Perceptual Quality Assessment of Digital Images in the Absence of Reference Information

Nisar Ahmed[1,2], Muhammad Usman Younus[3,4*], Kalsoom Safdar[5,6*]
[1]Department of Computer Engineering, University of Engineering and Technology Lahore, (54890), Pakistan
[2]Institute of Biomedicine, University of Turku, (20014), Finland
[3]Department of Computer Science & IT, Baba Guru Nanak University, Nankana Sahib, (39100) Pakistan
[4]Ecole Mathématiques, Informatique, Télécommunications de Toulouse, Université de Toulouse, (31400), France
[5]Department of Computer Science & IT, University of Jhang, Jhang, (35200), Pakistan.
[6]Faculty of Intelligent Computing, Universiti Malaysia Perlis, 02600 Arau, Perlis, Malaysia.
*Correspondence: usman1644@gmail.com

Citation | Ahmed. N, Younus. M. U, Safdar. K, "Perceptual Quality Assessment of Digital Images in the Absence of Reference Information", IJIST, Vol. 8 Issue 5 pp 2047-2066, August 2026.
DOI | https://doi.org/10.33411/IJIST/1996
Received | July 31, 2026 **Revised |** Aug 18, 2026 **Accepted |** Aug 20, 2026 **Published |** Aug 21, 2026.

Digital multimedia content experiences a variety of degradations during the process of acquisition, processing, and transmission. Automatic approaches are required to evaluate the performance of any of these processes regarding perceptual image quality assessment. Several human observers can judge an image based on its perceptual quality, and the results may then be averaged to obtain an image quality score, making it the most accurate way of assessing image quality. Because this approach requires the engagement of human resources, an automatic method based on machine learning modelling of the human visual system is necessary. Machine learning techniques may be trained in a supervised manner by providing a set of images along with their subjective quality score as ground truth. This subjective score is sometimes expressed as the Mean Opinion Score, which is the average of numerous human judgments (MOS). In this study, we have trained a deep Convolutional Neural Network (CNN) using labelled images. The study focuses on images that are naturally distorted during acquisition, processing, or transmission rather than on synthetically distorted images with discrete distortion levels. Using naturally distorted images will result in a more general-purpose image quality evaluation model can be obtained. Furthermore, collecting a significant number of labelled images for image quality is an expensive task, and no dataset with a sufficient number of images is available, thus several ways for training the CNN will be investigated. The trained model was evaluated by determining the correlation between the ground truth (MOS) and predicted quality scores, resulting in PLCC and SROCC values of 0.7729 and 0.7489, respectively of 0.7729 & 0.7489, respectively. However, when tested on the BIQ2021 dataset, the PLCC and SROCC values were 0.7845 and 0.7588, respectively.

Keywords: Image Quality, Image Quality Assessment, Deep Learning, Convolutional Neural Networks.

Introduction:

The Human Visual System (HVS) is a highly evolved sensory organ in the human body and is responsible for our ability to view the world around us. People absorb more than 70% of the information that they take in via their sense of vision. Processing visual images is the domain of the human brain's visual cortex, which is located deep inside the cerebral cortex. The visual system of a human being is the most complex and sophisticated among species in the animal world. As a consequence of this, visuals have superseded text as the major information source for humans. On the other hand, this data is very sensitive to both the image's characteristics and the observer's visual perception capabilities. The significance of visuals in communication and the presentation of information has significantly increased as a direct result of this phenomenon and the development of digital multimedia technologies. Digital images are representations of an object's visual characteristics that take the form of electrical impulses when taken by a digital camera. The quality of a digital image may be degraded by artefacts that are introduced into the signal during its acquisition, processing, storage, or transmission. In light of this, it is essential to evaluate the performance of different systems about the acquisition, processing, storage, and transfer of data by evaluating the quality of digital images.

The quality of an image as a human observer sees it is referred to as the perceptual image quality. Evaluation of the image's quality is essential, given that image artefacts result in data loss. Image Quality Assessment (IQA) may be performed in various ways, primarily through subjective and objective quality evaluation. Because humans are the ultimate consumers of visual content, subjective quality evaluation is the best and most conclusive method. Managing human resources is an expensive task that obviously cannot be undertaken in every circumstance. As a result, objective methods that use algorithms are used to evaluate image quality. Therefore, objective quality evaluation is simple and may be used in a wide variety of applications, such as broadcasting, the streaming of videos, the enhancement of image processing algorithms, and many circumstances involving real-time quality assessment. Conventional methods such as Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR) take the original image as a reference to perform the quality assessment of the processed image. Other ways of performing objective quality assessment depend on reference information availability. Even though these traditional methods have been shown to have a poor correlation with the perceptual quality of images, their more recent counterparts, such as the Structural Similarity Index Metric (SSIM) and its variants Multi-Scale SSIM (MS-SSIM), three-component SSIM (3-SSIM), Complex Wavelet SSIM (CW-SSIM), and Information Content Weighted SSIM (IW-SSIM), are better candidates for quality assessment in the presence of reference information. Despite this, full-reference image quality evaluation algorithms correlate highly with human judgement. Since it is either very difficult or impossible to make reference information accessible to end users, their practical applications in the real world of multimedia content dissemination are still limited.

Methods of quality assessment that do not require reference information can be used for quality assessment in many scenarios. The bulk of the research that has been done in the area of no-reference Image quality assessment has depended on small-scale image datasets that include a relatively low number of images and distortions. If an image quality assessment is trained on a database containing one type of distortion and it is used for a database containing another type of distortion, the algorithm will not achieve very good results from the algorithm. In addition, the performance of image quality evaluation algorithms trained on databases of simulated distortions could be better when applied to images with simultaneous genuine distortions. When contrasted with full-reference Image quality assessment, the performance of no-reference Image quality assessment algorithms has a long way to go before it can be

considered adequate. Since the prediction algorithm does not have access to an undistorted version of the image, no-reference image quality assessment solutions often perform below standard. In addition, most of the existing methodologies have tried to address the problem of no-reference Image quality assessment by training regression algorithms using features extracted from the statistics of real-world situations [1][2][3]. Therefore, there is a need to train deep learning models on large, representative datasets with genuine distortions, which is addressed in this research.

To model image quality assessment, we focus on Convolutional Neural Networks (CNN) that are inspired by the visual cortex of the human brain and have demonstrated implausible performance in many visual identification tasks. The end goal of this strategy is to replace subjective human evaluation with objective algorithmic assessment of image quality.

Training several CNN models on no-reference quality assessment datasets to identify the best-performing model.

Weakly supervised pre-training of CNN models and fine-tuning on natural distortion dataset to obtain exceptionally high performance.

Performing ensemble construction using meta-learning to improve the performance further and obtain competitive performance with existing approaches.

Materials and Methods:

Among the various categories of image quality assessment, no-reference quality assessment is the most challenging and requires further attention of the research community. Distortion-specific quality assessment deals with specific categories of image distortions and has specific application areas. This study is targeted to general purpose image quality assessment, which does not deal any specific distortion category and thus adds to the complexity of the problem. Some significant studies addressing the problem of image quality assessment are discussed in this section.

Natural Scene Statistics (NSS) is traditionally employed to address the problem of no-reference image quality assessment [4]. Various techniques based on natural scene statistics-based image quality assessment have been presented recently [1][3][5][6]. Since deep learning has outperformed NSS in image recognition, it has been the primary focus of IQA research. Because deep learning features are learned automatically from training images, they eliminate the need for laborious and error-prone feature engineering. The challenge with deep learning-based approaches is the need for more training data since the available training images for IQA are insufficient. On the other hand, researchers have created deep learning-based solutions to solve the issue of IQA.

[7] designed a CNN architecture by combining two architectures. The designed model is deeper and more complex but has demonstrated superior performance. They performed pre-training on this model using pseudo-labelled images with simulated distortion. The pre-trained model is fine-tuned on their self-collected dataset, having 12,000 naturally distorted images. The evaluation of the trained model is performed by performing several experiments. The model is compared with a few recently proposed methods and superior performance is obtained. They have performed experiments to check the dataset dependence and validate the generalization, which is also encouraging.

[8] used the VGG16 architecture for quality evaluation, reasoning, that different levels of convolution reflect image quality differently. After training an SVR on the features acquired at the various stages of convolution, they carried out average pooling to produce an image quality score.

Image quality can be predicted in two stages, according to a proposal by [9]. During the first step, a subjective quality score is predicted, and during the second stage, an objective error map is predicted. [10] has likewise chosen to handle this problem in a two-stage fashion. The first stage makes a prediction regarding the kind of distortion, and the second stage makes

use of a particular deep model to make a prediction regarding the subjective score and then performs an average pooling with the predictions made by other deep models. In addition, [11] developed a network with two stages: the first stage would determine the kind of distortion, and the second stage would do the subjective score prediction with the help of a quality prediction network that was specifically designed for each type of distortion.

In addition, [5] presented a two-step framework, the first stage of which detects the kind of distortion and the degree to which it is occurring. In the second step, a specific CNN model is used to make an assessment of the quality of ty. Their models were developed to be able to make quality judgments on distorted images that were either legitimately or artificially deformed. Additionally, [6] use a two-stage strategy, in which they first make a prediction about the kind of distortion, and then proceed to make a quality score prediction via several CNNs.

The results of using deep features to evaluate an image's quality have been surprisingly positive [12]. Although straightforward and efficient, these techniques suffer from the inability to be trained comprehensively from beginning to end. To evaluate the quality of an image, a regression method is developed using previously trained models, such as VGG [13]. [14] presented a on the use of deep deep-feature-based image quality assessment. They have used several pre-trained CNN models to perform feature extraction at different layers. These features are used to train regression algorithm for quality assessment. Similarly, fine-tuning of these models is performed followed by feature extraction for regression modeling. It is demonstrated that fine tuning significantly improves the predictive performance of deep features. The final model is constructed using Gaussian process regression and reasonable performance is obtained.

[15] proposed a method to evaluate the quality of an image using active inference. Active inference is what they call the process by which the human visual system's own generative mechanism extrapolates relevant information from images for understanding. With these presumptions in place, their suggested method employs active inference by means of main content prediction using a generative adversarial network. This basic material is then used to evaluate the image quality in the next phase. Image quality is affected by a number of factors, including the nature of the content, the kind of distortion, and the level of degradation. These three factors are calculated using a multi-stream convolutional neural network trained on the correlation between the original information and the distorted image. They used a technique based on active inference and multi-stream CNN to evaluate image deterioration comprehensively, and their method beat five benchmark datasets.

Ensemble learning is a popular approach to combine various learners to provide superior predictive performance. [16] proposed an ensemble learning method which use training snapshots obtained through cyclic learning rate schedule. The model snapshots are intelligently combined to construct an ensemble which provides superior performance to a single model. Similarly, [7] proposed an ensemble which is jointly trained for image quality assessment.

Although, various attempts are made towards the problem of no-reference quality assessment, but the problem is far from completion. Deep learning-based solutions are more successful in contrast to natural scene statistics-based approaches. Although, they have good performance in various benchmark datasets but most of the existing dataset are constructed by simulating various distortions on image with pristine quality. Designing and evaluating an image quality assessment model based on simulated distortions doesn't represent the problem sufficiently. Therefore, this research is directed to design a solution using image datasets containing naturally occurring distortions only. Live in the wild challenge dataset [17], KonIQ-10k [18] and BIQ2021 [19] are large scale dataset of images with naturally occurring distortions and are used for model building and evaluation.

Recent research demonstrates that no-reference image quality assessment (NR-IQA) has increasingly moved from conventional hand-crafted statistics toward deep feature learning, multi-scale representations, attention mechanisms, and hybrid CNN–Transformer architectures. In [20] investigated global and local image statistics for NR-IQA, while in [21], proposed a multibranch CNN that combines spatial- and gradient-domain information to better capture distortion characteristics. In [22], further demonstrated the usefulness of pretrained architectures and unified learning for NR-IQA. More recently, the researcher [23] introduced a reference-knowledge distillation strategy that enables NR-IQA models to exploit knowledge from non-aligned reference images, reporting strong performance across eight standard datasets. These studies indicate that exploiting complementary representations is increasingly important for accurately predicting perceptual quality without access to pristine reference images.

Further, the research direction has expanded toward distortion-aware learning, Transformer-based feature extraction, semantic information, and improved cross-dataset generalization. In [24], incorporated degradation priors and semantic features within a multi-task NR-IQA framework, while some researcher [25] developed an error-aware reconstruction network that jointly considers image errors and content information. In [26], proposed a dual-attention pyramid Transformer to capture richer multi-scale quality features. In 2025, recent approaches explored semantic-guided multi-scale feature extraction, dual-perception hybrid networks, and multi-layer cross-modal prompt fusion. More recent 2026 studies [27] have investigated multi-scale dynamic modulation and bidirectional feature fusion for authentic distortions, highlighting that robust generalization remains a major challenge because real-world distortions are heterogeneous and often spatially non-uniform

The problem of image quality assessment has a wide range of applications, from benchmarking to video streaming/encoding. It is a long-standing problem which is addressed through various approaches based on natural scene statistics or deep learning. Moreover, the application of image quality assessment based on the availability of reference information calls for a full-reference or no-reference approach. Full-reference algorithms have achieved widespread adoption and are increasingly being deployed, whereas the no-reference algorithms lack generalization. A no-reference algorithm trained on a specific image quality dataset generalizes poorly on other datasets or an independently collected test dataset. In order to address this challenge, we are targeting the problem of no-reference image quality assessment and verified its generalizability through cross-dataset experiments and a self-collected test set.

The problem of assessing image quality may be addressed using a variety of methods. However, to take a more systematic approach, we have searched for prominent techniques to solve these problems. The KDD, SEMMA, and CRISP-DM frameworks are examples of widely used methods of the most widely used methods for machine learning and data science projects in both academic and industrial settings. Each approach has unique advantages and disadvantages, but ultimately, it is up to the analyst to choose which one to put into practice. The CRISP-DM approach, a six-step machine learning project development process, has been our primary way of dealing with image quality assessment. In order to make it more suitable to our circumstances, some basic modifications are incorporated into the model. Figure 1 provides a graphical depiction of a six-stage process, and further information about each of these stages is provided below. The circular nature of the problem is illustrated by a continuous circle, which says that the creation, assessment, and deployment of a solution may reveal areas for improvement, resulting in the formation of new problem statements, and the process will continue.

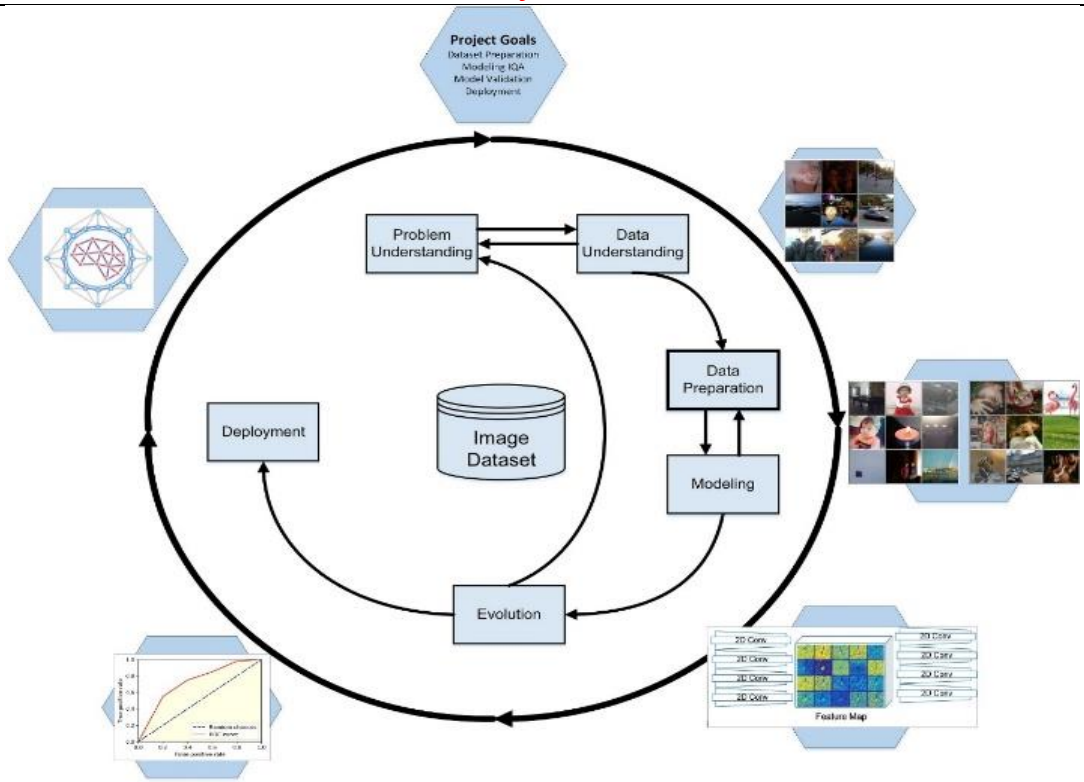

Figure 1. CRISP-DM adoption for the problem of image quality assessment

Problem Understanding:

The first step of the process is defined as problem understanding, and it is outlined in the section devoted to the problem statement. The purpose of the project is to investigate the no-reference quality assessment of digital images in such a manner as to ensure that the predicted quality score has a good correlation with the subjective quality ratings given by human observers. An algorithm for predictive modeling will be trained using an annotated dataset that contains subjective ratings in the form of Mean Opinion Scores (MOS). This algorithm will then be tested using a holdout dataset of images or a dataset of images that the user has gathered themselves. In addition, the assessment will be carried out based on correlation measures that will be defined in more detail later on in the evaluation stage.

Dataset Understanding:

Data is the most important component to the success of any machine learning algorithm, since the learning process depends on the training dataset to construct its decision boundaries. A dataset is considered to be representative when it correctly depicts the situation at hand, as shown by its wide variety of training images and its reliance on precise labelling. In order for the algorithm to learn the perceptual quality rather than the peculiarities of the data, the training examples should comprise pictures of different perceptual quality and have adequate content variety. We are not utilizing a dataset with particular sorts of distortions like the one that is simulated by image distortion models since we are creating the quality evaluation model for general purpose images. This is because we are designing the model for general purpose images. The datasets BIQ2021 and KonIQ-1K are the most extensive ones in this category; nevertheless, the Live in the wild challenge dataset, often known as LiveCD, is one of the most frequently used datasets. Here are some specifics about these three datasets:

LiveCD:

Researchers are able to analyze the effectiveness of their algorithms by using databases for objective image quality evaluation, which brings the algorithms closer to their ultimate goal of matching human perception. The great majority of image quality databases that are publicly

available were produced in laboratory conditions that were strictly controlled, with high-quality images being subjected to various simulated distortions at different levels. On the other hand, the synthetic distortions that have been detected in the most recent databases may not be able to accurately recreate the distortions that are seen in images that were collected using a typical real-world smartphone camera or other devices.

For this reason, LiveCD is a database consisting of naturally distorted images so that it could act as a benchmark for any future IQA databases. The LiveCD has a total of 1162 images that have organic distortions and were captured using a wide range of mobile devices. The images were captured, processed, and stored using the participants' mobile devices without any additional distortions being knowingly introduced on purpose. The dataset is scored via Amazon's crowdsourcing platform known as the Mechanical Turk, with the results of extremely extensive, multi-month-long image quality rating studies serving as the basis. The collection contains the opinions of a wide variety of participants, all of whom were asked to judge the overall quality of the images. An average of 175 different participants evaluated each image and evaluated it on a scale that goes from poor to excellent after seeing it online. There is a grand total of 350,000 ratings, which were collected from 8100 unique participants. In addition to the images, the database includes both the mean and the variance of the Mean Opinion Scores (MOS), which are obtained from the subjective ratings.

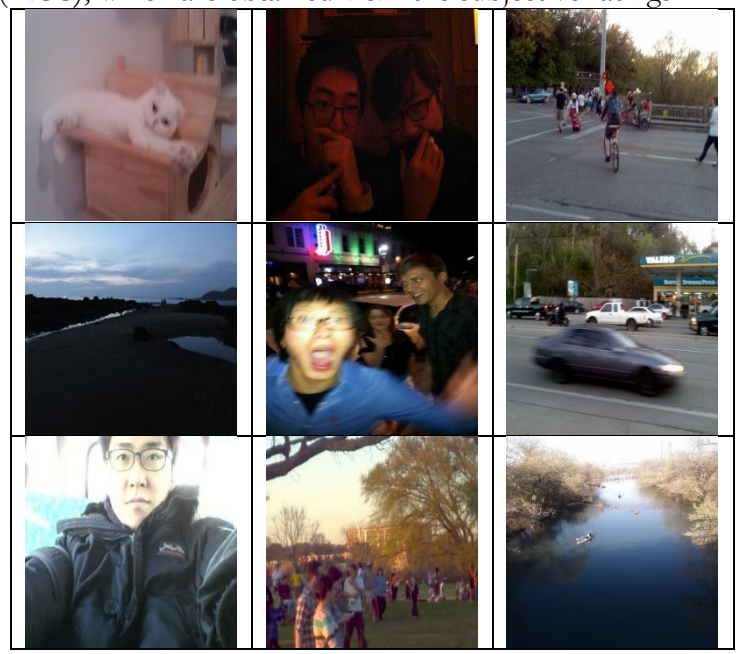

Figure 2. Randomly sampled images from LiveCD dataset

Figure 2 provides nine randomly sampled images from LiveCD dataset. These images indicate the content diversity and the range of perceptual qualities covered by the images in the dataset. It is to be noted that these images are captured using various mobile or handheld devices and are not post processed for the purpose of dataset preparation except cropping or resizing.

KonIQ-10K:

The KonIQ-10k image quality assessment dataset contains 10,073 images that have been rated for their quality. In order to build the dataset, a novel procedure that is both systematic and scalable is used. The result of this method is an IQA database that has ecological validity. The researchers chose 10,073 images at random from the estimated 4.8 million that were contained in the YFCC100m database. The researchers ensured that the images' distribution across seven quality measures, one content indicator, and machine tags were as uniform as possible to within a few percentage points. Experimentally, the analysis of

the dataset revealed that the KonIQ-10k had a much higher level of variety when compared to earlier datasets. An audience of 1,467 individuals each provided 120 ratings for each of their images, for a total of 1.2 million quality ratings that were gathered. It has been shown that the scoring method is of a high quality, and it has also been demonstrated that the results are trustworthy when compared to the assessments of experts. Because blind IQA is a challenging problem, it needs image quality assessment (IQA) databases that contains genuine cases such as these. As a consequence of this, the researchers were able to collect a large-scale dataset that contains a wide variety of naturally occurring distortions as well as reliable quality judgments from a diverse group of individuals.

Figure 3 provides nine randomly sampled images from the dataset to demonstrate the diversity of content and perceptual quality. It is to be noted that the images in the dataset are sampled from a large-scale dataset with certain selection criteria to ensure diversity of content and perceptual quality. The images are not post-processed except for cropping and resizing and unlike the other two dataset are not square images but contain 512 × 384 pixels dimension.

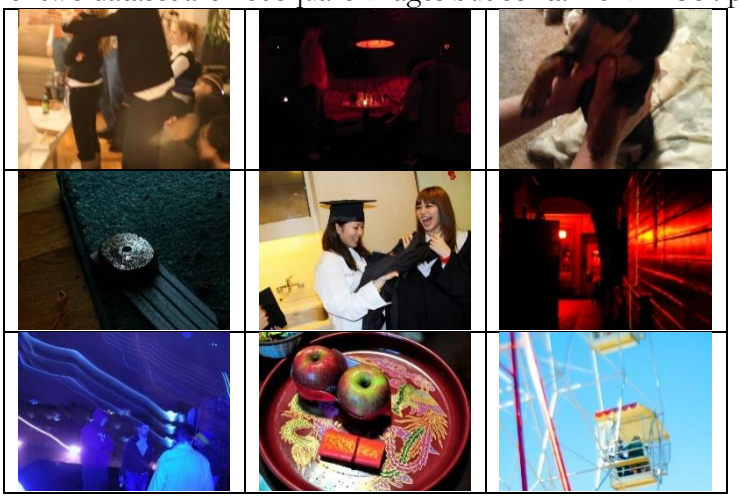

Figure 3. Randomly sampled images from KonIQ-10K datasets

BIQ2021:

Obtaining MOS based on a subjective evaluation and collecting representative images are two of the most challenging components in building a representative dataset with a sufficient number of images that have reliable quality ratings. This is because MOS is based on subjective evaluations, which require assessments from multiple participants. Because a significant number of human participants are required to evaluate the images, the collection of pictures and the acquisition of MOS are both time-consuming and expensive processes. This is due to the fact that the MOS will be utilized as a guide for the supervised training, and only a MOS that can be relied upon will result in a model that accurately replicates human judgment. The quality ratings included in the BIQ2021 dataset were compiled in a laboratory environment under controlled conditions. For the purpose of the subjective experiment, the same set of images, user interface, and instructions were provided to each participant. The level of conformance to the criteria that were stated will have a substantial impact on the reliability and repeatability of these experiments, as well as the MOS that was ultimately determined. The dataset was generated as a requirement to enrich existing datasets and offer a benchmark against which a trained model may be evaluated using images with intrinsic distortions. The development of the dataset took place as a result of this demand. The vast majority of databases currently on the market contain pictures that have been digitally manipulated to seem distorted, and they were developed for the purpose of analyzing full-reference IQA assignments. These datasets can be utilized for the evaluation of full-reference methods or scenarios in which the nature of the distortion is already known. The BIQ2021

database, on the other hand, contains 12,000 images with naturally occurring distortions that have been meticulously selected and vetted by experts. The tests are carried out in order to gather quality ratings from each of the 30 unique observers, and a total of 360,000 evaluations are obtained as a result. Images having a resolution of 512 by 512 pixels are included in the dataset, each of which is also accompanied by a mean opinion score, as well as a standard deviation.

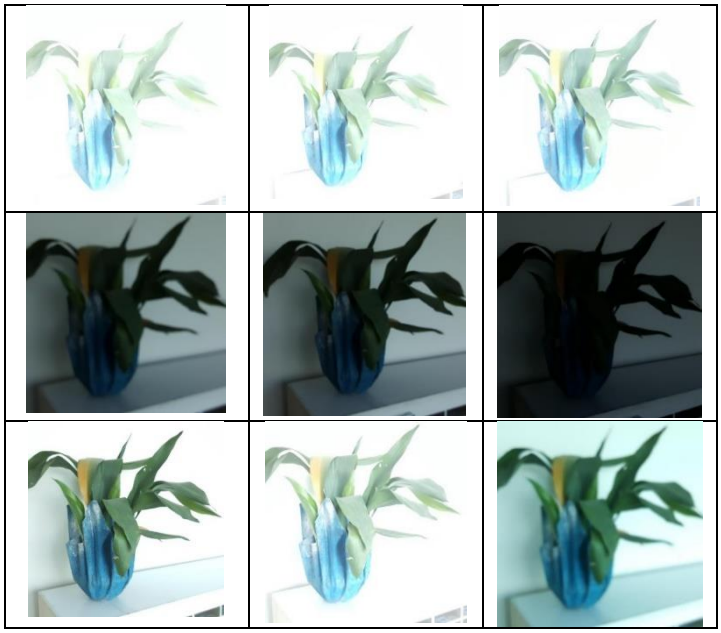

Figure 4. Randomly sampled images from first subset of dataset

Figure 4 provides nine randomly sampled images from the first subset of the dataset. These images contain various levels of perceptual distortion but are not post processed except cropping. Figure 5 provide the nine randomly sampled images from the second subset of the dataset and are images captured using various mobile devices and have a diversity of content and quality. Figure 6 contain nine randomly sampled images originally obtained from a large online collection to ensure diversity of content and images captured or processed by the photographic community.

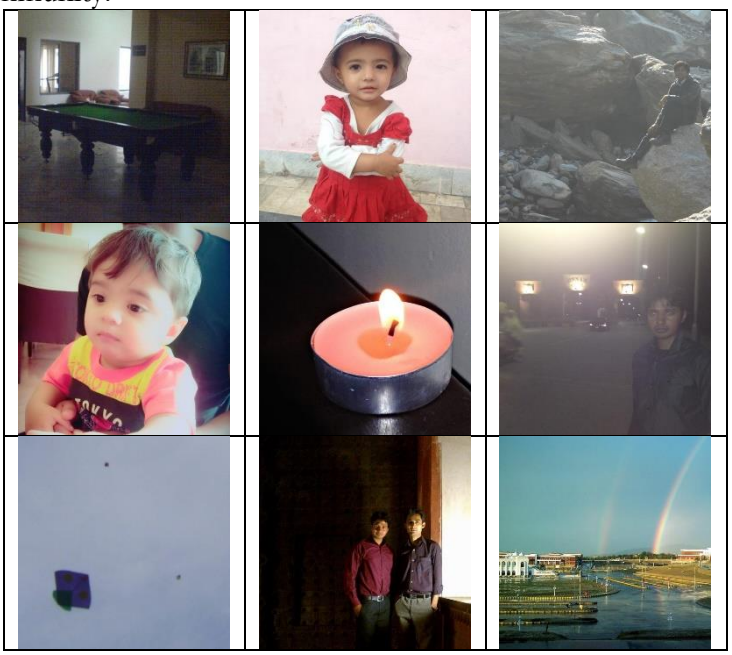

Figure 5. Randomly sampled images from the second subset of dataset

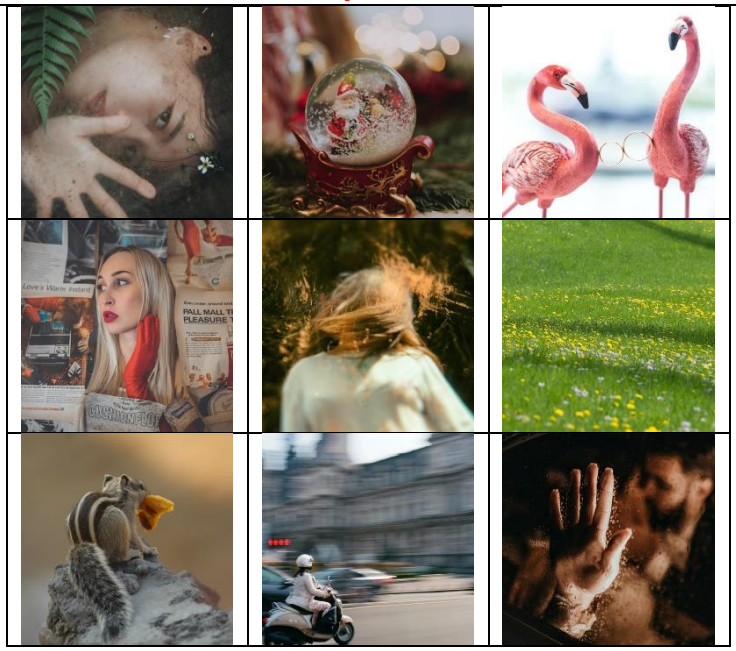

Figure 6. Randomly sampled images from the third subset of the dataset

Data Preparation:

Machine learning methods demand that data be prepared in an appropriate format prior to training, as poorly processed data have a significant impact on prediction performance. In this investigation, we are addressing the problem by employing Convolutional Neural Networks (CNNs), a type of artificial neural network that, in contrast to natural scene statistics-based methods, automatically learns the distinguishing features of labelled images. Our decision to use a CNN-based supervised learning was influenced by the research that was conducted and the excellent results that were obtained from using various learning strategies. Even while CNN doesn't require a significant amount of preprocessing of images before training, there are still a few things that may be done to improve the predictive performance.

Image Resizing and Normalization:

CNN are models which are designed for a fixed input image size and therefore any image which is larger or smaller than this size must be resized or cropped to fit the dimensions. In order to train various CNN models for quality assessment training, we have performed randomly cropped the images to the input size of these CNN models. Image resizing is not employed, as [citation] demonstrated that resizing affects the perceptual quality of an image and therefore cannot be used reliably. The use of random cropping instead of center cropping acts as a regularization method as the training algorithm encounters a different region of the image each time which has same perceptual quality but different content and therefore allow the modeling algorithm to learn the perceptual quality instead of the peculiarities of the content.

The datasets used for testing contain color images with three channels representing red, green, and blue. The images are of 24-bit with 256 gray levels in each layers making pixel values ranging from 0-255 for each layer of the image. The normalization or scaling of these pixel values allow the CNN model to converge faster and may provide a better convergence point. Therefore, the images are normalized by dividing each pixel value with 255 and making them have pixel values in the range of 0-1.

Rescaling MOS:

Mean Opinion Score (MOS) which is obtained by averaging the subjective scores assigned by multiple observers. It is important to note that image quality assessment dataset has their own scale of subjective scores and range of MOS. In some cases, the provided scores

are already scaled to a fixed range. As described in previous section that normalized scores allow the neural networks trained via gradient descent-based methods to converge faster therefore the scores are rescaled. Moreover, this rescaling will allow the scores of various datasets to have same range and can be used for comparison and prediction on other datasets. Therefore, the MOS is scaled in the range of 0-1 where minimum value of MOS is projected to zero and maximum value to 1 with all the intermediate values scaled uniformly in the range.

Image Augmentation:

Deep neural networks need large amounts of training data in order to function at their maximum potential. Deep neural networks often need the use of image augmentation in order to have their performance improved when dealing with a limited amount of data for training purposes. Image augmentation refers to the process of synthetically generating training images by using a variety of single or multiple processing techniques, such as random rotation, shifts, shears, flips, etc. Figure 7 provide the demonstration of the original image in Figure 7(b) and its augmented variants in Figure 7(a) by using these augmentations.

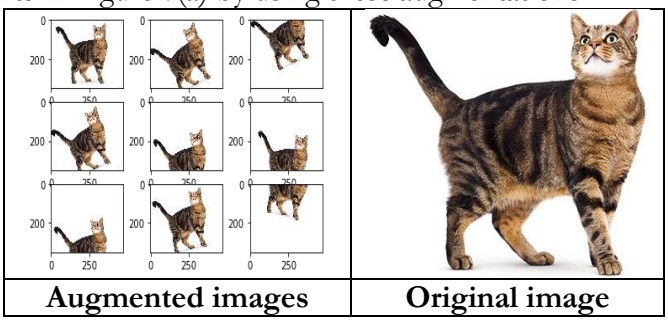

Figure 7. Original image and augmented variant generated by performing augmentation operations

Data augmentation through the use of these methods serves other purposes in addition to effectively increasing the size of training datasets. The model is given the opportunity to learn additional variations that can occur in an image but are not actually represented in the training dataset as a result of the perturbations that are introduced in these images during image augmentation. In addition, if the images are taken in a regulated setting (such as a laboratory), the model may begin to overfit by learning the peculiarities of the data (training data bias). As a result, the introduction of these perturbations acts as a regularize by introducing slightly different examples of the same class or category.

During the training of convolutional neural networks for image quality assessment, the data augmentation is implemented as an intermediate stage in order to facilitate the accomplishment of the outlined objectives. Furthermore, the perturbations that are used for data augmentation in visual classification challenges cannot be incorporated for image quality assessment without necessary modifications because it is acknowledged that few augmentation operations can alter the perceptual quality of these images. In the process of data augmentation, scaling, contrast adjustment, and other augmentation operations are not used. Instead, only translation, rotation, and random cropping are used because these operations did not cause any variation in the perceptual quality of the images.

Data Partitioning:

In order to perform model training and validation, the dataset is required to be portioned in training and testing set. The training set is used for model training whereas the validation set is used for validation the trained model. There are two popular approaches to model validation. The holdout validation use a split of data by providing one half of the data for model training and the other half for model validation. In case when the dataset size is smaller the training split is made usually larger than the validation split and same is the case for our problem. The other approach to model validation is cross-validation in which the data

is portioned into a number of splits and training and validation is repeated for each split of data. The model is tested on a unique split in each iteration whereas the rest of data is used for model training. The results obtained in each iteration are averaged to obtain cross-validation performance.

In case of our problem, due to computationally extensive nature of the problem and keeping in vie the trend of the literature, we have performed holdout validation. The model is trained on 80% data whereas the validation is performed on the 20% data. Moreover, in case of cross-dataset evaluation, the model is trained on 100% data of one dataset and evaluated on the 100% data of the other dataset in order to maximize the learning potential of the trained model.

Model Building:

Model building is perhaps the most important and decisive stage of the whole process. The prepared dataset is fed into the learning process, and the trained model's predicted performance is evaluated. In order to develop a model, we have conducted experiments using a number of different pre-trained models, and the assessment metrics are used to determine which strategies are the most effective. In order to train CNN models on the image quality evaluation dataset, we have relied on pre-trained models and carried out learning using a well-known method known as transfer learning. This has allowed us to successfully train CNN models on the dataset.

Transfer Learning:

The objective of transfer learning, which is a problem in machine learning, is to be able to take the information that was received from the solution of one problem and apply it to a new problem that is identical to the previous one. This research has some theoretical connections to the vast body of psychological literature that has been written on the subject of transfer of learning, despite the fact that there are relatively few direct connections between the two areas. Reusing or transferring knowledge from previously acquired tasks in order to learn new tasks can dramatically enhance the sample efficiency of a prediction problem. A classification model that was initially developed on a large-scale image classification dataset can serve as a starting point for further training on our image quality assessment problem if we use this approach. The pre-trained models, which are trained on an image classification dataset such as ImageNet, has already learned filters from images that are used for object recognition, and the relatively smaller size of the image quality assessment dataset will be compensated by leveraging the weights learned from the original dataset. ImageNet is a popular example of a dataset that is used for training pre-trained models.

[28] has provided some guidelines to look for when training a model using transfer learning. These three objectives can ensure that the transfer learning is a successful attempt and has resulted in improved model performance. Figure 8 provides the depiction of model performance trained with and without transfer learning.

The model should demonstrate high initial performance in comparison to training the model from scratch.

The training slope of the model via transfer learning should be steeper than training the model from scratch.

The asymptote (convergence performance) of the model should be higher than the model being trained from scratch.

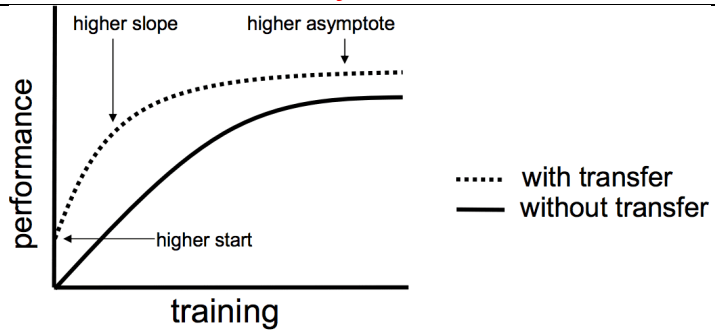

Figure 8. Depiction of model performance trained with and without transfer learning [28]

Proposed Model:

The proposed model is constructed as a deep ensemble of various CNN models. The model training and validation are carried out for each individual candidate model. Experiments are carried out in order to carry out transfer learning from previously trained CNN models, which were originally trained on the ImageNet dataset used in the ILSVRC 2012 competition. The CNN models that are utilized for the purpose of evaluating the performance of transfer learning can be found in Table 3. These models are a modification of the classification models, which are responsible for the visual classification of images into a thousand different categories. The last three layers of the pre-trained models are swapped out in order to transform the structure into a regression model during the modification that is performed. The fully connected layer, which formerly contained one thousand neurons, has been reduced to just one, and the softmax and classification layers have been swapped out for a regression layer. The mean squared error is the loss function that is utilized for the regression layer; however, the root mean squared error is the performance metric that is utilized for the validation performance of these models. The modified model is trained and validated using the 80% to 20% split, and the predictive performance for each of these models is reported in Table 3.

In order to carry out the process of building the final model, three candidate models are chosen from Table 1 to be incorporated into the building of the final model. When it comes to the accurate prediction of quality scores for images, NasNet demonstrated the best overall performance. In terms of PLCC and SROCC, InceptionResNet-V2 currently holds the position of second place for quality prediction. DenseNet201, which has provided the third highest performance, is the third model in terms of the quality prediction performance it has provided. These three models are altered from their initial architecture (which had already been pre-trained) by the addition of global average pooling, which is then followed by regression and the acquisition of quality scores for each individual predictor. These individual predictors are then given to a second stage learner, which in our case is stepwise linear regression in order for it to act as a Meta-learner. This meta-learner is responsible for performing learning based on the quality score predictions made by the other three predictors, as well as performing final quality score predictions. Figure 9 provides architecture of the model used to make quality predictions. The Meta-learner was built out of three separate models, and its performance as a regression analyzer is going to be reported in the following chapter based on the predictions that it came up with.

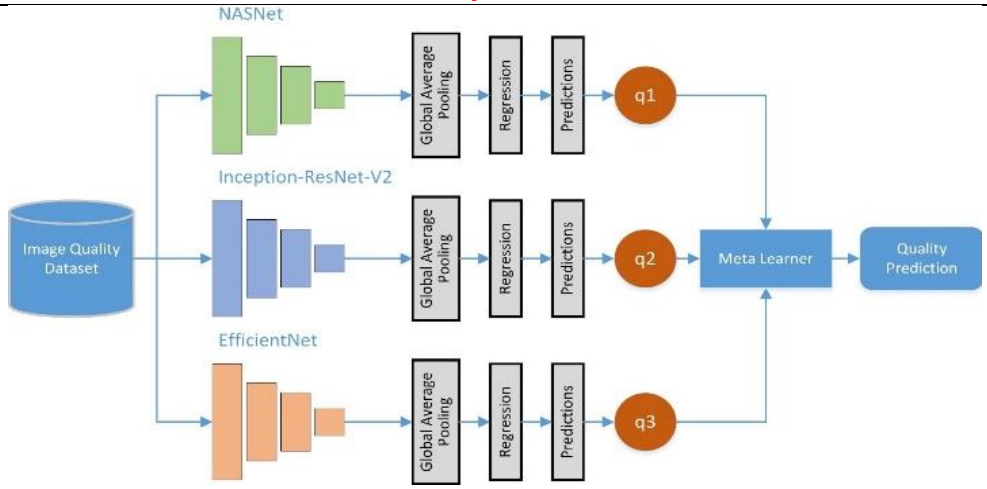

Figure 9. Block diagram of the proposed approach

Execution Environment:

Matlab ® 2022b was used to implement the proposed method and conduct the experiments that are reported in this study. The program is set up on a computer running the 64-bit version of Windows 10 as the operating system. Table 1 contains the system specifications that are utilized to carry out the experiments. The training parameters, such as the batch size, are selected based on the amount of memory that is readily available.

Table 1. System Specifications

Sr.	Attribute	Value
1	System	Dell T5600
2	Processor	2x Intel Xeon E5-2587
3	RAM	16 GB
4	SSD	512 GB
5	GPU	GTX 1070 Ti
6	GPU Memory	8GB GDDR5

Training Parameters:

In order to perform training of a CNN model, there are various parameters which are required to be provided to improve the convergence speed and terminal performance. Table 2 provide the training parameters which is selected by performing hyperparameter optimization using ResNet50 model on KonIQ-10K dataset. It is to be noted that batch size and validation frequency were selected differently for different models depending on the model size and the training images. Table 2 provides the various training parameters and their optimized value used for model training.

Table 2. Training parameters for training various CNN models

Sr.	Parameter	Value
1	Optimizer	Adam
2	Mini Batch Size	8-64
3	Epochs	100
4	Initial Learning Rate	0.001
5	Learning Rate Schedule	Piecewise
6	Learning Rate Drop Period	20
7	Learning Rate Drop Factor	0.1
8	Shuffle	Every Epoch
9	Validation Frequency	188-1500
10	Validation Patience	10
11	Output Network	Best validation loss

Model evaluation:

In order to determine how well a model predicts quality, one must first define evaluation metrics that are appropriate for the specific characteristics of the issue that is being researched. The challenge of assigning an accurate rating to the quality of an image can be conceptualized as a regression problem. The performance of regression algorithms can be assessed through the utilization of metrics such as mean squared error, mean absolute error, root mean squared error, and other comparable measurements. It is important to keep in mind that determining the absolute magnitude of the discrepancy that exists between the predicted and observed values is not the objective of quality control testing. The correlation between predicted quality scores and MOS is a more appropriate measurement method because it determines how well the algorithms are able to assign higher quality scores to higher-quality images. In other words, it determines how well the algorithms are able to predict the quality of an image. There are two correlation metrics that are widely used for evaluating different quality assessment models. Both of these correlation metrics are used in this study, and they are described further below:

Pearson Linear Correlation Coefficient (PLCC)

The Pearson correlation coefficient is a measure of linear correlation between two sets of data. It is also known as Pearson's r, the Pearson product-moment correlation coefficient, the bivariate correlation, or more colloquially just as the correlation coefficient. Other names for this statistic include Pearson's r and the bivariate correlation. It is the ratio between the covariance of two variables and the product of their standard deviations; consequently, it is essentially a normalized measurement of the covariance, in the sense that the result always has a value that falls somewhere between 1 and 1. The measure, just like covariance itself, can only reflect a linear correlation of variables and ignores a large number of other possible types of relationships or correlations. The sample correlation coefficient, also known as the sample Pearson correlation coefficient, is a common way to represent the Pearson correlation coefficient, and the two names are sometimes used interchangeably. It is possible to calculate the correlation coefficient by making use of the corresponding formula in the equation (1).

$$r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2}\sqrt{n\sum y_i^2 - (\sum y_i)^2}} \quad (1)$$

n is the sample size, $\bar{x}$ is the mean of samples and x_i, y_i are the individual sample points indexed with i

Spearman's Rank Order Correlation Coefficient (SROCC):

A nonparametric measure of rank correlation is known as the Spearman rank correlation coefficient, also known as Spearman's rho. This measure was named after Charles Spearman and is sometimes represented by the Greek letter rho. The statistically measurable aspect of the relationship between the ranks of two variables is the relationship itself. It examines the relationship between two variables and determines whether a monotonic relationship adequately describes that relationship.

Since the Pearson correlation between rank values of two variables is equivalent to the Spearman correlation between those same variables, we can conclude that the Pearson correlation evaluates linear relationships, whereas the Spearman correlation evaluates monotonic ones. A Spearman correlation of +1 or -1 indicates that the two variables are perfectly monotone functions of one another, provided that there are no recurring patterns in the data. This is the case even in the absence of such patterns. We anticipate a high Spearman correlation when the ranks, or relative positions, of the observations in the two variables are comparable to one another. On the other hand, we anticipate a low Spearman correlation when the ranks are dissimilar to one another. If the values of two variables are identical, then the correlation is said to be strong, and this is denoted by the number 1, but if the values are

very different from one another, then the correlation is said to be weak, and this is denoted by the number -1. In addition, the Spearman's rank coefficient can be applied to either continuous or discrete ordinal data in order to achieve the same results.

$$\rho_{\mathrm{R}(X),\mathrm{R}(Y)} = \frac{\mathrm{cov}(\mathrm{R}(X),\mathrm{R}(Y))}{\sigma_{\mathrm{R}(X)}\sigma_{\mathrm{R}(Y)}} \quad (2)$$

$R\ (f)$ is for variable ranking, ρ is Pearson correlation, cov is covariance matrix and σ denotes the standard deviation

Results:

The experimental evaluation of the proposed model and the CNN models that are utilized to perform transfer learning for the purpose of assessing image quality is presented in this chapter. In this chapter, we will report our findings from three distinct categories of experiments. In the first experiment, a holdout validation of the model is carried out on the dataset, and the performance of the experiment is reported in terms of two correlation measures: PLCC and SROCC. In the second experiment, a cross-dataset evaluation was carried out to gauge the generalization of the developed model. This involved testing the model that had been trained using one dataset and predicting the quality scores using another dataset. In the third experiment, model evaluation was carried out using a dataset that was collected by the researchers themselves.

Performance using Individual Models:

Transfer learning is performed using the pre-trained weights of the various CNN models so that the quality prediction performance of each individual model can be evaluated. ImageNet provides the training data for these models. The fundamental architecture of these pre-trained models is altered for the regression scenario, and model fine-tuning is carried out with the assistance of the transfer learning configurations and hyperparameter settings that are outlined in chapter 3. The PLCC and SROCC values for each of these models are provided in Table 3. These values were obtained by using 20% of the dataset as held-out data while the models were trained on 80% of the dataset.

Table 3. Quality prediction performance of individual CNN models

Sr.	Network	KonIQ-10K		LiveCD		BIQ2021	
		PLCC	SROCC	PLCC	SROCC	PLCC	SROCC
1	Densenet-201	0.9404	0.9268	0.9093	0.8957	0.7465	0.7329
2	Efficientnet-b0	0.8509	0.8265	0.8096	0.7852	0.5995	0.5773
3	GoogleNet	0.8867	0.8791	0.8165	0.8089	0.6249	0.6005
4	Inception-V3	0.5206	0.6874	0.5528	0.7196	0.6808	0.6732
5	InceptionResNet-V2	0.6622	0.7833	0.6796	0.8007	0.3379	0.5047
6	MobileNet-V2	0.8626	0.8445	0.8318	0.8137	0.4989	0.6200
7	NasNet_Large	0.9643	0.9403	0.9717	0.9477	0.6843	0.6662
8	NasNet_Mobile	0.8195	0.8329	0.7477	0.7611	0.8505	0.8265
9	Resnet18	0.7855	0.7896	0.7563	0.7604	0.6067	0.6201
10	Resnet50	0.8328	0.8123	0.8483	0.8278	0.6091	0.6132
11	Resnet101	0.7119	0.7114	0.7271	0.7266	0.7267	0.7062
12	ShuffleNet	0.7889	0.7843	0.7380	0.7334	0.5836	0.5831
13	SqueezeNet	0.7664	0.7583	0.7703	0.7622	0.5481	0.5435
14	Xception	0.9279	0.8997	0.9153	0.8871	0.5874	0.5793

Evaluation on benchmark datasets:

Experiments are performed on the three benchmark datasets discussed in section 3-(C) so that performance evaluations can be obtained for the final model so that its effectiveness can be evaluated. The performance evaluation is carried out using holdout

validation, in which 80% of the data is utilized for the training of the model, and 20% of the dataset is utilized for the validation of the model. In addition, in order to eliminate train-test bias, the data are shuffled at random without being replaced before the training and validation parts of the process are carried out. On each of the three datasets, the performance of the proposed model is reported in terms of the Pearson and Spearman correlation coefficients. Table 4 presents the findings that were derived from the analysis of the employee's performance.

Table 4. Model evaluation on three benchmark datasets with natural distortions

Sr.	Dataset	PLCC	SROCC
1	LiveCD	0.9885	0.9812
2	KonIQ-10K	0.9702	0.9658
3	BIQ2021	0.8840	0.8765

Cross-Dataset Evaluation:

In order to evaluate the generalizability of a model, it is necessary to test the model on a dataset collected independently of the training data. The predictive performance on independently collected datasets can be used to validate the ability of the model to perform prediction in actual scenario for images that are collected outside the experimental settings. This can be done by evaluating the model's performance on independently collected datasets. Cross-dataset evaluation is a method that is used to accomplish the goal of model generalizability. In this evaluation method, one type of dataset is used for training, and the other type of dataset is used for model evaluation. Because the datasets have reported their quality ratings within a variety of ranges, it is necessary to scale the quality scores in order to account for these differences. The relevant description can be found in section 3-(C). The MOS that is reported by each dataset is scaled to fall within the range of 0 to 1, and as a result, it is possible to avoid the variation in range, at least to some degree. In addition, the performance is evaluated based on correlation measures, which simply correspond to an increase or decrease in the quality that was predicted, as opposed to absolute change, which is typically measured via root mean squared error. After adjusting the models' MOS to fall within the range of 0–1, the Pearson and Spearman correlation coefficients are presented in Table 5 for models that were trained on the KonIQ-10k dataset and evaluated on the BIQ2021 and LiveCD datasets.

Table 5. Model evaluation on cross-dataset scenario

Sr.	Training Dataset	Testing Dataset	PLCC	SROCC
1	KonIQ-10K	LiveCD	0.7729	0.7489
2	KonIQ-10K	BIQ2021	0.7845	0.7588

Discussion:

Multiple factors influence how an image is perceived [14][7], making image quality assessment a challenging problem [3][5]. Additionally, it has been demonstrated that the visual content of an image has a significant impact on the viewer's perception of the quality of the image [7][18][19], which further complicates the scenario. The scenario that corresponds to the assessment of image quality known as reference less image quality or no-reference image quality occurs when the original version (the reference image) is not available to make a comparison with in order to determine the quality of the image [8][10][11][16][18]. However, as the number of applications that do not have access to reference information grows, the significance of this factor will only continue to grow. In addition, a wealth of research on no-reference quality assessment is currently available for scenarios in which the images are distorted by simulating the distortion on images that have not been altered in any other way [29][30]. Due to the fact that only particular degradations with discrete levels of degradations are included in such a dataset, it is relatively simple for the modeling algorithm to learn how

to handle this scenario. It is essential to emphasize the fact that this is not the scenario that occurs in actual use because the image degradations occur at a variety of granularities at the same time within the same image. Because it is possible for a natural image to have been distorted in a variety of ways during the process of acquiring, compressing, storing, or communicating it, it is necessary to train a prediction model for such images.

This research has therefore focused on the scenario of image quality datasets containing only naturally occurring distortions. This research combined the capabilities of three of the best performing predictors in order to achieve superior performance. The transfer learning was performed on CNN models that had been pre-trained on ImageNet. An algorithm for stepwise linear regression is given the quality scores that have been predicted for each of these three individual predictors. This allows the algorithm to optimally combine the results of the three models and combine these predictions to obtain the final quality scores. Extensive experiments were carried out in order to obtain quality prediction scores on each of these three datasets, in addition to cross-dataset settings, which are used to evaluate the performance of the model and its generalizability. The results of these experiments suggest that the model that was proposed offers an accurate prediction of image quality, and that the final model offers performance that is superior to that of the individual CNN models.

Conclusions:

The evaluation of the image's quality in the absence of reference information is difficult because there is no data to serve as a point of comparison, and there is no reference information. The problem has received much attention from the research community. However, there is still room for improvement in situations where the images are naturally distorted, and multiple distortions are happening simultaneously. The problem was effectively tackled by the research, which suggested a CNN-based model by combining three separate models using a Meta-learner. Three different benchmark datasets consisting of images with naturally occurring distortions are used to evaluate the proposed model's performance. The model's construction is carried out according to a method organized into six stages. Pearson and Spearman's correlation coefficients are utilized to evaluate the performance of the quality prediction model. We evaluate and compare the predictive performance of the proposed model with that of twelve individual CNN models, and the results show that the proposed model has superior predictive performance. In addition, experiments using multiple datasets are carried out to validate the generalizability of the proposed model, and satisfactory results are obtained from these experiments.

Funding: This research received no external funding.

Acknowledgment: The authors would like to thank the anonymous reviewers for their helpful remarks, which resulted in the work being improved. We also thank the original authors of the LiveCD, KonIQ-10K, and BIQ2021 datasets for making their data available for the research community to utilize for training and evaluating image quality assessment algorithms.

Conflicts of Interest: The authors declare no conflict of interest.

References:

[1] Y. Liu, H. Yu, B. Huang, G. Yue, and B. Song, "Blind Omnidirectional Image Quality Assessment Based on Structure and Natural Features," *IEEE Trans. Instrum. Meas.*, vol. 70, 2021, doi: 10.1109/TIM.2021.3102691.

[2] M. Vlachos and D. Skarlatos, "Self-Adaptive Colour Calibration of Deep Underwater Images Using FNN and SfM-MVS-Generated Depth Maps," *Remote Sens. 2024, Vol. 16, Page 1279*, vol. 16, no. 7, p. 1279, Apr. 2024, doi: 10.3390/RS16071279.

[3] N. Ahmed, H. M. S. Asif, and H. Khalid, "PIQI: perceptual image quality index based on ensemble of Gaussian process regression," *Multimed. Tools Appl. 2021 8010*, vol. 80, no. 10, pp. 15677–15700, Feb. 2021, doi: 10.1007/S11042-020-10286-W.

[4] N. Ahmad, H. M. S. Asif, G. Saleem, M. U. Younus, S. Anwar, and M. R. Anjum, "Leaf Image-Based Plant Disease Identification Using Color and Texture Features," *Wirel. Pers. Commun. 2021 1212*, vol. 121, no. 2, pp. 1139–1168, Aug. 2021, doi: 10.1007/S11277-021-09054-2.

[5] L. Li, C. Mu, J. Wang, X. Liu, and N. Weng, "BVIFormer: a binocular-vision-inspired transformer with binocular competitive fusion for single-image restoration," *Sci. Reports 2026 161*, vol. 16, no. 1, pp. 19862-, Apr. 2026, doi: 10.1038/s41598-026-46078-9.

[6] J. Chen, S. Chen, L. Wee, A. Dekker, and I. Bermejo, "Deep learning based unpaired image-to-image translation applications for medical physics: a systematic review," *Phys. Med. Biol.*, vol. 68, no. 5, p. 05TR01, Feb. 2023, doi: 10.1088/1361-6560/ACBA74.

[7] N. Ahmed, H. M. S. Asif, A. R. Bhatti, and A. Khan, "Deep Ensembling for Perceptual Image Quality Assessment," *Soft Comput.*, vol. 26, no. 16, pp. 7601–7622, May 2023, doi: 10.1007/s00500-021-06662-9.

[8] Y. Zhang, J. Qian, D. Liu, Z. Xiao, X. Zhou, and C. Shan, "Blind Light Field Image Quality Assessment via Cross-View Encoding and Dynamic Expert Mixture," *IEEE Trans. Multimed.*, 2026, doi: 10.1109/TMM.2026.3715326.

[9] X. Lv, T. Xiang, Y. Yang, and H. Liu, "Blind Dehazed Image Quality Assessment: A Deep CNN-Based Approach," *IEEE Trans. Multimed.*, vol. 25, pp. 9410–9424, 2023, doi: 10.1109/TMM.2023.3252267.

[10] K. Ohashi *et al.*, "Development of a No-Reference CT Image Quality Assessment Method Using RadImageNet Pre-trained Deep Learning Models," *J. Imaging Informatics Med. 2025 391*, vol. 39, no. 1, pp. 46–58, May 2025, doi: 10.1007/S10278-025-01542-2.

[11] 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 Trans. Image Process.*, vol. 27, no. 3, pp. 1202–1213, 2018, doi: 10.1109/TIP.2017.2774045.

[12] R. Zhang, P. Isola, A. A. Efros, E. Shechtman, and O. Wang, "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric," *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, pp. 586–595, Jan. 2018, doi: 10.1109/CVPR.2018.00068.

[13] X. Ma and X. Jiang, "Multimedia image quality assessment based on deep feature extraction," *Multimed. Tools Appl. 2019 7947*, vol. 79, no. 47, pp. 35209–35220, May 2019, doi: 10.1007/S11042-019-7571-Y.

[14] N. Ahmed and H. M. S. Asif, "Perceptual quality assessment of digital images using deep features," *Comput. Informatics*, vol. 39, no. 3, pp. 385–409, 2020, doi: 10.31577/CAI_2020_3_385.

[15] J. Ma *et al.*, "Blind Image Quality Assessment with Active Inference," *IEEE Trans. Image Process.*, vol. 30, pp. 3650–3663, 2021, doi: 10.1109/TIP.2021.3064195.

[16] N. Ahmed and H. M. S. Asif, "Ensembling Convolutional Neural Networks for Perceptual Image Quality Assessment," *MACS 2019 - 13th Int. Conf. Math. Actuar. Sci. Comput. Sci. Stat. Proc.*, Dec. 2019, doi: 10.1109/MACS48846.2019.9024822.

[17] Z. Yu, F. Guan, Y. Lu, X. Li, and Z. Chen, "SF-IQA: Quality and Similarity Integration for AI Generated Image Quality Assessment," *IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. Work.*, pp. 6692–6701, 2024, doi: 10.1109/CVPRW63382.2024.00663.

[18] V. Hosu, H. Lin, T. Sziranyi, and D. Saupe, "KonIQ-10k: An ecologically valid database for deep learning of blind image quality assessment," *IEEE Trans. Image Process.*, vol. 29, pp. 4041–4056, May 2020, doi: 10.1109/TIP.2020.2967829.

[19] N. Ahmed and S. Asif, "BIQ2021: A Large-Scale Blind Image Quality Assessment

Database," *J. Electron. Imaging*, vol. 31, no. 05, Jan. 2023, doi: 10.1117/1.JEI.31.5.053010.

[20] D. Varga, "No-Reference Image Quality Assessment Using the Statistics of Global and Local Image Features," *Electron. 2023, Vol. 12, Page 1615*, vol. 12, no. 7, p. 1615, Mar. 2023, doi: 10.3390/ELECTRONICS12071615.

[21] Z. Pan, F. Yuan, X. Wang, L. Xu, X. Shao, and S. Kwong, "No-Reference Image Quality Assessment via Multibranch Convolutional Neural Networks," *IEEE Trans. Artif. Intell.*, vol. 4, no. 1, pp. 148–160, Feb. 2023, doi: 10.1109/TAI.2022.3146804.

[22] J. Ryu, "Improved Image Quality Assessment by Utilizing Pre-Trained Architecture Features with Unified Learning Mechanism," *Appl. Sci. 2023, Vol. 13, Page 2682*, vol. 13, no. 4, p. 2682, Feb. 2023, doi: 10.3390/APP13042682.

[23] X. Li *et al.*, "Less is More: Learning Reference Knowledge Using No-Reference Image Quality Assessment," Dec. 2023, Accessed: Aug. 13, 2026. [Online]. Available: https://arxiv.org/pdf/2312.00591

[24] H. Zhang *et al.*, "Learning degradation priors for reliable no-reference image quality assessment," *J. Vis. Commun. Image Represent.*, vol. 102, Jun. 2024, doi: 10.1016/j.jvcir.2024.104189.

[25] Z. Zhou, Z. Zhou, X. Tao, H. Chen, Z. Yu, and Y. Cao, "EARNet: Error-Aware Reconstruction Network for no-reference image quality assessment," *Expert Syst. Appl.*, vol. 238, Mar. 2024, doi: 10.1016/j.eswa.2023.122050.

[26] J. Ma, Y. Chen, L. Chen, and Z. Tang, "Dual-attention pyramid transformer network for No-Reference Image Quality Assessment," *Expert Syst. Appl.*, vol. 257, Jan. 2024, doi: 10.1016/J.ESWA.2024.125008.

[27] Y. Zhao, Y. Zhang, T. Xia, T. Huang, X. Ben, and L. Chen, "No-reference image quality assessment based on multi-scale dynamic modulation and degradation information," *Displays*, vol. 91, Jan. 2026, doi: 10.1016/J.DISPLA.2025.103207.

[28] M. Shoaib, M. U. Younus, K. Safdar, and A. B. Dogar, "A Hybrid Approach for Efficient Database Fragmentation in Distributed Systems," *Int. J. Innov. Sci. Technol.*, vol. 8, no. 1, pp. 162–178, 2026, Accessed: Aug. 13, 2026. [Online]. Available: https://ideas.repec.org/a/abq/ijist1/v8y2026i1p162-178.html

[29] Z. Zhang, "Improved Adam Optimizer for Deep Neural Networks," *2018 IEEE/ACM 26th Int. Symp. Qual. Serv. IWQoS 2018*, Jan. 2019, doi: 10.1109/IWQOS.2018.8624183.

[30] N. Ahmed, H. M. S. Asif, and H. Khalid, "Image Quality Assessment Using a Combination of Hand-Crafted and Deep Features," *Commun. Comput. Inf. Sci.*, vol. 1198, pp. 593–605, 2020, doi: 10.1007/978-981-15-5232-8_51/SAVE-RESEARCH.