Use Visual Features From Surrounding Scenes to Improve Personal Air Quality Data Prediction Performance

Trung-Quan Nguyen[1], Dang-Hieu Nguyen[2], Loc Tai Tan Nguyen[3]
[1, 2, 3] University of Information Technology, Ho Chi Minh City, Vietnam
[1, 2, 3] Vietnam National University, Ho Chi Minh City, Vietnam
quannt.13@grad.uit.edu.vn,hieund.12@grad.uit.edu.vn,locntt.12@grad.uit.edu.vn

ABSTRACT

In this paper, we propose a method to predict the personal air quality index in an area by using the combination of the levels of the following pollutants: PM2.5, NO2, and O3, measured from the nearby weather stations of that area, and the photos of surrounding scenes taken at that area. Our approach uses the Inverse Distance Weighted (IDW) technique to estimate the missing air pollutant levels and then use regression to integrate visual features from taken photos to optimize the predicted values. After that, we can use those values to calculate the Air Quality Index (AQI). The results show that the proposed method may not improve the performance of the prediction in some cases.

1 INTRODUCTION

The need to know the personal air pollution data is vital because it is better to provide each individual with regional air quality data, which seems to be more accurate than the global data measured from far away weather stations. The problem is that the performance of personal air quality prediction mainly interpolated from public weather data is not good. This paper reports our solution to tackle this challenge by finding out whether pictures of places can improve the prediction results. To know more about this challenge and the dataset that we will use, you can refer to the overview paper of MediaEval 2020 - Insight for Wellbeing: Multimodal personal health lifelog data analysis [1].

2 RELATED WORK

The experiment on using surrounding images to predict the air quality has been conducted in several projects. For instance, analyzing the sky images [4] and integrating visual features [5] into the prediction model to predict the air quality rank are the most significant projects. Those two projects used neural network models to perform air quality rank prediction, which is a categorical variable. Unlike them, this paper will use the IDW method and the regression model to predict the numerical values of these air pollutants levels: PM2.5, NO2, and O3.

3 APPROACH

Because of the time limitation, we have to propose a method that does not require an incredible training time. At first, we will use the pure form of IDW technique [3] to predict pollutant levels. Then, the multiple linear regression will help us to combine these predicted values with an additional visual feature to produce new pollutant levels.

MediaEval'20, December 14-15 2020, Online

Table 1: Labels of image *GH030011_005250.jpg*

Description	Confidence Score
Waterway	0.841386
Sky	0.8220895
Morning	0.8061569
Tree	0.7989127
Road surface	0.7560913
Road	0.7502666
River	0.73584783
Walkway	0.73147374
Architecture	0.7263346
Thoroughfare	0.712235

3.1 Extract visual features

We use Google Cloud's Vision API to extract information about entities in images. Each image will have a maximum of 10 labels that have the highest confidence score. For example, Table 1 shows labels of the image *GH030011_005250.jpg*. We create a boolean feature from those labels to define whether that location is an open space or not. It means that if an image has one of the labels in Table 2, it will be a picture of an open space area, and therefore, the *is_open_space* feature has the value of 1 and vice versa. We believe that those areas usually have better air quality, so it is the reason why we use the *is_open_space* attribute as a supplemental input.

3.2 Produce the prediction

The first step is to use the IDW to predict pollutant levels of PM2.5, NO2, and O3 for each hourly time frame from the known values of pollution data provided by 26 weather stations surrounding Tokyo. These predicted values will be the first input of our regression model and the second one is the *is_open_space* attribute created when we extract visual features in section 3.1. We continue to fit the regression model with these two independent variables to make the prediction.

Our linear regression model has the following formula:

$$Y = \alpha X_1 + \beta X_2 \tag{1}$$

with Y is the value of the pollutant level needs to be predicted, X_1 is the value of the pollutant level predicted by IDW, X_2 is the *is_open_space* attribute, and α, β are the coefficients. Finding those coefficients means that the regression model will be fitted.

Table 2: Labels help to indicate whether a location is an open space or not

Label	Number of Occurrences in Dataset
Tree	38449
Sky	31284
Plant	15079
Cloud	9972
Water	6999
Woody plant	6237
Leaf	4661
Vegetation	2459
Natural landscape	1871
River	1700
Bridge	1666
Nature	1290
Grass	1147
Landscape	1009

Table 3: Evaluation of the PM2.5 prediction without using visual features

Running Course	MAE	RMSE	SMAPE
course1test1_20190415	3.03714	3.12928	0.15926
course1test2_20190415	1.682333	1.899772	0.09083
course2test_20190415	6.669283	7.157959	0.317238
course3test_20190418	16.13135	16.36637	0.756425
course4test_20190422	1.273104	1.378502	0.050883

Table 4: Evaluation of the O3 prediction without using visual features

Running Course	MAE	RMSE	SMAPE
course1test1_20190415	20.72737	22.16136	1.993319
course1test2_20190415	27.14771	27.2999	1.995775
course2test_20190415	5.835984	8.228739	1.97212
course3test_20190418	14.88066	16.05134	1.986563
course4test_20190422	11.70366	12.57533	1.98597

4 RESULTS AND ANALYSIS

The evaluation of PM2.5, NO2, and O3 prediction in the case of not using visual features and vice versa, provided by MediaEval task organizers are shown in Table 3, Table 4, Table 5, Table 6, Table 7, Trable 8, respectively.

In general, PM2.5, O3, and NO2 prediction results are improved, except for the case of NO2 levels of the two running courses course1test1, course1test2. The reason behind this could be because we did not cluster the images of each course separately.

5 DISCUSSION AND OUTLOOK

We are currently investigating more advanced algorithms, such as implementing the combination of IDW with multiple regression [2] and neural network models. Also, we plan to enrich our models with more weather data, such as wind direction, wind speed, temperature, to improve accuracy.

Table 5: Evaluation of the NO2 prediction without using visual features

Running Course	MAE	RMSE	SMAPE
course1test1_20190415	11.9742	12.22642	1.991634
course1test2_20190415	16.27428	16.83156	1.994024
course2test_20190415	35.23405	36.63051	1.99703
course3test_20190418	38.16392	47.97339	1.997812
course4test_20190422	41.49928	42.11436	1.996908

Table 6: Evaluation of the PM2.5 prediction using visual features

Running Course	MAE	RMSE	SMAPE
course1test1_20190415	1.161757	1.357502	0.055718
course1test2_20190415	2.295218	2.557532	0.112553
course2test_20190415	3.497192	3.840557	0.158891
course3test_20190418	13.292	13.5439	0.585743
course4test_20190422	7.318705	7.435286	0.260323

Table 7: Evaluation of the O3 prediction using visual features

Running Course	MAE	RMSE	SMAPE
course1test1_20190415	12.80536	14.91735	0.792462
course1test2_20190415	18.95374	19.16931	1.065872
course2test_20190415	5.488833	6.313631	0.889333
course3test_20190418	5.086703	6.326626	0.386353
course4test_20190422	4.624293	4.960847	0.45899

Table 8: Evaluation of the NO2 prediction using visual features

Running Course	MAE	RMSE	SMAPE
course1test1_20190415	32.63242	32.95355	1.154318
course1test2_20190415	29.67249	30.32006	0.961712
course2test_20190415	14.78879	16.70553	0.376085
course3test_20190418	26.92408	29.8781	0.760942
course4test_20190422	8.079853	10.48465	0.190828

REFERENCES

[1] Zhao P. J Nguyen N.T. Nguyen T.B. Dang-Nguyen D. T. Gurrin C. Dao, M. S. 2020. Overview of MediaEval 2020: Insights for Wellbeing Task - Multimodal Personal Health Lifelog Data Analysis. In *MediaEval Benchmarking Initiative for Multimedia Evaluation, CEUR Workshop Proceedings*.

[2] V. Roshan Joseph and Lulu Kang. 2011. Regression-Based Inverse Distance Weighting With Applications to Computer Experiments. *Technometrics* 53, 3 (2011), 254–265. http://www.jstor.org/stable/23210401

[3] Donald Shepard. 1968. A Two-Dimensional Interpolation Function for Irregularly-Spaced Data. In *Proceedings of the 1968 23rd ACM National Conference (ACM '68)*. Association for Computing Machinery, New York, NY, USA, 517–524. https://doi.org/10.1145/800186.810616

[4] Mohammadsaleh Vahdatpour, Hedieh Sajedi, and Farzad Ramezani. 2018. Air pollution forecasting from sky images with shallow and deep classifiers. *Earth Science Informatics* 11 (09 2018). https://doi.org/10.1007/s12145-018-0334-x

[5] P. Vo, T. Phan, M. Dao, and K. Zettsu. 2019. Association Model between Visual Feature and AQI Rank Using Lifelog Data. In *2019 IEEE International Conference on Big Data (Big Data)*. 4197–4200. https://doi.org/10.1109/BigData47090.2019.9005636