Conformer and Blind Noisy Students for Improved Image Quality Assessment

Marcos V. Conde, Maxime Burchi, Radu Timofte

https://github.com/burchim/IQA-Conformer-BNS

Motivation

Reference | (a) PSNR 29.22 SSIM 0.37 MOS 1222.33 | (b) PSNR 34.65 SSIM 0.82 MOS 1318.83 | (c) PSNR 29.43 SSIM 0.62 MOS 1390.17 | (d) PSNR 31.66 SSIM 0.90 MOS 1608.99
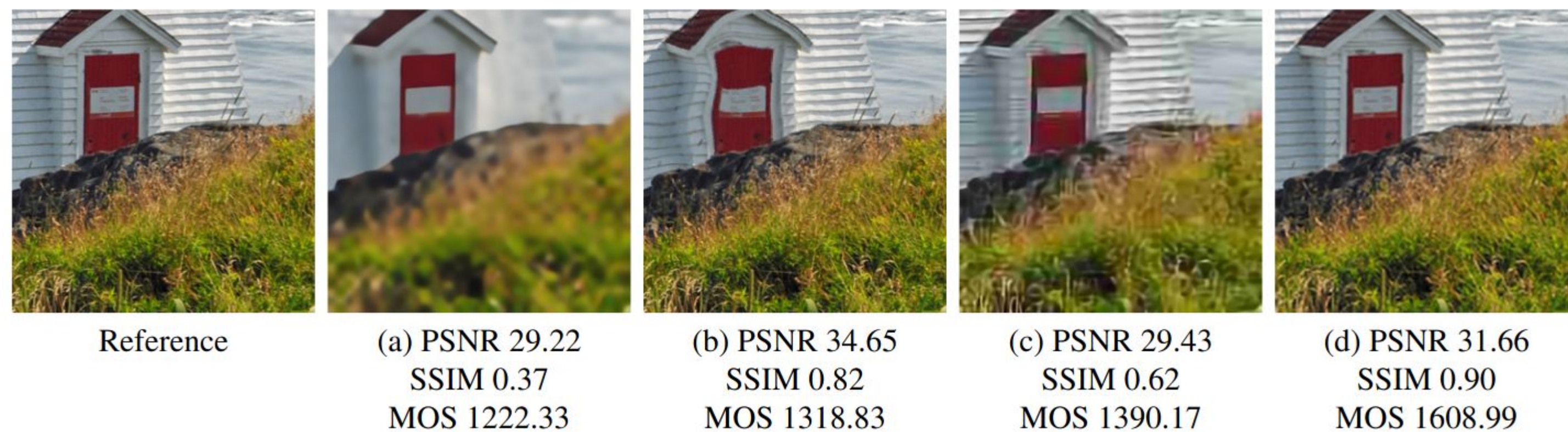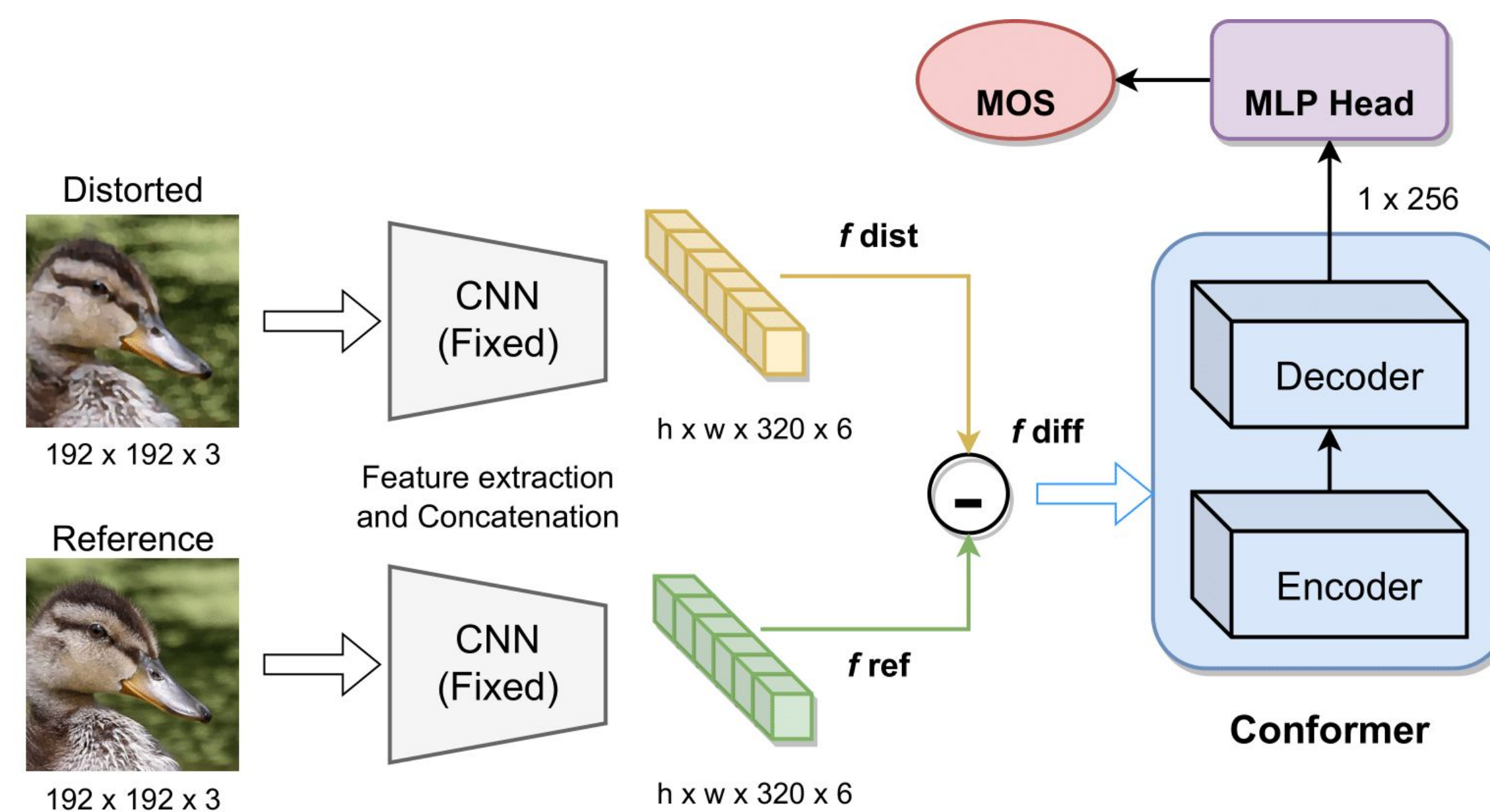
- Generative models for image restoration and enhancement produce more pleasant images to the human eye than other methods, yet, they may get a lower quality score using traditional **perceptual quality metrics** such as PSNR or SSIM.

- Therefore, we need to develop a quantitative metric to **reflect the performance of new algorithms**, which should be highly correlated with the person's mean opinion score (MOS).

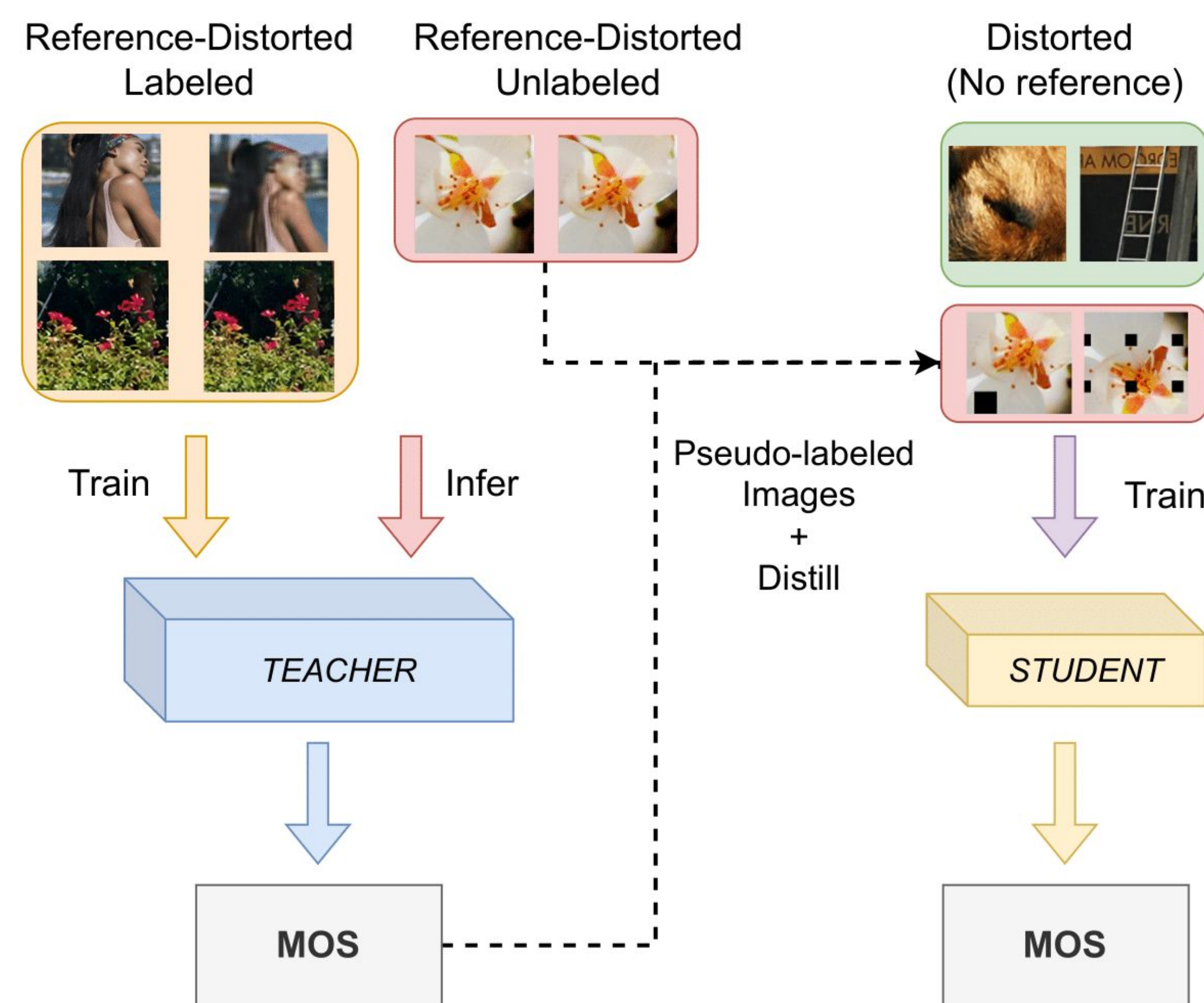### Our contribution in NTIRE 2022 Perceptual IQA Challenge

- Explore the performance of different transformer-based architectures like the **Conformer** for **full-reference IQA**. Our method was ranked 4th in this track of the challenge.

- Our method so-called "**Blind Noisy Student**", based on semi-supervised knowledge distillation from full-reference teacher models into blind student models using noisy pseudo-labeled data, was ranked 3rd in the **no-reference** track of the challenge.
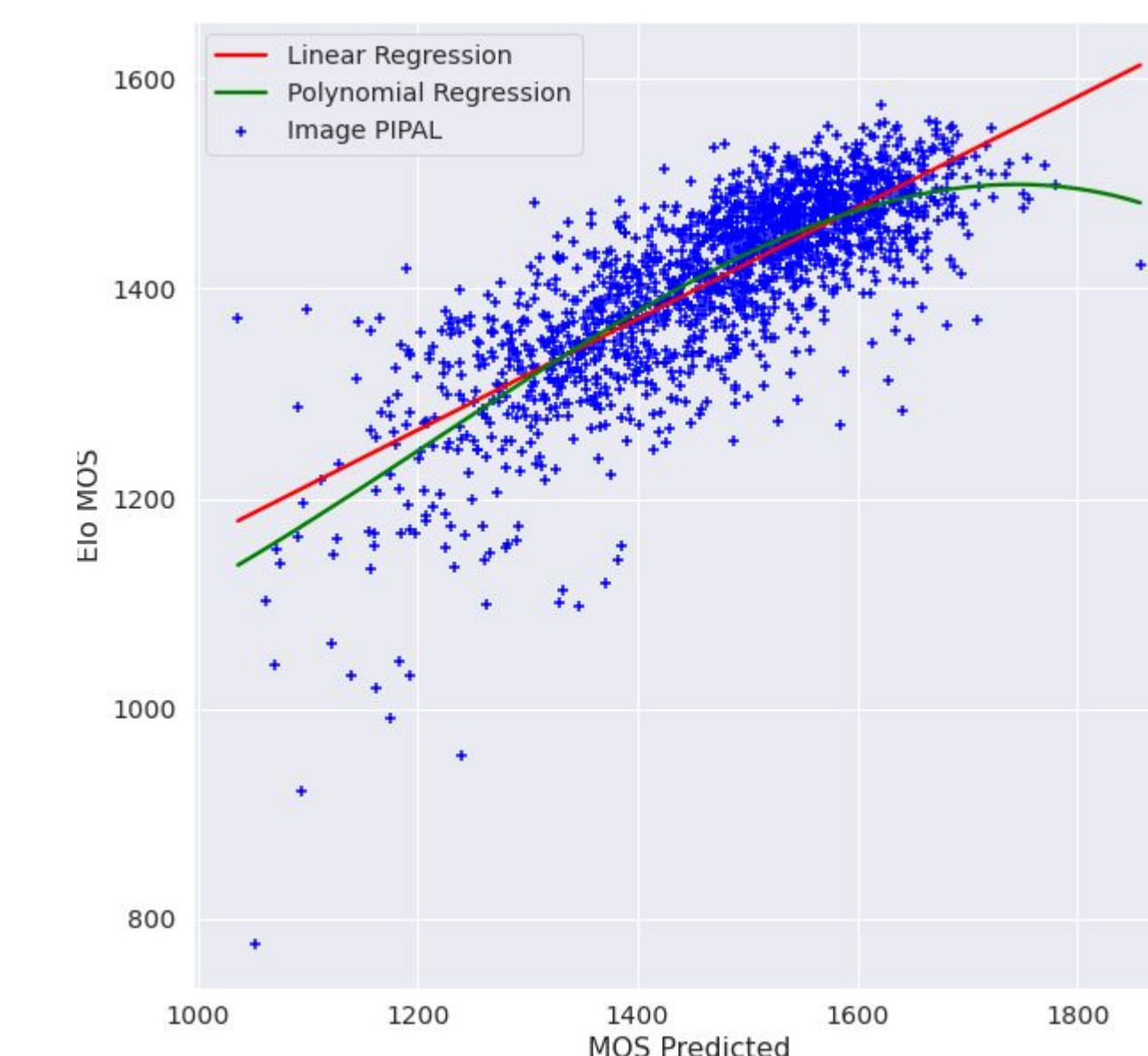
Our Methods

Full-Reference IQA Conformer

Blind Noisy Student

Results

Reference | PSNR 27.968 - SSIM 0.867 Ours MOS 1283.033 | PSNR 32.316 - SSIM 0.815 Ours MOS 1525.727 | PSNR 27.836 - SSIM 0.842 Ours MOS 1534.356

No-Reference MOS Results | Full-Reference MOS Results

Conclusions

- Knowledge distillation from FR models to NR models via semi-supervised learning looks promising.
- Transformer-based approaches are sensitive to the image features, yet very competitive.
- Models trained on PIPAL generalize great, in a zero-shot manner, to other datasets like LIVE or TID2013.