International Journal on Advanced Science Engineering Information Technology

Vol.14 (2024) No. 5
ISSN: 2088-5334

Clustering Urban Roads Using Local Binary Patterns to Enhance the Accuracy of Traffic Flow Prediction

Bagus Priambodo [a,*], Rabiah Abdul Kadir [b], Azlina Ahmad [b]

[a] *Faculty of Computer Science, Universitas Mercu Buana, Jalan Meruya Selatan, Jakarta, Indonesia*
[b] *Institute of Visual Informatics, Universiti Kebangsaan Malaysia, Bangi, Selangor, Malaysia*
*Corresponding author: *bagus.priambodo@mercubuana.ac.id*

***Abstract*—Many studies employ a dynamic statistical approach to model the relationship between roads for predicting traffic flow. Statistical relationships among roads pertain to the associations between road segments in nearby areas. Roads with similar traffic patterns usually define the relationship. Previous studies have shown that adjacent roads demonstrated a similar traffic pattern on the same day and time interval. Studying similar patterns between roads in surrounding areas provides information about the traffic state among roads in a cluster. Furthermore, the results of this finding may be considered to increase the performance of prediction of traffic conditions. In general, road segment roads are correlated with other road segments. They are connected upstream, downstream, or both; they are connected upstream, downstream, or both. To address this, we propose a Local Binary Pattern (LBP) to explore road connections by clustering congestion features between the target roads and their neighboring roads. The outcomes show that the roads with high connectivity obtained from the clustering process were utilized to predict traffic conditions using the SVM approach. The David-Bouldin Index of LBP with k-means shows the lowest compared with PCA k-means and k-means only. By evaluating the clustering result using the Davies-Bouldin Index and assessing the prediction results with SVM, we discovered that incorporating LBP with K-Means improved outcomes in identifying highly connected roads using K-Means.**

***Keywords*—Traffic state prediction; feature extraction; local binary pattern; spatial relationship; support vector machine.**

Manuscript received 7 Mar. 2024; revised 29 Jun. 2024; accepted 24 Aug. 2024. Date of publication 31 Oct. 2024.

I. INTRODUCTION

Traffic congestion occurs when the number of cars on the road outmatches its maximum capacity. Critical characteristics of road congestion include reduced average vehicle speeds, longer travel durations, and cars queuing up on the road. The effective identification of traffic congestion requires using a suitable method or model. One approach to assess congestion distribution within surrounding roads is by analyzing the correlation between congestion and increasing traffic congestion. As demonstrated in prior research, various factors can contribute to traffic congestion, including vehicle speeds, rain conditions, accidents, and special events. An inherent complexity of traffic congestion is its dynamic and interconnected nature, as it can spread from one road to a traffic network [1], [2]. Understanding these interconnections among roads in congested areas can offer valuable insights for assisting drivers in avoiding traffic jams or congestion. The outcomes of this study can potentially enhance the ability of the clustering algorithm to cluster road segments based on traffic state in day and time. Understanding the connections between roads in congested areas can offer valuable insights to help drivers avoid traffic congestion, and the findings of this research can enhance the efficiency of clustering road segments using the k-means algorithm.

Spatiotemporal factors have been considered for many years to predict traffic flow [3]. Surrounding roads, especially connected roads downstream upstream and adjacency roads, have been considered as spatial factors influencing traffic flow. A study by [4] uses spatial facts to predict traffic flows. A study by [5] stated that traffic congestion in a busy area can be anticipated quickly by utilizing traffic data from the connected upstream links.

Other studies extend connected roads with sub-connected roads. A study by [5] extends one level of sub-connected roads upstream and downstream as spatial factors. This study states that expanded connected roads improve traffic prediction accuracy compared to predictions using only connected roads. Past studies indicate a correlation between a road segment and others connected upstream, downstream, or in both directions using the regression method [6]. Several research studies identify the congestion relationship between two road segments [1], specifically from roads III to IV, covering distance d. If congestion is present on road III at t (time) and at t + T (time), Figure 1 illustrates that a congestion state occurs on road IV and road III. This situation will impact traffic flow on Road I and Road II.

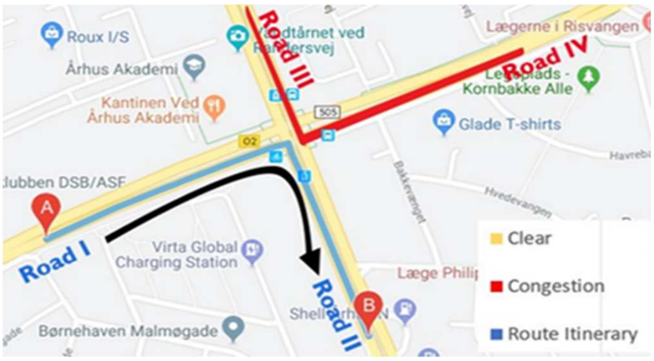


Fig. 1 Road III influences neighboring roads

Instead of predicting traffic flow based on linked roads, several studies employ a dynamic statistical approach to model the relationship between roads for predicting traffic flow. Statistical relationships among roads pertain to the associations between road segments in nearby areas. The prediction is usually based on roads with similar daily traffic patterns [7]. An example study by [1] established the relationship between road segments by examining the traffic conditions and congestion occurring on two roads within a specified distance.

In this study, we present the outcomes of our investigation, aiming to identify the correlation between roads in adjacent areas by examining traffic state extraction features concerning both day and time. Local binary pattern (LBP) is a popular technique for feature extraction [8], [9]. The principal idea of the LBP algorithm is the feature extraction of grayscale images [10]. Selecting the optimal pixel using the Local Binary Pattern (LBP) detector involves assessing the circle radius compared to neighboring pixels [10], which resemble the traffic conditions of neighboring roads surrounding the target road. The connection between roads is established by clustering the Local Binary Pattern (LBP) values between the target road and its surrounding roads by applying the k-means algorithm. The obtained relationship roads are subsequently employed to predict traffic conditions using machine learning. The choice of the best pixel by the detector (LBP) is based on the determination of the circle radius compared with the neighboring pixel [10]. This represents the traffic state of neighboring roads surrounding road targets. This study investigates the relationship between roads by clustering the LBP value between a target road and neighboring roads using k-means. The results are then used to predict the traffic state using the support vector machine (SVM) method. For evaluation purposes, we compare the results of traffic state prediction using the SVM method based on PCA k-means and k-means only.

II. MATERIAL AND METHOD

This study involves multiple phases, as shown in Figure 2. First, traffic data are collected, then congestion is determined, road segments are clustered, and, finally, the traffic state is predicted based on the clustering results.

Fig. 2 Diagram methodology

A. Sourcing Traffic Data

For the experiment, we employed a dataset sourced from IoT sensors situated in Aarhus, Denmark [11]. Figure 3 illustrates an approximate count of 449 sensors in this location.

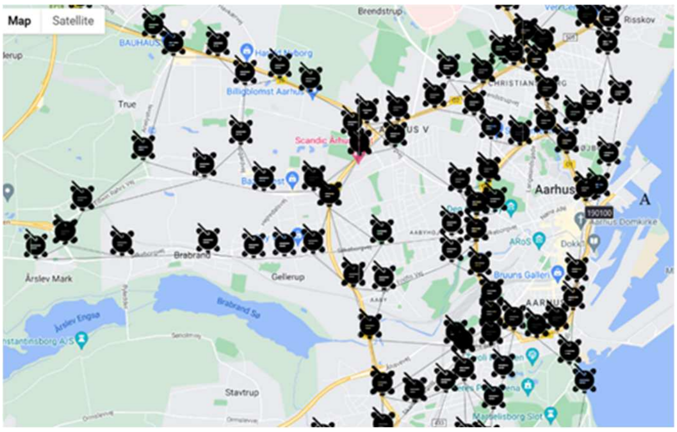


Fig. 3 Dataset location taken from the city of Aarhus, Denmark

B. Determine congestion

Various definitions and metrics for assessing road congestion have been developed. Road congestion occurs when the incoming traffic from upstream exceeds the downstream road capacity, causing a particular link to become congested after a certain period [1]. Some studies utilize an index of congestion level, considering factors such as saturation degree, travel speed, and a combination [12]. Another study introduced the traffic performance index[13]. The current research computed the congestion index for specific time intervals to measure the similarity between roads and calculate the congestion level. The congestion index is determined based on the traffic factor [12] and is calculated every 20 minutes following [14].

Finally, to determine road congestion, this study computed the congestion index over 20-minute intervals to gauge road similarity and determine the congestion level. According to traffic policy in Denmark [15], the average speed of regular town traffic is typically 50 km/hour. Based on this situation, traffic jams or congestion is defined in this study when the speed falls below 50 km/h. Therefore, we consider a road congested if the congestion index equals or exceeds 3, following [14].

C. *Clustering Road Segment*

K-means Clustering has found extensive application in intelligent transportation systems for grouping similar road segments and identifying daily traffic patterns [16]. In a prior investigation, road segments with comparable characteristics were clustered using K-means, considering factors such as volume, vehicle sensors, speed, and GPS trajectory [17]–[19]. This study clusters road segments based on the congestion index following [14].

D. *Local Binary Pattern*

This study uses a local binary pattern and a popular k-means algorithm to cluster surrounding roads. K-means clustering is a widely used unsupervised machine learning technique that aims to partition a set of data points into *k* distinct clusters [14]. Each cluster is characterized by its centroid [20], which is the average of the points in that cluster. K-means clustering aims to minimize the variance within each cluster, thereby making the points in the same cluster as similar as possible while keeping the clusters distinct. Local binary pattern (LBP) is a popular technique for feature extraction [8], [9]. The principal idea of the LBP algorithm is the feature extraction of grayscale images [10]. Selecting the optimal pixel using the Local Binary Pattern (LBP) detector involves assessing the circle radius compared to neighboring pixels [10], which resemble the traffic conditions of neighboring roads surrounding the target road. The connection between roads is established by clustering the Local Binary Pattern (LBP) values between the target road and its surrounding roads by applying the k-means algorithm, as shown in Figure 4.

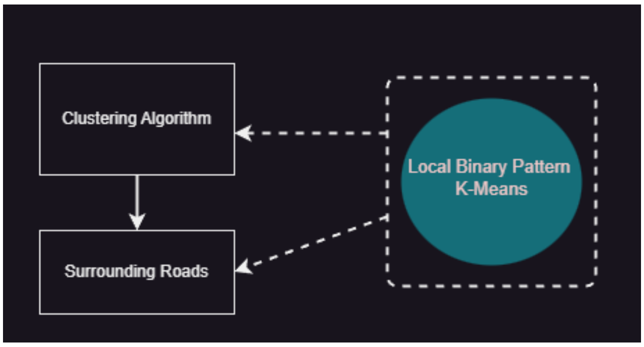


Fig. 4 Proposed Clustering Algorithm

The genuine local binary pattern algorithm [21] involves the pixel of an image representation using decimal numbers known as local binary pattern codes. After obtaining a list of these local binary patterns, the LBP method is applied to image recognition. This approach encodes the local structure surrounding each pixel, as depicted in Figure 7. For each pixel, a comparison is made with its eight neighbors in 3x3 surrounding pixels, determining whether each neighbor's pixel number is less than or equal to the middle or central pixel. The result is a binary code, where values less than the central pixel are encoded as 0 and others as 1. The final number (binary) is developed by merging all these binary codes in a clockwise behavior starting from the top-left, as shown in Figure 5.

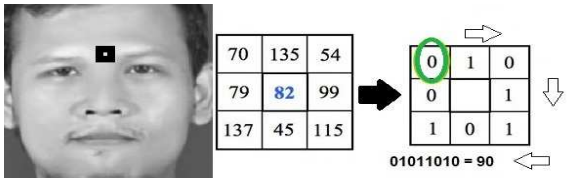
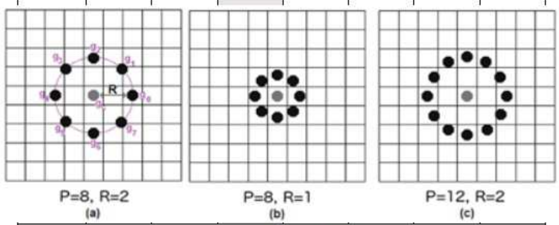


Fig. 5 The flow of the local binary pattern algorithm[22]

Ultimately, a histogram exists for each image (boundary) in this study. The boundary is the neighboring road within every training dataset. In simpler terms, for each of the N images in the training dataset, the Local Binary Pattern (LBP) extracts a corresponding histogram, which is stored for subsequent recognition processes. The algorithm keeps track of the association between the extracted histograms and the respective individuals. In the context of this study, LBP is applied to extract traffic flow patterns on each road. An illustration of the traffic on surrounding roads in the Local Binary Pattern is presented in Figure 6.

Road 158536							
Date	5:40	6:00	6:20	6:40	7:00	7:20	7:40
1	0	0	1	1	1	1	1
2	0	1	1	1	1	1	1
3	0	1	1	1	1	1	1



Road 158386							
Date	5:40	6:00	6:20	6:40	7:00	7:20	7:40
1	1	1	1	1	1	1	1
2	0	1	1	1	1	1	1
3	1	1	1	1	1	0	1

Fig. 6 The traffic of surrounding roads in the Local Binary Pattern

E. *David Bouldin Index*

In general, SSE, silhouette coefficient [23], and David Bouldin Index (DBI) [24], [25] are commonly employed to assess clustering results and calculate the optimal number of clusters (k). The David Bouldin Index (DBI) [24] is a metric used to evaluate the quality of clustering results in unsupervised machine learning. It measures how well the clusters are separated from each other and how compact the clusters are. The DBI helps assess a clustering solution's goodness by considering both clusters' separation and cohesion. A lower Davies-Bouldin Index (DBI) value suggests that the clustering outcomes demonstrate increased cohesion within individual clusters and more precise separation between distinct clusters. A smaller DBI indicates that items within the same group are closely related, while different groups are more distinctly separated. The formula is explained in (1).

$$DBI = \frac{1}{k}\sum_i^k max_{j\neq i}\left(\frac{S_i+S_j}{D_{ij}}\right) \quad (1)$$

Where:

- k is the number of clusters.
- S_i is the average distance between the points in cluster i and the centroid of cluster i, also known as the cluster's internal cohesion.
- D_{ij} is the distance between the centroids of clusters i and j, representing the separation between clusters.
- The expression $\frac{S_i+S_j}{D_{ij}}$ measures how close and similar clusters i and j are, where a lower value indicates better clustering

This study uses the David Bouldin Index to identify the optimal k, evaluating the average maximum similarity within each cluster. A lower Davies-Bouldin Index (DBI) value suggests that the clustering outcomes demonstrate increased cohesion within individual clusters and clearer separation between distinct clusters. A smaller DBI indicates that items within the same group are closely related, while different groups are more distinctly separated.

F. Compare with Principal Component Analysis

Principal Component Analysis (PCA) is a standard method for dimensionality reduction [26], [27]. The PCA method reduces linear projection, which maximizes the scatter of all projected samples. The principal idea is that correlated variables often explain a high-dimensional dataset and only keep the meaningful dimensions for most of the information. Further detail about the equation in [26]. In this study, we observed most of the variance in each road as shown in Figure 7. Figure 7 shows the average majority of variance on each road is below the first five components. However, in this study, we considered the first ten components to be clustered using k-Means following work by [28].

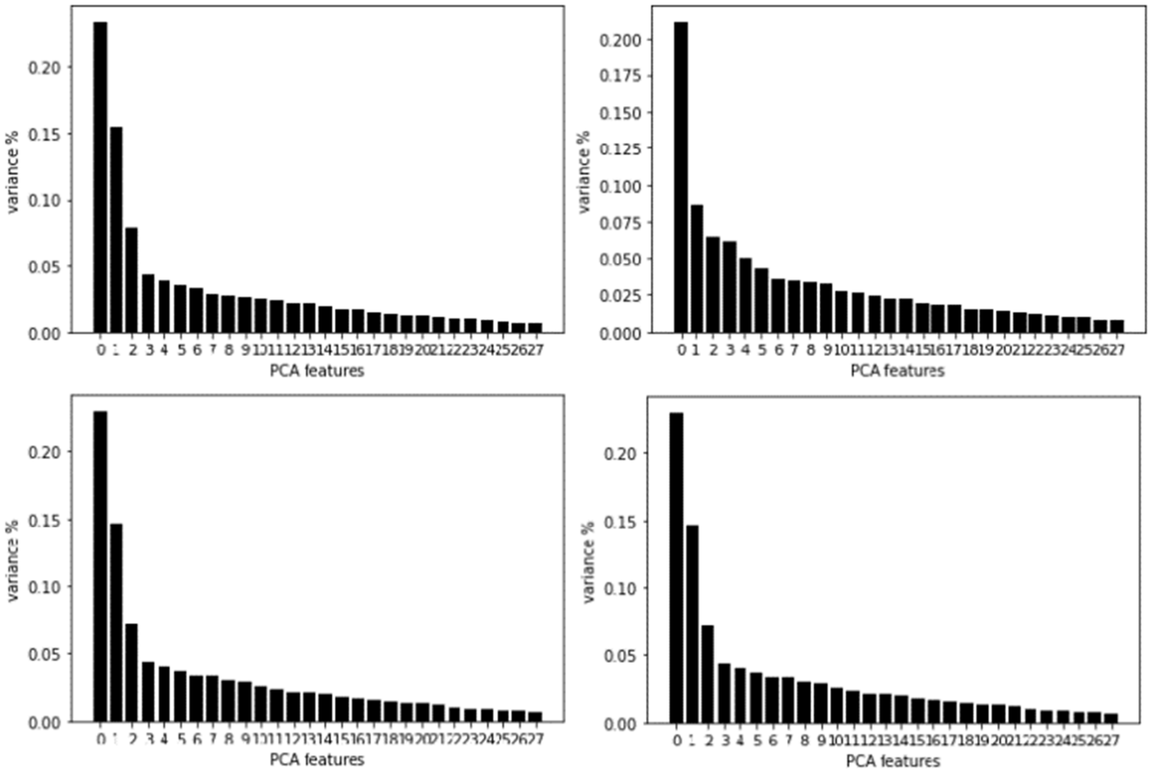


Fig. 7 The variance of principal components from four road segments

The proposed local binary pattern k-means was then compared with the principal component analysis (PCA) k-means algorithm [28].

G. Predict Traffic State

Many studies use the Support Vector Machine (SVM) algorithm to predict traffic flow [29]–[31]. In this study, the surrounding roads derived from the clustering method are utilized for prediction through the Support Vector Machine (SVM) method. The problem can be formally defined using the following formula (2).

$$\begin{gathered} D = \left\{(x_i, y_i), x_i \in R^n, y \in \{-1, 1\}\right\}_{i=1}^{m} \\ B = \min_{i=1\ldots m} |w \cdot x + b \\ H = \max_{i=1\ldots s}\{h_i | B_i\} \end{gathered} \quad (2)$$

D is the dataset, where x is surrounding roads and y is the road target (x and y are roads in the same cluster obtained using the clustering method). B is computed for each training data, and B is the smallest β obtained. Each H hyperplanes, each of them will have a B_i value, and finally, the hyperplane with the largest B_i Value is selected. The proposed algorithm for predicting traffic state is shown in Figure 8.

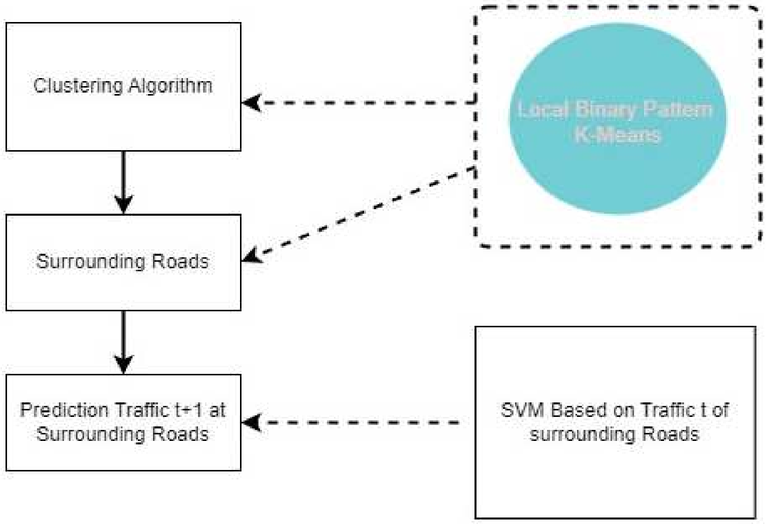


Fig. 8 Proposed Prediction Algorithm

III. Results and Discussion

A. Result

1) Clustering Road Segment: In this study, the David Bouldin index is used to find optimal numbers of *k*. This index assesses the average maximum similarity within each cluster. A lower Davies-Bouldin Index (DBI) value indicates that the clustering outcomes exhibit proximity within individual clusters and greater separation between distinct clusters. In simpler terms, a smaller DBI suggests that items within the same group are closely related, while different groups are more distinctly separated. The comparison of DB scores in Road 158324, road 158715, road 158536, and Road 1589595 is shown in Figure 9, and Table I. Figure 9 shows that using LBP with k-means has lower DBI than without LBP. Table II shows that combining LBP with k-Means filters more roads with a high relationship with the target road. It shows that LBP with k-Means performs better in finding high relationship roads with target roads.

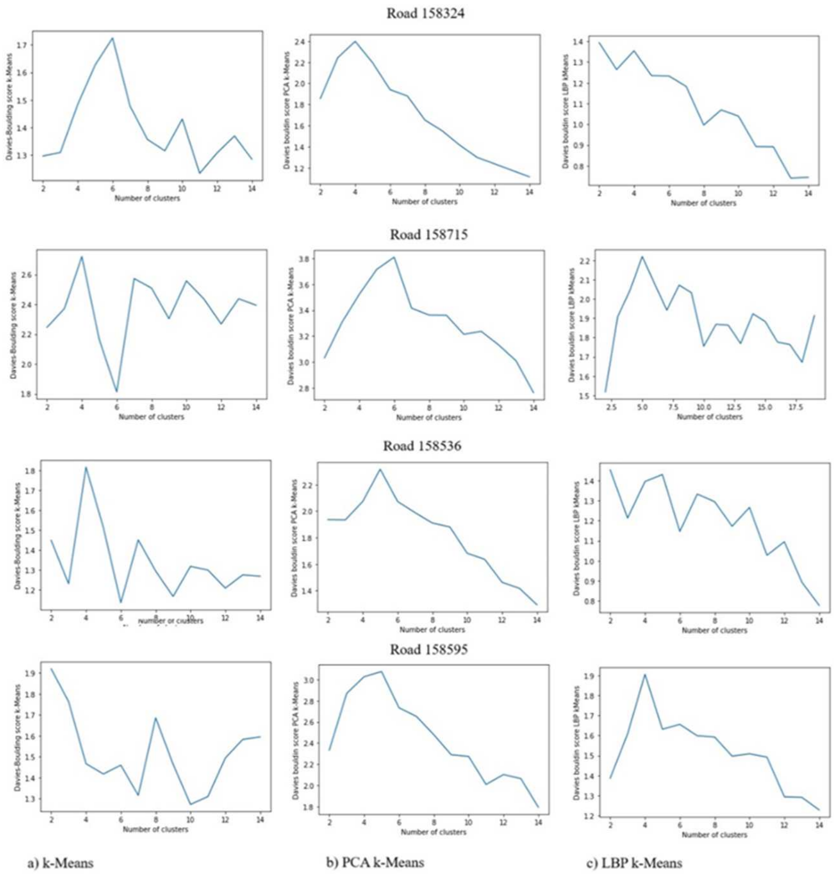


Fig. 9 Line chart of optimal k based on David Bouldin index using k-means, PCA k-means and LBP k-means

TABLE I
COMPARISON OF DBINDEX VALUE BETWEEN METHODS

Location	k-Means	PCA k-Means	LBP k-Means
158324	1.2	1.1	0.7
158715	1.8	2.8	1.7
158536	1.1	1.3	0.8
158595	1.3	1.8	1.2

TABLE II
COMPARISON MEMBER OF SURROUNDING ROADS IN FOUR LOCATIONS

Location	k-Means	PCA k-Means	LBP k-Means
158324	14	4	6
158715	14	8	8
158536	40	8	5
158595	16	18	3

2) Traffic Prediction: Table III shows the average accuracy of prediction results for all neighboring roads. It observes that the average accuracy of prediction utilizing LBP with k-means is better than without LBP.

TABLE III
ACCURACY OF PREDICTION OF TRAFFIC STATE USING SVM BASED ON THREE CLUSTERING METHODS

Location	k-Means	PCA k-Means	LBP k-Means
158324	0.88	0.91	0.93
158715	0.86	0.91	0.93
158536	0.88	0.92	0.92
158595	0.89	0.91	0.92

B. Discussion

In this research, our primary focus was identifying closely related roads within a neighboring region to predict traffic conditions. The findings underscore that applying LBP with k-Means consistently yielded superior results when investigating road relationships. DBI value confirmed that the LBP and K-Means combination effectively shows the lowest value than k-Means without LBP (PCA k-Means and k-Means only). This distinction is clearly illustrated in Figure 10, which provides a comparative overview of the three methods. Notably, predictions made using the K-Means with the LBP approach outperformed those based solely on the K-Means method. Furthermore, the combination of LBP with k-Means filtered more roads than without the LBP method, as shown in Figure 11.

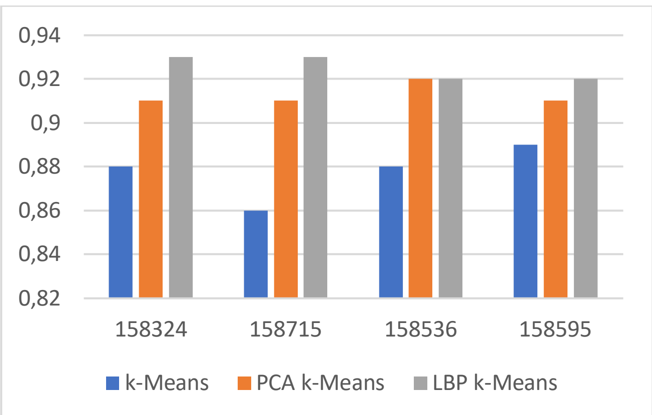


Fig. 10 Comparison of Average Prediction Accuracy based on three clustering algorithms

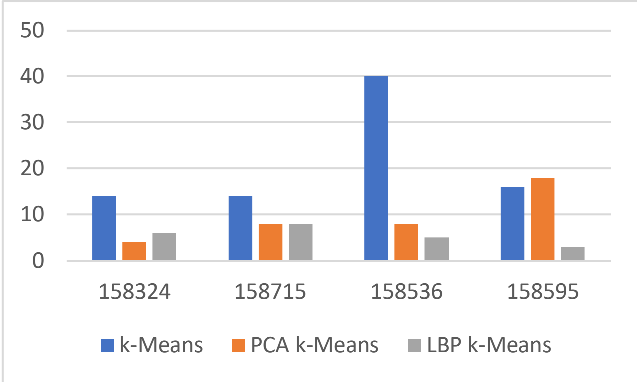


Fig. 11 Comparison of the number of neighboring roads obtained from clustering

IV. CONCLUSION

Our primary objective was to enhance the clustering of road segments using the k-means algorithm to investigate the relationships between these segments. It's recognized that the traffic flow on one road can impact the traffic flow on surrounding roads. However, not all streets exert a significant influence on neighboring traffic conditions. Only those roads exhibiting a solid relationship with the target road will likely impact traffic flow in the surrounding area. This study employed LBP to extract relevant features related to traffic flow at specific times of the day. LBP is a widely accepted technique for feature extraction due to its effectiveness. Combining LBP with the k-means clustering approach improved the results in identifying roads with high relationships when using k-means clustering. Consequently, the traffic state predictions using SVM, LBP and the k-means method consistently outperformed predictions made using the PCA k-means and k-means only. Our experimental findings provide compelling evidence that implementing LBP significantly enhances the performance of road segment clustering through the k-means method.

ACKNOWLEDGMENT

The authors thank Indonesia's Ministry of Education, Culture, Research, and Technology for funding this research work through the KDN Research Grant Scheme. They would also like to extend their gratitude for using the services and facilities of the Intelligent Big Data Analytics Lab at Universitas Mercu Buana and Institute of Visual Informatics Universiti Kebangsaan Malaysia.

REFERENCES

[1] W. Yue, C. Li, Y. Chen, P. Duan, and G. Mao, "What is the Root Cause of Congestion in Urban Traffic Networks: Road Infrastructure or Signal Control?," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 7, pp. 8662–8679, 2022, doi: 10.1109/TITS.2021.3085021.

[2] G. K. Befekadu, "Large deviations and rare events in the study of processes describing the spread of traffic congestions in urban network systems," *2021 55th Annu. Conf. Inf. Sci. Syst. CISS 2021*, no. 1, 2021, doi: 10.1109/CISS50987.2021.9400254.

[3] R. Zhou, H. Chen, H. Chen, E. Liu, and S. Jiang, "Research on Traffic Situation Analysis for Urban Road Network through Spatiotemporal Data Mining: A Case Study of Xi'an, China," *IEEE Access*, vol. 9, pp. 75553–75567, 2021, doi: 10.1109/ACCESS.2021.3082188.

[4] M. Jiang and Z. Liu, "Traffic Flow Prediction Based on Dynamic Graph Spatial-Temporal Neural Network," *Mathematics*, vol. 11, no. 11, pp. 1–16, 2023, doi: 10.3390/math11112528.

[5] T. Sun, Z. Huang, H. Zhu, Y. Huang, and P. Zheng, "Congestion Pattern Prediction for a Busy Traffic Zone Based on the Hidden Markov Model," *IEEE Access*, vol. 9, pp. 2390–2400, 2021, doi:10.1109/access.2020.3047394.

[6] S. Rajendran and B. Ayyasamy, "Short-term traffic prediction model for urban transportation using structure pattern and regression: an Indian context," *SN Appl. Sci.*, vol. 2, no. 7, pp. 1–11, 2020, doi:10.1007/s42452-020-2946-2.

[7] I. Taheri Sarteshnizi, M. Sarvi, S. A. Bagloee, and N. Nassir, "Temporal pattern mining of urban traffic volume data: a pairwise hybrid clustering method," *Transp. B*, vol. 11, no. 1, 2023, doi:10.1080/21680566.2023.2185496.

[8] R. Shaw, C. Mohanty, B. K. Patra, and A. Pradhan, "1D Multi-Point Local Ternary Pattern: A Novel Feature Extraction Method for Analyzing Cognitive Engagement of students in Flipped Learning Pedagogy," *Cognit. Comput.*, vol. 15, no. 4, pp. 1243–1256, 2023, doi:10.1007/s12559-022-10023-5.

[9] A. Chater, H. Benradi, and A. Lasfar, "New approach to similarity detection by combining technique three-patch local binary patterns (TP-LBP) with support vector machine," *IAES Int. J. Artif. Intell.*, vol. 12, no. 4, pp. 1644–1653, 2023, doi: 10.11591/ijai.v12.i4.pp1644-1653.

[10] R. A. Rahma, R. A. Nugroho, D. Kartini, M. R. Faisal, and F. Abadi, "Combination of texture feature extraction and forward selection for one-class support vector machine improvement in self-portrait classification," *Int. J. Electr. Comput. Eng.*, vol. 13, no. 1, pp. 425–434, 2023, doi: 10.11591/ijece.v13i1.pp425-434.

[11] S. Bischof, C.-S. Karapantelakis, Athanasios Nechifor, A. Sheth, A. Mileo, and P. Barnaghi, "Real Time IoT Stream Processing and Large-scale Data Analytics for Smart City Applications," in *Real Time IoT Stream Processing and Large-scale Data Analytics for Smart City Applications*, 2014.

[12] W. X. Wang, R. J. Guo, and J. Yu, "Research on road traffic congestion index based on comprehensive parameters: Taking Dalian city as an example," *Adv. Mech. Eng.*, vol. 10, no. 6, pp. 1–8, 2018, doi: 10.1177/1687814018781482.

[13] J. Zang, P. Jiao, S. Liu, X. Zhang, G. Song, and L. Yu, "Identifying Traffic Congestion Patterns of Urban Road Network Based on Traffic Performance Index," *Sustain.*, vol. 15, no. 2, 2023, doi:10.3390/su15020948.

[14] B. Priambodo, A. Ahmad, and R. A. Kadir, "Spatio-temporal K-NN prediction of traffic state based on statistical features in neighbouring roads," *J. Intell. Fuzzy Syst.*, vol. 40, pp. 9059–9072, 2021, doi:10.3233/JIFS-201493.

[15] T. Hels, A. Lyckegaard, and N. Pilegaard, *Evaluering af trafiksikkerhedstiltag - en vejledning*. 2011.

[16] P. Zhang, W. Ma, and S. Qian, "Cluster analysis of day-to-day traffic data in networks," *Transp. Res. Part C Emerg. Technol.*, vol. 144, no. September, p. 103882, 2022, doi: 10.1016/j.trc.2022.103882.

[17] Z. Sun, Y. Hu, W. Li, S. Feng, and L. Pei, "Prediction model for short-term traffic flow based on a K-means-gated recurrent unit combination," *IET Intell. Transp. Syst.*, vol. 16, no. 5, pp. 675–690, 2022, doi: 10.1049/itr2.12165.

[18] W. Lv, W. Tang, H. Huang, and T. Chen, "Research and Application of Intersection Clustering Algorithm Based on PCA Feature Extraction and K-Means," *J. Phys. Conf. Ser.*, vol. 1861, no. 1, pp. 0–7, 2021, doi: 10.1088/1742-6596/1861/1/012001.

[19] F. Liu, W. Bi, W. Hao, F. Gao, and J. Tang, "An Improved Fuzzy Trajectory Clustering Method for Exploring Urban Travel Patterns," *J. Adv. Transp.*, vol. 2021, no. 3, 2021, doi: 10.1155/2021/6651718.

[20] D. Witarsyah, M. A. Akbar, V. S. Praditha, and M. Sugiat, "Cluster Analysis of Japanese Whiskey Product Review Using K-Means Clustering," *Int. J. Informatics Vis.*, vol. 8, no. 1, pp. 395–400, 2024, doi: 10.62527/joiv.8.1.2601.

[21] N. T. A. Ramaha, R. M. Mahmood, A. A. Hameed, N. L. Fitriyani, G. Alfian, and M. Syafrudin, "Brain Pathology Classification of MR Images Using Machine Learning Techniques," *Computers*, vol. 12, no. 8, 2023, doi: 10.3390/computers12080167.

[22] B. Priambodo, Y. Jumaryadi, S. Rahayu, D. Firdaus, M. Sobri, and Z. P. Putra, "Prediction of Drug Users Based on Facial Scratching Pattern," *Int. J. online Biomed. Eng.*, vol. 17, no. 3, pp. 141–150, 2021, doi: 10.3991/ijoe.v17i03.17351.

[23] J. F. Zaki, A. Ali-Eldin, S. E. Hussein, S. F. Saraya, and F. F. Areed, "Traffic congestion prediction based on Hidden Markov Models and contrast measure," *Ain Shams Eng. J.*, no. xxxx, 2019, doi:10.1016/j.asej.2019.10.006.

[24] J. Chen, Q. Feng, and Q. Guo, "Multi-class freeway congestion and emission based on robust dynamic multi-objective optimization," *Algorithms*, vol. 14, no. 9, 2021, doi: 10.3390/a14090266.

[25] T. Si and A. Mukhopadhyay, "Breast DCE-MRI Segmentation for Lesion Detection Using Clustering with Fireworks Algorithm," in *Applications of Artificial Intelligence in Engineering. Algorithms for Intelligent Systems*, 2021, pp. 17–35. doi: 10.1007/978-981-33-4604-8_2.

[26] H. Santoso, I. Hanif, H. Magdalena, and A. Afiyati, "A Hybrid Model for Dry Waste Classification using Transfer Learning and Dimensionality Reduction," JOIV : International Journal on Informatics Visualization, vol. 8, no. 2, p. 623, May 2024, doi: 10.62527/joiv.8.2.1943.

[27] R. Meiyanti, D. I. Sensuse, and Y. G. Sucahyo, "The Main Component of Smart Tourism: A Principal Component Analysis Approach," *Webology*, vol. 18, no. 2, pp. 295–307, 2021, doi:10.14704/web/v18i2/web18322.

[28] B. Priambodo, B. Jokonowo, Samidi, A. Ahmad, and R. A. Kadir, "Predict Traffic State Based on PCA-KMeans Clustering of neighbouring roads," in *International Visual Informatics Conference*, 2023, pp. 430–442.

[29] M. Akhtar and S. Moridpour, "A Review of Traffic Congestion Prediction Using Artificial Intelligence," *J. Adv. Transp.*, vol. 2021, 2021, doi: 10.1155/2021/8878011.

[30] S. Dong, "Multi class SVM algorithm with active learning for network traffic classification," *Expert Syst. Appl.*, vol. 176, no. February, p. 114885, 2021, doi: 10.1016/j.eswa.2021.114885.

[31] T. D. Toan and V. H. Truong, "Support vector machine for short-term traffic flow prediction and improvement of its model training using nearest neighbor approach," *Transp. Res. Rec.*, vol. 2675, no. 4, pp. 362–373, 2021, doi: 10.1177/0361198120980432.