Department of
Computer Science
香港城市大學
City University of Hong Kong

Image Quality Assessment and Perceptual Optimization:

A Non-local Modeling Approach

Shuyue Jia

M.Phil. Thesis Defense

Supervisor: Dr. Shiqi Wang

Outline

1. Background

2. Related Works

3. Local and Non-local Analyses of Natural Images

4. Proposed Non-local Modeling Method

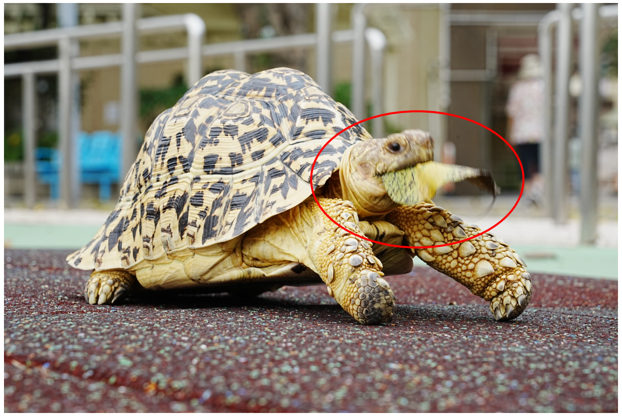
5. Conclusions

Part 1: Background

Image Quality Assessment (IQA) and Distortions

Reference/Pristine Image

Distorted Image
by Gaussian Noise



Motion Blur
by Low Shutter Speed

Visual Quality Assessment: Synthetic and Authentic Distortions

Part 1: Background

IQA Category and Problem Definition

香港城市大學
City University of Hong Kong

Full Reference IQA

Reference Image

Reduced Reference IQA

...
Features of Reference Image

No Reference (Blind) IQA

Distorted Image

IQA Model

Oh! This image looks **excellent**/**good**/**fair**/**poor**/**bad**!

Image Quality

"If you can't ***measure*** *it, you can't* ***improve*** *it."*

— William Thomson

Image Quality Assessment	Automatically measure the input image's visual quality
Synthetic Distortion	Synthetically generated distortions (mainly global uniform distortions)
Authentic Distortion	Images captured in the wild include various content and diverse types of distortions (global uniform distortions & local non-uniform distortions)

Image Credit: TID2013 Database; Quote Credit: British Physicist William Thomson

Part 1: Background Image Quality Assessment Roles

CityU
香港城市大學
City University of Hong Kong

IQA Roles

- Quality Assessment
- Performance Evaluation
- Optimization

Professional Generated Content (**PGC**)

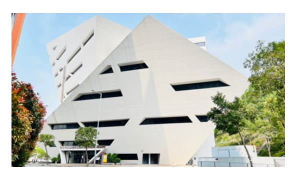
User Generated Content (**UGC**)

AI Generated Content (**AIGC**)

NR IQA Model → Image Quality

Blurred Image

→ Deblur System →



De-blurred Image

Evaluation
or
Optimization

PSNR
SSIM
LPIPS

Reference Image

Part 2: Related Works – Full-reference IQA

1. Signal Fidelity Approaches

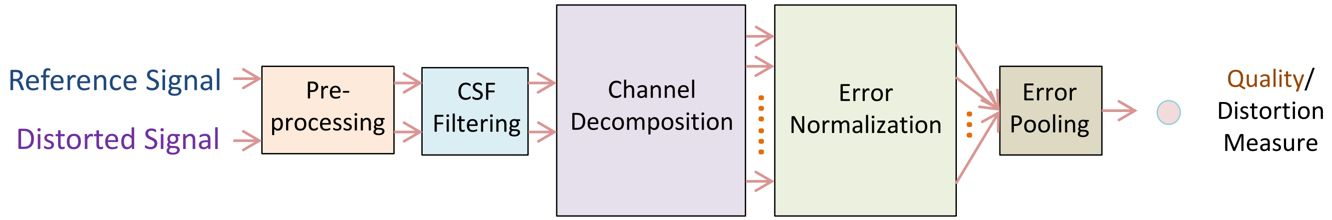
Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR)

I_p: the number of pixels in the image; x_i and y_i are the i^{th} pixels of the ref. and dis. images.

$$\mathrm{MSE} = \frac{1}{I_p}\sum_{i=1}^{I_p}(x_i - y_i)^2 \ , \ \ \mathrm{PSNR} = 10\times \log_{10}\left(\frac{255^2}{\mathrm{MSE}}\right).$$

2. Bottom-Up Approaches

Separately model each basic module of the Human Visual System (HVS)



A prototypical quality assessment system based on error sensitivity.
Image by courtesy of Wang *et al.*

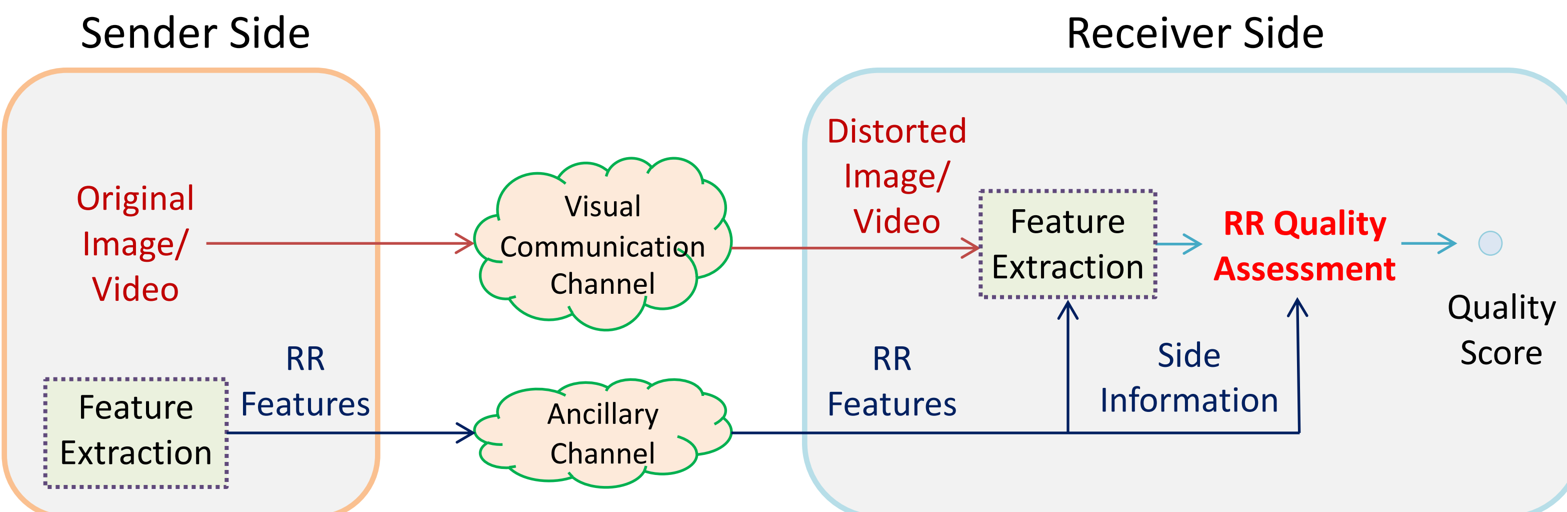
3. Top-Down Approaches

Directly imitate the function of HVS as a single model

Credit: Wang *et al.*, Image Quality Assessment: From Error Visibility to Structural Similarity, In IEEE T-IP'04

Part 2: Related Works

General Framework – Reduced-reference IQA



A general framework of an RR image or image QA system.
Image by courtesy of Wang *et al.*

Feature Extraction in the Spatial Domain and Transform Domain

Credit: Wang *et al.*, Reduced- and No-Reference Image Quality Assessment, In IEEE Signal Processing Magazine'11

Part 2: Related Works – No-reference (Blind) IQA

- **Distortion-Specific Modeling**

  Aware of image distortion types → build distortion-specific models

- **General NR-IQA Modeling**

Natural Scene Statistics (NSS) Modeling

Feature Extraction in the **Spatial Domain** and **Transform Domain**

Human Visual System (HVS) Modeling

CNNs Modeling Methods

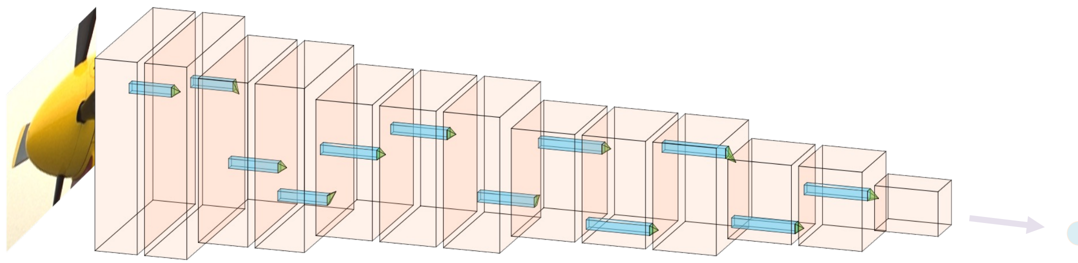
Assisted with visual importance information, reference image information during training, ranking-based methods, graph representation learning, etc.

Codebook-based Modeling

Constructing a **Codebook**

Part 3: Local Modeling Analyses of Natural Images

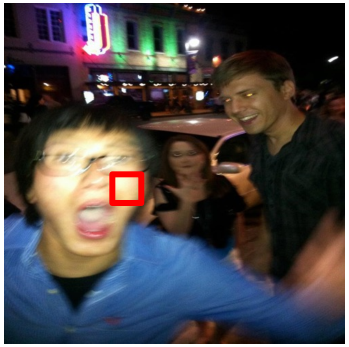
Taking **Convolutional Neural Networks (CNNs)** as an example



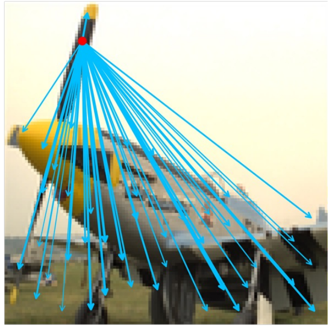
Input Image　　Convolutional Neural Networks　　Image Quality

- **Advantages**
  - ✓ Pooling → **Translation invariance**
  - ✓ Convolution → **Translation equivalence**
  - ✓ Weight sharing → **Sharable and fewer parameters**
- **Limitations**
  - ✓ Small-sized receptive field → **Extracted features are too local**
  - ✓ Parameters fixed across the whole image → **Image content is equally treated**
  - ✓ Lack of geometric and relational modeling → **Missing complex relations and dependencies**

Part 3: Non-local Modeling Analyses of Natural Images



Local Feature Extraction



Non-local Dependency

- **Advantages**
  - ✓ Non-local statistics (natural scene statistics) → **Perceptually relevant to HVS**
  - ✓ Non-local dependency and relational modeling → **Semantics and content understanding**
  - ✓ Effective image-specific prior → **Reflect the statistical property of the world**
- **Motivations**
  - ✓ HVS → **Adaptive to the local content**
  - ✓ HVS perceives image quality → **Long-range dependency constructed among different regions**

Image Credit: LIVEC and TID2013 Databases

Part 4: Proposed Non-local Modeling Method

NLNet Architecture

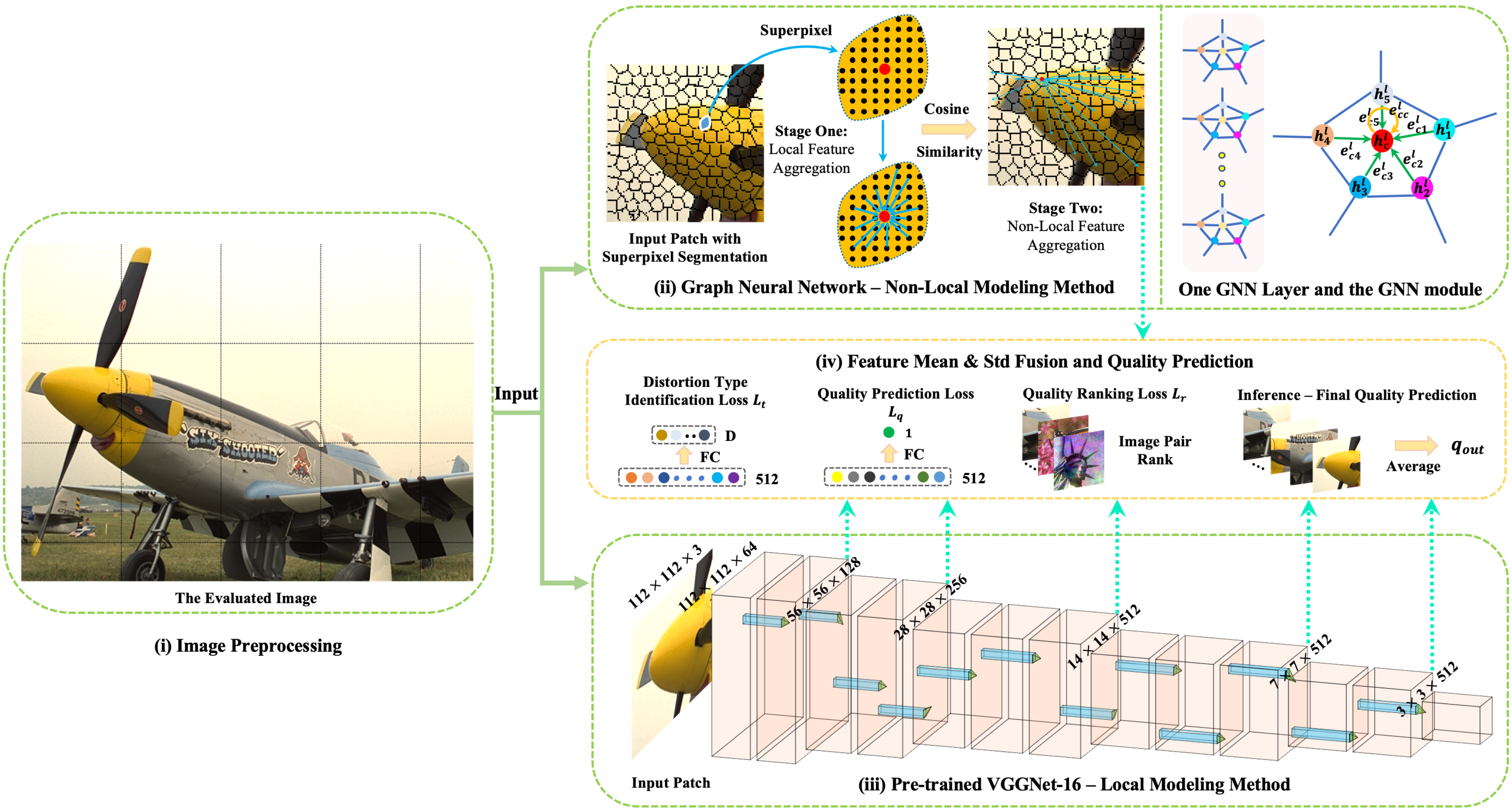
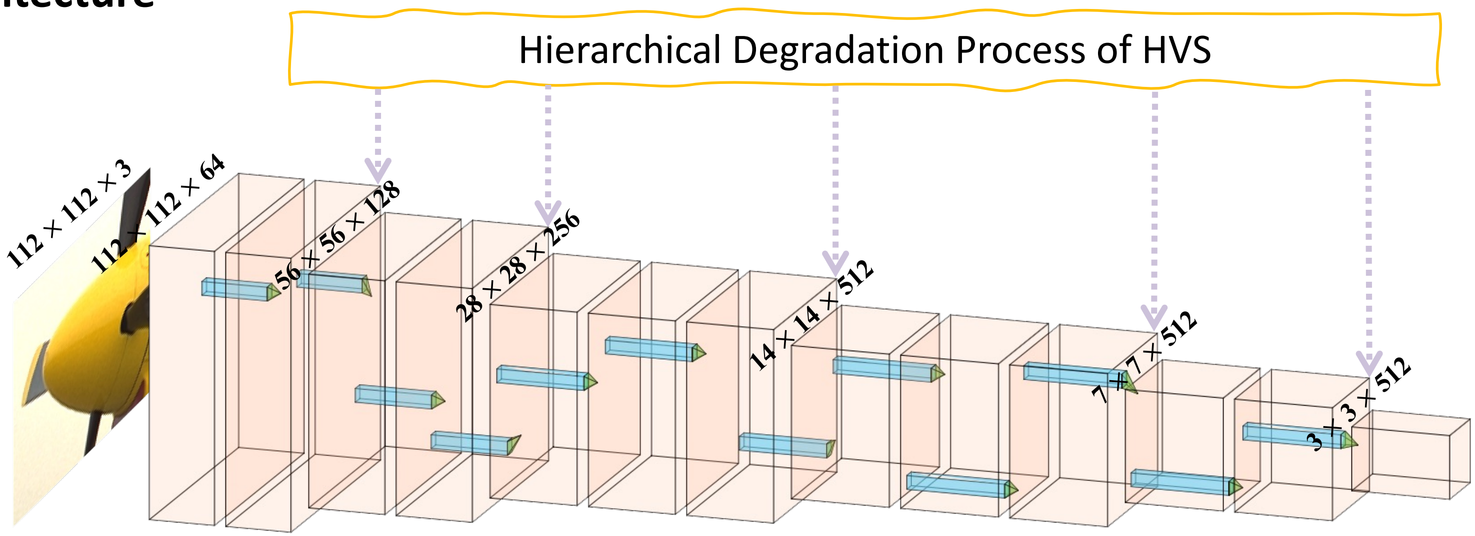


Image Credit: [TID2013 Database]

Part 4: Proposed Method

Local Modeling

1. Architecture



Input Patch

Pre-trained VGGNet-16

Pre-trained VGGNet features: “unreasonable” effectiveness in measuring **perceptual quality** [1]

2. Model Input

Randomly cropped image patches with the size of 112 × 112 × 3

3. Feature Extraction

Quality-aware features: hierarchical feature means and standard deviations [2]

Experimental Results: **0.936 PLCC** and **0.951 SRCC** on the CSIQ Database

Credit: [1] Zhang *et al.*, The Unreasonable Effectiveness of Deep Features as a Perceptual Metric, In CVPR’18
[2] Ding *et al.*, Image Quality Assessment: Unifying Structure and Texture Similarity, In IEEE T-PAMI’20

Part 4: Proposed Method

Non-local Modeling

Non-Local
Object-to-Pixel Modeling

Spatial Integration of Information

$$\mathbf{h}_i^l = \mathrm{ELU}\left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij}^l \mathbf{W}^l \mathbf{h}_j^l\right),$$

Spatial Weighting Functions

$$\alpha_{ij}^l = \frac{\exp(\mathrm{a}_{ij}^l)}{\sum_{k \in \mathcal{N}(i)} \mathrm{a}_{ik}^l},$$

$$\mathrm{a}_{ij}^l = \mathrm{LeakyReLU}\left(\mathrm{FC}\left(\left[\mathbf{W}^l \mathbf{h}_i^l \parallel \mathbf{W}^l \mathbf{h}_j^l\right]\right)\right).$$

Convolution
Pixel-to-Pixel Modeling

Local region feature extraction and non-local dependency feature extraction.

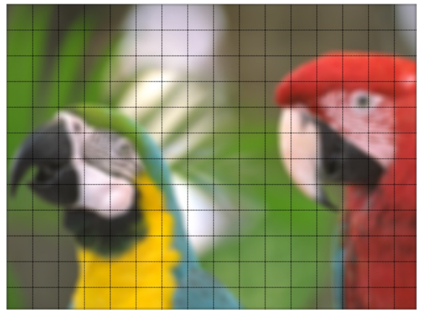
Local Modeling	Encodes **Spatially Proximate Local Neighborhoods**
Non-local Modeling	Establishes **Spatial Integration of Information** by **Long- and Short-Range Communications** with different **Spatial Weighting Functions**

Image Credit: TID2013 and LIVEC Databases

Part 4: Proposed Method

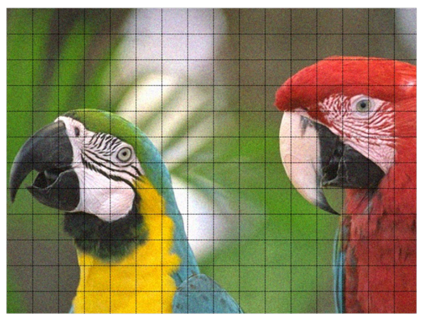
Superpixel Segmentation

(a)



(b)

(c)



(d)

The superpixel *vs.* square patch representation
(**a/b**: Gaussian Blur, **c/d**: Gaussian Noise)

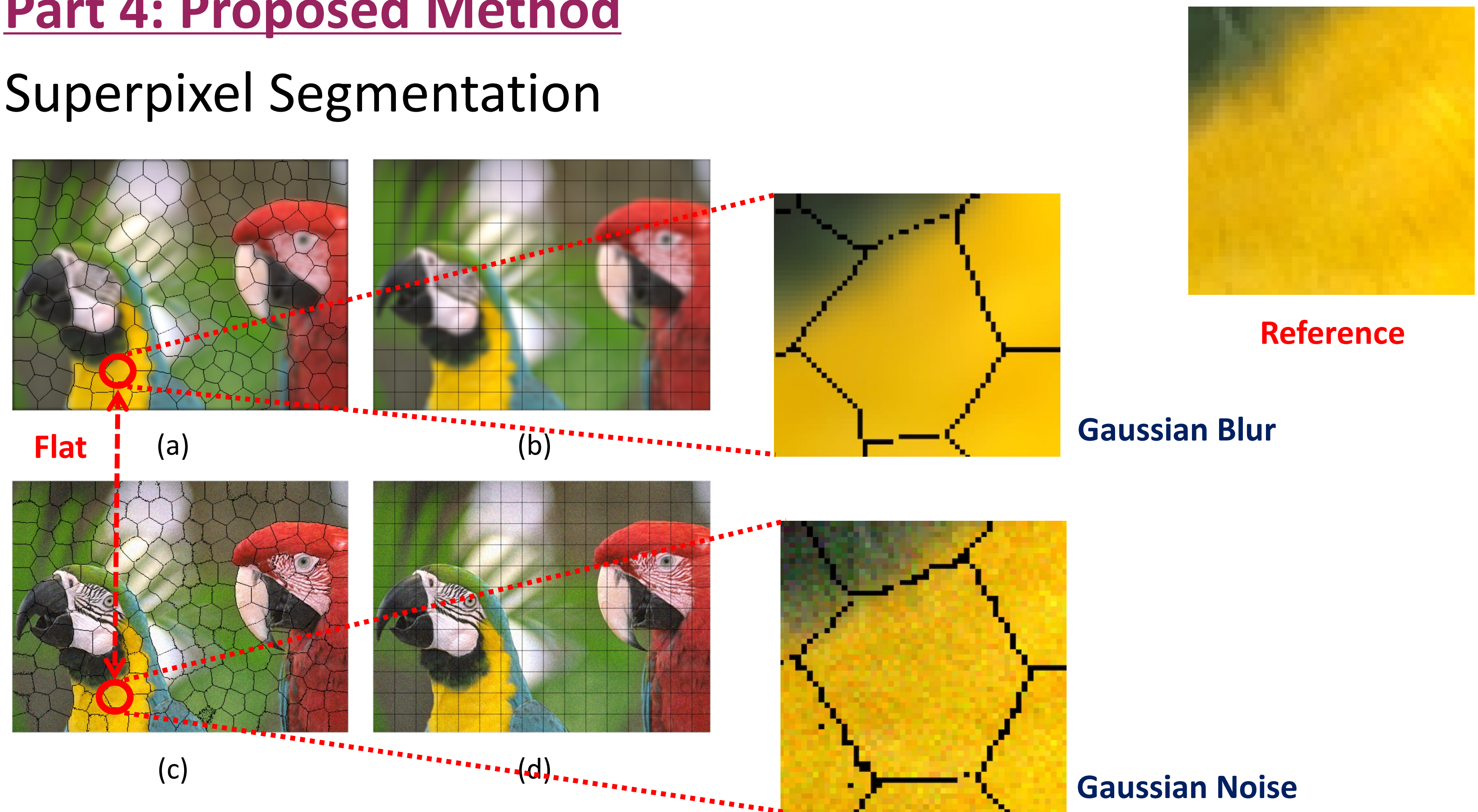
Superpixel *vs.* **Square Patch**

- ✓ Adherence to Boundaries
- ✓ Visually Meaningful
- ✓ Accurate Feature Extraction

Image Credit: TID2013 Database

Part 4: Proposed Method

Superpixel Segmentation

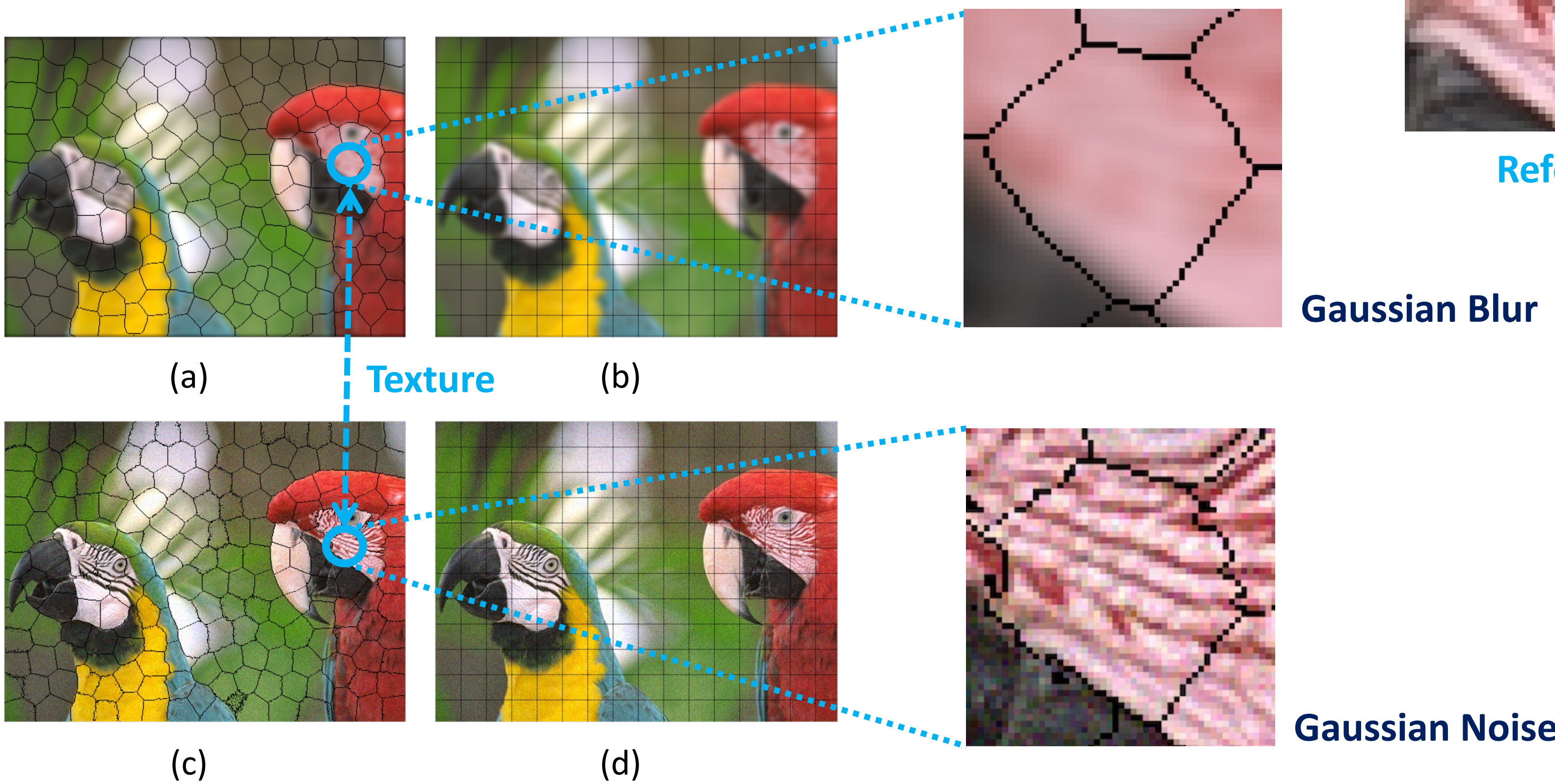


The superpixel *vs.* square patch representation
(**a/b**: Gaussian Blur, **c/d**: Gaussian Noise)

Image Credit: TID2013 Database

Part 4: Proposed Method

Superpixel Segmentation



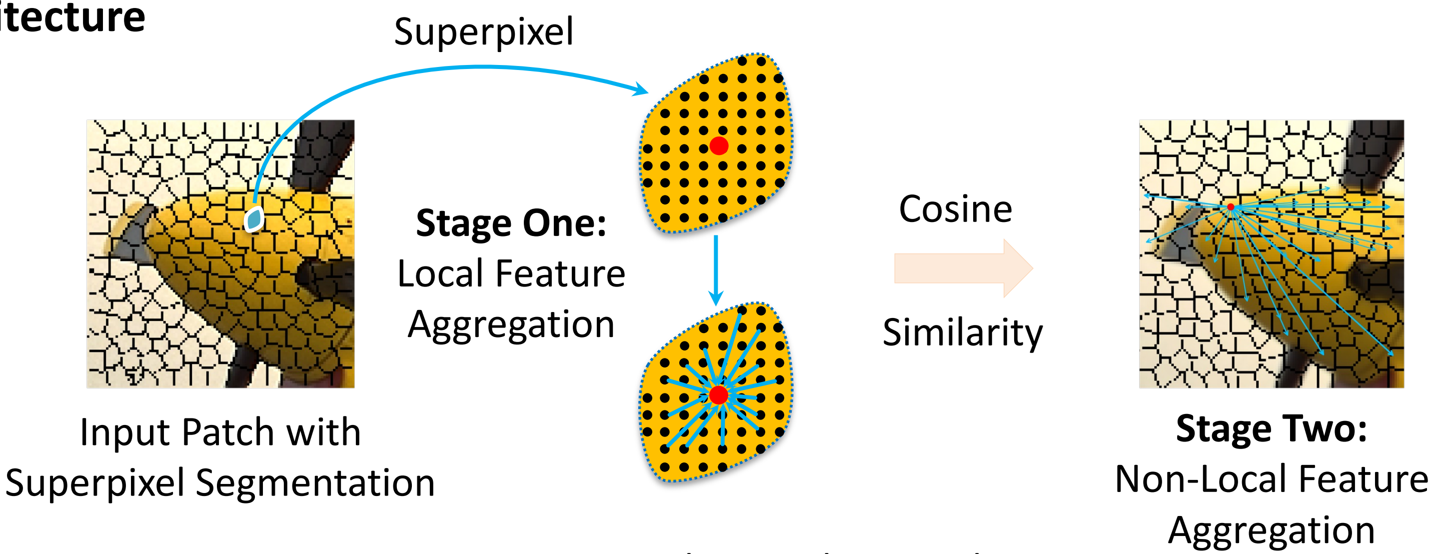
The superpixel *vs.* square patch representation
(**a/b**: Gaussian Blur, **c/d**: Gaussian Noise)

Image Credit: TID2013 Database

Part 4: Proposed Method

Non-local Modeling

1. **Architecture**



Two-stage Graph Neural Network

Graph Nodes Construction → **Non-local Feature Aggregation**

2. **Model Input**

Superpixel-segmented cropped patches with the size of 112 × 112 × 3 (superpixel size ≈ 8 × 8)

3. **Feature Extraction**

Local feature aggregation: **complete graph**

Non-local feature aggregation: **dense graph** measured by Cosine similarity

Experimental Results: **0.625 PLCC** and **0.577 SRCC** on the CSIQ Database

Image Credit: TID2013 Database

Part 4: Proposed Method

Non-local Behavior Visualization

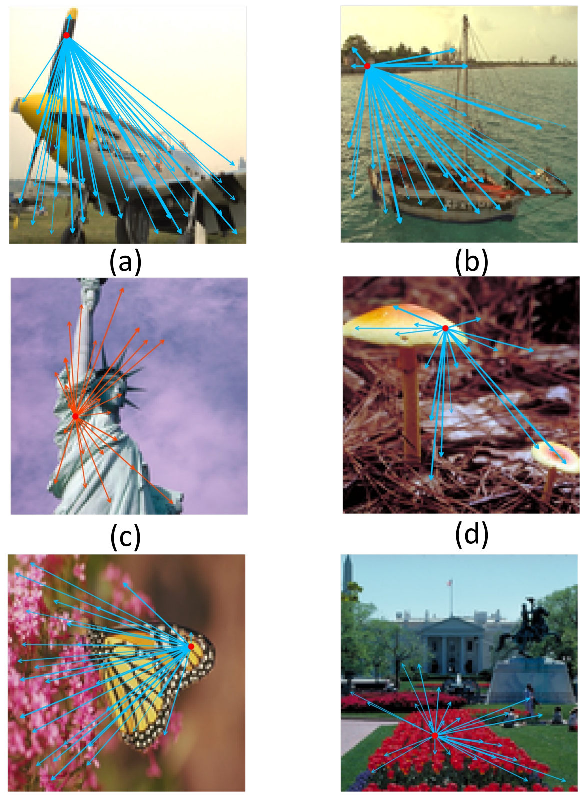
Object-to-Pixel Modeling
Region Feature Extraction

↑

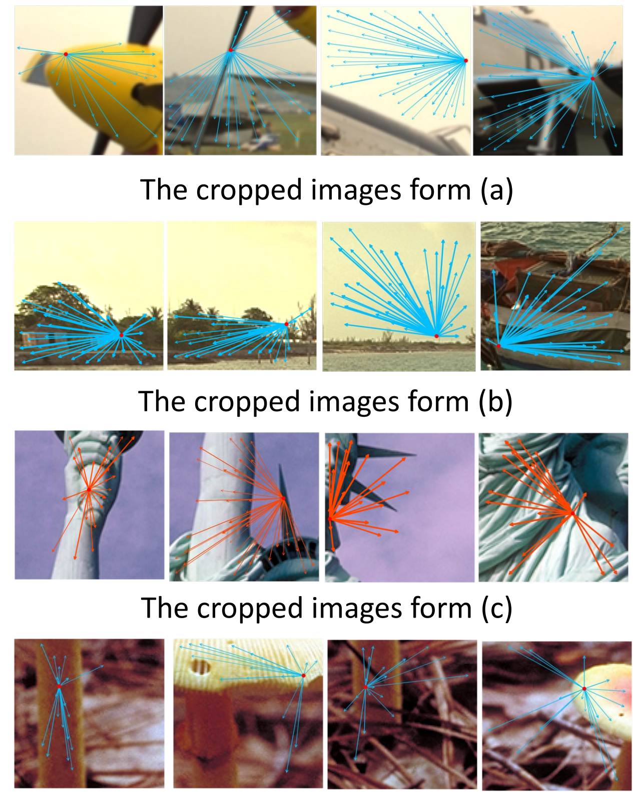
Non-local Dependency & Relational Modeling

↓

Semantics and Content Understanding



(a) (b) (c) (d) (e) (f)



The cropped images form (a)

The cropped images form (b)

The cropped images form (c)

The cropped images form (d)

The non-local behavior of the long-range dependency and relational modeling. **(a)** The plane image with a query on wings. **(b)** The boat image with a query on the nearby river bank. **(c)** The Statue of Liberty image with a query on the lady. **(d)** The shrooms image with a query on one shroom. **(e)** The butterfly image with a query on the wing. **(f)** The Lafayette Square, Washington, D.C. image with a query on flowers.

Image Credit: TID2013 and LIVE Databases

Part 4: Proposed Method

Training Objective Functions and Inference

- **Training**

  1. **Quality Prediction Loss**

     Huber Loss

$$L_q = \frac{1}{B}\sum_k \mathrm{HuberLoss}(\hat{q}_k - q_k)\,.$$

  2. **Distortion Type Classification Loss**

     Cross-entropy Loss

$$L_t = -\frac{1}{B}\sum_{i=1}^{B}\sum_{d=1}^{D} p_{i,d}\ln \hat{p}_{i,d}\,.$$

  3. **Quality Ranking Loss**

     Huber Loss

$$L_r = \frac{1}{B(B-1)/2}\sum_{j<k}\mathrm{HuberLoss}\Big(\big(\hat{q}_j - \hat{q}_k\big) - \big(q_j - q_k\big)\Big).$$

Notations	
B	Batch Size
$\hat{q}_k$	The predicted quality score
q_k	Mean Opinion Score (MOS)
D	The number of distortion types
$p_{i,d}$	The label probability of the d^{th} distortion type
$\hat{p}_{i,d}$	The predicted probability of the d^{th} distortion type

- **Inference**

  The average quality score of all the **non-overlapping patches**

average q_{out}

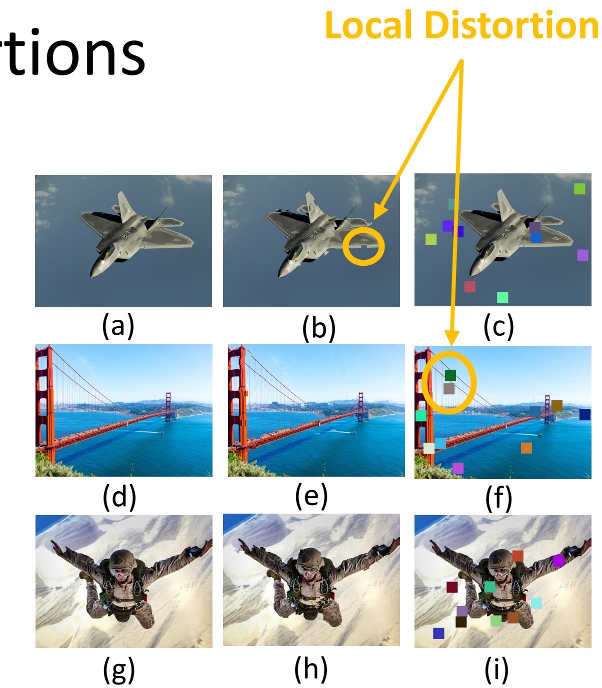
Image Credit: [TID2013 Database]

Part 4: Proposed Method

Definition of Global and Local Distortions

Non-Local Recurrence

Global Distortion
(**a/e**: reference image, **b/f**: Gaussian Blur, **c/g**: global contrast decrements, **d/h**: additive pink Gaussian noise)



Local Distortion
(**a/d/g**: reference image, **b/e/h**: non eccentricity patch, **c/f/i**: color block)

Global Distortion	**Globally and Uniformly Distributed** distortions with **Non-Local Recurrences** over the image
Local Distortion	**Local Nonuniform-Distributed** distortions **in a Local Region**

Image Credit: CSIQ and KADID-10k Databases

Part 4: Proposed Method

Experimental Setup

- **Databases**

  LIVE, CSIQ, TID2013, and KADID-10k

- **Experimental Settings**

  Intra-Database Experiments:

  → 60% training

  20% validation

  20% testing

  With "random" seeds from 1 to 10

  → **Median** PLCC and SRCC are reported.

  Cross-Database Evaluations:

  → **One database** as the training set

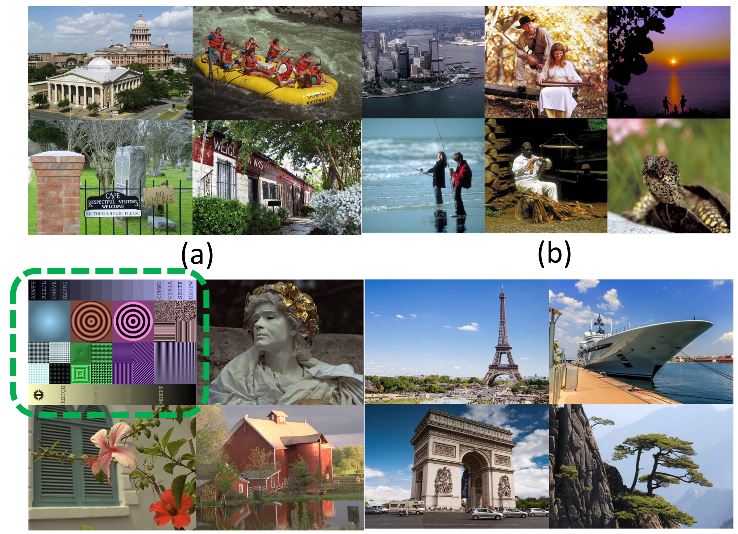
  The other databases as the testing set

  → Report the **last epoch**'s performance

- **Evaluation Metrics**

  PLCC (Prediction **Accuracy**)

  SRCC (Prediction **Monotonicity**)



(a) LIVE Database **(b)** CSIQ Database
(c) TID2013 Database **(d)** KADID-10k Database

A brief summary of the LIVE, CSIQ, TID2013, and KADID-10k databases.

Database	LIVE [13]	CSIQ [14]	TID2013 [15]	KADID-10k [16]
Num. of Reference Images	29	30	25	81
Num. of Distorted Images	779	866	3,000	10,125
Num. of Distortion Types	5	6	24	25
Num. of Distortion Levels	$5\sim8$	$3\sim5$	5	5
Annotation	DMOS	DMOS	MOS	MOS
Range	[0, 100]	[0, 1]	[0, 9]	[1, 5]

Image Credit: LIVE, CSIQ, TID2013, and KADID-10k Databases

Part 4: Proposed Method

Intra-Database Experiments

Performance comparisons on the LIVE, CSIQ, and TID2013 databases. The top two results are highlighted in bold.

Method	LIVE		CSIQ		TID2013	
	SRCC	PLCC	SRCC	PLCC	SRCC	PLCC
BRISQUE (2012) [10]	0.939	0.935	0.746	0.829	0.604	0.694
CORNIA (2012) [104]	0.947	0.950	0.678	0.776	0.678	0.768
M3 (2015) [105]	0.951	0.950	0.795	0.839	0.689	0.771
HOSA (2016) [103]	0.946	0.947	0.741	0.823	0.735	0.815
FRIQUEE (2017) [90]	0.940	0.944	0.835	0.874	0.68	0.753
DIQaM-NR (2018) [35]	0.960	**0.972**	-	-	0.835	0.855
DB-CNN (2020) [64]	0.968	**0.971**	**0.946**	**0.959**	0.816	0.865
HyperIQA (2020) [65]	0.962	0.966	0.923	0.942	0.729	0.775
GraphIQA (2022) [86]	**0.968**	0.970	0.920	0.938	-	-
TReS (2022) [87]	**0.969**	0.968	0.922	0.942	**0.863**	**0.883**
NLNet	0.962	0.963	**0.941**	**0.958**	**0.856**	**0.880**

SOTA Transformer

Fewer Training Data

↓ *20%* Total Data

↑ Highly Competitive Performance

Performance comparisons on the KADID-10k database. The top two results are highlighted in bold.

Method	BRISQUE [10]	CORNIA [104]	HOSA [103]	InceptionResNetV2 [16]	DB-CNN [64]	HyperIQA [65]	TReS [87]	**NLNet**
SRCC	0.519	0.519	0.609	0.731	0.851	**0.852**	**0.859**	0.846
PLCC	0.554	0.554	0.653	0.734	**0.856**	0.845	**0.858**	0.850

Part 4: Proposed Method

Cross-Database Evaluations

Cross-database performance comparisons.

Training	LIVE		CSIQ		TID2013	
Testing	CSIQ	TID2013	LIVE	TID2013	LIVE	CSIQ
BRISQUE (2012) [10]	0.562	0.358	0.847	0.454	0.790	0.590
CORNIA (2012) [104]	0.649	0.360	0.853	0.312	0.846	0.672
M3 (2015) [105]	0.621	0.344	0.797	0.328	0.873	0.605
HOSA (2016) [103]	0.594	0.361	0.773	0.329	0.846	0.612
FRIQUEE (2017) [90]	0.722	0.461	0.879	0.463	0.755	0.635
DIQaM-NR (2018) [35]	0.681	0.392	-	-	-	0.717
DB-CNN (2020) [64]	**0.758**	**0.524**	0.877	**0.540**	**0.891**	**0.807**
HyperIQA (2020) [65]	0.697	**0.538**	**0.905**	**0.554**	0.839	0.543
NLNet	**0.771**	0.497	**0.923**	0.516	**0.895**	**0.730**

Similar Distortions

TID2013: More Distortion Types & Levels

Part 4: Proposed Method

Single Distortion Type Evaluation on the LIVE Database

The **average SRCC and PLCC** results of the individual distortion type on the LIVE database. The top two results are highlighted in bold.

SRCC	Global Distortion				Local Distortion
	JPEG	JP2K	WN	GB	FF
BRISQUE (2012) [10]	0.965	0.929	0.982	**0.964**	0.828
CORNIA (2012) [104]	0.947	0.924	0.958	0.951	0.921
M3 (2014) [105]	0.966	0.930	**0.986**	0.935	0.902
HOSA (2016) [103]	0.954	0.935	0.975	0.954	**0.954**
FRIQUEE (2017) [90]	0.947	0.919	0.983	0.937	0.884
dipIQ (2017) [82]	0.969	**0.956**	0.975	0.940	-
WaDIQaM (2018) [35]	0.953	0.942	0.982	0.938	0.923
DB-CNN (2020) [64]	**0.972**	0.955	0.980	0.935	0.930
HyperIQA (2020) [65]	0.961	0.949	0.982	0.926	0.934
NLNet	**0.979**	**0.958**	**0.990**	**0.964**	**0.941**

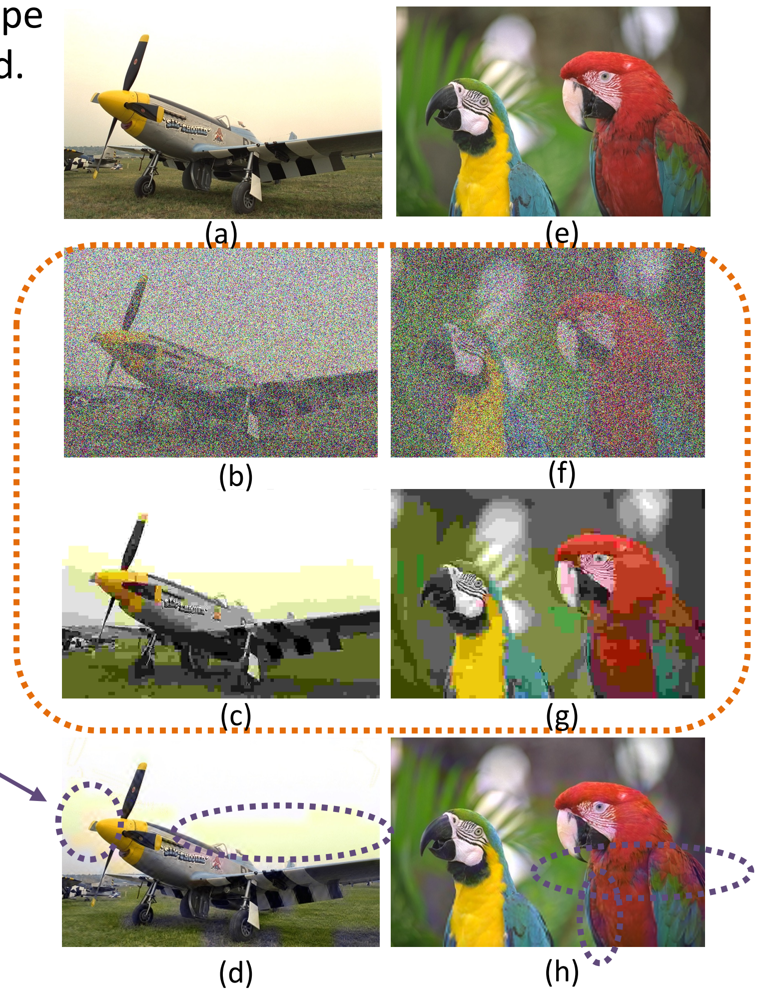
PLCC	Global Distortion				Local Distortion
	JPEG	JP2K	WN	GB	FF
BRISQUE (2012) [10]	0.971	0.940	0.989	**0.965**	0.894
CORNIA (2012) [104]	0.962	0.944	0.974	0.961	0.943
M3 (2014) [105]	0.977	0.945	**0.992**	0.947	0.920
HOSA (2016) [103]	0.967	0.949	0.983	**0.967**	**0.967**
FRIQUEE (2017) [90]	0.955	0.935	0.991	0.949	0.936
dipIQ (2017) [82]	0.980	**0.964**	0.983	0.948	-
DB-CNN (2020) [64]	**0.986**	**0.967**	0.988	0.956	**0.961**
NLNet	**0.986**	0.961	**0.993**	0.964	0.951

Noisy & Compressed Images

Global Distortion

Non-local Recurrence

Local Distortion



Demonstrations of the global distortions (**a/e**: reference image, **b/f**: white noise and **c/g**: JPEG) and local distortions (**d/h**: fast fading Rayleigh)

Image Credit: LIVE Database

Part 4: Proposed Method

Single Distortion Type Evaluation on the CSIQ Database

The **average SRCC and PLCC** results of the individual distortion type on the CSIQ database.
The top two results are highlighted in bold.

SRCC	JPEG	JP2K	WN	GB	PN	CC
BRISQUE (2012) [10]	0.806	0.840	0.723	0.820	0.378	0.804
CORNIA (2012) [104]	0.513	0.831	0.664	0.836	0.493	0.462
M3 (2014) [105]	0.740	0.911	0.741	0.868	0.663	0.770
HOSA (2016) [103]	0.733	0.818	0.604	0.841	0.500	0.716
FRIQUEE (2017) [90]	0.869	0.846	0.748	0.870	0.753	0.838
dipIQ (2017) [82]	0.936	0.944	0.904	0.932	-	-
MEON (2018) [71]	**0.948**	0.898	**0.951**	0.918	-	-
WaDIQaM (2018) [35]	0.853	0.947	0.974	0.979	0.882	**0.923**
DB-CNN (2020) [64]	0.940	0.953	0.948	**0.947**	**0.940**	0.870
HyperIQA (2020) [65]	0.934	**0.960**	0.927	0.915	0.931	0.874
NLNet	**0.972**	**0.963**	**0.965**	**0.955**	**0.969**	**0.968**

PLCC	JPEG	JP2K	WN	GB	PN	CC
BRISQUE (2012) [10]	0.828	0.887	0.742	0.891	0.496	0.835
CORNIA (2012) [104]	0.563	0.883	0.687	0.904	0.632	0.543
M3 (2014) [105]	0.768	0.928	0.728	0.917	0.717	0.787
HOSA (2016) [103]	0.759	0.899	0.656	0.912	0.601	0.744
FRIQUEE (2017) [90]	0.885	0.883	0.778	0.905	0.769	0.864
dipIQ (2017) [82]	0.975	**0.959**	0.927	0.958	-	-
MEON (2018) [71]	0.979	0.925	**0.958**	0.946	-	-
DB-CNN (2020) [64]	**0.982**	0.971	0.956	**0.969**	**0.950**	**0.895**
NLNet	**0.991**	**0.976**	**0.967**	**0.9746**	**0.966**	**0.969**

Noise-Related Distortions

Global Distortions

Demonstrations of the global distortions
(**a/e**: reference image, **b/f**: Gaussian Blur, **c/g**: global contrast decrements, **d/h**: additive pink Gaussian noise)

Image Credit: CSIQ Database

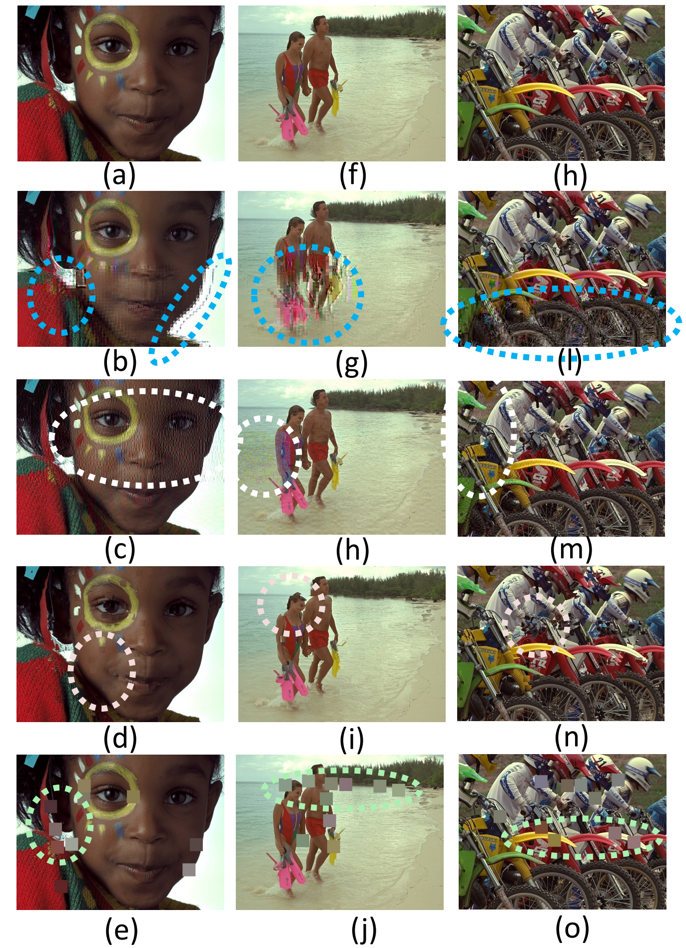
Part 4: Proposed Method

Single Distortion Type Evaluation on the TID2013 Database

The **average SRCC** results of the individual distortion type on the TID2013 database.
The top two results are highlighted in bold.

SRCC	Distortion Type	BRISQUE [10]	FRIQUEE [90]	HOSA [103]	MEON [71]	M3 [105]	DB-CNN [64]	CORNIA [104]	NLNet
Global Distortion	Additive Gaussian noise	0.711	0.730	**0.833** ↑8.4%	0.813	0.766	0.790	0.692	**0.917**
	Lossy compression of noisy images	0.609	0.641	0.838	0.772	0.692	**0.860** ↑7.5%	0.712	**0.935**
	Additive noise in color components	0.432	0.573	0.551	**0.722** ↑12.8%	[illegible]	0.700	0.137	**0.850**
	Comfort noise	0.196	0.318	0.622	0.406	0.353	**0.752** ↑11.8%	0.617	**0.870**
	Contrast change	-0.001	**0.585**	0.362	0.252	0.155	0.548	0.254	**0.793**
	Change of color saturation	0.003	0.589	0.045	**0.684**	-0.199	0.631	0.169	**0.827**
	Spatially correlated noise	0.746	0.866	0.842	**0.926** ↑3.2%	0.782	0.826	0.741	**0.958**
	High frequency noise	0.842	0.847	0.897	**0.911** ↑1.0%	0.900	0.879	0.815	**0.921**
	Impulse noise	0.765	0.730	0.809	**0.901** ↑1.2%	0.738	0.708	0.616	**0.913**
	Quantization noise	0.662	0.764	0.815	**0.888** ↑4.1%	0.832	0.825	0.661	**0.929**
	Gaussian blur	0.871	0.881	0.883	0.887	**0.896**	0.859	0.850	**0.912**
	Image denoising	0.612	0.839	0.854	0.797	0.709	**0.865** ↑1.7%	0.764	**0.882**
	JPEG compression	0.764	0.813	0.891	0.850	0.844	**0.894** ↑1.1%	0.797	**0.905**
	JPEG 2000 compression	0.745	0.831	**0.919** ↑1.1%	0.891	0.885	0.916	0.846	**0.930**
	Multiplicative Gaussian noise	0.717	0.704	0.768	**0.849** ↑5.5%	0.738	0.711	0.593	**0.904**
	Image color quantization with dither	0.831	0.768	0.896	0.857	**0.908**	0.833	0.683	**0.911**
	Sparse sampling and reconstruction	0.807	0.891	**0.909**	0.855	0.893	0.902	0.865	**0.940**
	Chromatic aberrations	0.615	0.737	0.753	**0.779**	0.570	0.732	0.696	**0.773**
	Masked noise	0.252	0.345	0.468	**0.728**	0.577	0.646	0.451	**0.700**
	Mean shift (intensity shift)	0.219	**0.254**	0.211	0.177	0.119	-0.009	0.232	**0.358**
Local Distortion	JPEG transmission errors	0.301	0.498	0.730	0.746	0.375	**0.772** ↑3.3%	0.694	**0.805**
	JPEG 2000 transmission errors	0.748	0.660	0.710	0.716	0.718	**0.773** ↑10.2%	0.686	**0.875**
	Non eccentricity pattern noise	0.269	0.076	0.242	0.116	0.173	**0.270** ↑34.6%	0.200	**0.616**
	Local bock-wise distortions with different intensity	0.207	0.032	0.268	**0.500**	0.379	0.444	0.027	**0.493**

Noise and Compression-related Distortions



Demonstrations of the local distortions (**a/f/h**: reference image, **b/g/l**: JPEG transmission errors, **c/h/m**: JPEG2000 transmission errors, **d/i/n**: non eccentricity pattern noise, **e/j/o**: local block-wise distortions of different intensity)

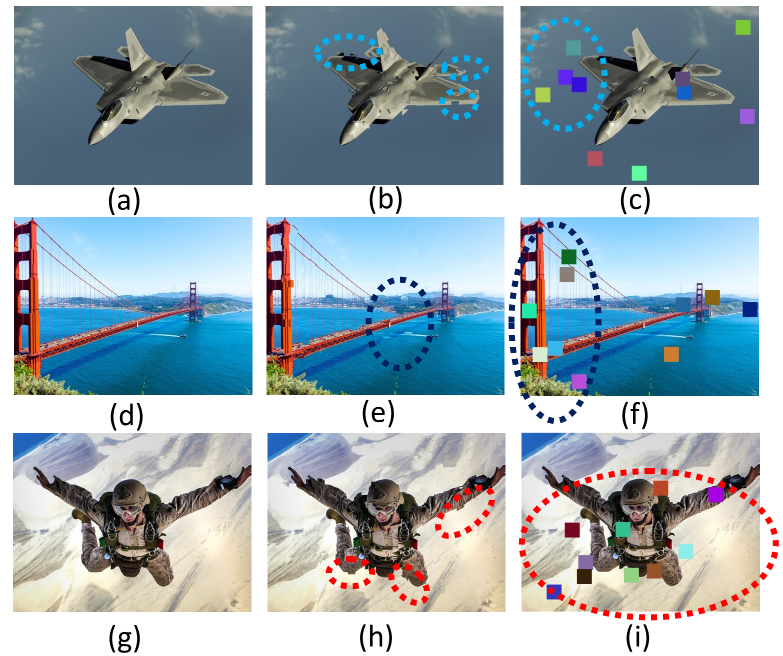
Image Credit: [TID2013 Database]

Part 4: Proposed Method

Single Distortion Type Evaluation on the KADID-10k Database

The **average SRCC** results of the individual distortion type on the KADID-10k database. The local distortions are highlighted in blue, and the top two results are highlighted in bold.

Distortion Type		BLIINDS-II [91]	BRISQUE [10]	ILNIQE [102]	CORNIA [104]	HOSA [103]	WaDIQaM [35]	**NLNet**
Blurs	Lens blur	0.781	0.674	**0.846**	0.811	0.715	0.730	**0.914**
	Gaussian blur	0.880	0.812	**0.883**	0.866	0.852	0.879	**0.914**
	Motion blur	0.482	0.423	**0.779**	0.532	0.652	0.730	**0.899**
Color distortions	Color diffusion	0.572	0.544	0.678	0.243	0.727	**0.833**	**0.916**
	Color saturation 2	0.602	0.375	0.677	0.120	**0.841**	0.836	**0.909**
	Color quantization	0.670	0.667	0.676	0.323	0.662	**0.806**	**0.853**
	Color shift	-0.139	-0.182	0.090	-0.002	0.050	**0.421**	**0.777**
	Color saturation 1	0.091	0.071	0.027	-0.019	**0.216**	0.148	**0.604**
Compression	JPEG compression	0.414	0.782	**0.804** ↑6.2%	0.556	0.582	0.530	**0.866**
	JPEG 2000 compression	0.655	0.516	**0.790** ↑6.3%	0.342	0.608	0.539	**0.853**
Noise	Denoise	0.457	0.221	**0.856** ↑9.7%	0.229	0.247	0.765	**0.953**
	White noise in color component	0.757	0.718	0.841	0.418	0.745 ↑1.1%	**0.925**	**0.936**
	Multiplicative noise	0.702	0.674	0.682	0.306	0.776 ↑5.0%	**0.884**	**0.934**
	Impulse noise	0.547	-0.543	0.808	0.219	0.254 ↑10.2%	**0.814**	**0.916**
	White Gaussian noise	0.628	0.708	0.776	0.357	0.680 ↑1.7%	**0.897**	**0.914**
Brightness change	Brighten	0.458	0.575	0.301	0.227	**0.753**	0.685	**0.822**
	Darken	0.439	0.405	0.436	0.206	**0.744**	0.272	**0.647**
	Mean Shift	0.112	0.144	0.315	0.122	**0.591**	**0.348**	0.335
Spatial distortions	Jitter	0.629	0.672	0.441	0.719	0.391	**0.778**	**0.899**
	Pixelate	0.196	0.648	0.577	0.587	**0.702**	0.700	**0.814**
	Quantization	**0.781**	0.714	0.571	0.259	0.681	0.735	**0.791**
	Color block	-0.020	0.067	0.003	0.094	**0.388**	0.160	**0.440**
	Non-eccentricity patch	0.083	0.191	0.218	0.121	**0.461**	0.348	**0.433**
Sharpness and contrast	High sharpen	-0.015	0.361	**0.681**	0.114	0.230	0.558	**0.932**
	Contrast change	0.062	0.105	0.072	0.125	**0.452**	0.421	**0.513**



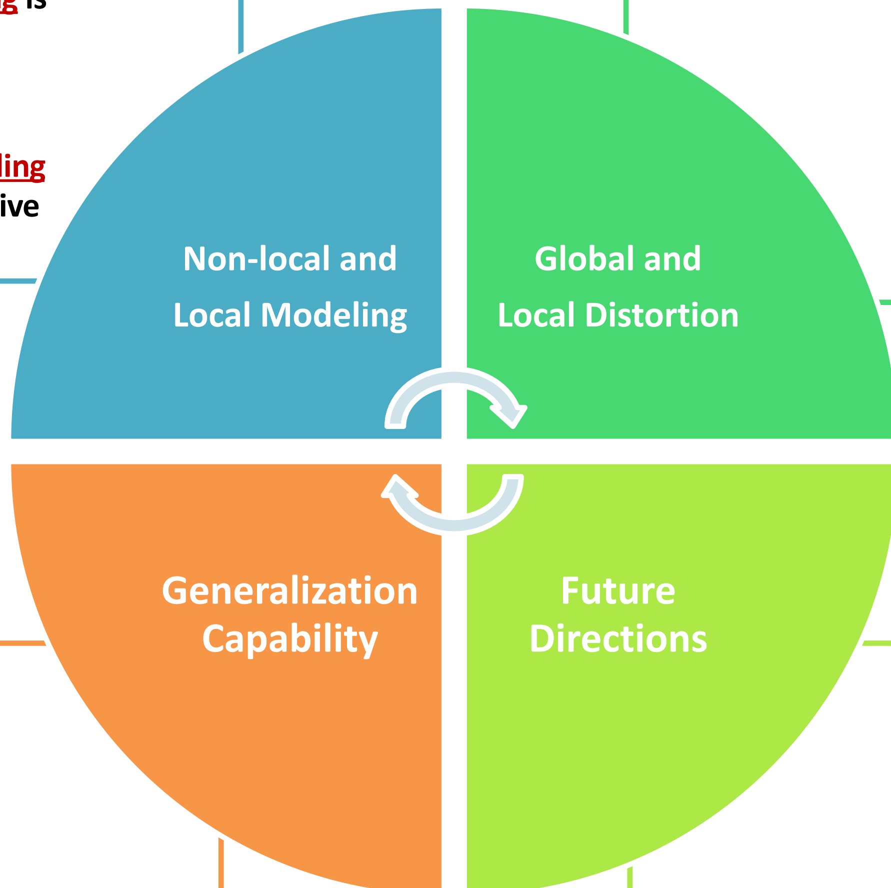
Demonstrations of the local distortions (**a/d/g**: reference image, **b/e/h**: non-eccentricity patch, **c/f/i**: color block)

Image Credit: KADID-10k Database

Part 5: Conclusions and Future Directions

- **Non-local Modeling is complementary** to traditional local methods
- CNNs' **Local Modeling features are effective and robust**

- Handle **Global Distortions**: globally and uniformly distributed with non-local recurrences
- Maintain sensitivity to **Local Distortions**: local nonuniform-distributed
- Better assess **Noisy and Compressed Images** Quality



- Cross-Dataset Setting → **High Generalization Capability**

- **Non-local Statistics** [1, 2]
- Non-local Modeling for Quality Assessment of **3D Point Clouds** [3]
- Quality Assessment of **AIGC** [4]

Credit:
[1] Zontak *et al.*, Internal Statistics of a Single Natural Image, In CVPR 2011
[2] Buades *et al.*, A Non-local Algorithm for Image Denoising, In CVPR 2005
[3] Zhou *et al.*, Blind Quality Assessment of Dense 3D Point Clouds with Structure Guided Resampling, Under Review In IEEE TCSVT'23
[4] Zhang *et al.*, A Perceptual Quality Assessment Exploration for AIGC Images, In arXiv

專業 創新 胸懷全球

Professional·Creative

For The World

Thank you very much for your Attention!

香港城市大學

City University of Hong Kong