Adele Kuzmiakova

Predicting Air Quality Values from Webcam Images

1 Introduction

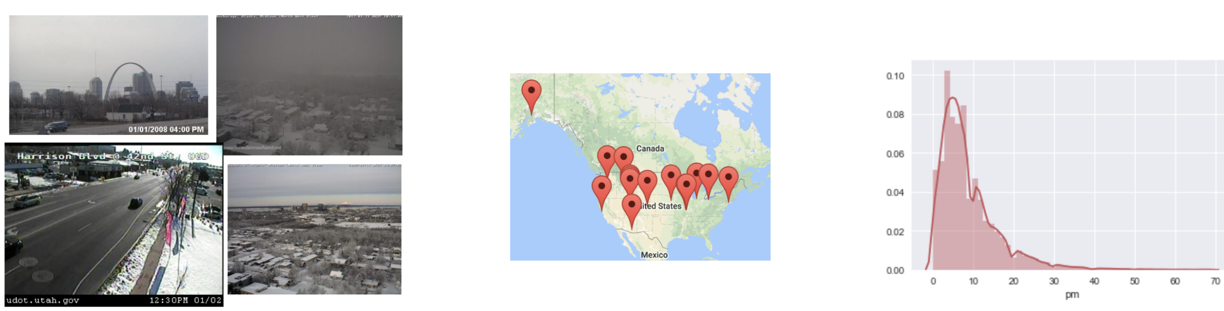
Air quality is becoming an increasingly significant problem in developed countries, such as India, China, or Bangladesh, which are currently working on implementing stricter regulations on air pollution values. However, many remote provinces in China or India lack quality sensors and data, and therefore, they do not have accurate data on every-day air pollution levels. To achieve this objective, consistent air quality data are necessary for effective monitoring of ambient air pollution and should provide robust recommendations to decision makers and legislators. As a result, there is a great need for widely-available and scalable sources of data. One of those sources are images from public webcams, which are ubiquitous and relatively inexpensive to install and maintain. Past works on using simple image feature selection include quantifying transmission features, power spectrum, atmospheric light, and dark channel prior as predictors for ambient air quality [1-6]. In terms of deep learning models, Zhang et al [7] developed a 9-layer convolutional neural network (CNN) architecture to estimate the air quality index from outdoor images taken in Beijing, China. Therefore, two main questions this project hopes to address are:

1) Do public images contain enough information to predict ambient air quality levels?
2) Are features used for dehazing purposes relevant for air quality predictions?

2 Methods

2.1 Simple and deep learning models

Overall, we have access to millions of public images collected by webcams from more than 800 locations across the US between 2008 and 2017. However, due to the relatively short timeline of this class, we don't envision using all of them for training and testing in the coming 3 weeks. Therefore, instead we restrict our current work to only 20,000 public images collected from 16 locations with the highest air quality values. Figure 1 below summarizes the data and the values of the air quality labels, which are represented as particulate matter (pm).



Public images of roads and skylines   800+ locations   pm label distribution

Figure 1. Left: public images, center: map with all 16 webcam locations, right: particulate matter (pm) label distribution.

To estimate the labels from the webcam images, we use simple baselines first, specifically weighted linear regression. Inputs into the weighted linear regression models include 1) raw images and various image features, such as 2) transmission features, 3) power spectrum, and 4) atmospheric light. The feature extraction process of these image features has been relatively well-documented in the literature [4 - 7] and we hypothesize that transmission matrix may be the most important feature for ambient air quality label prediction. Therefore, to achieve efficient and accurate air quality level inference, we extract multiple features from a single image and utilize multiple kernel learning to learn an adaptive classifier.

In terms of deep learning methods, one of the existing network architectures include DehazeNet [8], which has been used as an end-to-end framework for dehazing public or social media images. Although single image haze removal is a challenging and ill-posed problem, the DehazeNet is able to identify patches and features on a single image that are most predictive for the final transmission matrix, which indicates how hazy the image is. The original DehazeNet architecture is depicted in Figure 2 and was intended to give transmission matrices as outputs. However, in our problem, instead of outputting transmission values, we are more interested in predicting the air quality labels and thus, we make certain modifications to the final layer in the original network architecture. Modifications include removing the last layer and replacing it with a convolutional layer with inputs being 5x5x64 patches. These are then used to estimate the value of the pm label.

2.2 Error metrics

To evaluate the performance of our models, we use the mean absolute error, which relates the predicted labels with the actual ground truth with the following formula:

$$MAE = \sum_{i=1}^{n} \frac{predicted^{(i)} - actual^{(i)}}{n}$$

where n is a number of samples. The split between training, dev, and test set is 60%, 20%, and 20%.

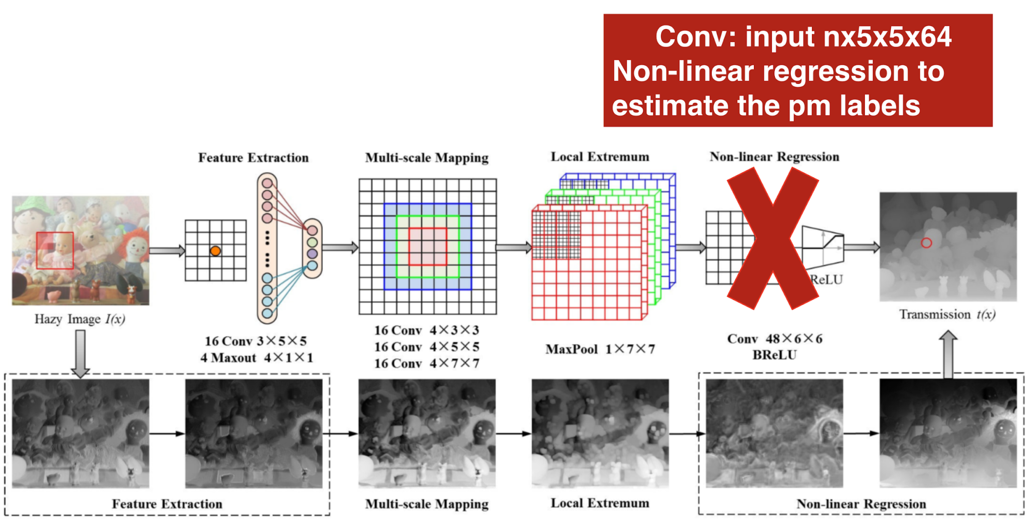


Figure 2: Original and modified DehazeNet architecture. Modified architecture includes removing the last layer and replacing it with a convolutional layer where inputs are 5x5x64 patches. These are then used to estimate the value of the pm label.

3 Results

Running the baseline model, which consists of weighted linear regression with 4 different inputs: 1) raw images, 2) transmission features, 3) power spectrum, and 4) atmospheric light, gave us the following results:

Baseline model (weighted linear regression with the following inputs):	**MAE**
Raw images	5.884
Transmission features	5.729
Power spectrum	5.775
Atmospheric light	6.049

As for clarification, we ran these models on a limited dataset (around 2,000 in training test and 673 in the development set) consisting of only pm values between 20 and 50. Since some of the images with pm values below 20 were corrupted, this slowed down the importing and modeling process a little bit. The best performance was obtained from regression on transmission features, where mean absolute error was 5.775. The scatter plot below in Figure 3 shows the comparison between ground truth (actual pm labels) and predicted pm labels using the transmission features as inputs. The correlation coefficient (r) is very close to 0 which indicates no association between predicted and actual labels. In this case and for the future reference, it also may be worthwhile to reframe the learning problem as a classification task, where pm labels are split into several bins, eg [1] – with pm between 0 and 10, [2] – with pm between 10 and 20, and so on. Thus, the goal of the learning algorithm would be to classify the predictions into one of these categories. The diagnostic plots would then include ROC curves and confusion matrix.

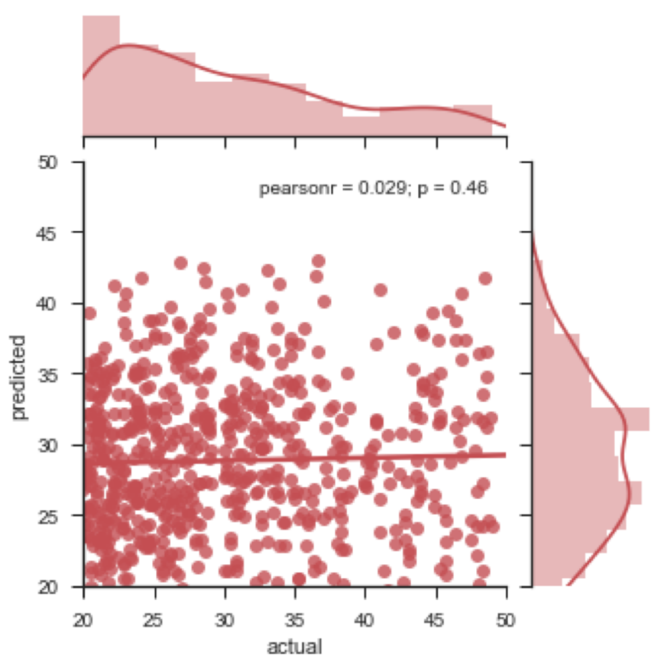


Figure 3: The scatterplot between actual and predicted pm labels using transmission features.

4 Next steps

Currently, we are working on implementing ResNet 50 and VGG as off-the-shelf architectures which have already been pre-trained to get some estimate on their performance. The final goal would be either to use these pre-trained networks for regression or classification or develop our own architecture from scratch. However, this may be a very long shot since the first paper proposing their own architecture for estimating pm labels [3] ended up with 9 customized layers.

In terms of image processing and importing, this task has been finished in the meantime so now we have access to images with full distribution of labels, ranging from 0 to 69.

5 References

[1] Zhang, Z., Ma, H., Fu, H., Liu, L., & Zhang, C. (2016). Outdoor Air Quality Level Inference via Surveillance Cameras. *Mobile Information Systems, 2016*, 1-10. doi:10.1155/2016/9825820

[2] https://github.com/linzhi/dehazing

[3] https://www.eecs.tu-berlin.de/fileadmin/fg144/Research/Theses/Thurow/DiplomaThesis_ChristianThrow_RealTimeImageDehazing_2011.pdf

[4] Makkar, D., & Malhotra, M. (2016). Single Image Haze Removal Using Dark Channel Prior. *International Journal Of Engineering And Computer Science*. doi:10.18535/ijecs/v5i1.12

[5] Qingsong, Z., Shuai, Y., & Yaoqin, X. (2013). An Improved Single Image Haze Removal Algorithm Based on Dark Channel Prior and Histogram Specification. *Proceedings of 3rd International Conference on Multimedia Technology(ICMT-13)*. doi:10.2991/icmt-13.2013.35

[6] Cai, B., Xu, X., Jia, K., Qing, C., & Tao, D. (2016). DehazeNet: An End-to-End System for Single Image Haze Removal. *IEEE Transactions on Image Processing, 25*(11), 5187-5198. doi:10.1109/tip.2016.2598681

[7] Zhang, C., Yan, J., Li, C., Rui, X., Liu, L., & Bie, R. (2016). On Estimating Air Pollution from Photos Using Convolutional Neural Network. *Proceedings of the 2016 ACM on Multimedia Conference - MM 16*. doi:10.1145/2964284.2967230