Enhancing Traffic Video Data Collection from Compromised Views with Adaptive Segmentation

Suman Kumar, Tyler McKinley, Bhimashankar Malbhage
Department of Computer Science, Troy University, Troy, AL, USA, 36082
Email: {skumar, tmckinley, bmalbhage }@troy.edu

***Abstract*—The compilation of high-quality traffic data is crucial for real-time traffic video surveillance, significantly impacting the accuracy and reliability of Intelligent Transportation Systems (ITS). This paper addresses the challenges posed by environmental factors such as dirt, dust, and obstructions on camera lenses, which degrade image quality and subsequently affect machine learning model performance. We propose a robust view segmentation and data collection method that ensures only high-quality segments of camera views are utilized for analysis, thereby improving the accuracy and efficiency of ITS. Our scheme dynamically divides road frames into sectors and merges adjacent sectors based on their similarity. The effectiveness of the proposed method in maintaining high data integrity and ensuring reliable traffic analysis is demonstrated by showing how the algorithm divides the video view into appropriate sectors under various conditions, considering different algorithm parameter values. The proposed approach significantly enhances the overall performance of computer vision systems in ITS.**

***Index Terms*—Traffic Data Mining, Computer Vision, Intelligent Transportation, Image Segmentation, Pattern Recognition**

I. Introduction and Motivation

The data-driven system has increasingly become the cornerstone of how our world operates [1] and how we interact with it. In today's digital age, advances in machine learning and artificial intelligence are creating new possibilities for uncovering insights, solving complex problems, transforming industries [2], driving innovation [3], and opening up unprecedented opportunities [4]. High-quality data is essential for the accuracy and reliability of machine learning models. Such data ensures that analyses are meaningful, predictions are accurate, and decisions based on this data are sound. Without clean data, even the most sophisticated algorithms can produce misleading results [5], leading to poor decision-making.

Among different kinds of data, video data poses significant challenges, particularly in outdoor surveillance systems. The performance of computer vision systems heavily relies on high-resolution, clean images, which are often compromised by environmental factors. Cameras installed in outdoor environments are subject to dirt, dust, and other obstructions, degrading image quality and subsequently affecting the accuracy of machine learning models. Issues such as dust particles on the lens, dirty lenses due to environmental changes, and obstructed views caused by wind or other factors such as poly bag or paper stuck or object obstructed lead to bad data capture. Such views are shown in Figure 1. This variability in image quality negatively impacts machine learning performance [6]. Therefore, addressing environmental factors that degrade image quality is crucial for computer vision in outdoor surveillance systems [7].

(a) Clear (b) Poly Bag (c) Paper (d) Obstructed

Fig. 1: Clear and Compromised Views

The collection of high-quality traffic data is fundamental for real-time traffic video surveillance. Traffic congestion and other traffic-related issues can be mitigated by enhancing the efficiency of the current transportation system through advanced vehicular traffic data collection techniques for Intelligent Transportation Systems (ITS). The development of ITS [8] requires high-quality traffic data in real-time [9], where minimal preprocessing is required to minimize system response time. Traditional applications rely on manual monitoring or on-road sensors (e.g., inductive loops), which are often expensive and labor-intensive [10]. Advances in computer vision have provided opportunities to design automated traffic data collection systems that can improve the development of existing intelligent transport systems [11]. The efficiency of traffic data collection depends on selecting the optimum region where data is clear and noise-free.

We propose a robust image segmentation and data collection method that enhances data quality before further preprocessing steps are applied, thereby improving overall accuracy and efficiency. Our automatic method dynamically divides road frames into sectors and utilizes data similarity between sectors to optimize sector merging for effective traffic data collection. The proposed algorithm effectively addresses issues caused by environmental disturbances such as dust particles, obstructions like paper or polybags stuck on the camera, and sensor noise.

Our specific contributions are summarized as follows:

- Dynamic Sectorization: An algorithm to dynamically divide road frames into minimum sector lengths from the far to the near end of the view based on vehicle length and lane width.
- Merge Sectors: An algorithm to optimize data collection by merging similar adjacent sectors.

- Computation of Minimum Initial Sector Length: Minimum starting sector length found to be approximately 1.5 times the average car width.

The remainder of this paper is organized as follows: In Section II, we review related work on traffic data collection and image segmentation techniques. Section III details our proposed algorithms. Section IV demonstrates the effectiveness of our approach under various environmental conditions. Section V concludes the paper with future research directions.

II. Related Work

Recent advancements in image processing techniques have shown significant promise in addressing issues related to suboptimal image capture conditions. Techniques such as deep learning-based image restoration and enhancement can substantially improve the quality of images captured under adverse conditions [12, 13]. Additionally, adaptive data collection strategies can selectively capture high-quality segments of images, thereby enhancing the overall performance of computer vision systems [14, 15].

Existing methods in traffic data collection include color image-based adaptive background subtraction preprocessing for accurate vehicle detection, blob detection for measuring vehicle speed under high traffic conditions, and video stabilization techniques for tracking vehicles from aerial videos. A notable method for real-time vehicle detection and classification using adaptive techniques was proposed in [16]. This method relies on real-time data streaming, dynamically adjusting parameters to match changing traffic conditions and environmental factors, thereby ensuring high accuracy and reliability in vehicle statistics [17].

Furthermore, a comparative review of video-based vehicular traffic data collection methods from previous studies to the present highlights the evolution and improvement of these methods over the years. The review [18] discusses how advancements in artificial intelligence, machine learning methods, and sensor fusion techniques have led to increased precision and scope in traffic data collection. Recent efforts to address data quality challenges in ITS have proposed robust image segmentation and data collection methods aimed at improving detection accuracy and efficiency by utilizing high-quality segments of the image [19–21].

Our proposed approach differentiates itself from existing methods by facilitating data collection from the cleanest and most appropriate segments of the camera views, thereby enhancing data quality before further preprocessing steps. Unlike traditional methods, our algorithm dynamically divides road frames into optimal sector lengths based on vehicle size and lane width. This ensures high-resolution data collection even in variable conditions of camera positions. The data collection process is further optimized by dividing the views into subviews of varying quality by grouping adjacent sectors with similar traffic characteristics.

Algorithm 1: Dynamic Sectorization

Input : frame, start_point, end_point, Road_Attr
Output: Sectors created within the frame

Procedure `CreateSector` (*frame, start_point, end_point*)
  $Standard\_lane\_width = Read\_Lane\_Width(Road\_Attr)$;
  $Car\_width = Standard\_Card\_Width$;
  $length \leftarrow |end\_point - start\_point|$;
  $lane\_scale \leftarrow \frac{Standard\_Lane\_Width}{length}$;
  $car\_width\_scale \leftarrow \frac{Standard\_car\_width}{lane\_scale}$;
  while *height of sector* $<$ *end of frame coordinate* **do**
    $height\_of\_sector \leftarrow start\_point + 1.5 \times car\_width\_scale$;
    Draw sector;
    CreateSector(frame, start_point_of_sector, end_point_of_sector);

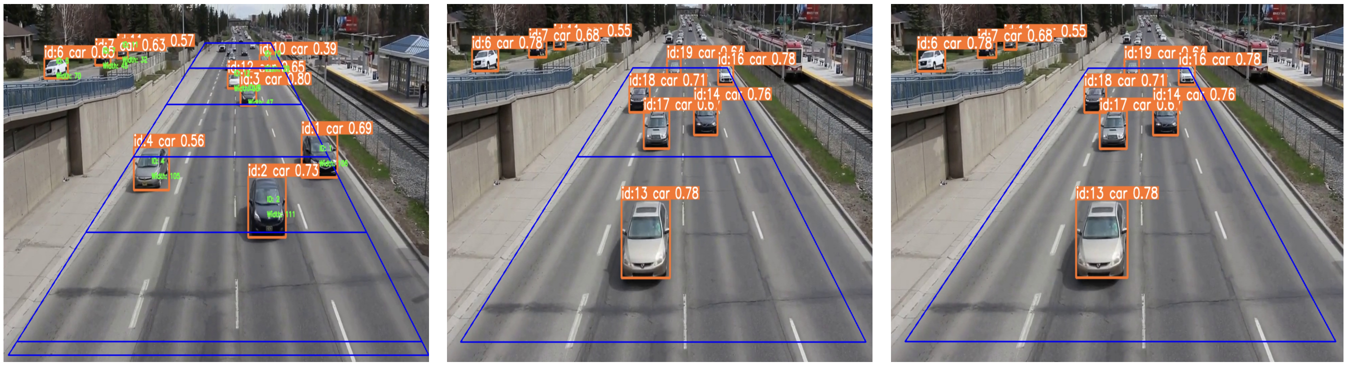
Algorithm 2: Merge Sectors

Input : car_count, vertices
Output: Merged arrays of car count and vertices

Procedure `MergeSectors` (*car_count, vertices*)
  Initialize merged_arr $\leftarrow$ [];
  Initialize merged_vertices $\leftarrow$ [];
  $i \leftarrow 0$;
  while $i <$ *len(car_count)* **do**
    if $Similarity(Adjacent\_Sectors) > Threshold$ **then**
      merged_arr.append((car_count[i] + car_count[$i+1$])/2);
      merged_vertices.append([vertices[i][0],
      vertices[i][1], vertices[i + 1][2], vertices[i + 1][3]]) i $\leftarrow i + 2$ ▷ Skip the next element if already merged;
    else
      merged_arr.append(car_count[i]);
      merged_vertices.append(vertices[i]);
      $i \leftarrow i + 1$;

III. Adaptive Segmentation Algorithm

This section presents two algorithms: Dynamic Sectorization (Algorithm 1) which aims to divide the road frame into smaller, manageable sectors based on the size of detected vehicles, and Merge Sectors (Algorithm 2), which merges adjacent sectors with similar traffic characteristics. These algorithms facilitate the utilization of high-quality segments of the image for ITS, optimizing data collection, reducing computational overhead, and maintaining high data integrity under varying conditions.



(a) Dynamic Sector Division (b) Count of vehicles in each sector (c) Merging of sectors

Fig. 2: Various stages of dynamic sectorization and merging.

A. Dynamic Sectorization

The Algorithm 1 takes the following parameters as input: the original video frame containing the road and vehicles, the starting point of the lane in the frame, the ending point of the lane in the frame, and attributes of the road, such as lane width. The road frame is split into sections, enabling division based on traffic flow and spatial layout. Lane detection algorithm is run to detect the lanes. It then read the standard lane width from the static data of the road. The working of the algorithm can be described in several steps.

First, the algorithm reads the standard lane width from the given road attributes and a standard car width based on typical vehicle dimensions. Next, the length of the lane between the start and end points is calculated. Using this length, the lane scale is computed by dividing a fixed lane width by the calculated length. The car width scale is then determined by dividing a fixed car width by the lane scale (line) .

The algorithm creates sectors within the frame (lines 7 - 10). The height of each sector is determined by adding 1.5 times the car width scale to the starting point. The explanation on 1.5 times width of car is explained in the result section. Sectors are drawn within the frame, and the process is recursively repeated for each new sector created. By dynamically adjusting the sector sizes based on the detected vehicle dimensions and merging similar adjacent sectors, the algorithm ensures efficient and high-quality traffic data collection.

B. Merge Sectors

The consistency of traffic patterns is evaluated by comparing the number of vehicles in adjacent sectors. Using traffic volume, the areas and traffic patterns are combined to create bigger, more coherent sectors for evaluation. The figure of merit for merging sector is chosen to be count of cars as a volume indicator since we want to observe as many vehicles as possible on the road.

The Algorithm 2 takes two primary inputs: an array of car counts in each sector and the vertices of these sectors. The output is a set of merged arrays for car counts and vertices, which represent the combined regions.

The algorithm starts by initializing two empty lists for storing the merged car counts and storing the merged sector vertices (lines 2-3). The algorithm runs until all sectors are processed. Inside the loop, it checks if the car counts in adjacent sectors are similar (i.e., their similarity exceeds a predefined threshold, see line 6). If the similarity condition is met, the algorithm calculates the average car count of the two adjacent sectors and appends this value to merged sector. It also merges the vertices of the two sectors and appends the new vertices to merged vertices (lines 7-9). If the car counts in adjacent sectors are not similar, the algorithm appends the car count and vertices of the current sector to their respective merged lists without any changes. The result is a set of merged sectors where each region has a more uniform car count, facilitating more efficient traffic data collection and analysis.

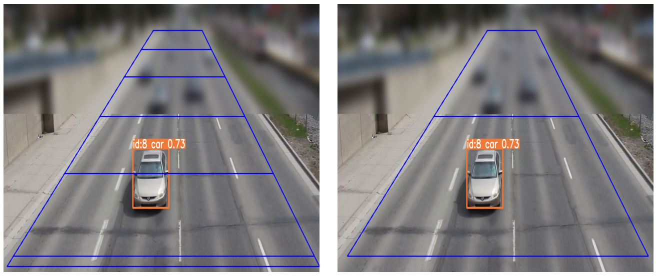
IV. Demonstration and Result

A. Parameter Setup and Implementation Environment

The object detection algorithm identifies vehicles in each frame of a video stream, leveraging the revolutionary You Only Look Once (YOLO) framework, which is renowned for its rapid object recognition capabilities. YOLO enhances efficiency by dividing an image into a grid and directly predicting bounding boxes and object classes from the network output. Our method employs YOLO for vehicle detection and uses OpenCV for reading video frames, image processing, and visualization. Geometrical computations and scale factors are applied to calculate the road width and standard car sizes, with the average lane width in the US being 12 feet and mid-size car width being 7 feet. This approach allows for dynamic responses to fluctuating road conditions. Additionally, a similarity threshold of 70% is used to determine when to merge adjacent sectors based on traffic data similarity, ensuring optimal data collection and analysis. The algorithm is ran for 5 minutes to ensure algorithm stabilizes and produces accurate result. However, the the duration depends on the traffic density.

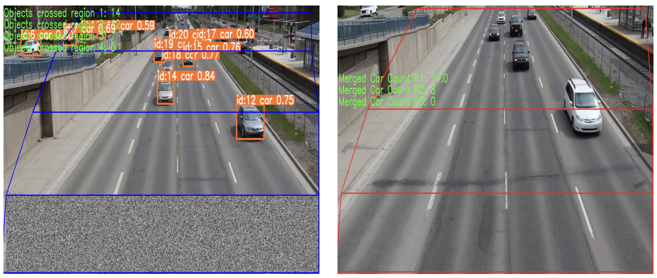
B. Demonstration of Algorithm

Dynamic sectorization breaks down the traffic scene into several sectors, simplifying the process of collecting a large



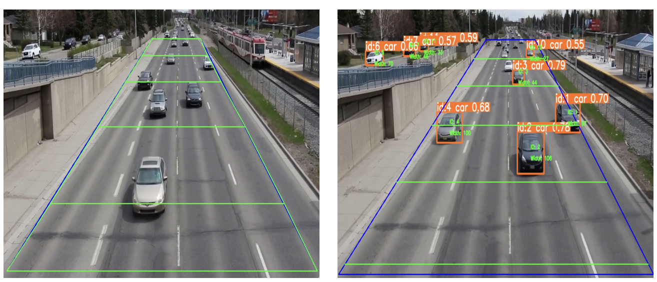
(a) Sector Division (b) Merging

Fig. 3: Paper or Polybag (blurring) on Camera view



(a) Sector Division (b) Merging

Fig. 4: View Obstruction



(a) Human Sector Division (b) Algorithm Sector Division

Fig. 5: Sector Division: Human Observation vs Algorithm

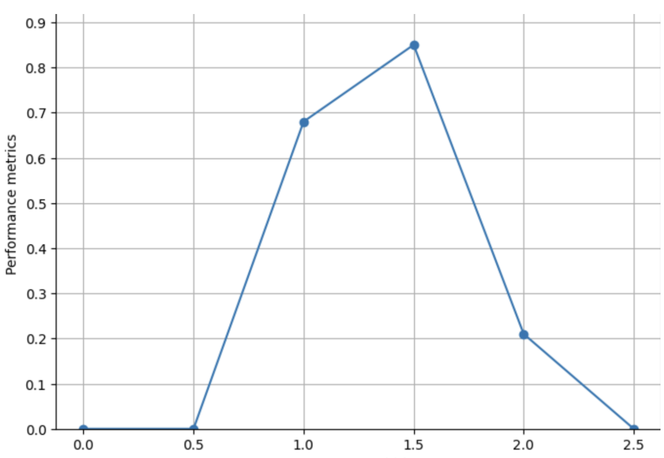


Fig. 6: Performance vs Segment Width $(XCar\_Width)$

amount of data. This is achieved by adjusting the borders of the sectors according to the patterns of traffic flow, ensuring accurate tracking of vehicle movements. Dynamic sectorization is particularly flexible in situations where traffic densities are not uniform. To ensure comprehensive coverage, sectors automatically adjust to accommodate increased traffic demand during peak hours. Figure 2 demonstrates the algorithm in action. Figure 2(a) illustrates the operation of Algorithm 1, where the road is divided into sectors based on the standard lane width and vehicle width. Figure 2(b) shows the intermediate stage of sector merging. Finally, Figure 2(c) depicts the merging of all sectors into one, as there is no noise obstruction and all sectors are similar in terms of the number of cars.

1) Paper/Polybag Stuck on Camera: This scenario simulates the situation when a paper or polybag is stuck to the camera lens, causing portions of the view to be blurry and unsuitable for data collection. This scenario is presented in Figure 3. Figure 3(a) shows the initial feed where dynamic sectorization is performed. The proposed method accurately separates the blurry sections from the non-blurry sections, as shown in Figure 3(b). This mechanism helps ITS prioritize sectors with clear views for data collection. Sectors with clear views and unaffected by blurring can be used to collect vehicle states, allowing for the development of a reliable system.

2) Environmental Obstruction: The case where there are partial obstructions or blockages in the view is shown in Figure 4. As shown in Figure 4(b), the obstructed view is clearly separated as segments that are similar to the segments which are part of the obstruction are combined to cover the obstructed area, separating the good area to collect the data from. In this case, the ITS can utilize the algorithm to prioritize the sector with the highest vehicle count, assuming it represents the region with the most significant traffic activity.

C. Evaluation of Pixel Accuracy

In evaluating the effectiveness of dynamic sectorization, pixel accuracy serves as a pivotal metric. For our research, we assessed the pixel accuracy between human-generated and model-generated sectorizations across a real-world traffic video dataset. Figure 5 illustrates the comparison between human observation-based sector division (Figure 5(a)) and algorithm-based sector division (Figure 5(b)). We asked students in our lab to observe the video and manually divide the image into several sectors. Subsequently, we performed the same task using the proposed algorithm and computed the pixel accuracy based on these sectors.

Our findings provided the following pixel accuracy results for the five sectors identified in the video: Sector 1: 87.17%, Sector 2: 82.48%, Sector 3: 85.67%, Sector 4: 61.49%, and Sector 5: 64.43%. Figure 6 shows the graph depicting the relationship between performance metrics and segment width, where segment width is proportional to the car width scale. Segment width is represented as a multiple of the car width scale. We observed that the performance metric peaks at 85% when the segment width is approximately 1.5 times the car width scale. This peak indicates the optimal segment width for maximum performance. Our observations suggest that performance is most similar to human assessment when the minimum sector width is 1.5 times the car length.

V. Conclusion

A robust method for enhancing traffic video data collection is presented by addressing the challenges posed by environmental factors such as polybag, paper stuck and other obstructions on camera lenses. The proposed dynamic view segmentation and data collection method enables ITS to utilze high-quality segments of camera views for analysis, thereby improving the accuracy and efficiency of the systems. By dynamically dividing road frames into sectors and merging adjacent sectors based on similarity, our approach maintains high data integrity and ensures reliable traffic analysis. The effectiveness of the method is demonstrated through its ability to accurately segment video views under various conditions and parameter settings, showcasing its adaptability and robustness. The proposed approach helps address the impact of environmental disturbances and optimizes the overall performance of computer vision systems in ITS.

Future work will focus on further refining the segmentation algorithm and exploring its application in other domains of computer vision and intelligent transportation. Additionally, integrating machine learning techniques and real-time data processing capabilities will be investigated for scalability and responsiveness. The performance of the algorithm on more complex situation of camera obstruction is in the works.

References

[1] Y. Fang, Z. Shan, and W. Wang, "Modeling and key technologies of a data-driven smart city system," *IEEE Access*, vol. 9, pp. 91 244–91 258, 2021.

[2] J. Wang, C. Xu, J. Zhang, and R. Zhong, "Big data analytics for intelligent manufacturing systems: A review," *Manufacturing Systems*, vol. 62, pp. 738–752, 2022.

[3] R. Majdzadeh, "Big data revolution: Transforming business landscapes through data-driven decision making," *Social Sc. Spectrum*, vol. 3, no. 1, pp. 115–125, 2024.

[4] Y. Zhan, P. Li, K. Wang, S. Guo, and Y. Xia, "Big data analytics by crowdlearning: Architecture and mechanism design," *IEEE Network*, vol. 34, pp. 143–147, 2020.

[5] A. Halevy, P. Norvig, and F. Pereira, "The unreasonable effectiveness of data," *IEEE intelligent systems*, vol. 24, no. 2, pp. 8–12, 2009.

[6] S. Minaee, Y. Boykov, F. Porikli, A. Plaza, N. Kehtarnavaz, and D. Terzopoulos, "Image segmentation using deep learning: A survey," *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 7, pp. 3523–3542, 2021.

[7] T. Brophy, D. Mullins, A. Parsi, J. Horgan, E. Ward, P. Denny, C. Eising, B. Deegan, M. Glavin, and E. Jones, "A review of the impact of rain on camera-based perception in automated driving systems," *IEEE Access*, 2023.

[8] S. Kaffash, A. T. Nguyen, and J. Zhu, "Big data algorithms and applications in intelligent transportation system: A review and bibliometric analysis," *Inter. journal of production economics*, vol. 231, p. 107868, 2021.

[9] Q. Liu, S. Kumar, and V. Mago, "Safernet: Safe transportation routing in the era of internet of vehicles and mobile crowd sensing," in *2017 14th IEEE Annual Consumer Communications Networking Conference (CCNC)*, Jan 2017, pp. 299–304.

[10] T. Gong, L. Zhu, F. R. Yu, and T. Tang, "Edge intelligence in intelligent transportation systems: A survey," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 9, pp. 8919–8944, 2023.

[11] A. Madani, S. Kumar, L. B. Nguyen, and J. Zhong, "A robust road region of interest identification scheme for traffic-video data mining," in *2019 International Conference on Computing, Networking and Communications (ICNC)*. IEEE, 2019, pp. 905–910.

[12] T. Azfar, J. Li, H. Yu, R. L. Cheu, Y. Lv, and R. Ke, "Deep learning-based computer vision methods for complex traffic environments perception: A review," *Data Science for Transportation*, vol. 6, no. 1, p. 1, 2024.

[13] A. Dumka, A. Ashok, P. Verma, and P. Verma, *Advanced digital image processing and its applications in big data*. CRC Press, 2020.

[14] H. Idrees, M. Shah, and R. Surette, "Enhancing camera surveillance using computer vision: a research note," *Policing: An International Journal*, vol. 41, no. 2, pp. 292–307, 2018.

[15] Z. Zha, X. Yuan, B. Wen, J. Zhang, J. Zhou, and C. Zhu, "Image restoration using joint patch-group-based sparse representation," *IEEE Transactions on Image Processing*, vol. 29, pp. 7735–7750, 2020.

[16] J. Doe and J. Smith, "An adaptive video-based vehicle detection, classification, counting, and speed-measurement system for real-time traffic data collection," *IEEE Transactions on Intelligent Transportation Systems*, 2019.

[17] S. Su, Q. Yan, Y. Zhu, C. Zhang, X. Ge, J. Sun, and Y. Zhang, "Blindly assess image quality in the wild guided by a self-adaptive hyper network," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 3667–3676.

[18] R. Roe and A. Johnson, "Comparative review on video based vehicular traffic data collection for intelligent transport system," *IEEE Transactions on Intelligent Transportation Systems*, 2017.

[19] L. Ke, M. Ye, M. Danelljan, Y.-W. Tai, C.-K. Tang, F. Yu *et al.*, "Segment anything in high quality," *Advances in Neural Information Processing Systems*, vol. 36, 2024.

[20] I. Castiglioni, L. Rundo, M. Codari, G. Di Leo, C. Salvatore, M. Interlenghi, F. Gallivanone, A. Cozzi, N. C. D'Amico, and F. Sardanelli, "Ai applications to medical images: From machine learning to deep learning," *Physica medica*, vol. 83, pp. 9–24, 2021.

[21] R. A. Manap, L. Shao, and A. F. Frangi, "Patch-iq: a patch based learning framework for blind image quality assessment," *Info. Sciences*, vol. 420, pp. 329–344, 2017.