Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.7546458

A STUDY OF THE EFFICIENCY OF DEEP CONVOLUTIONAL NEURAL NETWORKS FOR PICTURE QUALITY PREDICTION

Monelli Ayyavaraiah
Computer Science & Engineering Department, Chaitanya Bharathi Institute of Technology, Vidya Nagar, Proddatur, YSR Kadapa (Dist.), Andhra Pradesh ,India -516360

Abstract: Since the advent of smart phones, the act of capturing images has become ingrained in human behavior, evolving into an integral aspect of daily life. The human perception of image quality is a crucial area of research as individuals acquire and interpret images regularly. Automation in image processing systems plays a pivotal role in quantifying image quality, seamlessly addressing factors like blurriness, noise, and compression that can significantly degrade the visual experience. While image processing professionals can promptly identify such distortions, the incorporation of human visual perception into automated system design is often challenging. Consequently, the evaluation of image quality by a machine stands as a critical research domain. Image quality assessment algorithms are employed to gauge the quality of images, with the objective of approximating human visual judgments. Although deep learning has demonstrated remarkable success in addressing various real-world challenges, particularly in tasks like object detection and image classification, its application in image quality assessment is still evolving. The primary challenge lies in assessing image quality without explicit consideration of human perception, presenting a dynamic area of ongoing research. A key hurdle in this research is the scarcity of images accompanied by quality scores, specifically differential average feedback scores provided by image quality experts on a 1-100 scale. This limitation hinders the development of robust models for image quality assessment, emphasizing the need for innovative solutions in this dynamic research domain.
Keywords: Image quality assessment, Human visual perception, Image processing systems, Automation, Deep learning, Visual distortions, Quality scores, Differential feedback

1. INTRODUCTION

There will undoubtedly be advancements made if a significant amount of research and investment is made by large corporations. The evaluation of picture quality is considered to be one of the issues that will be investigated. Generally speaking, the quality of the pictures is reliant on the individuals who see them; nevertheless, the objective of Image Quality Assessment (IQA) is to reflect image quality in conjunction with individual outstanding comprehension [1]. Image quality assessment is generally broken up in to test free of reference with regard graphics. The gap between your evaluations is the fact that NR-IQA comes with a supreme quality reference image, where as a twisted image will there be from the NR-IQA. Generally, it's straightforward to appraise the standard tag of an image whenever there's because of the Structural Similarity Index, a standard to measure against. In general, the image alignment is what is required to get a no-reference picture, means that the samples were then calculated, and the scores have been generalized to individual comments. The shape of the image feature is the selection of a person to understand and it is a very labor-intensive process and should not be. If it comes to the countless for the graphics, it's practically impossible. Also, can leverage the potential of the deep learning and pattern recognition for the work to maintain the distortion may be the frustrating process automatically. As a result, extensive research has been conducted, leading to the proposal of various models employing deep neural networks to assess image quality. Both classification and regression paradigms have proven effective in the development of these neural networks. “Since classification is mainly done in graphics, this project intends to categorize the images in to four tags predicated on the superior scores which range from 0-100” [2]. The data set for it has been got by the University of Texas in Austin video and image production internet site and is made up of 982 graphics with five kinds of distortions. The image quality scores can also be daunted using by hand expressed border maps. The LIVE-IQA data set stated earlier was used. Additionally, TID 2008 and also CSIQ data sets were also used. They'd 1700 and 866 images, respectively, together using most of the 4 distortions JPEG, JPEG2000, and Gaussian Blur.

2. RELATED REVIEW

The key objective of image quality assessment will be always to look for certain kinds of algorithms which may gauge the quality of videos and images concerning human visual understanding. Image Quality Assessment (IQA) calculations generally translate image quality determined by the twisted image’s resemblance to a high-profile benchmark image [1]. Generally, the top high superior forecast is modeled in order the visual features are all near the optical apparatus [1]. They calculated that the image's caliber in reference and with reference utilizing a variety of metrics like Structural Similarity Index (SSIM) from the plasma and wavelet domain names. The predictions had been near to individual abstract remarks [1]. Xiang et al. proposed a deep neural model that utilized a modified version of the AlexNet architecture, incorporating fully connected layers and employing regression to predict distorted images [2]. They used a pair of thoughts; disperse the method of five kinds of twisted graphics, as well as the superior scores into these image quality descriptions. They used the CNN version on the ImageNet data set could be your bottom version [3].The tests were carried out with Gaussian Blur and JPEG images. Scale model of the sub-images of random pooling helps to avoid the problem. Contain maps are pulled just one time in the image. Arbitrary places are chosen in the image and so therefore are ready to offer fixed-length vectors. This had been modified to accommodate this dilemma of perfecting the jagged images dependent on their caliber. Talebi and Milanfar suggested neural image Assessment (NIMA) by which a profound CNN is a good-looking image perfect for the general user will be trained to assess [5]. To provide an understanding of the scores indicated in the network is connected to the video image is also used. Editing and processing of images is used to enhance photographs.

3. PROPOSED METHODOLOGY

Deep Learning: Operation of your brain may be the deepest essence of learning. Some of the objects in all of us, when training for specific tasks, in-depth learning and some of the things that will make the production of the master computer. Some of the statistical data of the means of education, then this can be achieved, which is in constant motion, [5]. Inflammation of the brain, which is not always due to individual neurons, the nervous systems operate in the set, in order to work with a specific weight matrix corrected and discover the best methods. Want to achieve human-level accuracy of deep learning, even though they have a lot of information and training to successfully perform the tasks taken [6].Deep Learning has turned into probably perhaps one of the very significant engineering in the 21st century also results in a lot of discoveries in comprehending pc vision shape. Depending on the task being performed by neural network, it will be composed of input coatings, outputs and some hidden layers also.

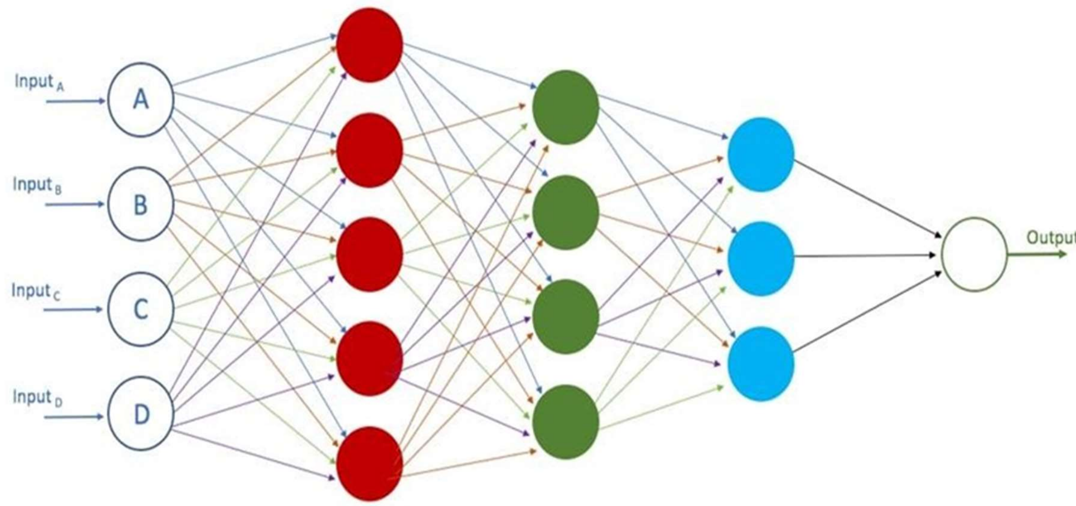

Figure 1: Deep Learning Neural Network Model [6]

3.1 Image Distortions

Reduction in graphics will be employed by an approach which is nothing but distortion in image processing. Sometimes, distortions might also be brought on by the lack of info whilst shifting information through a network as a result of packet losses. Generally, straight lines from the graphics are all made to distort a graphic. There are various sorts of image compression methods such as jpeg, jpeg 2000, white sound, fast evaporating, and blur. It uses both lossy and lossless compression, so with any jpeg, it just uses a lossy compression method. Gaussian Blur, Gaussian blur, and is a type of character from the movie, along with image reduce disturbance. When the rapid evaporation takes place in the distribution of wireless networks in the graphic.

Image Quality Prediction can be really actually a considerable area of study in computer perception, and, scientists are still researching to generate publication procedures for predicting image quality concerning individual comprehension. The projected job within this project targets accepting the top excellent forecast because classification and regression issues and betters the outcome of the prior works, especially in regression [3]. The images are arranged into categories according to the caliber tags that were created during one of the scoring periods. Within the context of the recession, the picture quality score is determined by extracting the features from the visuals and incorporating them into the DNN's.

“A few of the models suggested used architectures most useful in graphic design and object discovery, for example as for example for instance VggNet, ResNet, etc. [1], [2]”. Trained on supreme quality graphics, when we contemplate it tightly, generalizing a neural network for so various kinds of distortions is rather tricky. Additionally, the thing detection models possess the attributes observed for supreme quality graphics. Since the features are somewhat twisted in image quality forecast, the exact models don't carry out economically. The twisted graphics have twisted features, and show maps could comprise that advice which might not be helpful to predict an image's caliber [9].

3.2 Proposed architecture for classification

The architectures employed in the majority of experiments drew inspiration from the AlexNet [2] architectures, typically featuring 23 filtered convolution layers [12], along with pooling layers. These architectures also included a set of 23 repetitions and fully connected layers, subsequently shaping the input into a specified structure. The overall structure used is shown at Figure 3. This figure indicates the normal structure employed in this work [14].

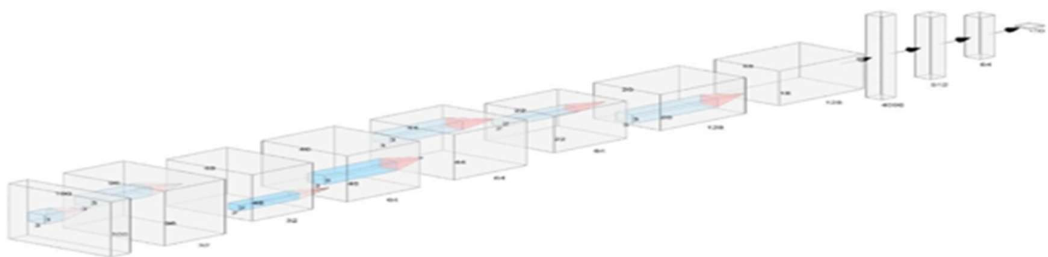

Figure 2: Proposed Deep CNN Architecture

3.3 Methods and Materials

"The projected work plays regression on most of the graphics from the 3 data sets Live-IQA, CSIQ, TID2008, along with respective distortions JPEG, JPEG2000, White Noise, and Gaussian Blur and accomplishes exceptional results. Generally, that the convolution neural networks have been stripped following a stringent learning process through a great number of iterations" [11]. "In issues such as object discovery and image classification, CNN's figure out how to identify patterning in steps" [14]. In situations when you have a significant quantity of information at your disposal, this method of learning is highly effective. Object identification and picture design are both activities that are associated with enormous data sets; accordingly, CNN is able to extract features during the process of teaching. In issues such as image quality forecast, you can find not many graphics with standard scores, so that consequently, profound neural networks fight to discover patterning.

In this task, a novel framework is proposed to extract edge features from an image using the Sobel filter and generating edge maps. Edge channels are generated by computing gradient approximations at each pixel in the image and determining the gradient magnitude. Sobel kernels are applied for image convolution while calculating the gradient. The edge maps are separately generated for portrait and landscape images, ensuring coverage of the entire image by considering different width and height orientations. The edge map is initially extracted for the left half of the image and subsequently for the right half in landscapes [16]. For portraits, the edge maps are first obtained for the top half and then for the bottom half. These edge channels are further divided into 2 to 8 intervals based on the magnitude of the original dataset. Because of this, the input is raised 24 times. The benefit channels are then passed to 2 stacked 3convolution neural networks with 16 filters together with standard metering. It is possible to make use of spatial pyramid pooling in an appropriate manner in order to deal with pictures of any size, and non-core regression may be carried out in order to make a prediction about the image quality score. Between 0 and 1 was the range of the findings. Before this, models that are currently in use did not make use of plasma volcano pooling.

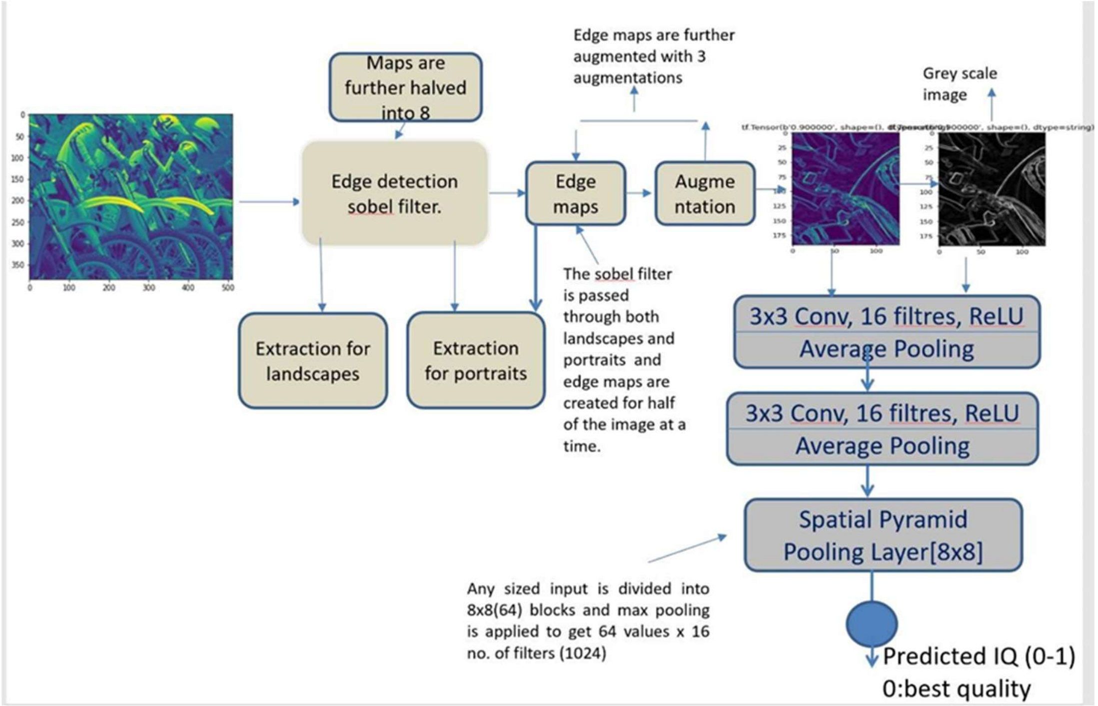

Figure 3: Regression Model Architecture

80 per cent of this data was employed for training, and 20 percentages is useful for testing, and also the version has been well trained for 40 epochs. The suggested version performs a lot much better compared to OG-IQA [1] model and the bottom research-paper [3]. In addition, it functions well compared to CURVLET- 2014, [4] model and the m 1 [14] version. The version out performs the no-reference image quality assessment such as BRISQUE, DEEP QA.

4. EXPERIMENTAL EVALUATION AND COMPARISON

4.1 Interpretations of Classification Models

As illustrated in the proposed model, comprehensive experiments were carried out with various variations. A kernel size of 3x3 was consistently chosen for the majority of the experiments conducted. The experiments involved dividing the images into patches, 2x and 4x up-scaled images, as well as full images. One factor in play was the up-scaling of image quality labels; for instance, in the case of a 2x up-scaled image, the quality label was up-scaled by 1 and for 4 x up-scaled images, it was up-scaled by 2.

Model Name	Model Details
CNN_Full_img	stacked 16 filtered convolution layers with spatial pyramid pooling and 4 fully connected layers
CNN_2x	2 stacked 32 filtered convolution layers with spatial pyramid

	pooling and 4 fully connected layers
CNN_4x	3 stacked 16 filtered convolution layers with spatial pyramid pooling and 4 fully connected layers.

Table 2: Proposed model variations

4.2 Full image model of CNN

This design is made with full graphics with three pile 16 filters using a modified adrenal pyramid pooling layer to take care of the different sizes facing the same reel finishing graphics precision and activation. The instrument has a management error of 87% with accuracy and 79.45% accuracy considering 4.2486.

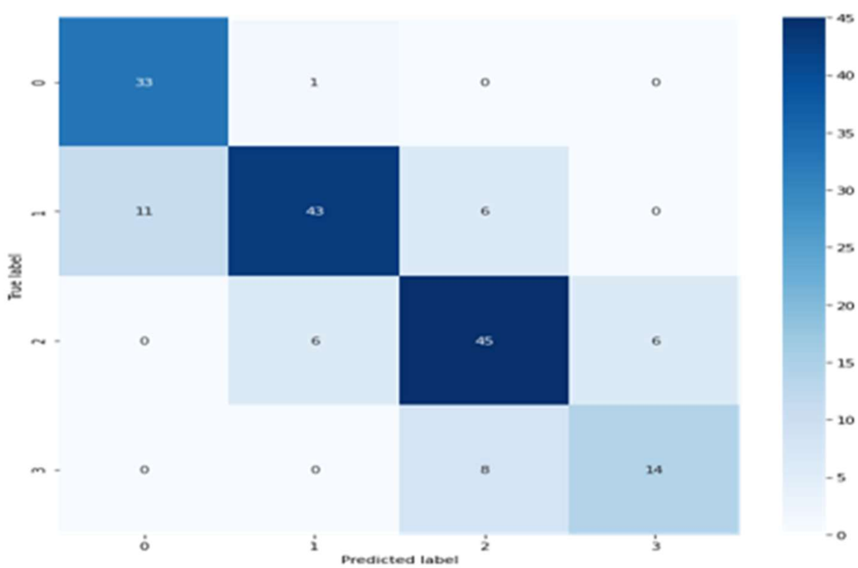

Figure 4: CNN_Full_imgConfusion Matrix Model

The outcomes are presented with respect to the confusion matrix, revealing weighted accuracy, recall, and F1 scores of 0.785, 0.780, and 0.778, respectively. The accompanying figure depicts the confusion matrix, illustrating the true positives, true negatives, false positives, and false negatives values for each label (0, 1, 2, 3).

4.3 CNN_1_Mvote

The architecture that is described in Table 4 has been used in order to extract a greater number of features by using an increasing number of filters. As a result of the fact that the majority of the distortions are collected concurrently, a comprehensive selection of characteristics is going to be extracted in the near future. An additional drop-out has been included in order to remove a few of those characteristics from the feature matrix. After then, the assessment set was divided into places, and each patch tag was given its own unique assignment, along with also a vast majority vote is extracted from all of the patch forecasts to acquire the superior tag to get a graphic. A precision of 82.4percent has been obtained. The weighted accuracy, recall, and F1 scores to this particular version were all 0.81182, 0.8292, and 0.8195.

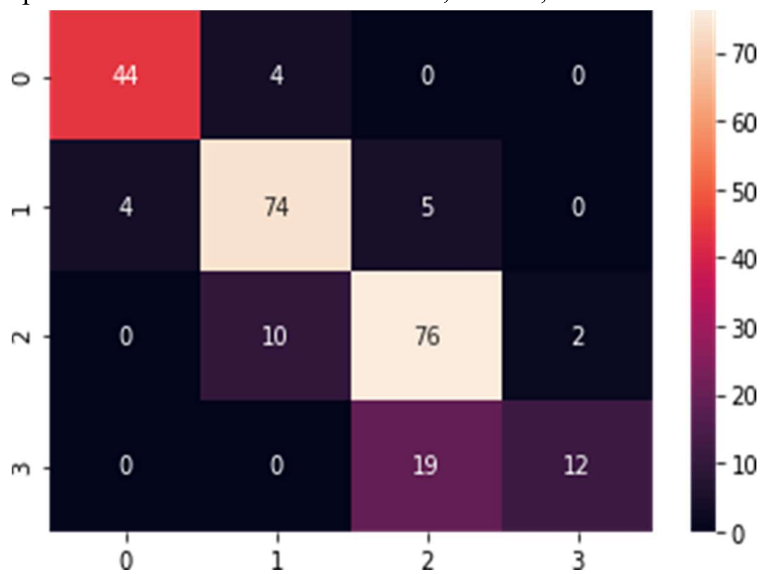

Figure 5: Confusion Matrix of CNN_1_Mvote

Figure 5 shows the confusion matrix results for CNN_Mvote_1. The results provide: 1) the number of correct and wrong predictions. 2) The total number of predictions.

4.4 Dataset Results Comparisons

The operation comparisons have been in relation to SROCC, KROCC, also PLCC.

Table 3: Comparison table of SROCC, KROCC, and PLCC for the LIVE IQA

	Model Name	BLUR	JPEG	WN	JP2K	ALL
Spearman Rank Order Correlation Coefficient (SROCC)	Deep QA	0.950	0.940	0.890	0.958	0.940
	BLISS-S	0.869	0.922	0.779	0.919	0.898
	Proposed	**0.972**	**0.979**	**0.944**	**0.984**	**0.947**
	CURVLET-2014	0.665	0.635	0.696	0.574	-
	M1	0.659	0.633	0.728	0.646	-

	Proposed	0.896	0.833	0.797	0.868	0.797
	BLISS-S	0.875	0.955	0.748	0.945	0.910
	SSIM	0.954	0.964	0.816	0.971	0.902
	Proposed	0.977	0.989	0.947	0.984	0.947
	Proposed	0.977	0.989	0.947	0.984	0.947

Table 3 shows the SROCC, KROCC, also PLCC comparisons to its suggested version and also the Existent IQA TID 2008 database distortions. The suggested version outperforms the present IQA versions on the regression models.

5. CONCLUSIONS AND FUTURE WORK

Image quality forecast continues to be an active subject of research. The units and processes suggested in this project make an effort to consider graphic quality forecast because of being a multi-class classification issue. Additionally, it will help decide if a specific image, based upon its quality, is of good use for your own image classification issue. The version investigates the up scaling of these graphics and alters that how the caliber labels so to generate a system which forecasts labels of categories well and uses them to enhance the complete system's results. Additionally, it utilizes a vast majority voting plan, unlike another image forecast issues [4]. "The evaluation set is analyzed on the image stains, and accuracy is figured on the stains in contrast to the entire image, usually the initial task" [6]. If it data is extensive, then neural systems might be trained, and also the outcomes could function improved. While collecting personal cognitive quality, the data could be difficult to examine to identify individual issues and datasets could be ordered to achieve specific image distortion to improve training patterns [17]. This is really just a time-consuming and feverish endeavor, though. Additionally, it could also be of use in the event the dozens can be accumulated through adverts online from arbitrary users for businesses analyzing the image quality appraisal [16]. Because of this, more diverse and broad datasets are available and analyzed for better outcomes.

6. REFERENCES

[1] Z. Wang, A. C. Bovik, H. Sheikh, and E. Simoncelli, "Image quality assessment: From error visibility to structural similarity," IEEE Transactions on Image Processing (TIP), vol. 13, no. 4, pp. 600–612, May 2004.

[2] A. Mittal, A. K. Moorthy, and A. C. Bovik, "No-reference image quality assessment in the spatial domain," IEEE Transactions on Image Processing (TIP), vol. 21, no. 12, pp. 4695–4708, December 2012.

[3] K. Ma, W. Liu, K. Zhang, Z. Duanmu, Z. Wang and W. Zuo, "End to end blind image quality assessment using deep neural networks," IEEE Transactions on Image Processing (TIP), vol. 27, no. 3, pp. 1202-1213, March 2018.

[4] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial Pyramid Pooling in Deep Convolution Networks for Visual Recognition," IEEE Transactions on Pattern Analysis and Machine Intelligence (TPAMI), vol. 37, no. 1, pp. 1904-1916, June 2014.

[5] H. Talebi and P. Milanfar, "NIMA:Neural Image Assessment," IEEE Transactions on Image Processing (TIP), vol. 27, no. 1, pp. 3998-4011, June 2018.

[6] B. Tierney, "Understanding and Building Neural Network Machine Learning Models", [Online]. Available: https://developer.oracle.com/databases/neural- network-machine-learning.html [Accessed: April 23 2020].

[7] A. Karpathy "CS231n: Convolution Neural Networks for Visual Recognition", [online] Available at: https://cs231n.github.io/convolutional-networks/ [Accessed 23 April 2020].

[8] A. Krizhevsky, I. Sutskever, and G. E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks.," In Proceedings of NIPS Conference on Advances in neural information processing systems, Lake Tahoe, Nevada, pp. 1097-1105, December 2012.

[9] K. Le, P. Ye, Y. Li, and D. S. Doermann. "Convolutional Neural Networks for No-Reference Image Quality Assessment," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (TIPAM), Columbus, Ohio, pp. 1733- 1740 April 2014.

[10] B. Sebastian, D. Maniry, T. Wiegand and W. Samek. "A deep neural network for image quality assessment.", in Proceedings of IEEE International Conference on Image Processing (ICIP), Buffalo, New York, pp. 3773-3777, September 2016.

[11] W. Xiaochuan, K. Wang, B. Yang, W. B. Frederick and L. Xiao-Hui. "Deep Blind Synthesized Image Quality Assessment with Contextual Multi-Level Feature Pooling." In Proceedings of IEEE International Conference on Image Processing (ICIP), Piscataway, New Jersey, pp. 435-439, May 2019.

[12] X. Yang, F. Li, and H. Liu. "A Comparative Study of DNN-Based Models for Blind Image Quality Prediction," in Proceedings of IEEE International Conference on Image Processing (ICIP), Taipei, Taiwan, pp. 1019-1023, September 2019.

[13] J. Kim, H. Zeng, D. Ghadiyaram, S. Lee, L. Zhang, and A. C. Bovik. "Deep Neural Models for Picture-Quality Prediction: Challenges and Solutions to Data-Driven Image Quality Assessment," IEEE Transactions on Signal Processing (TSP), Vol. 34, no. 6, pp. 130-141, Nov. 2017.

[14] L. Kong, A. Ikusan, R. Dai, and J. Zhu. "Blind Image Quality Prediction for Object Detection," in Proceedings of IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), San Jose, California, pp. 216-221, November 2019.

[15] W. Zhang, K. Ma, J. Yan, D. Deng, and Z. Wang. "Blind Image Quality Assessment Using Deep Bilinear Convolutional Neural Network," IEEE Transactions on Circuits and Systems for Video Technology (TCSVT), Vol. 30, no. 1, pp. 36-47, 5 January 2020.

[16] W. Szegedy, Y. Liu. "Going deeper with convolutions. In: Computer Vision and Pattern Recognition," in Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Boston, Massachusetts, pp. 1-9, June 2015.

[17] O. Messai, F. Hachouf, and Z. A. Seghir. "Deep Learning and Cyclopean View for No-Reference Stereoscopic Image Quality Assessment," in Proceedings of IEEE International

Conference on Signal, Image, Vision, and their Applications (SIVA), Guelma, Algeria, pp. 1-6, November 2018.