Image data visualization using t-SNE for Urban Pavement Disease Recognition

Siyu Huang[1]

[1] University of Wisconsin Madison, Madison WI 53715, USA
shuang353@wisc.edu

Abstract. Urban pavement disease recognition is for the most part, a mission performed manually. Recently, video analysis task has been one of the most important applications in various fields. Aims to renovate on the automated vision-based disease recognition and to experiment new methods of road disease detection, this paper analyzes the images took from a city and performed image data visualization of road issues. From image features extracted from histogram of oriented gradient, we perform principal components analysis to reduce the dimensionality of features. Results are presented by visualizing datapoints using t-distributed stochastic neighbor embedding. The experiment shows that the image data visualization using t-SNE is suitable for the growing field of urban road disease recognition.

Keywords: Image data visualization, dimensionality reduction, urban road disease recognition, HOG feature extraction, PCA feature reduction, data normalization

1 Introduction

In recent years, causing by factors like severe weather and heavy equipment, bad pavements conditions have been a common issue in city traffic. Issues like potholes, cracks, and manholes caused traffic accidents which brings negative effects to the safety of people. Therefore, it is important to have detection methods of road problems [1].

With the development of Artificial Intelligence, there are increasing problems in society that being solved by algorithms developed by researchers. For the road issues, it is also possible to apply computer vision techniques on the detection of road problems in cities. From extracting the characteristics and applying machine learning models to analyze road conditions [2].

Data visualization is important for data analysis, as it provides data representation of that originating from different sources. People that need to make decisions can see analytics in a visual format, making the information easier to understand. Specifically, data visualization promotes the discovery of patterns, the understanding of information, and the creation of an opinion [3]. To solve this issue, we involve three methods of

image data processing and visualization for vision based urban road disease recognition: Histograms of Oriented Gradients (HOG) [4] for image data extraction, Principal Component Analysis (PCA) [5] for image data reduction, and T-SNE [6] for image data visualization.

As known, there are previous works that utilize Histograms of Oriented Gradients (HOG) for image detection. *An HOG-LBP Human Detector with Partial Occlusion Handling* [7] propose a human detection method that deals with partial occlusion. Constructed the feature set by combining Histograms of Oriented Gradients and Local Binary Pattern, Wang et al. used two detectors for image scanning. From the result of this project, it is found that that the HOG and LBP combined feature outperforms other state-of-the-art detectors on the INRIA dataset.

PCA is used for feature selection and dimension reduction in this project because of complex multivariate data. Principal Component Analysis by Bro and Smilde [5] provides a comprehensive understanding and usage to PCA. It is noticeable that PCA is good at revealing relationships between samples and connections between variables. In addition, the feature of quantifying features also provides assistance to the feature selection in this project. T-SNE plays important role for the visualization of image features in this project. Visualizing Data using t-SNE [6], presented by the author of t-SNE technique, illustrates the development of this method, shows the comparison with other visualization techniques, and explain the performance of t-SNE on a wide variety of data sets.

In this paper, we first introduce the framework of method using HOG and PCA in Section 2. And then, in Section 3, we also present the visualization tool T-SNE in details. In Section 4, we discuss the experimental implementation and result analysis. Finally, we summarize the conclusion in Section 5.

2 Method Introduction

2.1 Image feature extraction using HOG（Histogram of Oriented Gradient）

Histogram of Oriented Gradient (HOG) is used in this project for image feature extraction. The first step of HOG is gradient calculation, and we can compute image derivative using 1-D derivative mask. Then a cell histogram is created where each pixel within a cell has a weigh corresponding to the values found in the gradient computation. Based on the orientation of the gradient element centered on it, each pixel calculates an edge orientation histogram channel a weighted vote, which are aggregated into orientation bins across local spatial regions that we call cells. Cells might be rectangular or radial in shape. Votes are interpolated bilinearly between neighboring bin centers in both direction and position to reduce aliasing. The third step is histogram normalization [8]. It is necessary to normalize the gradient strengths locally to make up the changes in contrast and illumination, which necessitates categorizing the cells together into larger,

geographically connected blocks. For the HOG descriptor, it is the accumulated vector of block regions' cell histograms' normalization. The overlapping of these blocks means the final descriptor has each cells contributes more than once. Then we can normalize blocks with normalization factor like L1-norm:

$$f = \frac{v}{||v||_1 + e} \tag{1}$$

where v is the non-normalized vector containing all histograms in a given block, $||v||_1$ is 1-norm, and e is a small constant.

2.2 Image feature selection and reduction using PCA

Principal Component Analysis (PCA) is used in this project for image feature selection and reduction. PCA is a way of that transforms the data to a new coordinate system in orthogonal linear transformation. In that way the scalar projection of the data's greatest variance places on the first principal component, the second coordinate is lie by the second greatest variance, etc. [9]

For an $n \times p$ data matrix X with sample mean of each column equals to zero, a different repetition of the experiment is represented by each n rows, and a particular feature is represented by each of the p columns. P-dimensional vectors with weights W maps each vector of matrix X to a new vector of principal component scores t.

$$t_{k(i)} = x_i \cdot w_k \; for \; i = 1, \ldots, n \;\; k = 1, \ldots, l \tag{2}$$

We only keep the first L principal components produce the truncated transformation, $T_L = XW_L$, where T_L is the matrix with n rows whereas only L columns. The score matrix below maximizes the preserved variance in the original data, but minimizes the squared reconstruction error in total with all the transformed data matrices having L columns:

$$\|TW^T - T_L W_L^T\|_2^2 \tag{3}$$

After applying PCA, the road data matrix is dropped to 10-dimensional.

Employing HOG and PCA, there are previous works [10] that works on image reconstruction. It demonstrates that weighing the dimension of PCA with eigenvalue can help minimizing the HOG feature error.

3 Visualization Tool

3.1 t-SNE

The dataset is high dimensional. t-Distributed Stochastic Neighbor Embedding is used to visualizes high-dimensional data in the project. It gives each datapoint a position in the two-dimensional map, where similar objects have close distance and distinct objects have distant distances. [6]

To start with, SNE converts high-dimensional Euclidean distances between datapoints into similarities represented by conditional probabilities [11]. For two datapoints x_j and x_i, the similarity between them is $p_{j|i}$. For datapoints that are close to each other, $p_{j|i}$, is high, while for separated datapoints, this value will be infinitely small. $p_{j|i}$ is represented by

$$p_{j|i} = \frac{exp(-\|x_i - x_j\|^2 / 2{\sigma_i}^2)}{\sum_{k \neq i} exp(-\|x_i - x_k\|^2 / 2{\sigma_i}^2)} \tag{4}$$

where σ_i is the variance of the Gaussian of datapoint x_i.

For the low-dimensional counterparts y_i and y_j of x_i and x_j in high-dimensional space, the conditional probability $q_{j|i}$is similar. $q_{j|i}$ is represented by

$$q_{j|i} = \frac{exp(-\|y_i - y_j\|^2)}{\sum_{k \neq i} exp(-\|y_i - y_k\|^2)} \tag{5}$$

To minimize the distance between $p_{j|i}$ and $q_{j|i}$ by low-dimensional data representation, SNE can use Kullback- Leibler divergence, whose sum of all datapoints can be minimized by using a gradient descent method.

The cost function of SNE is

$$C = \sum_i KL(P_i||Q_i) = \sum_i \sum_j p_{j|i} \, log \frac{p_{j|i}}{q_{j|i}} \tag{6}$$

where P_i stands for the conditional probability distribution over datapoints except x_i, and Q_i represents the conditional probability distribution over all other map points except y_i. The gradient for minimizing this cost function is

$$\frac{\delta C}{\delta y_i} = 2 \sum_j (p_{j|i} - q_{j|i} + p_{i|j} - q_{i|j})(y_i - y_j) \tag{7}$$

T-SNE is a variation of Stochastic Neighbor Embedding but it avoids crowding points together in the map center, which produces better visualization. The significant difference between t-SNE and SNE is the cost function. T-SNE uses Student-t distribution instead of normal distribution to compute the similarity of two points in the low-dimensional space; t-SNE uses a symmetrized version of the SNE cost function with simpler gradients. [6]

In symmetric SNE, q_{ij}, the pairwise similarities in low-dimensional space is:

$$q_{ij} = \frac{exp(-\|y_i - y_j\|^2)}{\sum_{k \neq l} exp(-\|y_k - y_l\|^2)} \tag{8}$$

and p_{ij}, the pairwise similarity in high-dimensional space is:

$$p_{ij} = \frac{exp(-\|x_i - x_j\|^2 / 2\sigma^2)}{\sum_{k \neq l} exp(-\|x_k - x_l\|^2 / 2\sigma^2)} \tag{9}$$

The major benefit of symmetric SNE over asymmetric SNE is the simpler form of the gradient.

$$\frac{\delta C}{\delta y_i} = \sum_j (pij - qij)(yi - yj) \tag{10}$$

For this project, t-SNE is implemented in python. The first step is PCA (Principal Component Analysis) mentioned in Section 2.2, where we initialize the variables and reduce the dimensionality to desired. Local similarities of points are going to be measured. By running iteratively and compute the pairwise affinities, we can keep updating data points locations in *Y*. The result is two-dimensional array *Y* and it can be visualized by scatter plot, as shown in Figure 1.

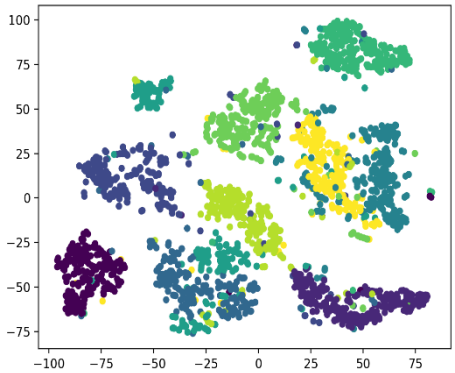

Fig. 1. An example visualization of t-SNE

3.2 A Subsection Sample

The inputs into the t-SNE model are *X* and labels, where *X* is a two-dimensional array represent the image features and labels is a one-dimensional array represent the road disease corresponding to each sample in *X*. Running t-SNE in the NxD array X on the dataset can bring down the dimensionality of NxD array X to no_dims dimensions. The syntax of the function is Y = tsne.tsne(X, no_dims, perplexity), where X is an NxD NumPy array. Figure 2 shows a simplified demonstration of this process. To do that, the NxD array X can be reduced to no_dims by performing Principal Components Analysis.

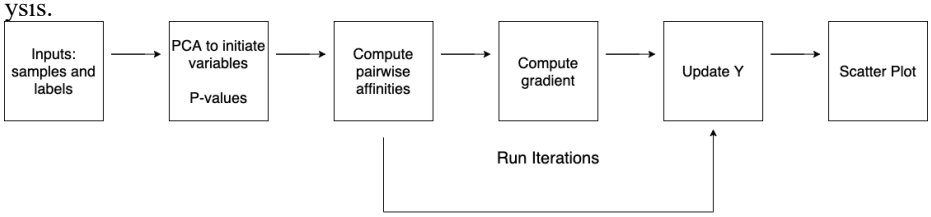

Fig. 2. The Overall Process of t-sne visualization

After initializing variables, P-values that make each conditional Gaussian having the equal perplexity can be get by performing a binary search. Then pairwise affinities, gradient, and the cost of p-value can be computed by running iterations. Lastly, after iterations the outcome is Y, which can be used to visualize the scatter plot.

4 Experimental Implementation and Result Analysis

4.1 Experimental data

The data of this project are photos captured by car camera in different roads in Shenzhen, China. The size of each photo is 1600 x 1184. There are 6000 of them in total and they were captured from March 1, 2019, to June 30, 2019. The goal is to capture the pavements as main targets.

From 6000 photos captured, we summarized 10767 samples from them because each photo can contain multiple road issues or same issues with multiple spots. We summarized 8 types of road problems from image data: crack, manhole, net, others, patch-crack, patch-net, patch-pothole, pothole. For each type of the road issues, there are 1656, 4164, 322, 47, 2543, 1035, 755, 254 samples respectively as shown in Figure 3. For each road disease, several representatives are selected in Figure 4 for illustration.

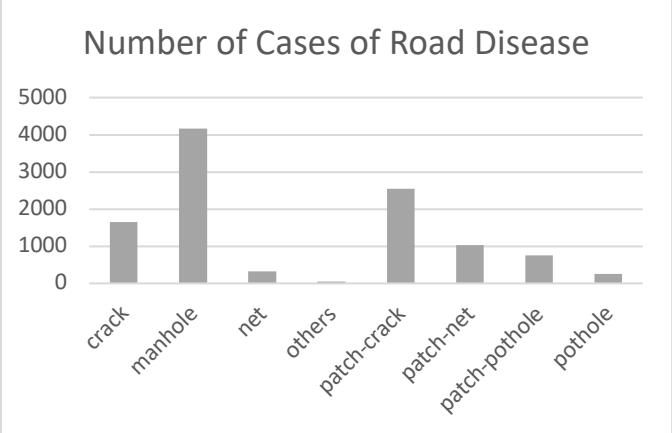

Fig. 3. Classification of 10767 samples

Fig. 4. Examples of road diseases

4.2 Visualization result analysis

For the first attempt of visualization, we include all 10767 samples with two-dimensional and complexity of 5.0, and the outcome in Figure 5 shows that classifications are roughly scattered but there are some types mixed. For example, orange dots representing “other” and black dots representing “patch-net”, blue dots representing “net” and yellow dots representing “patch-crack”, and yellow dots representing “patch-crack” and purple dots representing “pothole”. To distinguish them, we decide to separate these three groups and perform new visualization on each pair of them [12].

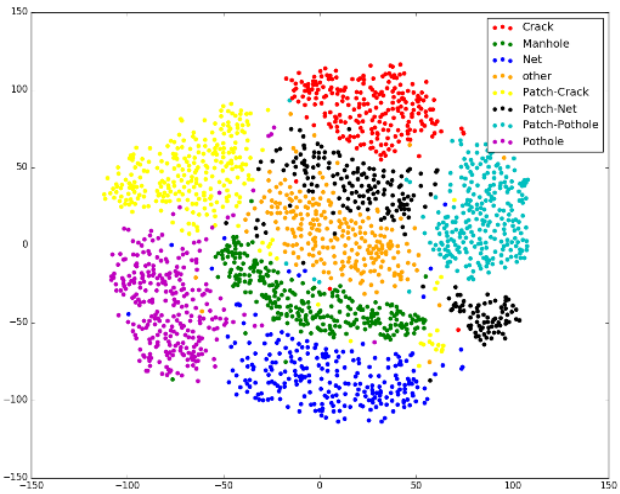

Fig. 5. Visualization of all seven road diseases

The first group is “other” and “patch-net”. From the new visualization of Figure 6 of these two types of road disease, it is found that there are still multiple points that are hard to distinguish though most are separated. They have similar t-sne characteristics and dimensional vectors. It is analyzed that the “other” group is defined as the road diseases that are hard to classify and many of them have the similar pattern to “patch-net”. In the future research, the classification of “other” can be improved by further research to road disease inspection.

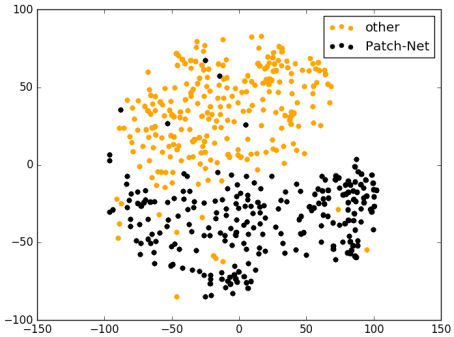

Fig. 6. Visualization of “other” and “Patch-Net”

The second group is “net” and “patch-crack”. It can be seen in figure 7 that they are now well scattered except for one “net” point at the top right corner that is far from its cluster. Analyzing its t-sne value, it is found that this “net” point has Euclidean distance 9.61 with 10 near “patch-crack” points, which means that they have the similarities in dimensional vectors. Extracting the road image of this point and it looks like “patch-crack”.

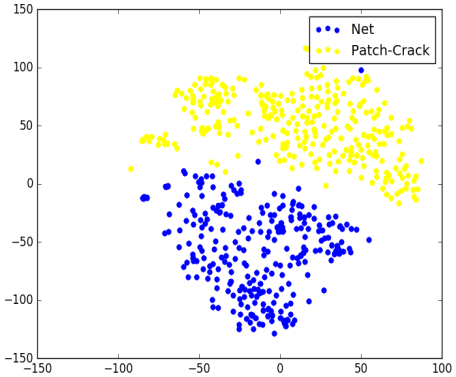

Fig. 7. Visualization of "Net" and "Patch-Crack"

The third group is "patch-crack" and "pothole". Although most points are clustered with correct points in Figure 8, there are 7 purple "pothole" points that are mixed with "patch-crack" points. This is likely caused by insufficient iterations in t-sne algorithm. Thus, we increase the iteration from 1000 to 10000 and regenerate new visualization as shown in Figure 9. The result is much better than the first time, which shows the clear distinction.

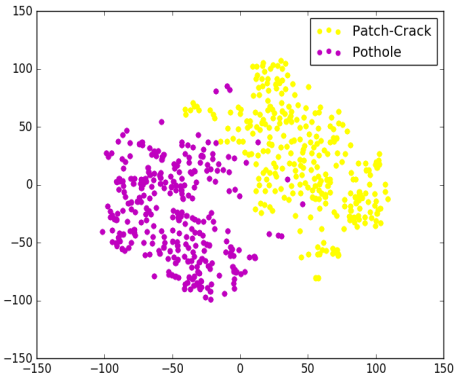

Fig. 8. Visualization of "Patch-Crack" and "Pothole"

The third group is "patch-crack" and "pothole". Although most points are clustered with correct points in Figure 8, there are 7 purple "pothole" points that are mixed with "patch-crack" points. This is likely caused by insufficient iterations in t-sne algorithm. Thus, we increase the iteration from 1000 to 10000 and regenerate new visualization as shown in Figure 9. The result is much better than the first time, which shows the clear distinction.

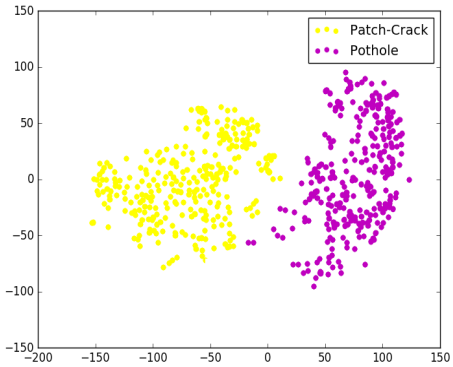

Fig. 8. Visualization of "Patch-Crack" and "Pothole"

5 Conclusion

This paper first analyzes the road images took from a city and the potential diseases within them. Image features are extracted from HOG. Adapting from existing methods about PCA and t-sne, we reduce the dimensionality of features and performed data visualization on these features of road issues. It is found that most of the road diseases can be distinguished from each other, while there are pairs of road diseases that are similar for a portion. Further research is needed to better classify road diseases that have multiple characteristics like "patch-net" and "net". The image data visualization work [13-16] will also be applied other fields in the future.

References

1. J.-p. Li, Y. Sheng, and J. Zhang, "Study on diseases of cement concrete pavement in permafrost regions," *Cold Regions Science and Technology*, vol. 60, no. 1, pp. 57–62, 2010.
2. M. D. Jenkins, T. A. Carr, M. I. Iglesias, T. Buggy, and G. Morison, "A deep convolutional neural network for semantic pixel-wise segmentation of road and pavement surface cracks," in *2018 26th European Signal Processing Conference (EUSIPCO)*, pp. 2120–2124, IEEE, 2018.
3. M. Sadiku, A. E. Shadare, S. M. Musa, C. M. Akujuobi, and R. Perry, "Data visualization," *International Journal of Engineering Research And Advanced Technology (IJERAT)*, vol. 2, no. 12, pp. 11–16, 2016.
4. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05)*, vol. 1, pp. 886–893, Ieee, 2005.
5. R. Bro and A. K. Smilde, "Principal component analysis," *Analytical methods*, vol. 6, no. 9, pp. 2812–2831, 2014.
6. L. Van der Maaten and G. Hinton, "Visualizing data using t-sne.," *Journal of machine learning research*, vol. 9, no. 11, 2008.
7. X. Wang, T. X. Han, and S. Yan, "An hog-lbp human detector with partial occlusion handling," in *2009 IEEE 12th international conference on computer vision*, pp. 32–39, IEEE, 2009.

8. N. B. Kar, K. S. Babu, and S. K. Jena, “Face expression recognition using histograms of oriented gradients with reduced features,” in *Proceedings of International Conference on Computer Vision and Image Processing*, pp. 209–219, Springer, 2017.
9. J. Lever, M. Krzywinski, and N. Altman, “Points of significance: Principal component analysis,” *Nature methods*, vol. 14, no. 7, pp. 641–643, 2017.
10. H. G. Jung, “Analysis of reduced-set construction using image reconstruction from a hog feature vector,” *IET Computer Vision*, vol. 11, no. 8, pp. 725–732, 2017.
11. G. Hinton and S. T. Roweis, “Stochastic neighbor embedding,” in *NIPS*, vol. 15, pp. 833–840, Citeseer, 2002.
12. M. Wattenberg, F. Vi′egas, and I. Johnson, “How to use t-sne effectively,” *Distill*, 2016.
13. T. Walter, D. W. Shattuck, R. Baldock, M. E. Bastin, A. E. Carpenter, S. Duce, J. Ellenberg, A. Fraser, N. Hamilton, S. Pieper, et al., “Visualization of image data from cells to organisms,” *Nature methods*, vol. 7, no. 3, pp. S26–S41, 2010.
14. J. N. Kather, A. Weidner, U. Attenberger, Y. Bukschat, C.-A. Weis, M. Weis, L. R. Schad, and F. G. Z¨ollner, “Color-coded visualization of magnetic resonance imaging multiparametric maps,” *Scientific reports*, vol. 7, no. 1, pp. 1–11, 2017.
15. J. C. Caicedo, S. Cooper, F. Heigwer, S. Warchal, P. Qiu, C. Molnar, A. S. Vasilevich, J. D. Barry, H. S. Bansal, O. Kraus, et al., “Data-analysis strategies for image-based cell profiling,” *Nature methods*, vol. 14, no. 9, pp. 849–863, 2017.
16. K. B¨orner, A. Bueckle, and M. Ginda, “Data visualization literacy: Definitions, conceptual frameworks, exercises, and assessments,” *Proceedings of the National Academy of Sciences*, vol. 116, no. 6, pp. 1857–1864, 2019.