethery, B. M. Sabath, D. Braun, and F. Dominici, "Exposure to air pollution and COVID-19 mortality in the United States: A nationwide cross-sectional study," *medRxiv*, 2020.

[17] M. Coccia, "Two mechanisms for accelerated diffusion of COVID-19 outbreaks in regions with high intensity of population and polluting industrialization: the air pollution-to-human and human-to-human transmission dynamics," *medRxiv*, 2020.

[18] S. Ruder, "An overview of multi-task learning in deep neural networks," *CoRR*, vol. abs/1706.05098, 2017.