LAR-IQA: A Lightweight, Accurate, and Robust No-Reference Image Quality Assessment Model

Nasim Jamshidi Avanaki, Abhijay Ghildyal[§], Nabajeet Barman[$], Saman Zadtootaghaj[$]

[§]Portland State University, OR, USA, [$]Sony Interactive Entertainment, UK/Germany

n.jamshidi.avanaki@gmail.com, abhijay@pdx.edu, {nabajeet.barman,saman.zadtootaghaj}@sony.com

EUROPEAN CONFERENCE ON COMPUTER VISION MILANO 2024

AIM 2024

Introduction

- NR-IQA is vital for mobile and real-time applications, but existing models are too large and computationally intensive for resource-limited devices.
- **LAR-IQA** tackles these challenges with a lightweight solution, achieving SOTA performance on the UHD-IQA challenge dataset, while being **5.7 times faster** than the fastest SOTA model and robust against color transformations.

Problem Statement

Existing NR-IQA models are computationally expensive, requiring significant resources, making them impractical for deployment on mobile or embedded systems. The need for a model that balances speed, accuracy, and robustness is clear.

Proposed Model

- **LAR-IQA** introduces a dual-branch architecture that processes authentic and synthetic distortions independently.
- Each branch utilizes a MobileNetV3 encoder, with their outputs combined and fed into a KAN for the final quality score.
- The use of lightweight encoders and the Efficient KAN improves both accuracy and computational efficiency.
- Incorporation of multiple color spaces for improved robustness.

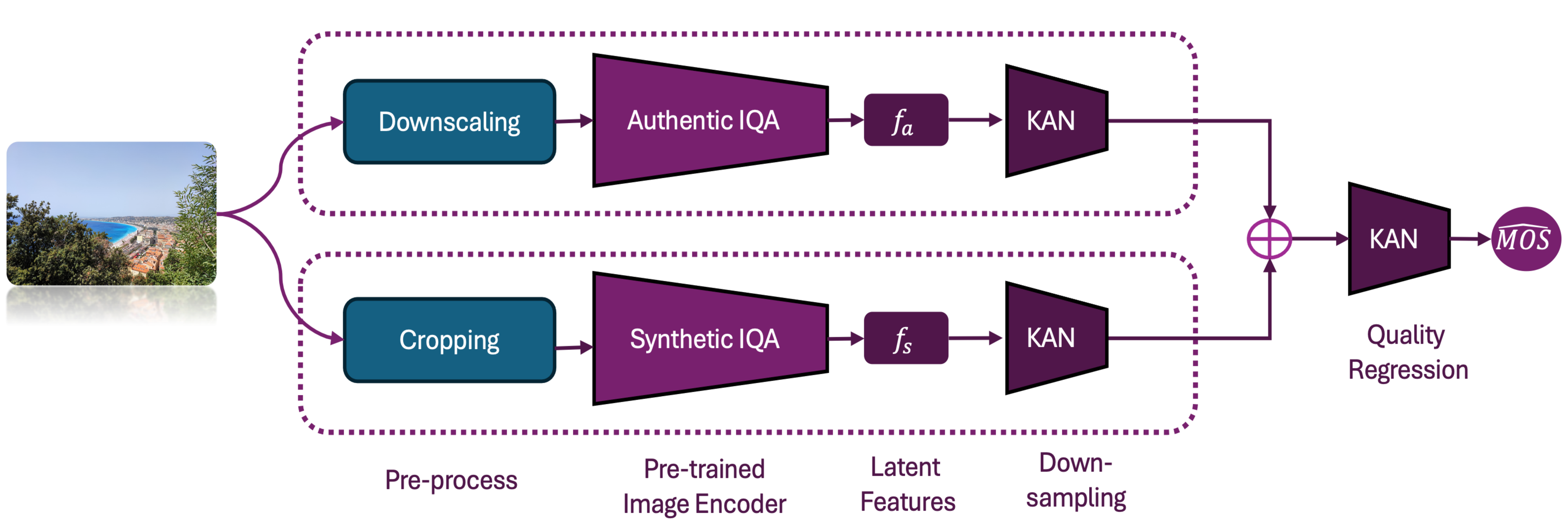

Fig. 1: Proposed model architecture.

Training Stages

- Each branch is trained separately on a specific task: authentic or synthetic.
- A multi-head model is used to train each branch on multiple datasets, following a multi-task training approach where each dataset represents a new task.
- The two pretrained models are merged into one using a KAN head, followed by fine-tuning.

Datasets and Evaluation Metrics

- **Datasets:** MSL-IQA was evaluated on seven publicly available datasets, including:
  - Authentic: KONIQ-10K, SPAQ, BID, UHD-IQA, PIPAL
  - Synthetic: KADID-10k, TID2013, PIPAL
- **Evaluation Metrics:** to evaluate the model's performance, we use three main common criteria:
  - SRCC for prediction monotonicity
  - PLCC and KRCC for rank consistency

Results

Method	PLCC↑	SRCC↑	KRCC↑	RMSE↓	MAE↓	MACs(G)↓
HyperIQA	0.103	0.553	0.389	0.118	0.070	<u>211</u>
Effnet-2C-MLSP	0.641	0.675	0.491	0.074	0.059	345
CONTRIQUE	0.678	0.732	0.532	0.073	0.052	855
ARNIQA	0.694	0.739	0.544	0.074	0.052	855
CLIP-IQA+	0.709	0.747	0.551	0.111	0.089	895
QualiCLIP	0.725	0.770	0.570	0.083	0.066	901
LAR-IQA (MLP head)	<u>0.774</u>	<u>0.809</u>	<u>0.616</u>	**0.058**	<u>0.042</u>	**≤37**
LAR-IQA (KAN head)	**0.787**	**0.836**	**0.642**	<u>0.061</u>	**0.041**	**≤37**

Table 1: Evaluation of the performance of the baselines on the test set of UHD-IQA. ↑ means that higher values are better, ↓ means that lower values are better.

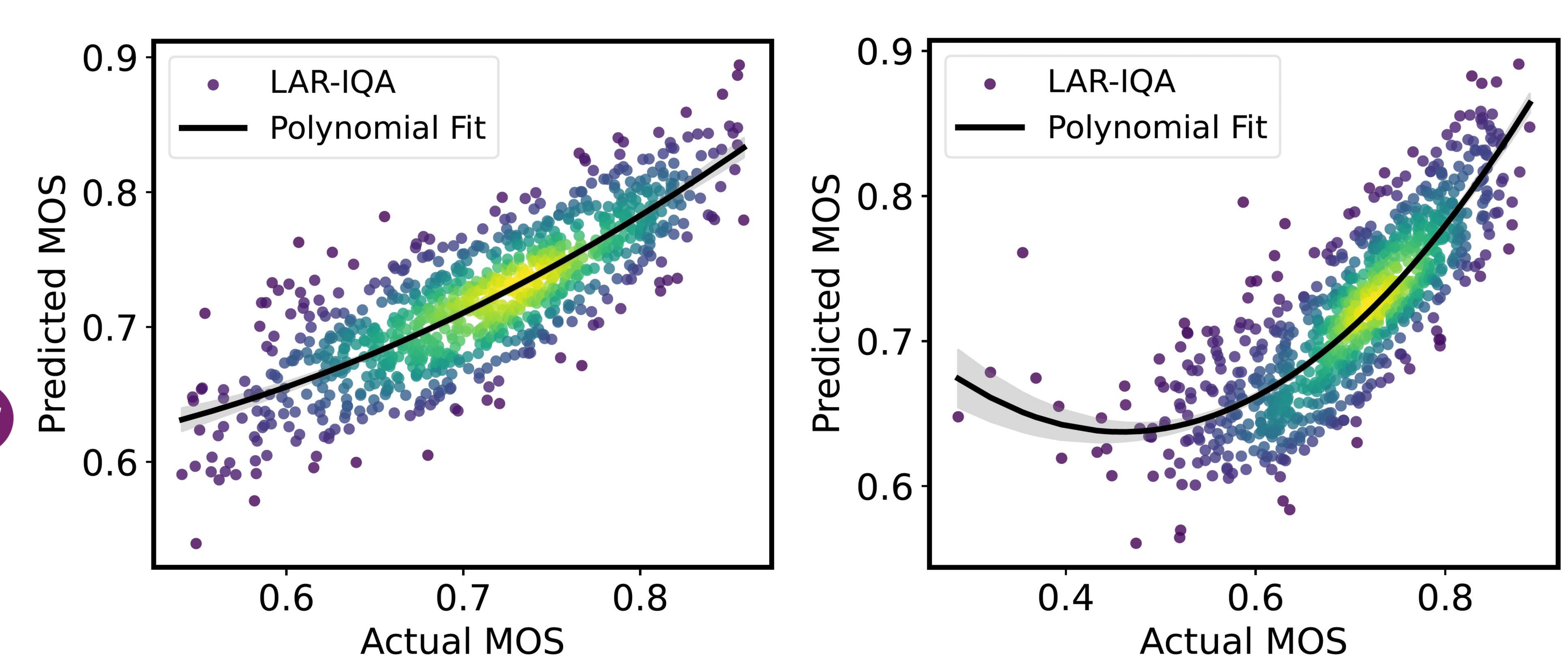

Fig. 2: Performance of LAR-IQA on validation (left) and test (right) sets of UHD-IQA Dataset.

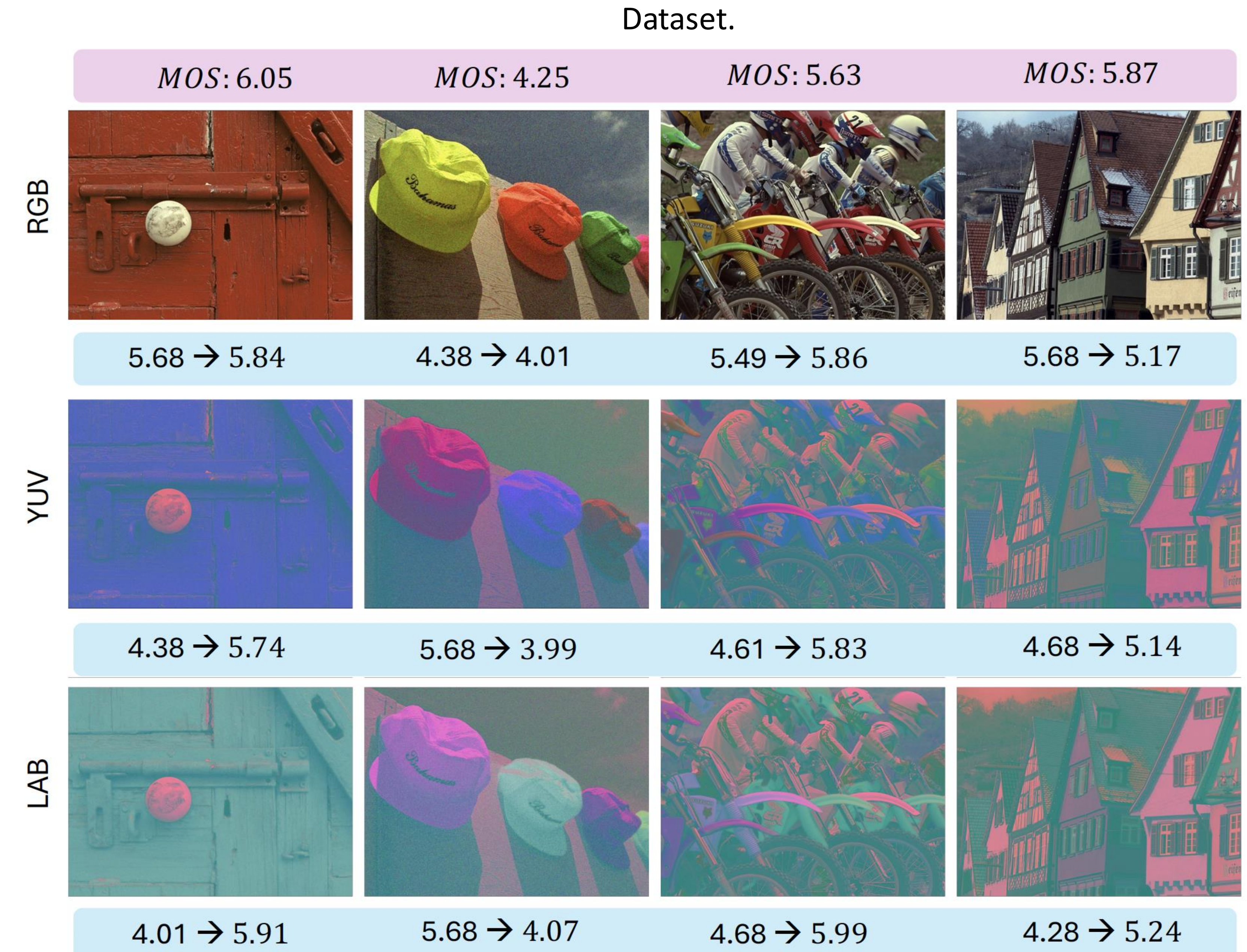

Fig. 3: MOS predictions of the Synthetic model using MobileNetV3, trained on KADID-10K and tested on sample images from the TID2013 dataset across different color spaces.

Conclusion

LAR-IQA offers an efficient solution for NR-IQA on mobile and <u>resource-constrained devices</u>, combining high accuracy with fast computation.

Sony Interactive Entertainment

Portland State UNIVERSITY