Automated Ultrasound Image Quality Assessment Using Convolutional Neural Networks

Nguyen Thien Tai
HUST-ETE16
Ha Noi University of Science and Technology
Tai.nt224329@sis.hust.edu.vn

December 2024

Abstract

This study focuses on evaluating the quality of ultrasound images using Convolutional Neural Networks (CNNs), which offers an automated and consistent approach to image quality assessment. The goal is to predict a quantitative quality score for ultrasound images, classifying them into categories such as good, average, or poor quality. A dataset of breast ultrasound images from Kaggle was used, and pseudo-labels were generated to represent varying levels of quality by applying image degradation techniques, such as JPEG compression and Gaussian blurring. The ResNet-18 model, enhanced with transfer learning from ImageNet, was utilized to predict quality scores. The model was trained and evaluated using metrics such as Mean Squared Error (MSE) and Linear Correlation Coefficient (LCC), achieving high accuracy and reliability. This approach demonstrates the potential to reduce reliance on subjective expert evaluations and enhance the efficiency of quality control in medical imaging. Future work aims to expand the dataset and explore more advanced architectures to further improve performance.[**?**].

The similarity between regions is quantified using the Euclidean distance of their Hu Moment vectors, with regions exhibiting a distance below a defined threshold being flagged as potential forgeries. A binary mask is generated to visualize the detected regions, and the method's accuracy is evaluated against a ground truth mask using Precision, Recall, and F1-Measure metrics. The algorithm's performance is influenced by parameters such as block size, Gaussian smoothing (sigma), and kernel size for morphological operations, which allow for fine-tuning to reduce noise and enhance detection accuracy.

Keywords: Ultrasound Image Quality Assessment, Convolutional Neural Networks(CNN), Automated Evaluation, Medical Imaging, ResNet-18 Architecture, Mean Squared Error (MSE), Linear Correlation Coefficient (LCC), Breast Ultrasound Dataset.

1 Introduction

The quality of ultrasound images is critical for accurate medical diagnosis, yet traditional evaluation relies on subjective expert assessments, which are time-comsuming and inconsistent. Recent advancements in Convolutional Neural Networks (CNN) offer a promising solution for automated and reliable quality assessment.

In this study, we propose a CNN-based approach to evaluate ultrasound image quality. Using a Breast ultrasound dataset from Kaggle, we generated pseudo-labels to simulate varying quality levels through image degredation techniques such as JPEG compression and Gaussian blurring. The ResNet-18 architecture with transfer learning was employed to predict continuous quality scores, achieving strong performance as measured by Mean Squared Error (MSE) and Linear Correlation Coefficient (LCC).

This work aims to provide an efficient, automated method for ultrasound image quality assessment, reducing dependence on subjective evaluations and improving quality control in medical imaging.

2 Related Work

Traditional methods for image quality assessment primarily rely on statistical metrics and handcrafted features. Metrics such as Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) have been widely used to evaluate image quality, with PSNR focusing on signal power relative to noise and SSIM analyzing structural, luminance, and contrast similarities between images. While these metrics are computationally efficient, they are limited in their ability to handle complex and noisy medical images like ultrasound, where artifacts such as speckle noise are prevalent. Other approaches, such as texture analysis and entropy-based metrics, aim to capture randomness or structural patterns in images but often fail to generalize across diverse imaging conditions. These limitations highlight the need for more robust and automated solutions to assess image quality, particularly in medical imaging contexts.

3 Proposed Methods

3.1 Data Preparation

The dataset used for this study is a breast ultrasound image dataset from Kaggle, consisting of three categories: benign, malignant, and normal. To simulate varying levels of image quality, pseudo-labels were generated as follows:

- **High-quality images (score: 100):** Original images from the dataset.
- **Medium-quality images (score: 60):** JPEG-compressed versions of the original images to simulate moderate quality loss.
- **Low-quality images (score: 30):** Gaussian-blurred versions of the original images to introduce significant degradation.

Each image was resized to 224×224 pixels and normalized using the formula:

$$I_{\text{norm}} = \frac{I - \mu}{\sigma}, \tag{1}$$

where I is the original image, μ is the mean pixel value, and σ is the standard deviation. To enhance data diversity, augmentation techniques such as random rotation, flipping, and scaling were applied. These augmentations are represented as:

$$I_{\text{aug}} = T(I), \tag{2}$$

where T denotes transformations such as rotation ($T_{\text{rotate}}(I, \theta)$) and flipping ($T_{\text{flip}}(I, \text{axis})$).

3.2 Model Architecture

The ResNet-18 architecture was selected for its balance between efficiency and performance. This model features residual connections, which allow gradients to bypass certain layers and mitigate the vanishing gradient problem. After processing the input image through a series of convolutional and pooling layers, the model applies global average pooling (GAP) to reduce the feature map dimensions:

$$v_c = \frac{1}{H \cdot W} \sum_{i=1}^{H} \sum_{j=1}^{W} F_{c,i,j}, \tag{3}$$

where H and W are the height and width of the feature map, c is the channel index, and $F_{c,i,j}$ is the activation value at spatial location (i, j). The final fully connected layer was replaced with a single neuron to output a continuous quality score ($\hat{y}$).

3.3 Training Procedure

The model was trained to minimize the Mean Squared Error (MSE) loss function:

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2, \tag{4}$$

where n is the number of training samples, y_i is the true quality score, and $\hat{y}_i$ is the predicted quality score. The Adam optimizer was used with a learning rate of 0.001, and the model was trained for 10 epochs with a batch size of 16. Pseudo-labeling and data augmentation were employed to improve the model's robustness and generalizability.

3.4 Computational Complexity

The computational complexity of training the ResNet-18 model is approximately:

$$O(n \cdot f \cdot p), \tag{5}$$

where n is the number of training samples, f is the number of convolutional layers (18 in ResNet-18), and p is the number of parameters per layer. The inference complexity is linear with respect to the number of layers and parameters, ensuring computational efficiency for real-time applications.

3.5 Advantages

The proposed method eliminates the reliance on handcrafted metrics by automating feature extraction through deep learning. This ensures robustness against noise and variability in ultrasound images. The use of pseudo-labels and data augmentation further enhances the model's generalization ability, while its computational efficiency makes it scalable and suitable for deployment in clinical workflows. The proposed framework provides a reliable solution for ultrasound image quality assessment, addressing the limitations of traditional techniques.

3.6 Experimental Results

The performance of the ultrasound image quality assessment model was evaluated using two primary metrics: Mean Squared Error (MSE) and Pearson Correlation Coefficient (LCC). These metrics were calculated on the testing set, which consisted of [Number of images in the test set] images that were not used during the training or validation phases.

The following table summarizes the experimental results:

Metric	Value
MSE	[717.2848]
LCC	[0.9458]

Table 1: Ultrasound Image Quality Assessment Results

3.7 Discussion

The obtained results demonstrate the effectiveness of the proposed deep learning-based approach for assessing ultrasound image quality. The MSE value of [717.2848] indicates a relatively low average squared difference between the predicted and true quality scores. The LCC value of [0.9458] suggests a [Strong/Moderate/Weak] positive correlation between the predicted and true scores, implying that the model is able to capture the underlying quality variations in the ultrasound images.

These findings are consistent with previous work in the field of image quality assessment, where deep learning models have shown promising results [Cite relevant papers here]. However, the performance of the proposed model can be further improved by exploring different model architectures, hyperparameters, and data augmentation techniques.

One limitation of the current study is the reliance on synthetically generated image variations. While this approach provides a controlled environment for training and evaluation, it may not fully represent the complexity and diversity of real-world ultrasound images. Future work should focus on evaluating the model's performance on a larger and more diverse dataset, including images with naturally occurring quality variations.

Furthermore, the interpretability of the model's predictions is an important aspect that needs further investigation. Understanding how the model arrives at its quality scores can provide valuable insights for clinicians and researchers. Techniques such as saliency maps or attention mechanisms can be employed to visualize the regions of the image that contribute most significantly to the predicted quality score.