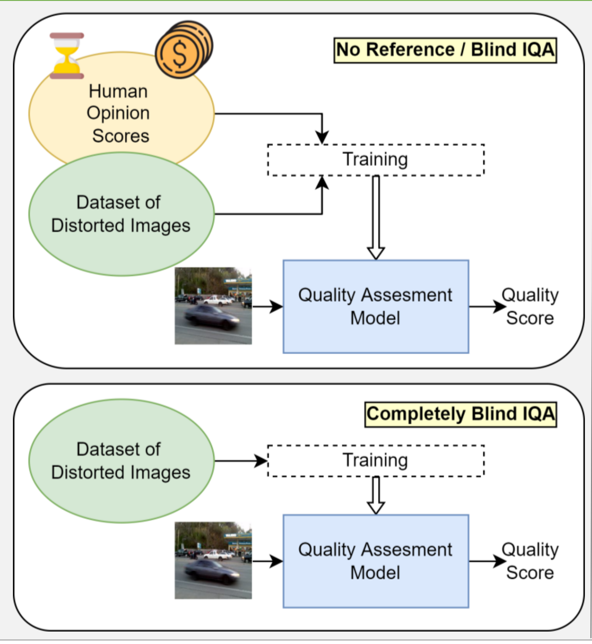
No Reference Opinion Unaware Quality Assessment of Authentically Distorted Images

Nithin C Babu, Vignesh Kannan and Rajiv Soundararajan

Background and motivation

- **Image Quality Assessment (IQA)** - Analysis of the perceptual quality of an image affected by distortions.
  - Applications - Image enhancement, restoration, and compression.



- We focus on designing **deep quality features** in a **completely blind IQA** framework. Our main **contributions** are:
  - A **two-stage self supervised quality feature learning** approach
  - The use of **only positives** in contrastive learning while training on **authentically distorted images**.
  - **Mutual information based loss** function to mitigate content dependence.
  - **Improved** version of **variational approximation** used while estimating the mutual information loss.

Quality feature learning

1 Synthetic data pre-training

- Start with a **pre-trained network**
  - M-SCQALE [1] – performs **contrastive learning** on **synthetically distorted images**.
  - Push and pull features from differently distorted images.

2 Authentic data finetuning

- Authentically distorted images: **no** way of finding **negatives**.
- We adopt the **SimSiam** [2] framework which works with **only positives**.

$$\mathcal{L}_c = \sum_{n=1}^{N} \frac{D(\mathbf{p_1^{(n)}}, \text{stopgrad}(\mathbf{z_2^{(n)}}))}{2} + \frac{D(\mathbf{p_2^{(n)}}, \text{stopgrad}(\mathbf{z_1^{(n)}}))}{2}$$

3 Reducing Content dependence

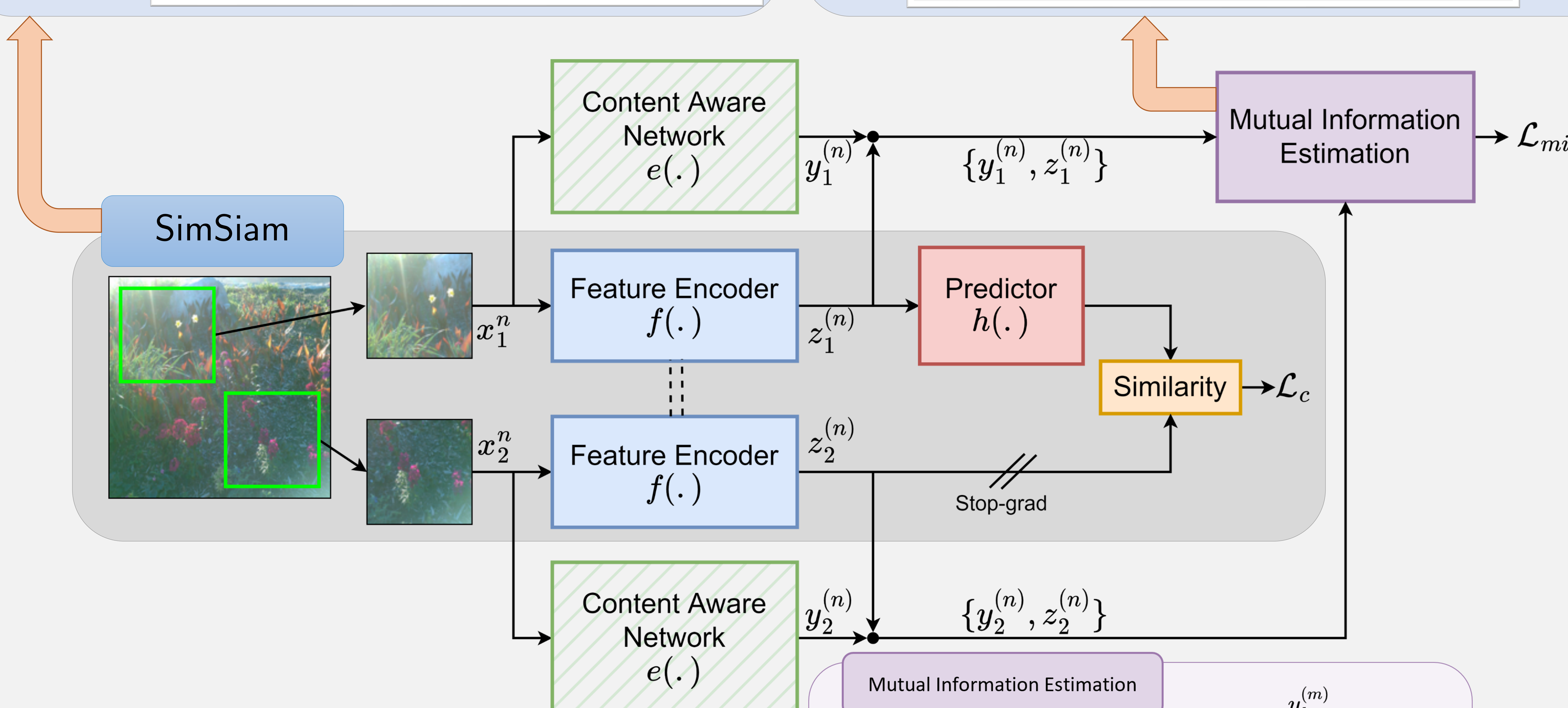
- Working with only positives – **can learn** the **content correlation** between two patches.
- **Minimize mutual information** between quality features and content information using **Contrastive Log-ratio upper bound (CLUB)** [3].

$$I_{CLUB}(Y;Z) = \mathbb{E}_{p(Y,Z)}\left[\log p(Y|Z)\right] - \mathbb{E}_{p(Y)}\mathbb{E}_{p(Z)}\left[\log p(Y|Z)\right]$$

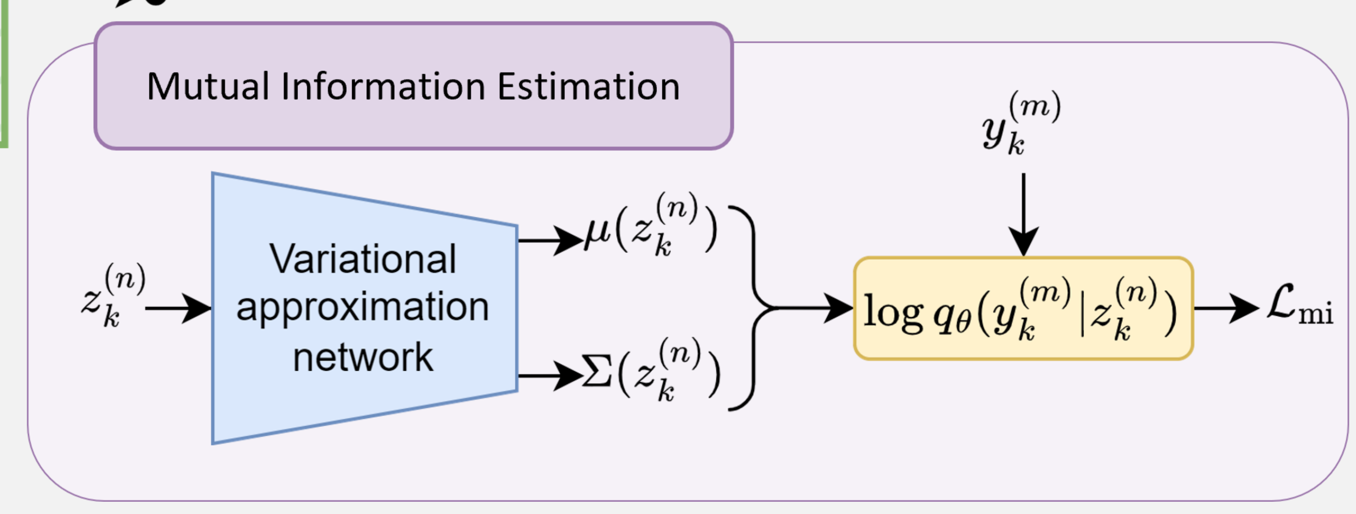
$$I_{CLUB}(Y;Z) \geq I(Y;Z)$$

- Estimated using a variational approximation $q_\theta(.)$ as:

$$\mathcal{L}_{mi} = \frac{1}{N}\sum_{i=1}^{N}\left[\log q_\theta\left(y^{(i)}|z^{(i)}\right) - \frac{1}{N}\sum_{j=1}^{N}\log q_\theta\left(y^{(j)}|z^{(i)}\right)\right]$$



Steps in one iteration		
	Update	**Optimization**
Step 1	$q_\theta(.)$	$\max\{\mathcal{L}_\theta\}$
Step 2	$f(.)$, $h(.)$	$\min\{\mathcal{L}_c + \lambda_{mi}\mathcal{L}_{mi}\}$



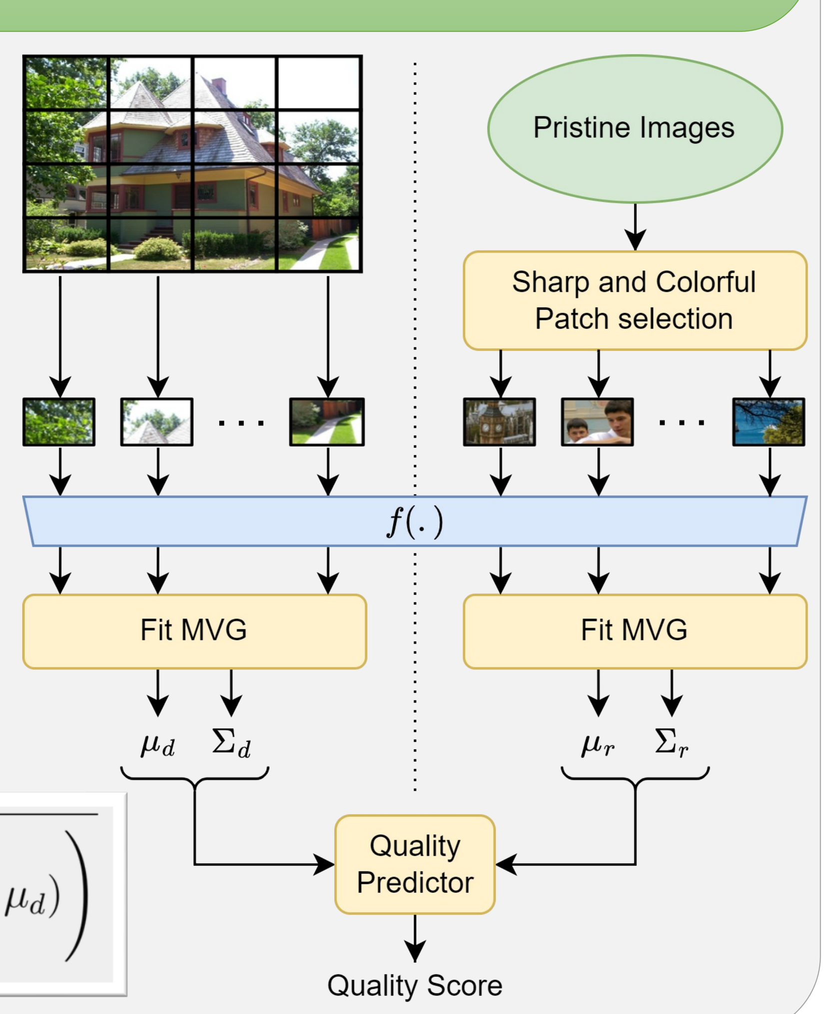
4 Improved variational approximation

- According to the authors of CLUB, θ should be updated such that $q_\theta(Y,Z)$ is **similar to** $p(Y,Z)$, **than to** $p(Y)p(Z)$.
- Does not guarantee if $p(Y)p(Z)$ is different from $q_\theta(Y,Z)$.

CLUB (Log-Likelihood Loss)	$\min_\theta \{\mathrm{KL}(p(Y,Z)\Vert q_\theta(Y,Z))\}$
Our work (Contrastive Likelihood Loss)	$\min_\theta \{\mathrm{KL}(p(Y,Z)\Vert q_\theta(Y,Z)) - \mathrm{KL}(p(Y)p(Z)\Vert q_\theta(Y,Z))\}$

Completely blind quality prediction

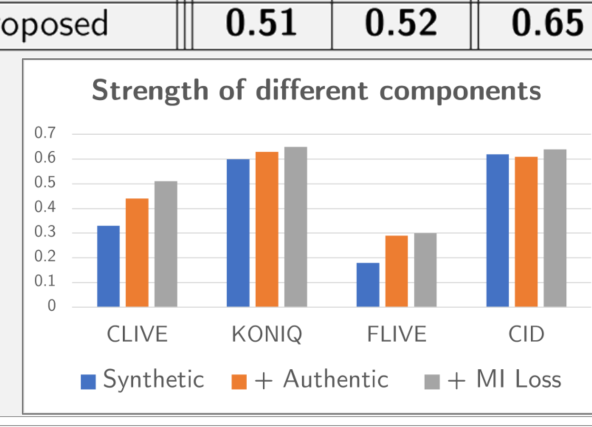
- **Fit MVG models** to quality features of:
  - input image patches
  - corpus of sharp and colorful pristine patches.
- Quality score computed as a **distance between distributions**.

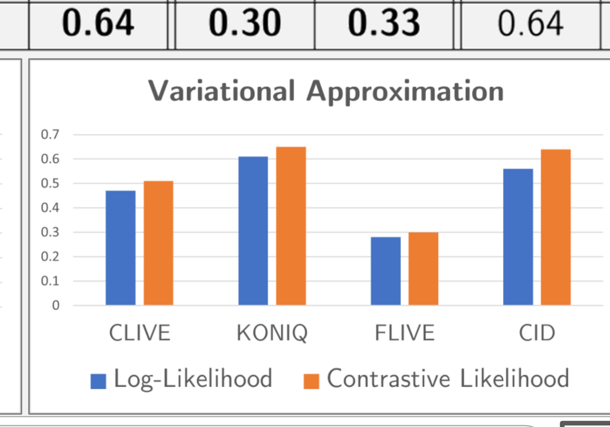


$$Q = \sqrt{\left((\mu_r - \mu_d)^T\left(\frac{\Sigma_r + \Sigma_d}{2}\right)^{-1}(\mu_r - \mu_d)\right)}$$

Results and performance comparisons

Datasets	CLIVE		KonIQ		FLIVE		CID	
Method	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC
NIQE	0.46	0.48	0.53	0.54	0.21	0.29	0.23	0.22
IL-NIQE	0.44	0.49	0.51	0.53	0.22	0.27	0.31	0.40
CORNIA*	0.07	0.07	0.04	0.02	0.05	0.13	0.27	0.29
CONTRIQUE*	0.38	0.42	0.63	0.61	0.26	0.29	**0.74**	**0.76**
Proposed	**0.51**	**0.52**	**0.65**	**0.64**	**0.30**	**0.33**	0.64	0.66





References

[1] Kannan *et al.* "Quality Assessment of Low Light Restored Images: A Subjective Study and an Unsupervised Model." *arXiv preprint 2022*

[2] Chen *et al.* "Exploring simple siamese representation learning." *CVPR 2021*

[3] Cheng *et al.* "CLUB: A contrastive log-ratio upper bound of mutual information." *ICML 2020*