NANYANG TECHNOLOGICAL UNIVERSITY SINGAPORE

School of Computer Science and Engineering
College of Engineering

Automated Image Quality Assessment
and its Applications in Computer Vision

Student: Phoebe Zhou Huixin

SCSE Supervisor: Dr Sourav Sen Gupta
GovTech Supervisor: Dr Chua Teck Wee

Abstract

The performance of Convolutional Neural Networks (CNNs) in computer vision tasks depends on image quality. Hence, input quality monitoring is needed to ensure that CNN outputs are reliable. This can be achieved via Image Quality Assessment (IQA) algorithms, which predict quality in the form of Mean Opinion Scores (MOS). In this project, we implement learning-based IQA models from literature and demonstrate their generalizability on a wide range of practical domains. We extend the study by using IQA to assess the relationship between image quality and object detection performance.

Research Gap

Most IQA models were trained and evaluated using artificially distorted datasets.


Prediction: 1.02


Prediction: -1.30

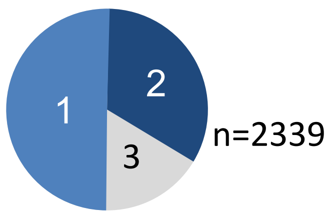

Prediction: -1.66

As expected, predicted MOS decreases with increasing severity of applied distortion. However, ideal software-generated distortions are not equivalent to authentic distortions.

→ **Must verify model generalizability using a wider range of testing domains**

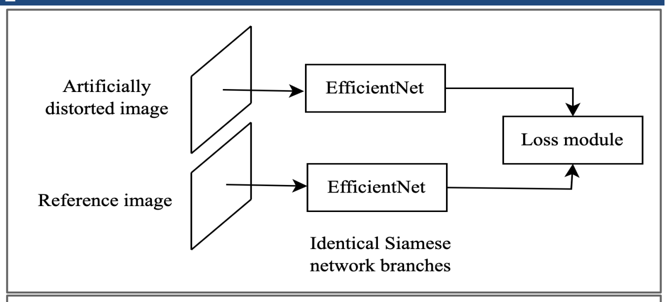
Datasets and Implementations

Mixed-dataset training and testing

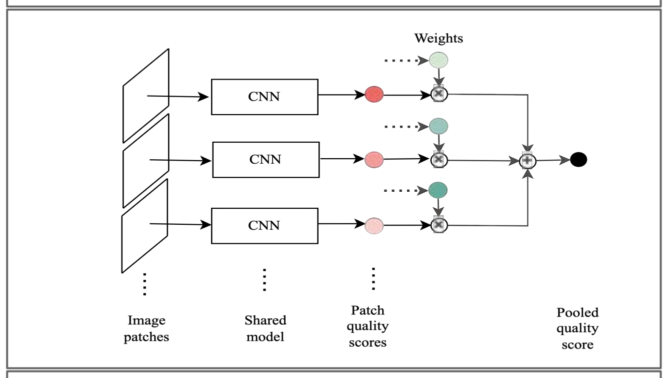


1. Authentically distorted benchmark (LIVE in the Wild)
2. Artificially distorted benchmark (LIVE)
3. GovTech CCTV + COCO object detection challenge

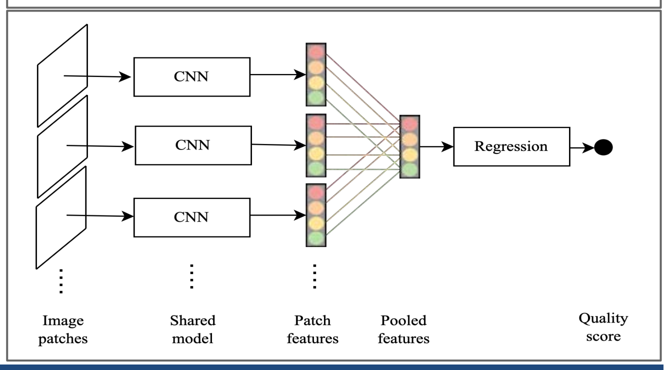
Rank IQA



Deep IQA

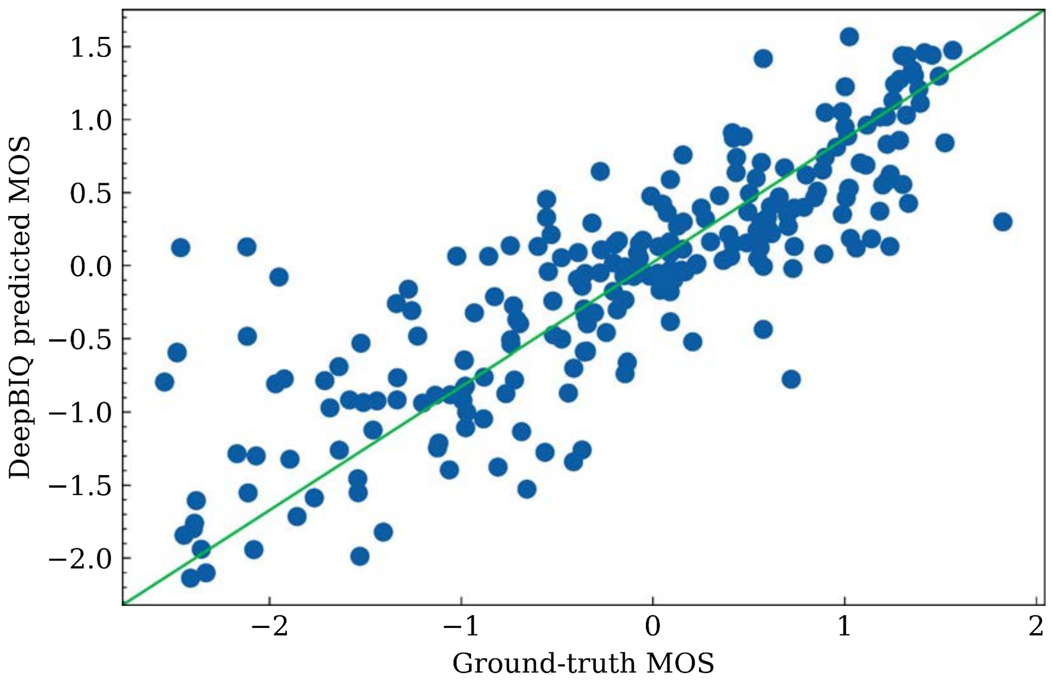


Deep BIQ



Results

Selected model: DeepBIQ



- **Spearman correlation between predicted MOS and ground-truth: 0.863**
- Improvement over control model (AlexNet): 14.6%
- Correlation between predicted MOS and object detector performance (mAP): 0.924 for artificial distortions; 0.562 for authentic distortions

✓ Support Vector Regressor (SVR) provides regularization, allowing training on small datasets

✓ Feature extractor is trained via weakly supervised learning, reducing reliance on intensive data labelling

✓ Accounts for local features by using image patches


Predicted MOS: -2.49


Predicted MOS: -1.62


Predicted MOS: 0.62


Predicted MOS: 0.73