Image Quality Assessment for Computer Vision based Perception Algorithms using Edge and Structural Features

V G Akshaya, Rithika S, Lanciya Angel A,
K S M Amirtha Pravin, Karthi Balasubramanian*
Department of Electronics and Communication Engineering
Amrita School of Engineering, Coimbatore,
Amrita Vishwa Vidyapeetham, India
*b_karthi@cb.amrita.edu

Suganthi Srinivasan, Rakesh Rajegowda,
Payal Bhattacherjee
Comfort Driving Assistance Group,
R & D-Computer Vision Systems,
Valeo India Pvt.ltd, Chennai
suganthi.srinivasan@valeo.com,
rakesh.rajegowda@valeo.com,
payal.bhattacherjee@valeo.com

TENCON 2023 - 2023 IEEE Region 10 Conference (TENCON) | 979-8-3503-0219-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/TENCON58879.2023.10322345

***Abstract*—Image Quality Assessment (IQA) is a critical task in image processing, computer vision, and other related fields, as it helps in evaluating the effects of various impairments on the Quality of Experience (QoE) of consumers. However, current IQA metrics may have limitations in evaluating image quality accurately for different types of images or under different conditions. To address this, a proposed approach involves calculating an overall image quality score based on linear combination of individual image quality metrics. This method considers multiple image features such as sharpness, contrast, color, and noise, and assigns weights to each feature based on its relative importance in determining overall image quality. By combining and weighting multiple image features, the proposed approach aims to provide a more comprehensive and accurate evaluation of image quality compared to using individual metrics alone. In this study, we focus on the importance of edge features and structure-based metrics in object detection and analyze existing No-Reference (NR) IQA metrics, including Just Noticeable Blur (JNB), Cumulative Probability Blur Detection (CPBD), Visual Quality Assessment (VQA), Blind/Referenceless Image Spatial Quality Evaluator (BRISQUE) and No-Reference Low- Light Image Enhancement Evaluation (NLIEE), and propose a linear combination formula that combines these metrics to evaluate image. We evaluate the performance of the proposed metric on several datasets by calculating the linear combination of two, three, and all the five metrics. The weights assigned to each metric in the linear combination formula are determined through experimental analysis. The combined metric scores and MOS scores are then used to compute the Spearman's rank order correlation coefficient, which measures the monotonic relationship between two variables. A higher SROCC value indicates a stronger correlation between the combined metric and MOS scores and is used to evaluate the performance of the metric. Our results show that the proposed metric, which combines VQA, BRISQUE and JNB , provides the highest correlation with MOS scores, over all the datasets. This suggests that the proposed approach can provide an effective and comprehensive way of evaluating image quality in object detection algorithms. Our findings can have significant implications in improving the overall performance of object detection algorithms and enhancing the QoE of consumers.**

***Index Terms*—Perception, IQ metric, IQ analysis, JNB, CPBD, VQA, BRISQUE, NLIEE**

I. INTRODUCTION

Recent studies by the World Health organization (WHO) has indicated that road traffic accidents are the cause of more than 1.3 million deaths and close to 50 million injuries every year [1]. ADAS (Advanced Driver-Assistance Systems) and autonomous vehicles have gained significant importance in the progression towards autonomous driving, primarily due to their role in enhancing passenger safety. In such systems, perception algorithms play a vital role through which the complete environment is perceived. Vision sensors are primarily used in extracting the surrounding information. In general, computer vision algorithms uses the features such as edges, structures etc., in the captured images for efficient training in order to achieve the desirable performance [2]–[4]. The same algorithm may result in degraded performance when such features are not prominent in a image. For example, a different imager tuned to different IQs, may result in degraded performance which need to be retrained to achieve the same performance. Therefore, Image Quality Assessment (IQA) of computer vision algorithms are imperative to ensure the performance of perception algorithms such as objection detection.

The significance of Image Quality Assessment (IQA) extends to a diverse range of multidisciplinary topics, including image and signal processing, computer vision, information theory, machine learning, and the design of image acquisition, communication, and display systems. During various stages of processing and communication, distortions are often introduced, and their visibility can significantly affect the Quality of Experience (QoE) of end-users. Therefore, it is important for computer vision systems to have a means of evaluating the impact of such distortions on visual quality. Towards that goal, several IQA metrics have been extensively studied and developed over the past few decades [5], [6]. These metrics can be categorized into three types based on the amount of reference information required from the original source. These include full-reference, reduced-reference, and no-reference metrics. Full-reference (FR) metrics use the complete reference information, while reduced-reference (RR) metrics use certain features which are extracted from the reference information to calculate the quality score. On the other hand, no-reference (NR) metrics do not rely on any reference information, making them the most suitable for applications especially for situations where the origin source is not available [7]. This work focuses on the use of NR-IQA for deriving the image quality metric of an efficient computer vision algorithm.

NR-IQA metrics offer a convenient and efficient means of evaluating image quality in diverse applications [8]. As the performance of most of the object detection algorithms depends on the important features present in the images such as edges, structures, sharpness and contrast, image quality can be measured based on the presence of these features in the images. The IQ metric are grouped as structure-based, edge-based, feature- based and training based metrics. Some examples of these are given below:

1) Structure-based metrics:
   - NLIEE [9]
   - TDMEC [10]
   - VQA [11]
   - BLIND IQA [12]
2) Edge-based metrics:
   - JNB [13]
   - CPBD [14]
3) feature-based metrics:
   - NO REFERENCE SHARPNESS METRIC [15]
   - BRISQUE [16]
   - BNB METRIC [17]
   - NR METRIC [18]
4) Training based metrics
   - Hallucinated-IQA [19]

These metrics are limited in evaluating the image quality accurately with respect to the specified features. Any object detection approaches depends on one or more combination of features present in the images. It is important that all relevant features be retained in the image in order to achieve the expected performance of the perception algorithms. For example there are instances where existing metrics exhibit good performance for natural images but may have difficulty in evaluating images that have undergone processing. To overcome these limitations, a combination of metrics can be advantageous as they can consider a diverse set of image features and can provide a more robust evaluation of image quality across different conditions [20]–[23]. These studies have been done on full Reference (FR-IQA) metrics and have used methods including Support Vector Regressor (SVR), metric multiplication with optimized parameters and simulated annealing. In this work, five no reference IQ metrics that include JNB, CPBD, NLIEE, BRISQUE and VQA have been chosen considering the sharpness, contrast, color and noise as important features for analysis. A new metric is proposed by the linear combination of these metrics. It can be further tuned for any specific applications or preferences by adjusting the weights assigned to each feature. By combining and weighting multiple image features, the proposed approach aims to provide a more comprehensive and accurate evaluation of image quality compared to using individual metrics alone. The proposed metric is evaluated based on the performance of an object detection algorithm. The performance of the algorithm and thereby the IQ of the images using KITTI dataset is considered to be the gold standard. The individual metrics and proposed combined metric are evaluated for NuScenes dataset and compared against the performance using KITTI dataset. Also, it has been evaluated after enhancing the images in NuScenes dataset.

The paper is organized as follows. Section II describes the methodology used for the analysis including an overview of the chosen IQA metrics and the image datasets while section III explores the proposed metric. The results are presented in section IV and the paper concludes in Section V with discussions and future possible directions.

II. Materials and Methods

In this section, the description of five IQ metrics and object detection algorithm are presented. Also, the details of the dataset used for evaluating the performance of the proposed IQ metric are presented.

NR-IQA metrics

A brief description of the metrics considered for analysis is given below.

JNB: It provides an objective measure of image sharpness/blurriness and is based on the perception of Just Noticeable Blur (JNB). It is computed as :

$$S = L/D \tag{1}$$

where, L is the total number of processed blocks in the image and D is the perceived blur distortion measure given by equation 2, where β is a parameter chosen for least-square fitting analysis and R_b is the perceived blur distortion within an edge block.

$$D = \sum_{R_b} \left(|DR_b|^{\beta}\right)^{\frac{1}{\beta}} \tag{2}$$

CPBD: Cumulative Probability Blur Detection (CPBD) is an image analysis technique that detects blur by evaluating the sharpness of edges and their likelihood of being affected by blurring. It involves initial edge detection followed by the assessment of blur detection probability at each identified edge. The local contrast of edges is analyzed to distinguish between blurry and sharp edges. The probability of blur detection is determined based on edge contrast, orientation, and the likelihood of blur impact.

$$P_{blur}(e_i) = 1 - e^{w(e_i)/w_{JNB}(e_i)} \tag{3}$$

where, $w(e_i)$ is the width of edge e_i, $w_{JNB}(e_i)$ which depends on JNB width, the local contrast C and β and is obtained by means of least-square fitting. A probability density function is computed as a normalized histogram, and the cumulative probability of blur detection is derived from it. The proposed CPBD measure uses a threshold value to determine the level of blur in an image: if the cumulative probability exceeds the threshold, the image is considered blurry, while a lower value indicates sharpness.

$$CPBD = P_{(P_{blur} \leq P_{JNB})} = \sum_{P_{blur}=0}^{P_{blur}=P_{JNB}} P_{P_{blur}} \tag{4}$$

where, P_{JNB} denotes the probability of Just Noticeable Blur and $P_{P_{blur}}$ is the value of the probability distribution function at a given P_{blur}.

VQA: Most existing Visual Quality Assessment (VQA) techniques are designed for natural photos and struggle to accurately predict visual quality for synthetic compound images (SCIs), which contain diverse visual content. To address this, statistical luminance and texture features are extracted from SCIs using techniques like local normalization, local binary patterns (LBPs), and Sobel filters. Support vector regression (SVR) is then employed to map these features to human subjective assessments.

BRISQUE: BRISQUE (Blind/Referenceless Image Spatial Quality Evaluator) is an objective Image Quality Assessment (IQA) metric that evaluates image quality without a reference image. BRISQUE operates in two stages, extracting Natural Scene Statistics (NSS) features that capture statistical properties of the image and mapping these features to a quality score using a Support Vector Regression (SVR) model. The SVR model is trained on a set of images rated by human observers. BRISQUE eliminates the need for a reference image, captures comprehensive visual properties through NSS features, and provides accurate evaluations.

NLIEE : This metric utilizes two sets of features to objectively evaluate the enhanced image. The first set of features captures statistical properties of the image, while the second set focuses on luminance characteristics. These features are extracted using specific techniques, and a support vector regression (SVR) model is used to predict the quality score.

Dataset

For the image quality analysis two datasets namely KITTI [24] [25] [26] [27] and NuScenes [28] were used for evaluation. These datasets consists of real world free drive captures that are most widely used for perception algorithm development and validation exclusively for ADAS and AV applications. The capture vehicle of KITTI dataset include stereo cameras and LiDARs whereas NuScenes include 1 LiDAR, 5 Radars, 6 cameras, GPU and IMU. In this work, only Front view camera images are collected from both the datasets. In total, around 400 images from each KITTI and NuScenes dataset are collected at various environments such as clear sky, cloudy, dark and daylight conditions.

A. *Object detection algorithm*

To evaluate the performance of the IQA metrics over the different datasets, You Only Look Once (YOLOV5) algorithm is employed. It is a commonly used one-stage object detection technique that apportions images into a grid, where each cell in the grid has the responsibility of detecting objects within its own area [29].

III. COMBINED METRIC FOR IMAGE QUALITY ASSESSMENT

The quality of an image can be influenced by various factors such as brightness, contrast, sharpness, noise, and color. In an image with all these image characteristics, an individual metric to define the specific feature may not be sufficient. The performance of any computer vision algorithm depends on the presence of almost all features. Towards this endeavor, a combined IQ metric based on the linear combination of individual IQA metrics is proposed as given in equation 5.

$$Combined_{IQ} = \sum_{i-1}^{n} w_i IQ_{mi} \tag{5}$$

where $IQ_{m1}, IQ_{m2}, ..., IQ_{mn}$ are the different NR IQA metrics that are considered and $w_1, w_2, ..., w_n$ are the corresponding weights. The weights can be tuned according to the expected performance of the object detection algorithm. In this work, the weights are tuned according to the image quality of KITTI dataset. This approach seems to be more comprehensive and effective way of evaluating image quality in object detection algorithms.

The performance of the combined metrics is assessed using the images from the databases KITTI & NuScenes. The combined metric scores and Mean Opinion Scores (MOS) are then used to compute the Spearman's rank order correlation coefficient (SROCC) that calculates the monotonic relationship between two variables. Linear combination of two, three, and all five metrics are evaluated and correlated with the Mean Opinion (MOS) score. A higher SROCC value indicates a stronger correlation between the combined metric and MOS scores, and is thus used to evaluate the metric's performance. The weights of each metric is chosen to maximize the correlation.

IV. RESULTS AND ANALYSIS

In this section, the analysis on the results obtained using individual IQ metrics and proposed combined IQ metric are presented. The evaluation is performed considering an object detection algorithm using YOLOv5 architecture. The network is trained using the images from KITTI dataset and evaluated using NuScenes dataset separately.

The object detection performance are tabulated in Table I from which it can be observed that the precision, recall and F1-score values are better for KITTI dataset compared to NuScenes dataset.

TABLE I: Performance of YOLOv5 - Object detection algorithm

Performance metric	KITTI	NuScenes
Precision	0.81	0.51
Recall	0.57	0.29
F1-Score	0.67	0.378

The network is trained using KITTI dataset and hence, when evaluated on the images from different imager (NuScenes), the performance of the network is degraded. Here, the network needs to be re-trained using NuScenes images to improve the detection performance. It will be challenging to undergo the process of capturing of images and re-training cycle whenever there is a change/up-gradation in the imager, especially for vehicle deployed perception systems. A simple approach could be to define an IQ metric for the images that are used for training. Whenever there is a change in the imager, its IQ parameters can be tuned to the same value in order to retain the detection performance. Therefore, there is a need to define a IQ metric as proposed in this work.

Around 400 images each from KITTI and NuScenes datasets are collected as a set of images from two different imagers. It includes images captured at different lighting conditions such as daylight, dark, cloudy and artificial lights.

TABLE II: SROCC Values of individual IQ metric for KITTI Images

IQA Metric	SROCC
JNB	0.8476
CPBD	0.4479
NLIEE	0.0928
VQA	0.6351
BRISQUE	0.6592

All five metrics are computed for each of the images and the average metric value is calculated for each imager. The calculated individual IQ metric values are shown in Fig. 1.

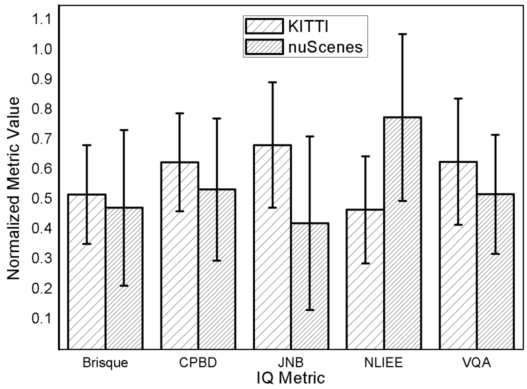

Fig. 1: Bar Chart representation of individual IQ metrics using KITTI & NuScenes images

It can be seen that JNB, CPBD, BRISQUE, and VQA resulted in better performance for the imager used in KITTI dataset whereas NLIEE shows better for NuScenes. It indicates that the KITTI images have better sharpness, low blur content and good texture & edge features that resulted in better object detection performance. The difference in NLIEE metric value may be due to less samples in low light scenarios in KITTI compared to NuScenes or the statistical features present during low light condition are not prominent. As the individual metrics are not good enough to define the imager characteristics which is already tuned to some IQ value, a combined IQ metric is proposed and evaluated.

The proposed metric is arrived at by analyzing the SROCC score between all possible combinations of metrics and the corresponding MOS score over the KITTI images. MOS is obtained by scoring each image used in the evaluation visually by around 50 people. The images are evaluated subjectively based on the presence of edge information, contrast, blur and sharpness on a scale of 1 (poor) to 10 (good) by each individual. Average MOS score is computed for each image used in the evaluation.

Tables II - V presents the SROCC score obtained between MOS and individual metrics, linear combination of two, three and all, five IQA metric values respectively for KITTI images. Weights for each metric value is obtained on trial and error method. The combination for which the maximum correlation is achieved are listed in the tables.

It is observed that in almost all combinations and as individual metric value, JNB, VQA and BRISQUE shows the highest correlation with MOS scores.

TABLE III: SROCC Values of linear combination of two metrics for KITTI Images

w1	w2	IQA1	IQA2	SROCC
0.5	0.5	CPBD	BRISQUE	0.6767
0.5	0.5	NLIEE	BRISQUE	0.4842
0.6	0.4	NLIEE	VQA	0.4506
0.4	0.6	NLIEE	CPBD	0.4205
0.5	0.5	NLIEE	JNB	0.6491
0.6	0.4	CPBD	JNB	0.7703
0.6	0.4	CPBD	VQA	0.799
0.6	0.4	JNB	VQA	0.8761
0.5	0.5	JNB	BRISQUE	0.8601
0.4	0.6	BRISQUE	VQA	0.8332

TABLE IV: SROCC Value of linear combination of three metrics for KITTI Images

w1	w2	w3	IQA1	IQA2	IQA3	SROCC
0.4	0.4	0.2	CPBD	JNB	BRISQUE	0.8732
0.4	0.3	0.3	CPBD	JNB	VQA	0.9393
0.4	0.3	0.3	CPBD	JNB	NLIEE	0.7201
0.4	0.3	0.3	JNB	VQA	BRISQUE	0.9433
0.3	0.4	0.3	JNB	VQA	NLIEE	0.7867
0.4	0.3	0.3	VQA	BRISQUE	NLIEE	0.7421
0.4	0.2	0.4	VQA	BRISQUE	CPBD	0.9251
0.3	0.2	0.5	BRISQUE	NLIEE	CPBD	0.6058
0.3	0.3	0.4	BRISQUE	NLIEE	JNB	0.7733
0.2	0.5	0.3	NLIEE	CPBD	VQA	0.7556

It indicates that the KITTI images have prominent sharpness, texture and statistical features. Table VI shows the summary of the maximum SROCC value obtained from individual and all the possible combinations. For KITTI images, it resulted in maximum correlation for the linear combination of JNB, VQA and BRISQUE metrics.

Table VII shows the combined IQ metric evaluated for NuScenes images and compared against the KITTI images.

It can be seen that the combined IQ metric value for KITTI images are high compared to NuScenes images. This indicates that the sharpness, texture and statistical features that are present in NuScenes images may not be sufficient to achieve the equivalent detection of performance which was achieved using KITTI images. This shows that there is need to further tune the IQ setting in NuScenes imager in order to achieve the same or equivalent detection performance.

The observed results on combined IQ metric for NuScenes images are further validated on NuScenes enhanced images. NuScenes images are subjected to simple image enhancement techniques such as sharpness, contrast and blur removal. This is carried out in order to replicate the process of tuning IQ setting from an imager perspective. This may not represent the exact method of IQ setting happening from the hardware setup, but nevertheless it is being used for showing the proof of concept. After enhancing the NuScenes images, the combined IQ metric is generated and compared against the NuScenes and KITTI images which is shown in Fig. 2.

The proposed combined IQ metric seems to show improved

TABLE V: SROCC Value of linear combination of five metrics for KITTI Images

w1 (JNB)	w2 (CPBD)	w3 (NLIEE)	w4 (BRISQUE)	w5 (VQA)	SROCC
0.2	0.2	0.2	0.2	0.2	0.9278

TABLE VI: Summary of different linear combination of metrics resulting in maximum correlation for KITTI Images

IQA Metric	SROCC Value
JNB	0.8476
JNB and VQA	0.8761
JNB,VQA and BRISQUE	0.9433
JNB, CPBD, VQA,NLIEE and BRISQUE	0.9278

TABLE VII: Proposed Combined IQ metric value evaluated for KITTI and NuScenes Images

w1	w2	w2	IQA1	IQA2	IQA3	KITTI	NuScenes
0.4	0.3	0.3	JNB	VQA	BRISQUE	0.62 ± 0.16	0.46 ± 0.11

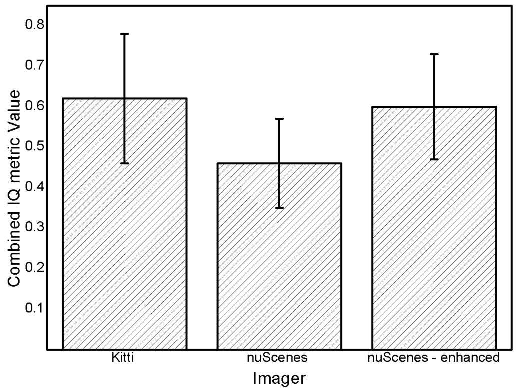

Fig. 2: Bar Chart representation of Proposed Combined IQ metrics using KITTI, NuScenes & NuScenes enhanced images

value when NuScenes images are subjected to image quality enhancement techniques. This implies that the necessary features similar to KITTI images has been improved in NuScenes images also. To validate further, the performance of the object detection algorithm using NuScenes enhanced images are also generated and the comparative performance analysis of KITTI, NuScenes and NuScences enhanced is shown in Fig. 3.

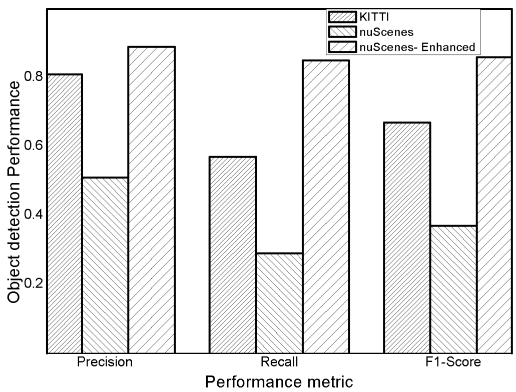

Fig. 3: Bar Chart representation of Object detection performance using KITTI, NuScenes and NuScenes-enhanced images

This shows that the detection performance is improved using NuScenes-enhanced images. This also indicates that the detection algorithm performance is dependent on the features present in the images. Therefore, It indicates that when a combined IQ metric value is defined for imagers, it is possible to ensure the performance of the any computer vision algorithms irrespective of the imagers.

V. Conclusions and Discussions

Computer vision algorithms used in order to perceive the surrounding environment especially in automotive applications demands to be retrained whenever there is change in the imager characteristics. Therefore, there is a need to define image quality metrics for perception algorithms. In this work, an attempt is made to derive a combined IQ metric as a linear combination of IQ metrics representing sharpness, texture and statistical features present in the images. The performance of the proposed metric is evaluated using the images from KITTI and NuScenes dataset. The same has been validated using the performance achieved using an object detection algorithm after enhancing the NuScenes images. The results show that the proposed combined IQ metric is able to identify the main features present in the KITTI images. After enhancing the NuScenes images, the object detection performance has been improved to 89% which indicates that the proposed combined IQ metric is effective. Also, it will be helpful to further fine tune the imager based on the expected algorithm performance. The proposed metric is achieved based on trial and error method. Therefore, future scope of this work will be focused to further generalize the derivation of the weights & metric and validate the same using the imagers tuned to actual IQ parameter settings

References

[1] Road traffic injuries — who.int. https://www.who.int/news-room/fact-sheets/detail/road-traffic-injuries. [Accessed 18-May-2023].

[2] AK Sandeep, S Nithin, and KI Ramachandran. An image processing based pedestrian detection system for driver assistance. *2016 International journal of Control theory and Application*, pages 7369–7375, 2016.

[3] S Raviteja and R Shanmughasundaram. Advanced driver assistance system (ADAS). In *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, pages 737–740. IEEE, 2018.

[4] Saksham Srivastava, S Adarsh, Binoy B Nair, and KI Ramachandran. Driver's face fetection in poor illumination for ADAS applications. In *2021 5th International Conference on Computer, Communication and Signal Processing (ICCCSP)*, pages 1–6. IEEE, 2021.

[5] Damon M Chandler. Seven challenges in image quality assessment: past, present, and future research. *International Scholarly Research Notices*, 2013, 2013.

[6] Hima T Suseelan, V Sowmya, and KP Soman. Image dehazing using variational mode decomposition. In *2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, pages 200–205. IEEE, 2017.

[7] Zhou Wang and Alan C Bovik. Reduced-and no-reference image quality assessment. *IEEE Signal Processing Magazine*, 28(6):29–40, 2011.

[8] Wufeng Xue, Lei Zhang, and Xuanqin Mou. Learning without human scores for blind image quality assessment. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 995–1002, 2013.

[9] Zicheng Zhang, Wei Sun, Xiongkuo Min, Wenhan Zhu, Tao Wang, Wei Lu, and Guangtao Zhai. A no-reference evaluation metric for low-light image enhancement. In *2021 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6. IEEE, 2021.

[10] Karen Panetta, Arash Samani, and Sos Agaian. A robust no-reference, no-parameter, transform domain image quality metric for evaluating the quality of color images. *IEEE Access*, 6:10979–10985, 2018.

[11] Ke Gu, Jun Zhou, Jun-Fei Qiao, Guangtao Zhai, Weisi Lin, and Alan Conrad Bovik. No-reference quality assessment of screen content pictures. *IEEE Transactions on Image Processing*, 26(8):4005–4018, 2017.

[12] Qiaohong Li, Weisi Lin, and Yuming Fang. No-reference image quality assessment based on high order derivatives. In *2016 IEEE International Conference on Multimedia and Expo (ICME)*, pages 1–6. IEEE, 2016.

[13] Rony Ferzli and Lina J Karam. A no-reference objective image sharpness metric based on the notion of just noticeable blur (JNB). *IEEE transactions on image processing*, 18(4):717–728, 2009.

[14] Niranjan D Narvekar and Lina J Karam. A no-reference image blur metric based on the cumulative probability of blur detection (CPBD). *IEEE Transactions on Image Processing*, 20(9):2678–2683, 2011.

[15] Nabil G Sadaka, Lina J Karam, Rony Ferzli, and Glen P Abousleman. A no-reference perceptual image sharpness metric based on saliency-weighted foveal pooling. In *2008 15th IEEE International Conference on Image Processing*, pages 369–372. IEEE, 2008.

[16] Anish Mittal, Anush Krishna Moorthy, and Alan Conrad Bovik. No-reference image quality assessment in the spatial domain. *IEEE Transactions on image processing*, 21(12):4695–4708, 2012.

[17] Ruigang Fang and Dapeng Wu. No-reference image quality assessment based on BNB measurement. In *2013 IEEE China Summit and International Conference on Signal and Information Processing*, pages 528–532. IEEE, 2013.

[18] Zhang Taojia, Zhang Hong, and Yuan Ding. A no-reference quality metric for blur image. In *2013 IEEE International Conference on Green Computing and Communications and IEEE Internet of Things and IEEE Cyber, Physical and Social Computing*, pages 1813–1816. IEEE, 2013.

[19] Kwan-Yee Lin and Guanxiang Wang. Hallucinated-IQA: No-reference image quality assessment via adversarial learning. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 732–741, 2018.

[20] Domonkos Varga. Full-reference image quality assessment based on an optimal linear combination of quality measures selected by simulated annealing. *Journal of Imaging*, 8(8):224, 2022.

[21] Krzysztof Okarma and Mateusz Kopytek. Improved combined metric for automatic quality assessment of stitched images. *Applied Sciences*, 12(20):10284, 2022.

[22] Vladimir V Lukin, Nikolay N Ponomarenko, Oleg I Ieremeiev, Karen O Egiazarian, and Jaakko Astola. Combining full-reference image visual quality metrics by neural network. In *Human Vision and Electronic Imaging XX*, volume 9394, pages 172–183. SPIE, 2015.

[23] Anustup Choudhury and Scott Daly. Combining quality metrics for improved hdr image quality assessment. In *2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR)*, pages 179–184. IEEE, 2019.

[24] Andreas Geiger, Philip Lenz, and Raquel Urtasun. Are we ready for autonomous driving? the KITTI vision benchmark suite. In *2012 IEEE conference on computer vision and pattern recognition*, pages 3354–3361. IEEE, 2012.

[25] Andreas Geiger, Philip Lenz, Christoph Stiller, and Raquel Urtasun. Vision meets robotics: The KITTI dataset. *The International Journal of Robotics Research*, 32(11):1231–1237, 2013.

[26] Yiyi Liao, Jun Xie, and Andreas Geiger. KITTI-360: A novel dataset and benchmarks for urban scene understanding in 2d and 3d. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[27] Jens Behley, Martin Garbade, Andres Milioto, Jan Quenzel, Sven Behnke, Cyrill Stachniss, and Jurgen Gall. Semantickitti: A dataset for semantic scene understanding of lidar sequences. In *Proceedings of the IEEE/CVF international conference on computer vision*, pages 9297–9307, 2019.

[28] Holger Caesar, Varun Bankiti, Alex H Lang, Sourabh Vora, Venice Erin Liong, Qiang Xu, Anush Krishnan, Yu Pan, Giancarlo Baldan, and Oscar Beijbom. Nuscenes: A multimodal dataset for autonomous driving. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 11621–11631, 2020.

[29] Glenn Jocher, Ayush Chaurasia, Alex Stoken, Jirka Borovec, Yonghye Kwon, Kalen Michael, Jiacong Fang, Zeng Yifu, Colin Wong, Diego Montes, et al. ultralytics/yolov5: v7.0-YOLOv5 sota realtime instance segmentation. *Zenodo*, 2022.