Proceedings of the International Conference on Inventive Computation Technologies (ICICT-2025)
DVD Part Number: CFP25F70-DVD; ISBN: 979-8-3315-1223-1

Enhancing Blind Image Quality Scores using Multi-Scale feature extraction

1st Akhil Kambhatla
Artificial Intelligence & Data Science
Chaitanya Bharathi Institute of Technology
Hyderabad, India
akhilkambhatla.edu@gmail.com

2nd Smt. S. Shobarani
Artificial Intelligence & Data Science
Chaitanya Bharathi Institute of Technology
Hyderabad, India
shobarani_aids@cbit.ac.in

3rd Likhita Golukonda
Artificial Intelligence & Data Science
Chaitanya Bharathi Institute of Technology
likhitaa04@gmail.com

4th Renu Prasad
Artificial Intelligence & Data Science
Chaitanya Bharathi Institute of Technology
renuprasad703@gmail.com

***Abstract*—The application of BIQA is necessary in cases where no references on images are present, yet the existing techniques are insufficient to catch the whole range of distortions. This failure is especially frequently marked when distortions cover more than one scale. Based on multiscale analyses of image information, recent prospects have been promising for algorithms based on feature extraction being able to break through this limitation and provide a more detailed insight into local and global degradations. This review of the existing literature discusses state-of-the-art multiscale BIQA models in terms of their capabilities to generalize across datasets and how well they work with distortion types. A broad, overall analysis of recent contributions reveals the significant weaknesses of existing techniques that fail to reliably describe complex and mixed distortions and maintain cross dataset robustness. Using Laplacian Pyramid Networks fill these gaps, we present a new multiscale feature extraction framework. Using LPN, an image breaks up into its frequency bands, capturing more details precisely on the higher end and structure on the lower so that quality prediction accuracy attempts are done by improving and enhancing predictions. The current survey gives special attention to the role played by multiscale techniques in BIQA and introduces LPN as a potential solution for quality assessment performance over a large spectrum of visual distortions ..**

***Index Terms*—BIQA, Multi-Scale Feature Extraction, Laplacian Pyramid Networks (LPN), Perceptual Quality Evaluation, Deep Learning, Frequency Decomposition, Generalization, Quality Prediction**

I. Introduction

A very rapidly growing field is known as Blind Image Quality Assessment (BIQA) dealing with assessment of image quality without recourse to a reference image. Unlike full reference IQC approaches which usually require images to be compared to versions of themselves, BIQA models are developed under conditions where the reference image is not available. This gives them real time assessment capabilities with applications ranging from video streaming surveillance to medical imaging. This kind of application will often host image manipulations that reduce appeal, yet it is in the greater interest that the call for reliable quality assessment is on the rise with the prolific dissemination of digital content across a plethora of interfaces.

According to Wang's research [1], with the increasing supremacy of visual data, BIQA turns out to be a necessity in various industries. The questionnaire clarifies that the HVS can spot flaws that the model fails to notice. Thus, the main task for BIQA is that the perceptions should coincide with the evaluation made by the automated tool. It is crucial for industries like entertainment and telemedicine as the quality of the image directly impacts the user experience.

There are two main categories as shown by the study in [2]: distortion-specific and non-distortion-specific approaches to BIQA. Most of the research that is being conducted up to now falls into the category of non-distortion-specific which aims to defeat a variety of distortions regardless of what type of distortion is applied without known distortion type. The distortion-specific approach predicts the quality using known types of image degradation. They drive home the persistent challenges of obtaining reliable quality scores on various datasets, thus the main reason this industry is still challenging to operate in. This review shares that though new BIQA methods have been developed, they are yet unable to handle the sheer number of distortions found in the real world. This is a gap this project hopes to bridge.

From this, Nizami et al. (2019) [3] presented a distortion-specific feature selection method that overcame the pivotal weakness in traditional BIQA models that is, its inability to adapt to the various types of image degradation. In the work, a three-step feature selection approach was introduced with the purpose of detecting and fighting the element designed specifically for targeted distortions. This contributes strongly to the improvement of the model's ability to predict the quality of images in the presence of different distortion mechanisms. Although it's an important feature of the algorithm that places such a huge emphasis on distortion-specific adaptation, it unfortunately has the consequence of drawing attention to one or two weaknesses that single-scale methods do have. They often fail to capture the richness of image distortion at different scales; extracting multi scale features can address this. Most of the existing BIQA models are single-scale methodology techniques, though significant progress made in this field is still limited.

According to Yang et al. (2023) [2] and Wang (2023) [1], the inability of single scale strategies to detect distortions appearing at different spatial frequencies means that single-scale methods cannot be useful enough to determine high frequency distortions such as noises and others, which have a low frequency of blurring, possibly occurring in a single image. Here, multi-scale feature extraction is necessary. Our goal is to make the ability of the model in recognizing a more general class of distortions better, and we perform this by analyzing features of the images across different scales.

II. RELATED WORKS

A. Multi-Scale Feature Extraction for Capturing Image Distortions

This technique entails image analysis at scales of variations ranging from fine details to panoramic features by zooming in and out, allowing for the capturing of both fine details and more panoramic features of an image. Indeed, the process of feature extraction, which is usually finding significant visual components such as edges, textures, or colors, is very advanced when carried out across multiple scales. That's why models can detect all wide ranges of distortions in images, including blur noise and compression effects, often overlooked by one-scale based approaches. When we look at images from many perspectives, we get a more complete view of their quality.

A multi-scale analysis systematically deals with the richness of distortion in images, respecting the regions and the levels of detail. The traditional BIQA models that estimate the quality at a single resolution are often unable to capture the subtle distortions that become visible at specific scales. Multi-scale techniques, as in [4], extract features across different zoom levels, that is, large and small distortions. This enables higher holistic quality assessment, based on global and local distortions. The model decomposes images into spatial frequency bands, which allows the detection of broad issues like blurriness and finer anomalies like noise or compression artifacts. This technique provides multi-scale methods with more accurate quality estimations: macro-level distortions as well as micro- level distortions. Macro distortions affect the overall structure, while the micro distortions relate to the smaller details such as the presence of noise or sharpness. By a combination of these two aspects, multi-scale feature extraction enhances understanding of image degradation. Recent studies show that these methods outperform the traditional single-scale model in terms of robustness and hold great potential in promoting the development of BIQA technology across a myriad of applications.

It is further confirmed, then, that the multi-scale feature extraction in the BIQA framework is important, by proposing the PMT-IQA method that incorporates multi-scale feature extraction with progressive multi-task learning, as put forward by Pan et al. (2023) [5]. The feature extraction module across various scales helps to learn the complex distortion patterns at various scales, and those are learned progressively in the way humans perceive the visual spectrum. This type of progressive learning boosts the ability of a model to address the diversity of distortion types found in real-world images.

Another very important example of multi-scale BIQA is the method, which applies KLT [6] for an unsupervised feature extraction of images at different scales. The KLT-based approach has proved to be more efficient than traditional deep learning models which require big training datasets, since it eliminates the need for annotated data. The data intrinsic statistical characteristics-based method reduces computational costs and boosts generalization in various types of distortion. Moreover, the KLT-based multi-scale framework reverses all limitations of traditional BIQA models due to their learning procedure, making it a better adaptable strategy for distortion estimation.

Another remarkable contribution towards the research of multi-scale BIQA focuses on the use of opinion-unaware techniques [7]. One way this model shines is through the removal of a reliance on subjective human opinion scores for benchmarking. This model instead relies on objective deep feature statistics, derived at various scales, to judge the image quality. Such methodology tends to provide effective generalization across different types of distortion without requiring opinion-based training data. Deep feature statistics can be used in applications to capture more comprehensive distortions, making the system subject to a more complete evaluation of the quality of an image.

It significantly improved the accuracy and consistency of the models of BIQA by strategies of multi-scale feature extraction. Compared to single-scale models, these techniques provide more generalization and flexibility: KLT-based and deep feature statistics. That is one of the major turning points in BIQA research, where multi-scale approaches are established as necessary for further improvement in image quality evaluation in the far future.

B. Feature Fusion and Aggregation Techniques

taken into consideration for assessing image quality. Two main approaches used include feature fusion, or the integration of many different attributes about an image, and aggregation, or the compiling into a final quality metric. Feature fusion is akin to solving a puzzle. An image has details at various scales, from the finest pixel elements to higher level features like broad patterns. With this, the model captures a holistic view of the image by incorporating both subtle and significant distortions. Aggregation then goes ahead to summarize all the integrated details into one score indicating the quality of an image. It is critical that each detail is properly considered at this step.

In [8], authors discuss two approaches to feature fusion: concatenation and score averaging with Multilevel Feature Fusion (MFF). The feature concatenation of the distorted image with its AI-restored counterpart enables the comparison of the ”before” and ”after” states in order to spot distortions. MFF builds on this with its use of a network that combines fine details to more general patterns, thus giving room to the model for the evaluation of quality in as much perspective as possible-that is from micro textures up to macro shapes. They then aggregate using score averaging, which means the scores obtained from various feature sets ensure no important information is missed for an overall summation quality evaluation. The authors introduce AFF coupled with element-wise multiplication using cosine similarity in reference [9]. AFF offers an alternative flexible method of feature integration, scanning the image at various resolutions allowing zooming in and out, focusing on all important features. That means it might be able to drive the model towards important details. The cosine similarity is next used to analyze the relationship between the image features and a textual description thereby ensuring that the image is visually precise but also reconciles with the intended content or meaning

Furthermore, in paper on Deep Response feature decomposition and aggregation [10], the authors have used graph-based approach using Graph Attention Networks. In this model, it had split the images into smaller features by considering each of the features as a ”node” within a graph. Here, its attention mechanisms have been utilized to point out those nodes which are the most significant ones towards quality assessment. Finally, this information is passed through a regression to convert it into the appropriate score that represents the quality of the image. Thus, it helps in relating different parts of the image, leading to an increasingly correct final score.

C. Handling Diverse Image Distortions with Statistical Models

Statistical models are important because they serve as analytical tools that dig deeper into the characteristics of an image, such as the structure of pixels, distribution of colors, and detail depth. Through the analysis of those features, these models can identify various kinds of distortions-affecting images such as blur, noise, and compression artifacts against the ideal expectation of undistorted images. One important aspect of this technique is derived from the concept of Natural Scene Statistics, or NSS, that typically describe the common patterns present in high-quality natural images. Statistical models con- sider the deviation from observed patterns as a measure for generalizing the quality of an image. Thus, statistical models form an appropriate base for a broad range of distortions within the BIQA paradigm.

(NSS), which encapsulates the patterns typically observed in high-quality natural images. When an image diverges from these established patterns, statistical models leverage this discrepancy to assess its overall quality. Thus, the application of statistical models provides a robust framework for effectively handling diverse image distortions in the BIQA context.

Authors of [11] proposed an approach that emulates how the visual cortex processes visual information. This innovative model uses NSS to detect aberrations in images since natural, undistorted images clearly obey some statistics conformity. When an image contradicts these norms, the model is poised to pinpoint the problems. The technique Spatial Feature Ex- traction is based on the ability of a human eye to detect detailed view complexities and structures along with features of texture and energy to detect changes brought by blur or noise. In contrast, the Spatial-Frequency Features model makes use of log-Gabor filters to simulate the recognition capabilities of a brain of different types of frequency changes in an image to highlight areas which distortions differently affect the frequency characteristics of the image. Additionally, the Color Features model checks the color distribution that simulates the process of color perception by retinal cells. This should automatically help it to detect most distortions like chromatic aberrations. This method emulates brain processing and so is efficient enough to identify a wide range of distortions without even needing a reference image.

The authors emphasize the use of Natural Scene Statis- tics in combination with multiscale feature extraction techniques in their paper ”Blind Image Quality Assessment Using Naturalness-Aware Multiscale Features” [12], in which it seeks to effectively address various distortions pertinent to images. As human vision perceives both the minute particulars as well as the general view, this technique evaluates the images at different levels of detail. The application of MSCN allows the model to identify multiple distortions without having been trained for separately, enhancing the statistical properties of the image. The inclusion of Inception Modules and Pyramid Pooling in Multiscale Feature Extraction will allow full con- sideration to be given to the analysis of the image at different scales just like the human visual system, where fine details as well as contextual larger information can be processed. This approach combines statistical analyses of natural images with a multiscale viewpoint and hence provides a holistic understanding of image distortions.

D. Deep Learning Architectures for Multi-Scale BIQ

DL has thoroughly transformed BIQA in contemporary research by designing architectures that help in the extraction of multi-scale features. This gave rise to enhanced accuracy in BIQA regarding the quality of images. The traditional BIQA methods primarily try to check the images at one resolution, and more often than not, they are inadequate in dealing with distortions operating variably with different scales. Integration of multi-scale feature extraction with advanced deep learning architectures effectively neutralizes this as the images are evaluated at various resolutions, which allows models to pick up on overhead distortions as well as localized distortion.

Among the architectures of modern deep learning, widely applied in BIQA are the hybrids of many sophisticated methods to enhance the accuracy and precision of quality estimates. CNNs are often applied since they display outstanding abilities in feature extraction. Such networks learn hierarchical representations of images, which enables the identification of low-level and high-level distortions. The work of Huang et al. (2024) [13] shows a representation model that uses a multi-layer approach. Such a model interacts with images at different resolutions, thereby creating a hierarchical perception of distortions. Further, by the utilization of deep feature extraction, the architecture encapsulates even the most subtle image characteristics at different levels of granularity. It is quite effective in handling both coarse and fine distortions. It is the multi-layer methodology that enables the model to focus on several other classes of distortions as they come through different scales, hence leading to more robust quality evaluations. Additionally, the extra insight by this work is focusing on hierarchical deep feature extraction that enhances the abilities of the model in understanding and processing images at different levels of complexity.

Another technique that is found to be supported by the previous findings [5] and falls under the category of vast deep learning is MTL. MTL constructs, like PMT-IQA, build competence to detect and remove distortions through first interacting with simple and more noticeable alterations before handling the tough ones. Through this adaptive learning process, the model achieves improvement in understanding distortions at each stage of training, just like the human visual system adapts. A multi-scale module of PMT-IQA analyzes images at different magnifications, meaning that it actually captures distortions affecting minute details as well as expansive structures. Such methodology differs from the traditional models that tackle all the distortions on the same scale and are often wanting in terms of adaptability to the various degrees of complexity. It provides important insights into how deep learning architectures can incrementally improve their understanding of distortions and thus represent a more dynamic approach towards quality assessment, rather than the static evaluations seen with the traditional BIQA models.

In addition to this, along with MTL and CNNs, models based on transformers have gained significant interest in the BIQA research area due to their capacity to pay attention to very long dependencies in images. The so-called ViTs, which rely on a mechanism-based focus on the most informative areas influenced by distortions, outperform CNNs in global distortions. Zhou et al. (2023) illustrate in the Separate Representations and Adaptive Interaction of Content and Distortion study [14] how transformers can even surpass localization and global distortions in terms of representation. This is because with the better utilization of attention mechanisms enabled by the transformers, they are able to focus more acutely on those areas of the image where the distortion has an extreme deviation. This architecture improves the capability of the model to capture spatial interrelations between pixels, and it becomes especially good at explaining global distortions that CNNs cannot represent straightforwardly.

Above, there are several advantages of deep learning with multi-scale feature extraction presented over the traditional BIQA methods. The models that rely upon deep learning and the decomposition of images at different resolutions can have a more extensive perception in the recognition of distortions, resulting in more accurate and reliable quality assessments. Moreover, contemporary approaches to deep learning incorporate progressive learning and adaptive mechanisms that allow continual refinement for better estimations. Advanced feature extraction capabilities in deep learning models allow them to provide abilities involving the recognition of more complex and subtle features of images that traditional approaches fail to capture. An advancement in both deep learning and multi- scale techniques is highly expected, which will contribute a lot toward improving the accuracy and reliability of BIQA over a wide range of applications.

E. Natural Scene Statistics (NSS) in Multi-Scale BIQA

The abbreviation NSS was used for the statistical features of a natural image. This is well aligned with human vision; some specific regularities of the characteristics depicted can be used to effectively assess the quality of images. Traditional BIQA approaches rely on the use of NSS to derive perception-related features in order to evaluate quality without reference images. State-of-the-art work includes the Natural Image Quality Evaluator (NIQE) [15]. Based on the analysis of local image patches, it models the distributions in these patches to find their statistical properties using a framework of multivariate Gaussian. NIQE evaluates the feature distributions that can come from such distorted images against corresponding reference high-quality images allowing objective quality assessments. The critical issue here is that NSS can indeed capture perceptual attributes with a high degree of reliability, thus forming a trustworthy criterion for evaluating the quality of an image independently of subjective input.

This integration then can further improve the BIQA model's robustness and accuracy by combining the present model with natural scene statistics NSS, which multi-scale feature extraction represents. This type of multiple techniques, naturally enough, has proven necessary for considering image features at various resolutions when it comes to considering effects from distortion on perceived quality. In this direction, Ni et al. have ”Opinion-Unaware Blind Image Quality Assessment using Multi-Scale Deep Feature Statistics” [7] in 2024 that shows a framework that integrates deep learning and statistical analysis for an opinion unaware BIQA system on NSS; the authors extract patch-wise multi-scale features from pre- trained architectures like ResNet or VGG that are then fed into a Multivariate Gaussian model for statistical representation. The final quality score is obtained as a measure of the distance between the MVG model of the test image and a reference MVG model for high-quality images. As we saw from the MDFS model, MSS could improve the predictive capability of

BIQA with deep features of large-scale data sets and statistical integrity with MVG fitting.

Further improvements in [12] extend the conventional NIQE by adding color, gradient, and frequency characteristics to its feature extraction process. As such, this extension is based on the NSS principles to cover more statistical features with reliability and accuracy when applied under different contexts. Another notable contribution is found in quality-aware clustering [16], which proposes a method to learn a set of quality centroids to estimate the quality of each patch within an image. The work highlights how NSS may be adapted to complex models that consider different kinds of distortions across the image in diverse areas.

It is noticeable from the investigation of Natural Scene Statistics in multi-scale BIQA frameworks that the impact of statistical properties significantly sheds light on the enhancement of existing image quality assessment methodologies. Integration of NSS with modern deep learning techniques continues to inspire newer versions and robust models reflecting the latest developments in the field.

F. Transformers and Vision-Language Models in Multi-Scale BIQA

Transformers have been widely used in BIQA, their attention mechanisms can model rich relations and dependencies of long-range regions with high effectiveness for images. The ability is essential to help determine subtle variances that are otherwise impossible to identify with conventional techniques. Self-attention of the Transformer lets it look at different pixels and find out which are most relevant to one another such that when training, we focus only on certain regions with high impact on perceived quality. This model can naturally learn to focus on elements matching human quality assessments, something particularly promising for quantifying spatially scattered distortions. Moreover, employing vision- language models enhances the BIQA domain by allowing multimodal data — visual content with textual summaries. This amalgamation allows the model to enhance accuracy and verifiability of assessments.

[17] introduces a multitask learning framework that receives shared insights from related tasks into better improving the accuracy of BIQA. Auxiliary tasks such as forecasting types and magnitudes of distortions are adapted into the primary quality evaluation task to make for better developments in the image's representation so as to prevent overfitting. The experiments on a variety of test collections confirm their approach, with considerable gains in accuracy using the Pearson Linear Correlation Coefficient (PLCC) as well as Spearman Rank Order Correlation Coefficient (SROCC). Using the precision of BIQA systems in this multitask learning framework indicates the ways through transformers and vision-language models, where the information can be captured effectively from these complementary knowledge bases.

Another paper [18] introduces KGANet - one of the latest architectures tailored specifically for BIQA, where the knowledge of distortion is considered in multiple aspects, combined with the use of multiscale feature extraction techniques. In the architecture of the KGANet model, it passes images through several scales to capture both high-level and low-level characteristics that are important for the understanding of various forms of distortions. Adaptive technique-based feature fusion across different layers is used to achieve maximum image quality representation by fusing features from different layers. The kind of distortions can be detected through facilities within KGANet, so the quality evaluation process can be tailored to each distortion's specific features. With state-of- the-art performances over KADID-10K benchmark datasets, KGANet exceeds existing methods with a 7.7% advance in SROCC and demonstrates excellent generalization capability over the test data and types of distortion. This further enhances the significance of transformers and vision-language models in advancing BIQA.

Transforming and vision-language models are investigated in the multi-scale BIQA frameworks to explicitly validate the essential contribution of these towards advancing IQA.

III. COMPARISON OF EXISTING MODELS

S.no	Method	Datasets Used	Metrics
[1]	Multimodal quality representation	L + T + C + K + Ko	Perceptual quality metrics; Statistical Characteristics
[2]	CONTRIQUE (2022), SFA (2018), MEON (2018), CNN (2014)	Lc + C + T13 + LC + Bid + Ko	Lc-SRCC: 0.463 - 0.631, Lc-PLCC: 0.507 - 0.654, Bid SRCC: 0.539 - 0.573, Bid-PLCC: 0.576 - 0.598, Ko-SRCC: 0.700 - 0.894, Ko PLCC: 0.704 - 0.906
[3]	Feature Extraction → Distortion Specific Feature Selection → Support Vector Regression	L+ C + T13 + L-WIQCD	SROCC, LCC, KCC, RMSE
[4]	Multi-scale Filtering → CNN → MSCN operation	L + C + T13	L-SRCC: 0.969, L PLCC: 0.978, T13-SRCC: 0.835, T13 PLCC: 0.859
[5]	PMT-IQA	BID + L + Lc + C	SRCC: 0.856 (LIVE C), 0.929 (BID), 0.971 (LIVE), 0.942 (CSIQ), PLCC: 0.893 (LIVE-C),

			0.969 (BID), 0.971 (LIVE), 0.951 (CSIQ)
[6]	Multi-scale Karhunen-Lo´eve Transform (MsKLT)	L + T13 + C + Toyoma	MsKLT: 0.798
[7]	Multi-scale Deep Feature Statis- tics (MDFS) model using Multi-variate Gaussian (MVG) model	C13 + Ko + L + C + T13 + SPAQ + KADID + MDL + MIDL	SROCC: 0.8571 (C13), 0.9361 (L), RMSE: 11.0931 (C13), 14.1344 (L), PLCC: 0.8717 (C13), 0.8558 (L)
[8]	Multilevel Feature Fusion → Generative Adversarial Network	L + C + T13 + KADID10K	SROCC: 0.953 (C), 0.865 (T13), 0.965 (L), PLCC: 0.956 (C), 0.883 (T13), 0.971 (L)
[9]	AMFF-Net	AGIKA-3K + AIGCQA2023 + PKU- IQA	SRCC: 0.7980, KRCC: 0.5914
[10]	Deep Response Feature Decom-position and Aggregation	L + C + T13 + KADID10K	SROCC: 0.920 - 0.958, KTAU: 0.740 - 0.860
[11]	Support Vector Regression	L + C + T13	10%, L-SROCC: 0.935, L-PLCC: 0.932, C-SROCC: 0.912, C PLCC: 0.951, T-SROCC: 0.583, T PLCC: 0.705
[12]	Inception and Pyramid Pooling + NSS → FC layers to predict scores	T08 + C + T13 + KADID10K	CSIQ, PLCC: 0.955, SRCC: 0.953
[13]	GRU-based Fusion Encoder Perception Oriented → Quality Regression Network	L + C + T13 + KADID + Lc + Ko + Lm	C-PLCC: 0.960, C SROCC: 0.952, T13-PLCC: 0.911, T13-SROCC: 0.896, Lm-PLCC: 0.950, Lm SROCC: 0.942
[14]	Collaborative Autoencoder Features Adaptive combined Weighting in → Self- based Quality Predictor	L + C + T13 + KADID10K + Lc + Ko	T13-PLCC: 0.986, T13-SRCC: 0.984, Ko-PLCC: 0.966, Ko-SRCC: 0.963
[15]	Multivariate Gaussian Model + NSS	L + C + T13 + MD1 + MD2	TID2013: SRCC = 0.898, PLCC 0.903, = CSIQ: SRCC = 0.815, 0.854, 0.891, 0.905
[16]	Convolutional DAE-aware deep architecture	L + T08 + T13	PLCC: 0.942 – 0.970
[17]	Language-Image Quality Evaluator	L + C + KADID10K + BID + CLIVE + Ko + Mean	SRCC, L: 0.970, C: 0.936, KADID10K: 0.930, Ko: 0.919
[18]	Pyramid Vision Transformer → Cross-Layer Information Fusion + Knowledge-Guided Attention	L + C + T13 + KADID10K	C-PLCC: 0.963, C SROCC: 0.954, T13-PLCC: 0.933, T13-SROCC: 0.927, L-PLCC: 0.966, L SROCC: 0.963

IV. PROPOSED WORKFLOW

The growing complexity and variety of image distortions in modern datasets require a more sophisticated multiscale feature extraction that fully captures the fine and global distortion characteristics, which is often not achieved by current methods. Moreover, the limitations of existing BIQA models can be overcome by new approaches that advance our current understanding of perceptual quality, accuracy and generalization in partial reference degradations.

Work to date on BIQA, while providing useful multi-scale feature extraction for IQA models, the new method of decomposition introduced in our Laplacian Pyramid Network (LPN) may provide benefits by better separating fine textures and gross structural distortions vs. prior art was first to work. Given the increasing complexity of image degradations, LPN's hierarchical approach could enhance the accuracy and generalization of quality predictions across diverse datasets. Although the exact extent of improvement is yet to be fully verified, LPN shows promising potential as a complementary or even superior multiscale feature extraction method in BIQA.

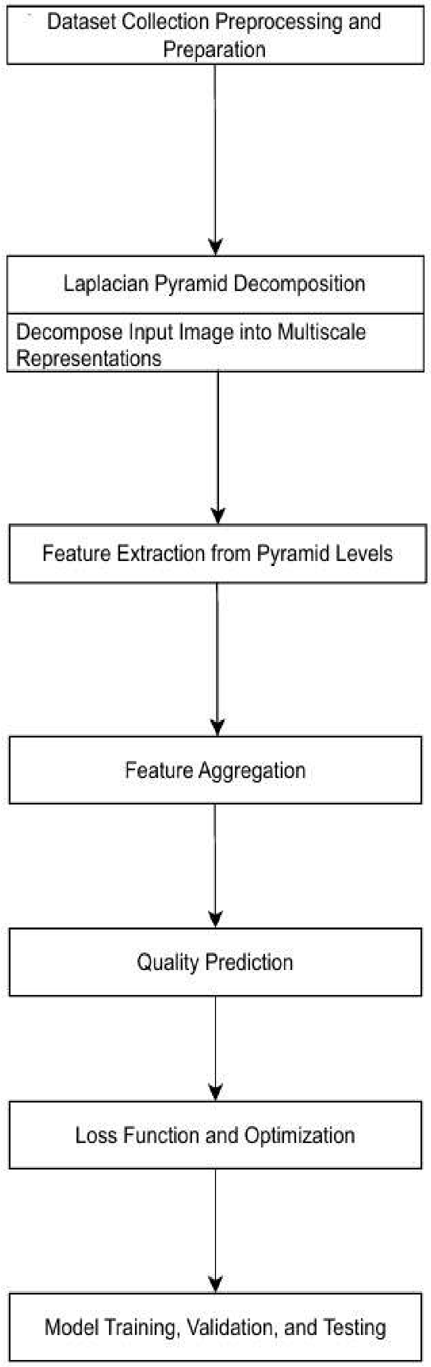

Fig. 1. Laplacian Pyramid Networks-Based BIQA

V. CONCLUSION AND FUTURE SCOPE

In conclusion, this project highlights the effectiveness of multiscale feature extraction techniques, in enhancing Blind Image Quality Assessment (BIQA) scores. By leveraging statistical characteristics and deep learning approaches, we can improve the model's ability to detect and quantify diverse image distortions without requiring reference images. Future work will focus on optimizing the model for computational efficiency, exploring the integration of additional multimodal datasets, and refining the evaluation metrics to ensure robust generalization across varied image quality scenarios.

REFERENCES

[1] M. Wang, "Blind Image Quality Assessment: A Brief Survey," arXiv preprint arXiv:2312.16551, 2023.

[2] P. Yang, J. Sturtz, and Q. Li, "Progress in blind image quality assessment: a brief review," Mathematics, vol. 11, no. 12, p. 2766, 2023.

[3] I. F. Nizami, M. Majid, W. Manzoor, K. Khurshid, and B. Jeon, "Distortion-specific feature selection algorithm for universal blind image quality assessment," EURASIP J. Image Video Process., vol. 2019, pp. 1–13, 2019.

[4] M. Liu, J. Huang, D. Zeng, X. Ding, and J. Paisley, "A multiscale approach to deep blind image quality assessment," IEEE Trans. Image Process., vol. 32, pp. 1656–1667, 2023.

[5] Q. Pan, N. Guo, Q. Li, J. Zhang, and P. Yang, "PMT-IQA: progressive multi-task learning for blind image quality assessment," in *Pacific Rim International Conference on Artificial Intelligence*, Singapore: Springer Nature Singapore, Nov. 2023, pp. 153–164.

[6] C. Yang, X. Zhang, P. An, L. Shen, and C. C. J. Kuo, "Blind image quality assessment based on multi-scale KLT," IEEE Trans. Multimedia, vol. 23, pp. 1557–1566, 2020.

[7] Z. Ni, Y. Liu, K. Ding, W. Yang, H. Wang, and S. Wang, "Opinion- unaware blind image quality assessment using multi-scale deep feature statistics," IEEE Trans. Multimedia, 2024.

[8] X. Lan, M. Zhou, X. Xu, X. Wei, X. Liao, H. Pu, and Z. Shang, "Multilevel feature fusion for end-to-end blind image quality assessment," IEEE Trans. Broadcast., vol. 69, no. 3, pp. 801–811, 2023.

[9] T. Zhou, S. Tan, W. Zhou, Y. Luo, Y. G. Wang, and G. Yue, "Adaptive mixed-scale feature fusion network for blind AI-generated image quality assessment," IEEE Trans. Broadcast., 2024.

[10] H. Wang, G. Wang, W. Xia, Z. Yang, H. Yu, L. Fang, and Y. Zhang, "Blind image quality assessment via deep response feature decomposition and aggregation," IEEE J. Sel. Top. Signal Process., vol. 17, no. 6, pp. 1165–1177, 2023.

[11] R. Cai and M. Fang, "Blind image quality assessment by simulating the visual cortex," *The Visual Computer*, vol. 39, no. 10, pp. 4639–4656, 2023.

[12] N. C. Lynn, Y. Sugiura, and T. Shimamura, "Blind image quality assessment using naturalness-aware multiscale features," *J. Signal Process*, vol. 28, no. 2, pp. 45–55, 2024.

[13] Q. Huang, B. Fang, X. Ai, and T. Nie, "Perceiving multi-layer representations for no-reference image quality assessment," in *ICASSP 2024- 2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Apr. 2024, pp. 3945–3949.

[14] Z. Zhou, F. Zhou, and G. Qiu, "Blind image quality assessment based on separate representations and adaptive interaction of content and distortion," IEEE Trans. Circuits Syst. Video Technol., 2023.

[15] L. Zhang, L. Zhang, and A. C. Bovik, "A feature-enriched completely blind image quality evaluator," IEEE Trans. Image Process., vol. 24, no. 8, pp. 2579–2591, 2015.

[16] T. Bian, "An ensemble image quality assessment algorithm based on deep feature clustering," Signal Process. Image Commun., vol. 81, p. 115703, 2020.

[17] W. Zhang, G. Zhai, Y. Wei, X. Yang, and K. Ma, "Blind image quality assessment via vision-language correspondence: a multitask learning perspective," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit*, pp. 14071–14081, 2023.

[18] T. Zhou, S. Tan, B. Zhao, and G. Yue, "Multitask deep neural network with knowledge-guided attention for blind image quality assessment," IEEE Trans. Circuits Syst. Video Technol., 2024.