2024 International Conference on Computer Communication and Informatics (*ICCCI*), Jan. 23 – 25, 2024, Coimbatore, INDIA

A Comprehensive Comparative Analysis BetweenTEED and DexiNed for Enhanced Image Edge Detection

Sanket Kothari
School of CSE
KLE Technological University
Hubli, India
sanketkothari1925@gmail.com

Dhruv Ostawal
School of CSE
KLE Technological University
Hubli, India
dhruvostawal@gmail.com

Anupama P Bidargaddi
School of CSE
KLE Technological University
Hubli, India
anupamapb@gmail.com

Saisamarth Udikeri
School of CSE
KLE Technological University
Hubli, India
saisamarthudikeri@gmail.com

Yoginath G Lakhaman
School of CSE
KLE Technological University
Hubli, India
ylakhaman13@gmail.com

***Abstract*—This paper presents a comparative analysis between two advanced edge detection models Tiny Efficient Edge Detection (TEED) and Dense Extreme Inception Network (DexiNed). The study evaluates the models across diverse image datasets using quantitative metrics like Mean Squared Error, Peak Signal to Noise Ratio, Structural Similarity Index, Non shift Edge Based Ratio, Edge-Based Image Quality Assessment, and Sobel-based Reduced Reference method. It examines the architectural designs of TEED and DexiNed, highlighting differences in components like the upsampling network, loss functions, and skip connections. While TEED prioritizes efficiency with only 58K parameters, DexiNed focuses more on maximizing accuracy. The analysis reveals strengths and limitations of both models across factors like speed, precision, adaptability to realtime applications, and generalization capabilities. Quantitative results showed TEED achieving lower error metrics like MSE and higher edge preservation scores like NSER compared to DexiNed. However, DexiNed produced visually sharper edges due to its multi-scale processing strategy. The average EBIQA values for TEED and DEXINED in our dataset, which stand at 0.3304 and 0.5717, respectively. Additionally, the average NSER values for TEED and DEXINED are reported as 13.028 and 9.159. The insights from this comparative study can inform the selection and future development of edge detection techniques for computer vision tasks.**

***Index Terms*—Edge detection, model comparison, dual loss strategy, mean squared error (MSE), Peak Signal to Noise Ratio (PSNR), Structural Similarity Index (SSIM), Non-shift Edge Based Ratio (NSER), Edge Based Image Quality Assessment (EBIQA), Tiny and Efficient Edge Detector (TEED), Dense Extreme Inception Network for Edge Detection (DexiNed), Computer Vision Image Processing.**

I. INTRODUCTION

Edge detection is a critical element in tasks such as image segmentation, object recognition, and scene analysis. The emergence of deep learning has given rise to sophisticated neural architectures for edge detection, as seen in designs like the tiny and efficient edge detector (TEED) [10] and the dense extreme inception network (DEXINED) [12]. This study conducts a comparative analysis of these models, thoroughly examining their capabilities and scrutinizing their contributions.

Traditional edge detection methods, utilizing filters like Sobel, Prewitt, and Canny, based on gradients and local pixel data, are effective but lack higher-level contextual understanding, often grappling with noise and ambiguity. Deep convolutional neural networks (CNNs), exemplified by models like Deep Edge, have showcased superior performance by leveraging hierarchical visual features. However, many CNN models face challenges due to computational demands, limiting their real-time applicability, particularly on embedded devices.

Addressing this challenge, TEED introduces a swift compact, yet accurate detector with only 58k parameters. In contrast DexiNed sacrifices efficiency to achieve maximum precision through multi-scale feature fusion, pushing the boundaries of state-of-the-art results. The paper meticulously examines the merits and limitations of both approaches through a comparative analysis, employing metrics such as mean squared error and assessments of edge quality [8]. It sheds light on nuances in the loss functions and upsampling strategies employed by TEED and DexiNed.

Our investigation into edge detection models has yielded meaningful insights. The EBIQA averages of 0.3304 for TEED and 0.5717 for DEXINED indicate that DEXINED excels in delivering superior image quality. In contrast, the NSER averages of 13.028 for TEED and 9.159 for DEXINED reveal a higher degree of edge loss or distortion in TEED. These meticulous measurements, executed on our dataset, serve as a pivotal guide in the selection of edge detection models. The considerations extend beyond image quality, incorporating a delicate balance with the preservation of edges.

The primary objective is to elucidate the strengths of distinct edge detectors, facilitating the selection process based on factors like speed, accuracy, or parameter budgets. The insights derived from this study aim to guide future innovations, striking an optimal balance between efficiency and performance. The paper unfolds across various sections, with Section II delving into the background study, Section III is the proposed approach, Section IV delving into assessment criteria, Section V delves into the results, and Section VI neatly summarizing conclusions drawn from the preceding sections.

II. BACKGROUND STUDY

In the realm of edge detection, TEED (Tiny Edge Enhancement Device) and DexiNed (Dense Extreme Inception Network for Edge Detection) stand out as two contrasting yet impactful models. TEED, a lightweight CNN, prioritizes efficiency and real-time processing, accurately predicting over 80 percent of edges. It introduces a compact, diverse dataset for comprehensive evaluation.

In contrast, DexiNed focuses on accuracy, leveraging advanced network architectures such as DenseNet and Inception modules. Achieving state-of-the-art results on established datasets, it redefines the boundaries of edge detection precision.

Both models contribute significantly TEED facilitates benchmarking with its new dataset, while DexiNed advances the state-of-the-art. Together, they showcase diverse approaches to edge detection, addressing the evolving needs of computer vision applications. Below subsections gives architectural background of TEED and DexiNed.

A. Architecture

1) TEED Architecture: TEED is a computer vision system that draws inspiration from successful models such as ResNet [3], Xception [2], and EfficientNet [13]. These models are renowned for their effectiveness in enhancing operations across different depths of CNN layers through dense and skip connections. Similar principles are employed in DexiNed [12] and LDC [11] architectures, both of which have achieved state-of-the-art accuracy in edge detection. TEED's architecture consists of three green blocks, each containing two standard CNN layers. VGG16 [9] was excluded due to its lack of skip connections, which could compromise

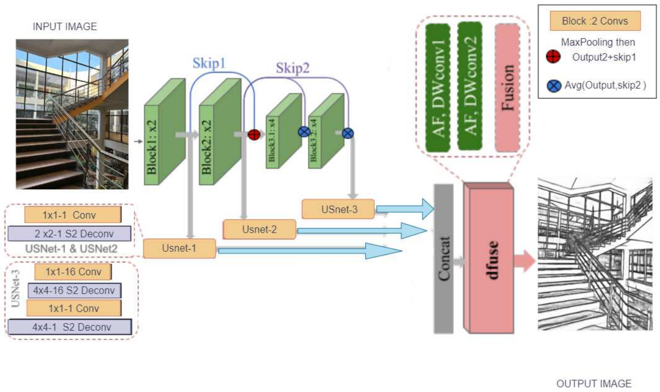

Fig 1: TEED Architecture

effectiveness in deeper layers. To reduce convolutional layers and omit batch normalization, TEED utilizes only 58k parameters, a significant reduction compared to DexiNed's 35m and LDC's 674k. While this reduction may impact accuracy, it is addressed by introducing an activation function called "smish" [14]. The function $\zeta - [$ conv $1 +$ smish $+$ conv $2 +$ smish $]$ helps with efficient training optimization in the absence of batch normalization. The three backbone blocks consist of 16, 32, and 48 layers, respectively. Skip connections [11], [12] (Skip1 and Skip2) are strategically used to combine outputs and facilitate information flow. Skip1 is applied post-max-pooling, while Skip2 fuses outputs from Block 3. The combined outputs feed into the upsampling network (USNet), showcasing a thoughtful integration of connectivity and fusion strategies. The convolutional layers used for skip-connections have a 1x1 kernel size.

2) Dexined Architecture: DexiNed's architectural sophistication is evident in its intricately designed DexiNet, which is composed of six interconnected modules known as encoders. Each encoder contains sub-blocks housing specialized processing layers, , and these units seamlessly connect through skip connections, ensuring a fluid exchange of crucial information and feature maps throughout the system.

Within the sub-blocks, meticulous transformations refine the feature maps. Initially, they traverse through pairs of convolutional layers, each armed with a 3x3 kernel for analyzing local image details. The initial encoder employs convolutions that span the entire image, capturing broader features to prevent information loss. Each convolutional layer is succeeded by normalization and ReLU activation, except for the final layer in the third encoder.

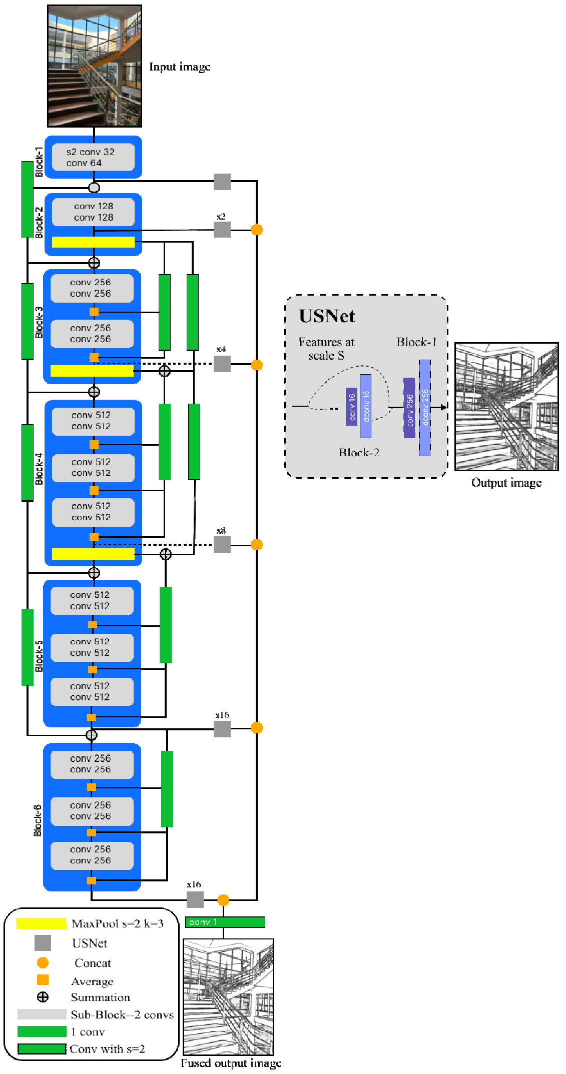

Fig 2: Dexined Architecture

DexiNed transcends conventional methods by deploying parallel skip [3] connections, artfully preventing the erosion of vital edge details during processing. These connections act as conduits, bridging earlier information-rich stages with later ones, preserving intricate edge features. Two types of skip connections, namely Second Skip Connections (SSCs) and First Skip Connections (FSCs), work harmoniously, merging outputs from different stages for a comprehensive representation of the image's edges.

Moreover, DexiNed utilizes max-pooling to condense information while retaining key details, balanced by a sophisticated averaging mechanism, preserving fine-grained edge structures. Figure 1 visually represents the intricate interplay between DexiNet and the upsampling network (USNet), providing a comprehensive overview of DexiNed's robust edge detection architecture [3].

B. USNet:

The Upsampling Network (USNet) is a key element in both DexiNed and TEED, contributing to refined edge detection. While sharing a common structure, each model introducesunique features.

1) TEED's USNet Module: The usnet module within teed reflecting dexined [12] integrates xavier initialization and aligns activation functions with the backbone structured with a single conv layer activation and a deconv layer teed tailors its architecture for down-sampled input in usnet-3 the anticipated edge-maps y_i are produced by applying the sigmoid function to the usnet output.

2) Dexined's USNet: DexiNed utilizes a conditional two-block USNet, with Block-2 up-sampling feature-maps until twice the size of the ground truth (GT). Transitioning to Block-1, a 1 × 1 kernel and ReLU activation process precedes deconvolution, ensuring alignment with the GT size. DexiNed explores bi-linear interpolation, sub-pixel convolution, and transpose convolution for upsampling, impacting thin edge generation.

In the comparative analysis of DexiNed, TEED, and traditional methods (Sobel, Canny, Laplacian), understanding the nuances of their USNet components sheds light on their distinctive edge detection capabilities.

III. Proposed Approach

In this section, we delineate the methodology employed for the comparative analysis between the edge detection algorithms TEED and DexiNed. Our procedural approach encompasses the selection of input data, algorithm implementation, and the utilization of evaluation metrics to yield substantiated results.

For input data selection, a deliberate effort was made to assemble a diverse set of images. This inclusivity, spanning various domains, aimed to holistically evaluate the robustness and generalizability of both algorithms across a spectrum of image types and scenarios.

Implementation of both TEED and DexiNed algorithms was executed in a controlled environment, deploying consistent parameters to upload fairness in the comparison. This implementation involved the application of algorithms to the chosen input images, resulting in the generation of respective edge maps.

To assess the performance of TEED and DexiNed quantitatively, we leveraged several key evaluation metrics. NSER [15] served as a metric to quantify dissimilarity between generated edge maps and the ground truth. EBIQA [1] gauged the quality of the edge maps, while Sobel-based-RR [6] scrutinized the algorithms' adeptness in capturing intricate details and local features. The metrics, including NSER, EBIQA, and Sobel- based-RR, underwent thorough statistical analysis.

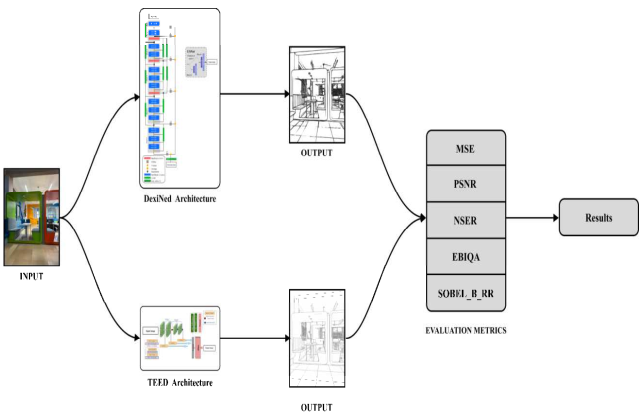

Fig 3: Flow of Approach

In tandem with quantitative metrics, our methodology em- braced qualitative analysis through the presentation of visualizations. These visual aids, showcasing the generated edge maps for select input images, provided an intuitive grasp of the algorithms' performance.

The comprehensive discussion of the findings from the application of TEED and DexiNed, coupled with the evaluation metrics, facilitated a nuanced understanding of each algorithm's strengths and weaknesses across diverse image- processing scenarios. This deliberation laid a robust foundation for subsequent sections of the paper, ultimately culminating in a well-informed and substantiated conclusion.

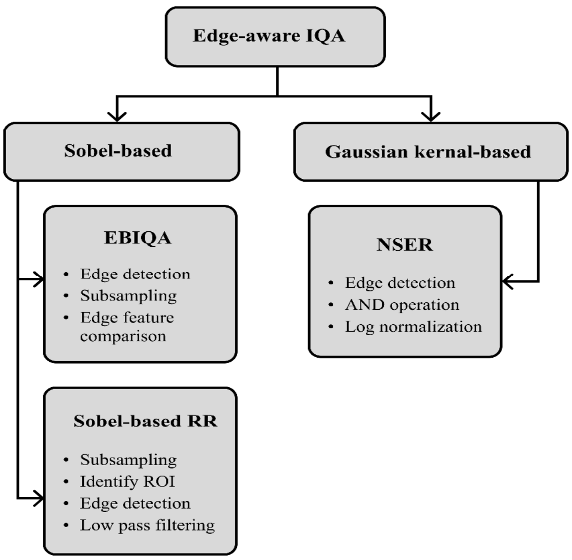

Figure 4:Overview of edge based IQA

A. Evaluation Metrics

Assessing the effectiveness of edge detection methods is crucial for identifying objects in images. Evaluation metrics play a significant role in this regard. Traditional filters often lead to blurred edges due to non-uniform pixel intensity values. Edge-aware filters, such as anisotropic diffusion, bilateral filter, domain transform filter, guided filter, wavelet transform based filtering, and G-neighbor classification filtering, are designed to address this issue.

There are several widely used metrics for evaluating filter efficiency, including Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), and Structural Similarity Index (SSIM) [7]. However, while these metrics measure over-all image quality, they may not fully capture subtleties in edge quality. This study rigorously examines promising edge-oriented evaluation metrics, highlighting the significance of comprehensive assessments beyond traditional measures.

1) Standard Evaluation Metrics: Evaluation Metrics Stan-dard: The most commonly utilized metrics for conventional image assessment include Mean Squared Error (MSE) [5], Peak Signal to Noise Ratio (PSNR), and the Structural Simi-larity (SSIM) Index. MSE, the simplest pixel-wise difference $M \times N$:

$$MSE = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N}[\mathbf{I_1}(i,j) - \mathbf{I_2}(i,j)]^2 \tag{1}$$

While MSE is informative about average error energy, interpreting it as an absolute quality score can be challenging. PSNR expresses MSE as a logarithmic decibel value relative to the maximum possible pixel value MAX_I. It is more easily comprehensible and is commonly used for comparing compression performance or filtering results between images. PSNR represents the power ratio between maximum image intensity and mean squared error:

$$PSNR = 10\log\left(\frac{\max(I)^2}{MSE}\right) \tag{2}$$

However, MSE and PSNR provide limited insight into perceived visual quality, primarily relying on pixel-level nu-merical errors that may not align well with human observer judgments. This limitation led to the development of metrics like SSIM [7], which incorporates aspects of visual perception. Instead of focusing on pixel differences, SSIM evaluates the similarity of luminance, contrast, and image structures between two images. It separates the tasks of similarity and error measurements, using statistics like sample means μ, variances σ^2, and cross-covariance σ_{xy}. SSIM is calculated on small windows, and the overall score is a weighted average over all windows:

$$\text{SSIM}\ (\mathbf{x},\mathbf{y}) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)} \tag{3}$$

The constants c_1 and c_2 stabilize divisions with small denominators. SSIM tends to align better with human assessments across various image processing applications.

2) NSER (Non-shift Edge Based Ratio): The twisted picture and the reference images edge maps are measured using the NSER [15] (Non-shift Edge Based Ratio). In order to detect edges in the images across various size spaces, this measure uses a Gaussian kernel to modify the deviation. All of the discovered edges come together to generate each edge map. Next, NSER calculates the ratio of shared common edges between the reference and distorted edge maps in respect to the total number of edges in the reference picture. The results of this computation indicate the degree to which the edges of the distorted image reference are preserved.

To compare the edge maps, their intersection is considered, and the common edges are counted. The ratio p_i at scale i is calculated as:

$$p_i = \frac{\|\mathbf{I_1} \cap \mathbf{I_2}\|}{\|\mathbf{I_1}\|} \tag{4}$$

Here, $\mathbf{I_1}$, $\mathbf{I_2}$ denote edge maps at scale i, and represents the count of common edge pixels. Subsequently, NSER is determined by taking the negative logarithm and summing over the scales:

$$NSER(\mathbf{I_1},\mathbf{I_2}) = -\sum_{i=1}^{N} \log_{10}(1-p_i) \tag{5}$$

Implementing the logarithm improves correlation with subjective ratings. Smaller values in the NSER range, which goes from 0 to infinity, indicate greater edge preservation.

3) EBIQA (Edge-Based Image Quality Assessment): Ebiqa edge-based picture quality evaluation [1],Ebiqa first compares edge-based features taken from fixed 16x16 windows in both images the features utilized include edge orientation and total number of edge pixels it next utilizes sobel edge detection to identify edges in refrence and distorted images number of pixels with comparable intensity and average edge length vectors i1 and i2 for both images are created by combining the pl number of edge pixels and the n number of horizontal-vertical edge pixels next we compute ebiqa which is the average euclidean distance between these vectors throughout the whole image.

$$EBIQA = \frac{1}{MN}\sum_{i=1}^{M}\sum_{j=1}^{N} \|\mathbf{I_1} - \mathbf{I_2}\| \tag{6}$$

When comparing the deformed image to the reference smaller ebiqa values signify higher preservation of edge features.

4) Sobel-Based Reduced Reference Method: The sobel-based reduced reference method [6] employs sobel filtering on selected regions of interest roi blocks instead of the entire image to minimize the transmitted data the image is partitioned into 16x16 blocks of which 12 symmetrically chosen blocks are selected for comparison sobel filtering is individually applied to these chosen blocks to generate an edge map the count of edge pixels in each block is determined by applying two distinct thresholds namely t_1 and t_1. The information from all blocks is combined to get a single metric value I given by*:*

$$I = \frac{1}{MN}\sum_{i=1}^{MN} \frac{(w_1 t_1 + w_2 t_{1,2})}{MN} \tag{7}$$

where w_1, w_2 are weighting constants, and MN is the total number of pixels in the 12 blocks. Lower values of I indicatebetter edge preservation.

IV. Results

This section articulates the results emanating from our meticulous analysis, systematically comparing the efficacy of TEED and DEXINED in the realm of image edge detection. The primary objective of this study was to rigorously examine and evaluate the performance of these well-established methodologies, delineating their inherent strengths and limitations in discerning intricate details within images. Our findings not only contribute to a heightened understanding of the capabilities inherent in TEED and DEXINED but also furnish valuable insights for professionals and researchers operating in the domains of computer vision and image processing. These insights serve as a guide for the selection of optimal solutions within these fields.

The Table 1 showcases the performance of the TEED and DexiNed models through various evaluation metrics, including Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Normalized Structural Edge Response (NSER), Edge-Based Image Quality Assessment (EBIQA), and Sobel-based Relative Recall (Sobelbased-RR). The metrics offer insights into the models' abilities to generate accurate and visually appealing edge maps.

TABLE I
AVERAGE RESULTS OF TEED AND DEXINED ON OUR DATASET

IMAGES	TEED	DEXINED
MSE	0.072	0.096
PSNR	11.6	10.8
SSIM	0.5932	0.5611
NSER	13.028	9.159
EBIQA	0.3304	0.5717
Sobel-Based-RR	1.8484	0.6933

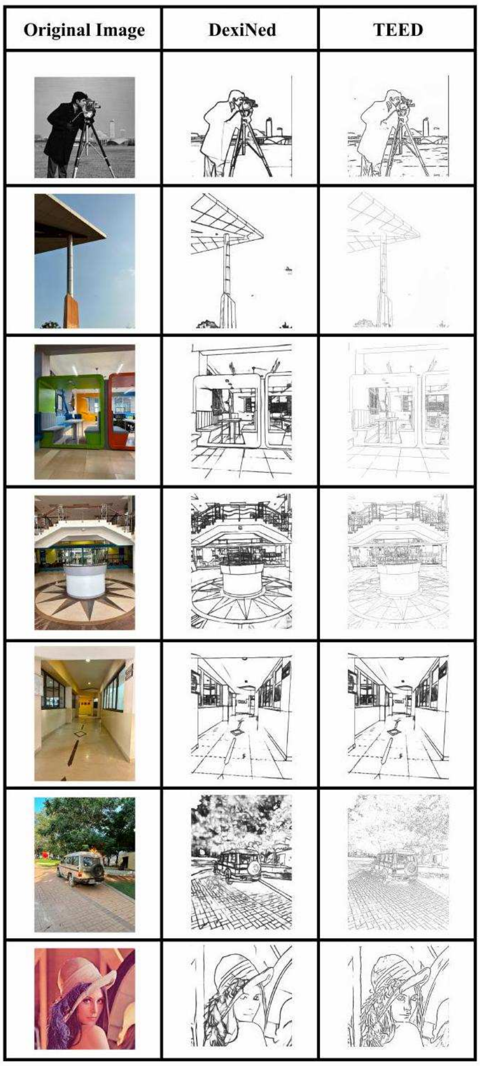

Figure 5. Output of TEED and DexiNed

Figure 5 visually presents examples of input images and their corresponding output edge maps generated by the TEED and DexiNed model.

V. CONCLUSION

In this comparative study, we analyzed two state-of-the-art deep learning-based edge detectors, namely TEED and DexiNed. The quantitative evaluation revealed that DexiNed achieved superior accuracy, with the lowest NSER score of 9.159 and the highest similarity in the Sobel comparison (0.6933). However, TEED demonstrated a competitive MSE (0.072 vs. 0.096) and efficient performance with only 58K parameters.

Our analysis also revealed that DexiNed's effectiveness lies in capturing finer edge details and textures compared to TEED. However, it was found that DexiNed tends to have more false positives in complex regions. On the other hand, TEED provided smoother, cleaner edges, albeit with less sharpness.

In conclusion, this study provides a blueprint for selecting the appropriate model based on constraints and use cases. The analysis also highlights the limitations around optimizing for pixel accuracy rather than perceptual quality. By advancing the understanding of deep learning techniques for edge detection, this study aims to spur progress in this integral capability underpinning multiple computer vision tasks. Therefore, it is important to consider both the advantages and limitations of each algorithm to determine the most suitable one for a specific use case.

REFERENCES

[1] Abdolrahman Attar, Reza Moradi Rad, and Asadollah Shahbahrami Ebiqa: An edge based image quality assessment. In 2011 7th Iranian Conference on Machine Vision and Image Processing, pages 1-4. IEEE, 2011.

[2] François Chollet. Xception: Deep learning with depthwise separable convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 1251-1258, 2017.

[3] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770-778, 2016.

[4] Quan Huynh-Thu and Mohammed Ghanbari. Scope of validity of psnr in image/video quality assessment. Electronics letters, 44(13):800-801, 2008.

[5] Javier López-Randulfe, César Veiga, Juan J Rodríguez-Andina, and José Farina. A quantitative method for selecting denoising filters, based on a new edge-sensitive metric. In 2017 IEEE International Conference on Industrial Technology (ICIT), pages 974-979. IEEE, 2017.

[6] Maria G Martini, Chaminda TER Hewage, and Barbara Villarini. Image quality assessment based on edge preservation. Signal Processing: Image Communication, 27(8):875-882, 2012.

[7] Sameeulla Khan Md and Sumohana S Channappayya. Multiscale-ssim index based stereoscopic image quality assessment. In 2016 Twenty Second National Conference on Communication (NCC), pages 1-5. IEEE, 2016.

[8] Diana Sadykova and Alex Pappachen James. Quality assessment metrics for edge detection and edge-aware filtering: A tutorial review. In 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI), pages 2366-2369. IEEE, 2017.

[9] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014

[10] Xavier Soria, Yachuan Li, Mohammad Rouhani, and Angel D Sappa. Tiny and efficient model for the edge detection generalization. In Proceedings of the IEEE/CVF International Conference on Computer Vision, pages 1364-1373, 2023.

[11] Xavier Soria, Gonzalo Pomboza-Junez, and Angel Domingo Sappa. Ldc: lightweight dense cnn for edge detection. IEEE Access, 10:6828168290, 2022.

[12] Xavier Soria, Angel Sappa, Patricio Humanante, and Arash Akbarinia. Dense extreme inception network for edge detection. Pattern Recognition, 139:109461, 2023.

[13] Mingxing Tan and Quoc Le. Efficientnet: Rethinking model scaling for convolutional neural networks. In International conference on machine learning, pages 6105-6114. PMLR, 2019.

[14] Xueliang Wang, Honge Ren, and Achuan Wang. Smish: A novel activation function for deep learning methods. Electronics, 11(4):540, 2022.

[15] Min Zhang, Xuanqin Mou, and Lei Zhang. Non-shift edge based ratio (nser): An image quality assessment metric based on early vision features. IEEE Signal Processing Letters, 18(5):315-318, 2011.