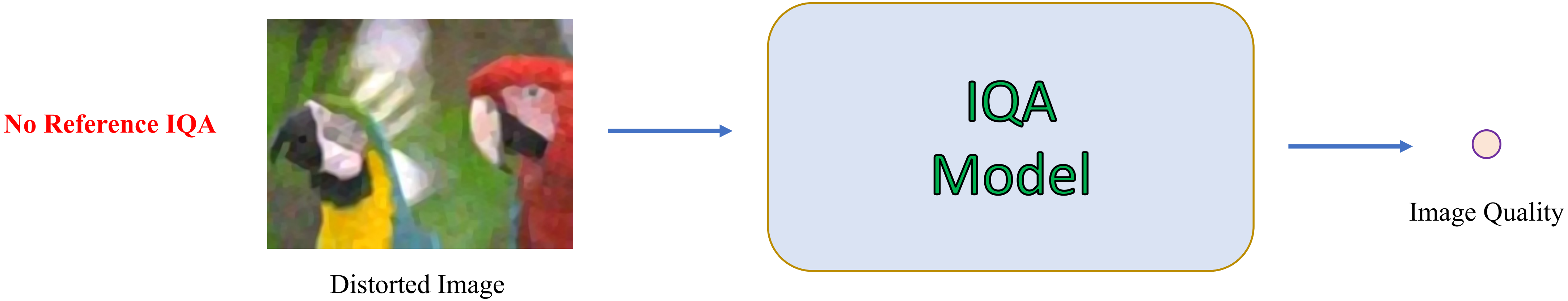
A Brief Summary of Three Selected Projects

Tasks

- ✓ Image Quality Assessment (IQA) **Regression**



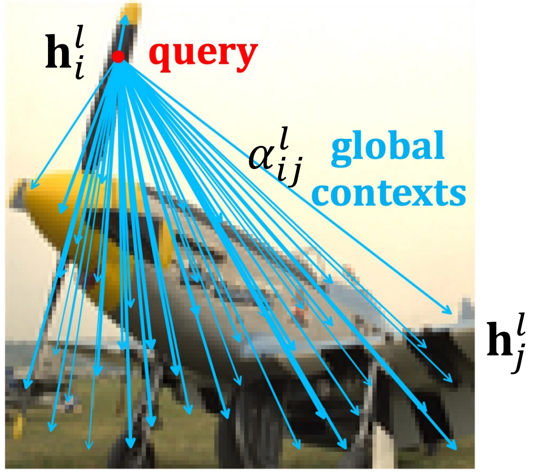
No-reference Image Quality Assessment Framework

A Brief Summary of Three Selected Projects

IQA Research Novelty

- ✓ [**Definition**] Non-local Modeling and Local Modeling
- ✓ [**Definition**] Global Distortions and Local Distortions
- ✓ [**Motivation**] Human Visual System (HVS) perceives Image Quality:
  Adaptive to local content + Long-range dependency constructed among different regions
- ✓ [**Method**] Superpixel-based Graph Neural Network to explore Non-local Interactions

Local Modeling



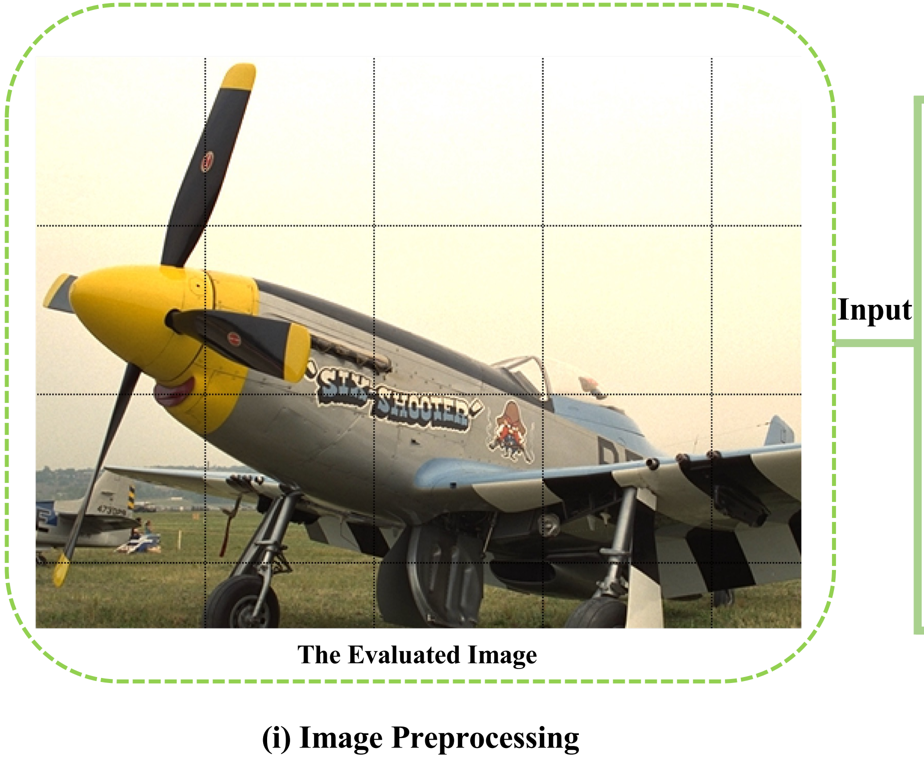
Non-local Modeling

Global Distortions

Local Distortions

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

NLNet Architecture



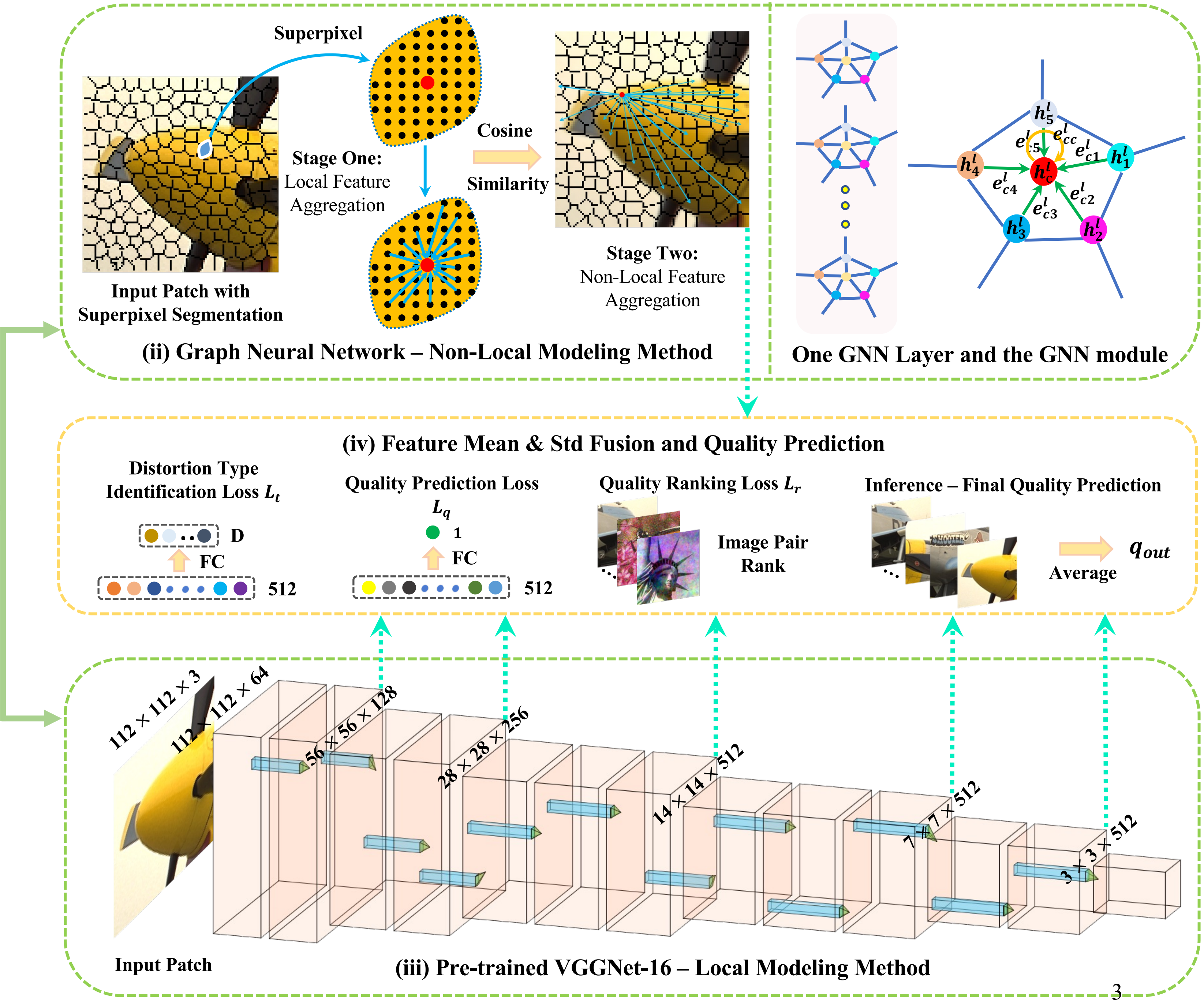


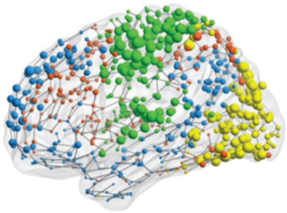
Image Credit: TID2013 Database

A Brief Summary of Three Selected Projects

Tasks

✓ Electroencephalogram (EEG) Tasks **Classification**

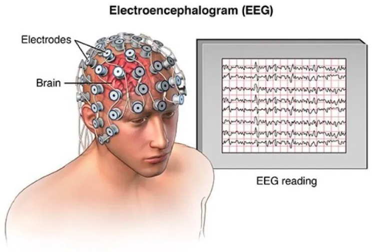
Control a wheelchair via EEG

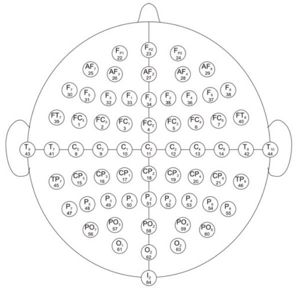


Functional Networks

mapping

interpret

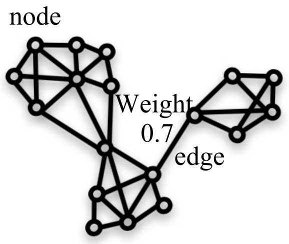




International 10-10 EEG System

topology

model



Graph

Interpret Functional Networks and better understand human brain

Image Credit: in the public domain.

A **Brief Summary** of Three Selected Projects

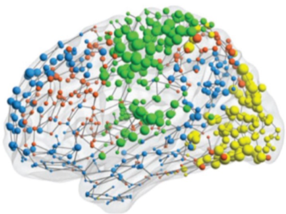
EEG Research Novelty

- ✓ [**Motivation**] Graph Modeling for EEG Electrodes System
- ✓ [**Method**] Graph Representation Learning of EEG Signals
- ✓ [**Motivation**] Spatial-Temporal Analysis of EEG Signals
- ✓ [**Method**] Deep Feature Mining of EEG Signals

A Brief Summary of Three Selected Projects

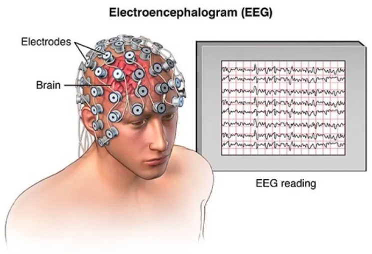
EEG Research Novelty

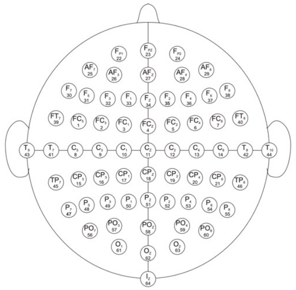
- ✓ [**Motivation**] Graph Modeling for EEG Electrodes System
- ✓ [**Method**] Graph Representation Learning of EEG Signals
- ✓ [**Motivation**] Spatial-Temporal Analysis of EEG Signals
- ✓ [**Method**] Deep Feature Mining of EEG Signals



Functional Networks

mapping



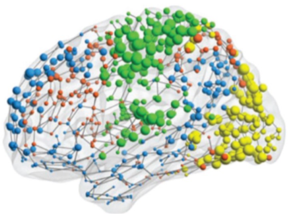


International 10-10 EEG System

A Brief Summary of Three Selected Projects

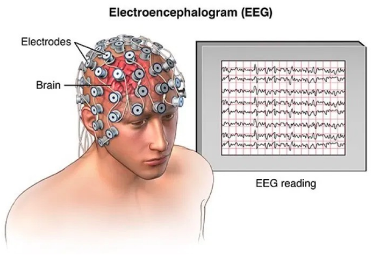
EEG Research Novelty

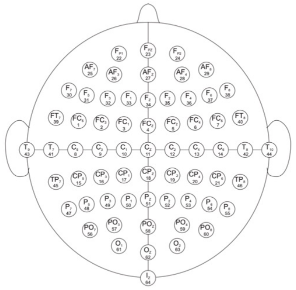
- ✓ [**Motivation**] Graph Modeling for EEG Electrodes System
- ✓ [**Method**] Graph Representation Learning of EEG Signals
- ✓ [**Motivation**] Spatial-Temporal Analysis of EEG Signals
- ✓ [**Method**] Deep Feature Mining of EEG Signals



Functional Networks

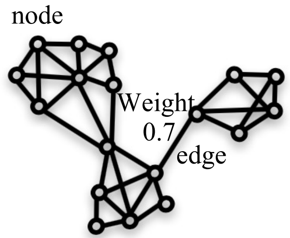
mapping





International 10-10 EEG System

topology

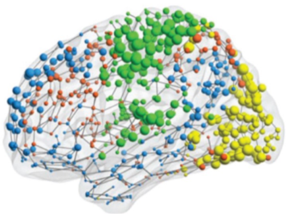


Graph

A Brief Summary of Three Selected Projects

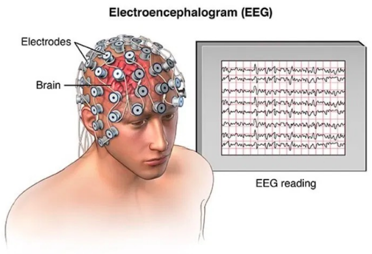
EEG Research Novelty

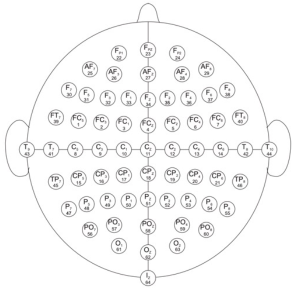
- ✓ [**Motivation**] Graph Modeling for EEG Electrodes System
- ✓ [**Method**] Graph Representation Learning of EEG Signals
- ✓ [**Motivation**] Spatial-Temporal Analysis of EEG Signals
- ✓ [**Method**] Deep Feature Mining of EEG Signals



Functional Networks

mapping

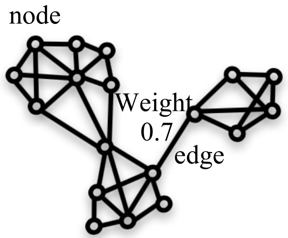




International 10-10 EEG System

topology

model

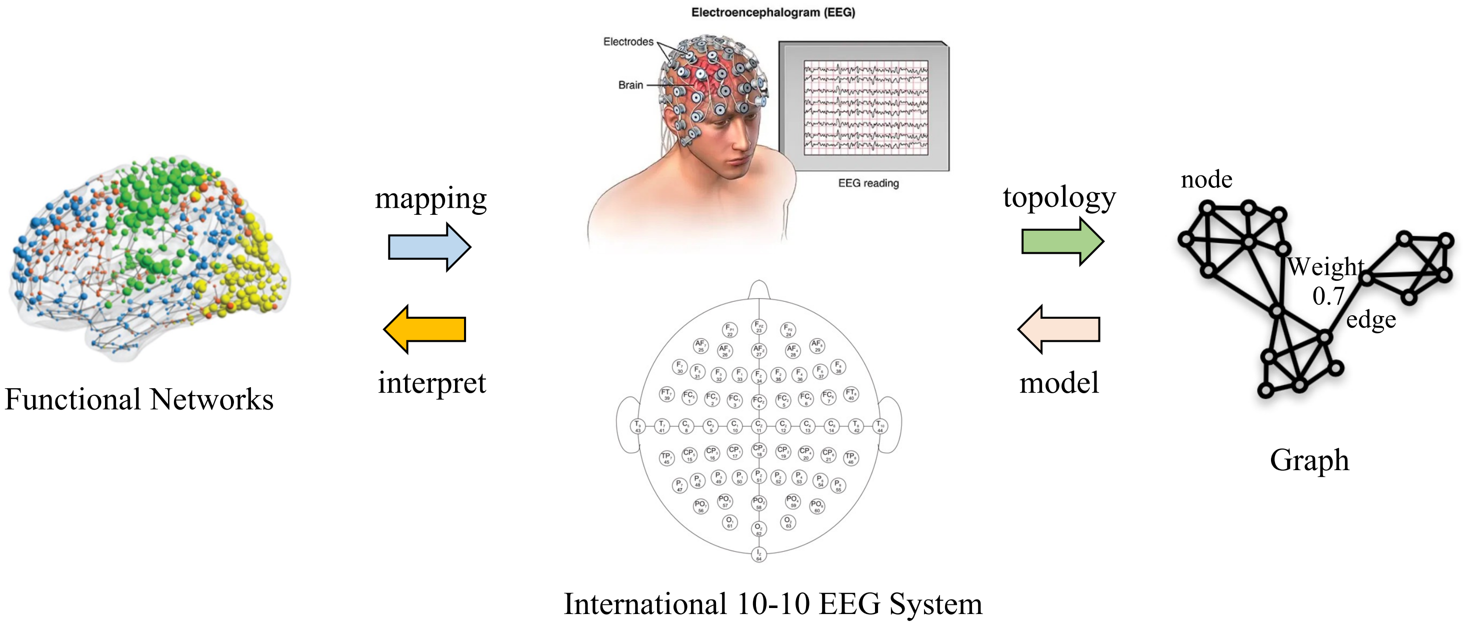


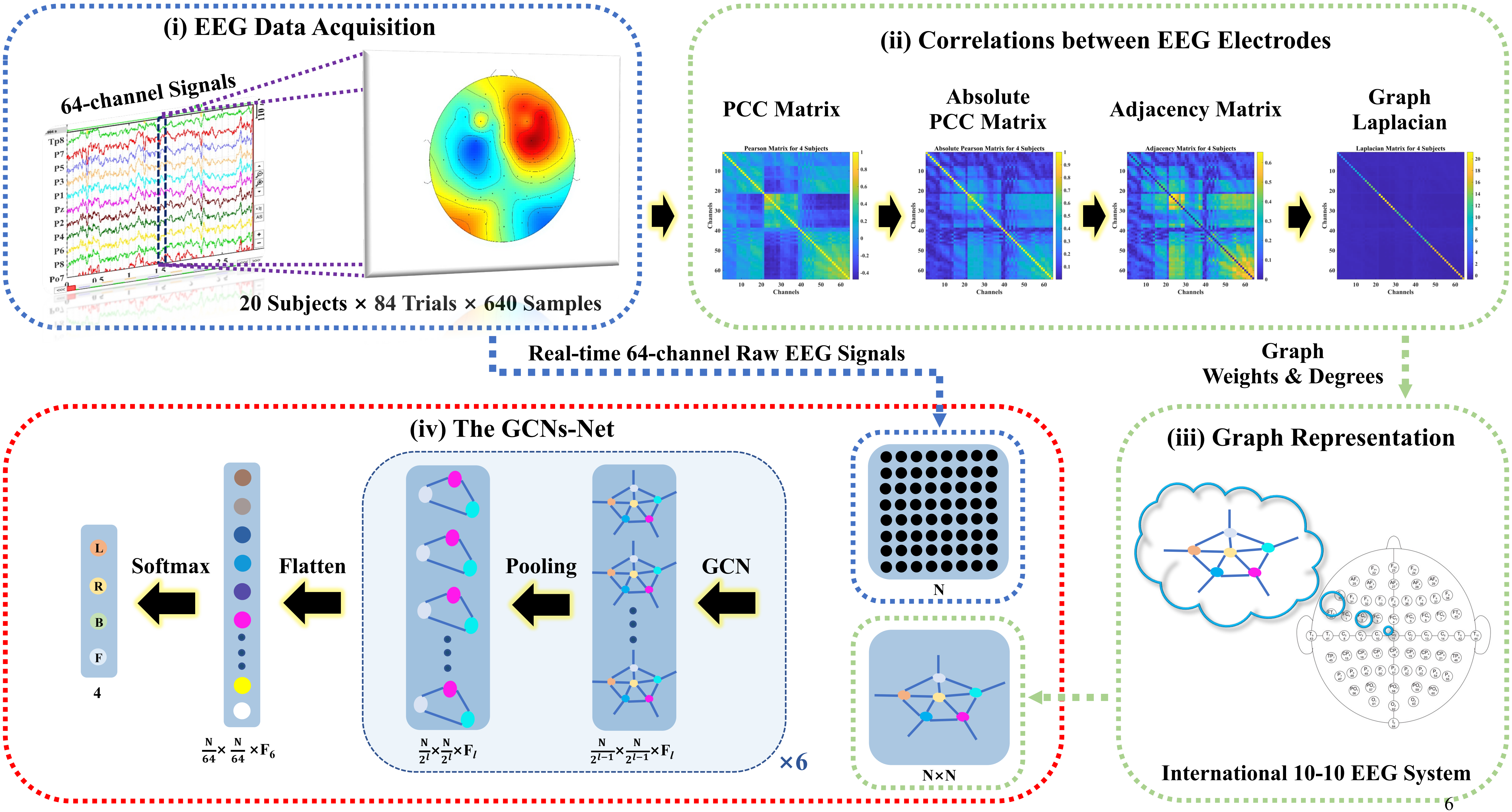
Graph

A Brief Summary of Three Selected Projects

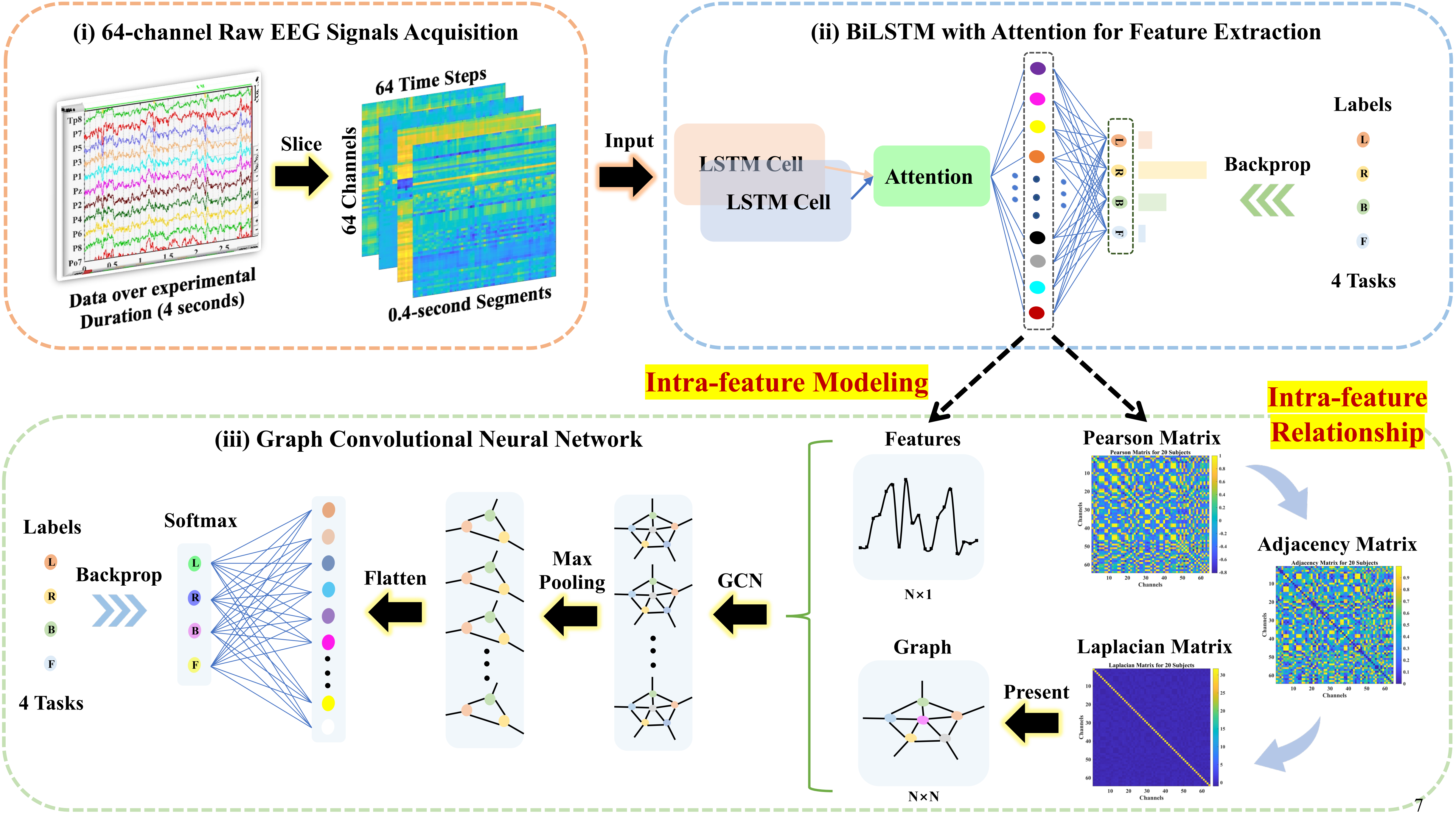
EEG Research Novelty

- ✓ [**Motivation**] Graph Modeling for EEG Electrodes System
- ✓ [**Method**] Graph Representation Learning of EEG Signals
- ✓ [**Motivation**] Spatial-Temporal Analysis of EEG Signals
- ✓ [**Method**] Deep Feature Mining of EEG Signals



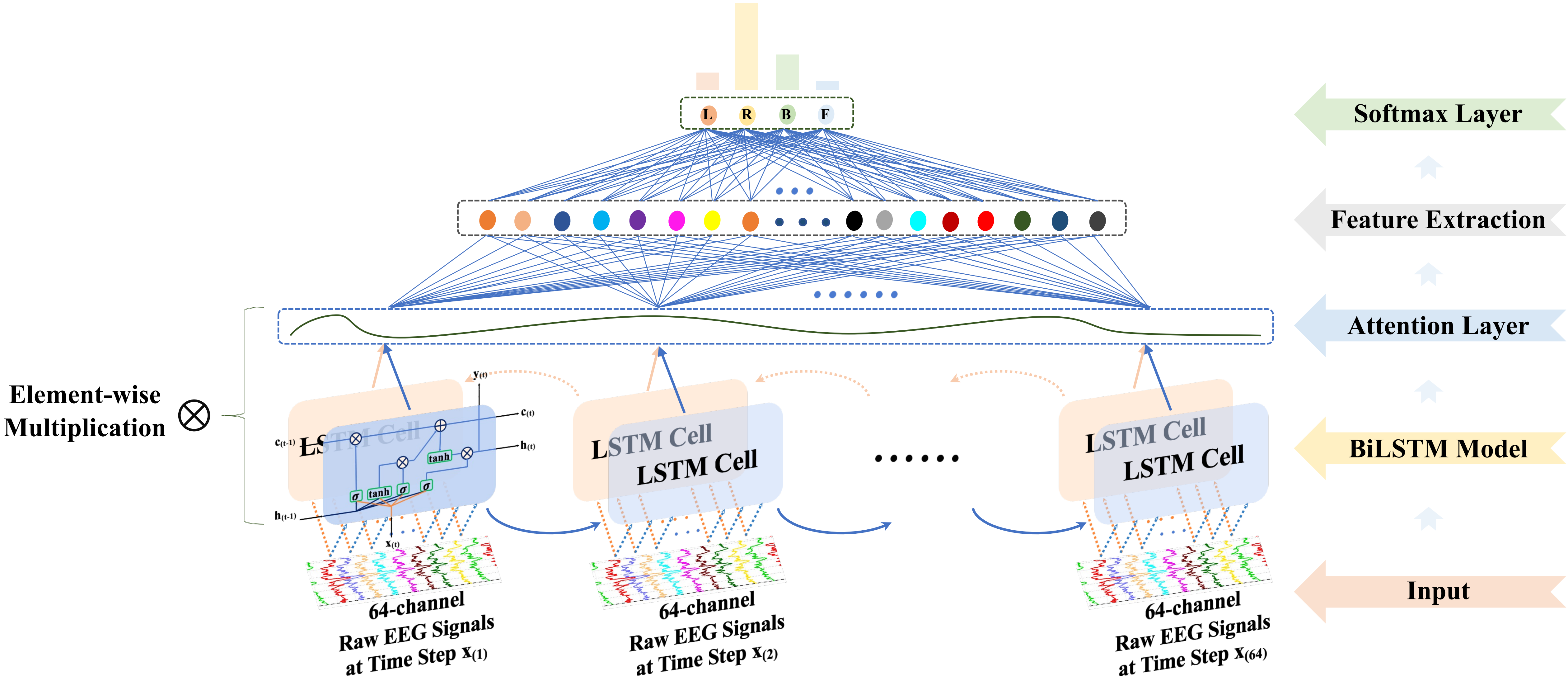


(i) EEG Data Acquisition
64-channel Signals
20 Subjects × 84 Trials × 640 Samples
(ii) Correlations between EEG Electrodes
PCC Matrix
Absolute PCC Matrix
Adjacency Matrix
Graph Laplacian
Pearson Matrix for 4 Subjects
Absolute Pearson Matrix for 4 Subjects
Adjacency Matrix for 4 Subjects
Laplacian Matrix for 4 Subjects
Channels
Real-time 64-channel Raw EEG Signals
Graph Weights & Degrees
(iii) Graph Representation
International 10-10 EEG System
(iv) The GCNs-Net
Softmax
Flatten
Pooling
GCN
N
N×N
×6
4
L
R
B
F



(i) 64-channel Raw EEG Signals Acquisition
Slice
64 Time Steps
64 Channels
Data over experimental Duration (4 seconds)
0.4-second Segments
Input
(ii) BiLSTM with Attention for Feature Extraction
LSTM Cell
LSTM Cell
Attention
Labels
L
R
B
F
Backprop
4 Tasks
Intra-feature Modeling
Intra-feature Relationship
(iii) Graph Convolutional Neural Network
Features
N×1
Pearson Matrix
Adjacency Matrix
Graph
N×N
Present
Laplacian Matrix
GCN
Max Pooling
Flatten
Softmax
Labels
Backprop
4 Tasks

Attention-based Bidirectional Long Short-term Memory (Bi-LSTM)



A Brief Summary of Three Selected Projects

Tasks

- ✓ **Regression** (IQA) and **Classification** (EEG)

IQA Research Novelty

- ✓ [**Definition**] Non-local Modeling and Local Modeling
- ✓ [**Definition**] Global Distortions and Local Distortions
- ✓ [**Motivation**] Human Visual System (HVS) perceives Image Quality:
  Adaptive to local content + Long-range Dependency constructed among different regions
- ✓ [**Method**] Superpixel-based Graph Neural Network to explore Non-local Interactions

EEG Research Novelty

- ✓ [**Motivation**] Graph Modeling for EEG Electrodes System
- ✓ [**Method**] Graph Representation Learning of EEG Signals
- ✓ [**Motivation**] Spatial-Temporal Analysis of EEG Signals
- ✓ [**Method**] Deep Feature Mining of EEG Signals

No-reference Image Quality Assessment via Non-local Modeling

GCNs-Net: A Graph Convolutional Neural Network Approach for Decoding Time-Resolved EEG Motor Imagery Signals

Deep Feature Mining via Attention-based BiLSTM-GCN for Human Motor Imagery Recognition

Selected Research Projects

Shuyue Jia

January 10th, 2023

https://github.com/SuperBruceJia

No-reference Image Quality Assessment via Non-local Modeling

Shuyue Jia [1], Baoliang Chen [1], Dingquan Li [2], and Shiqi Wang [1] *

[1] Department of Computer Science, City University of Hong Kong

[2] Peng Cheng Laboratory

Project: https://github.com/SuperBruceJia/NLNet-IQA

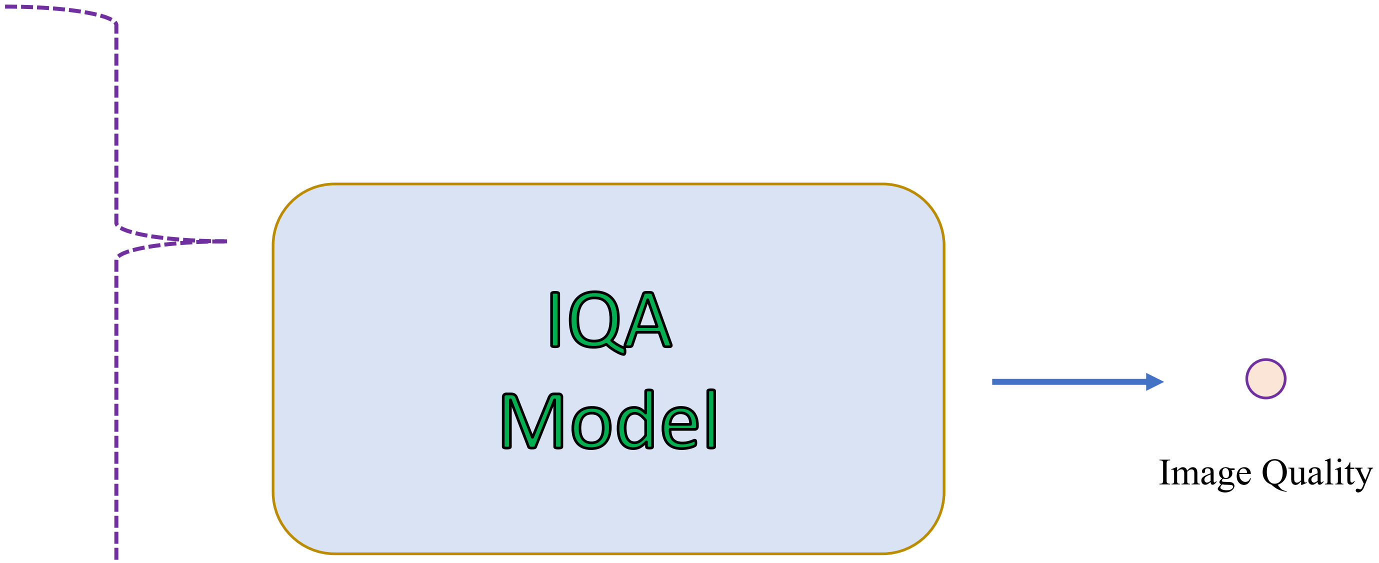
Image Quality Assessment (IQA)

Image Credit: TID2013 Database

Image Quality Assessment (IQA)

Full Reference IQA

Reference Image



Distorted Image

Image Credit: TID2013 Database

Image Quality Assessment (IQA)

IQA
Model

Image Quality

Distorted Image

Image Credit: TID2013 Database

Image Quality Assessment (IQA)

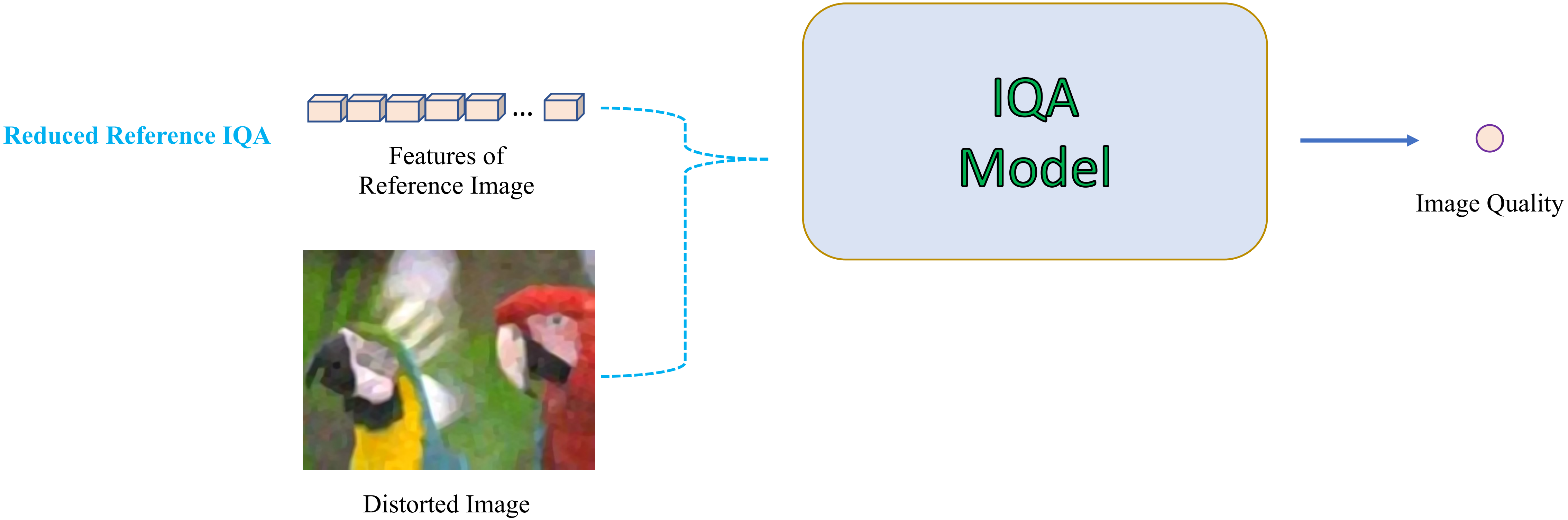


Image Credit: TID2013 Database

Image Quality Assessment (IQA)

IQA
Model

Image Quality

Distorted Image

Image Credit: TID2013 Database

Image Quality Assessment (IQA)

IQA
Model

Image Quality

No Reference IQA

Distorted Image

Image Credit: TID2013 Database

Image Quality Assessment (IQA)

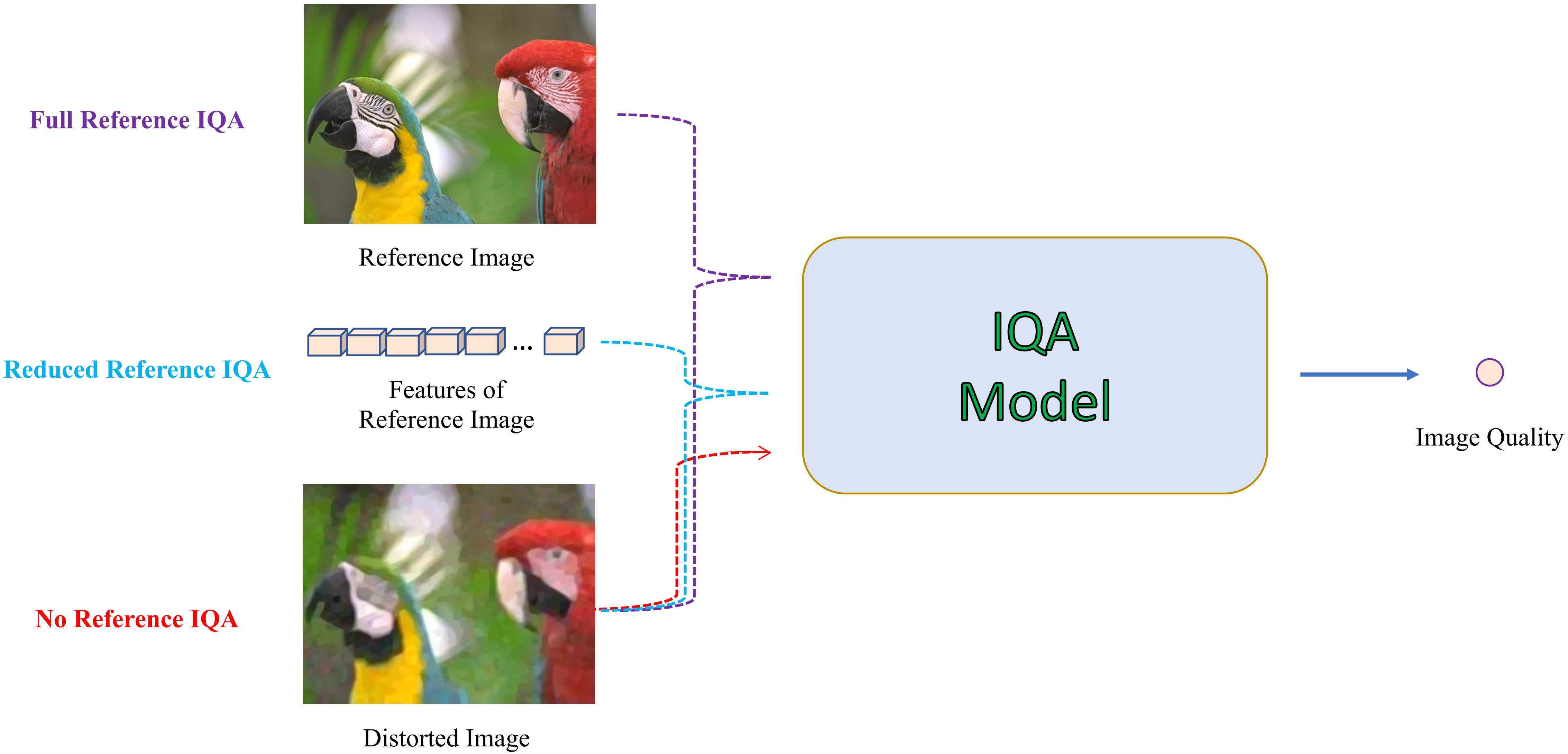
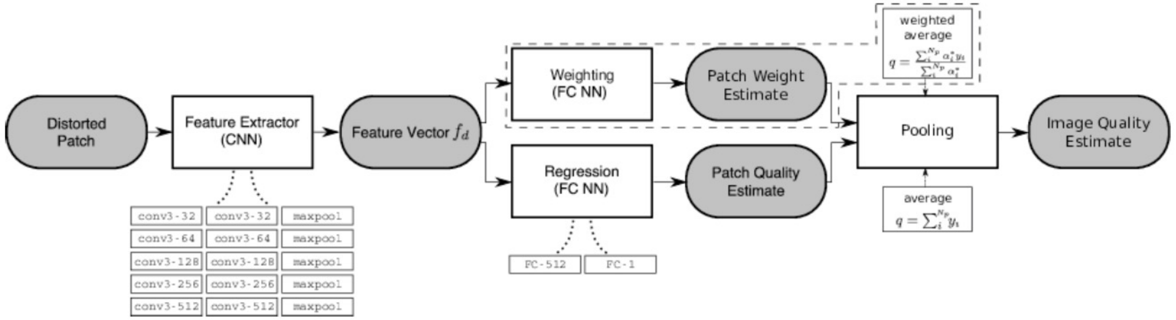
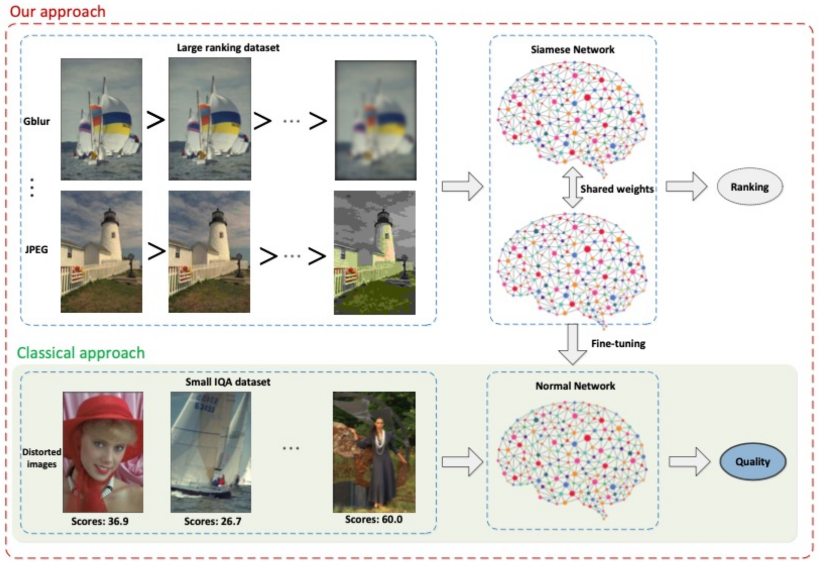


Image Credit: TID2013 Database

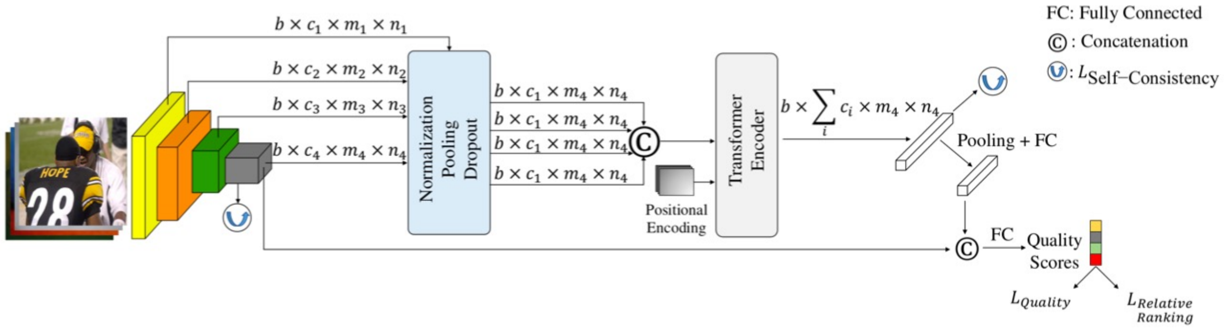
Recent Progress on No-reference IQA



CNN-based Methods [1]



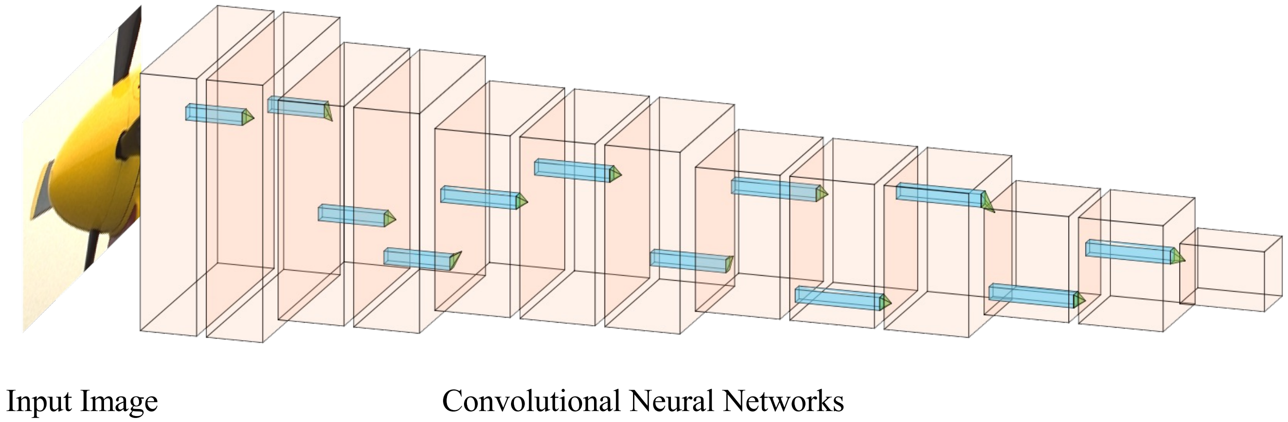
Ranking-based Methods [2]



Transformer-based Methods [3]

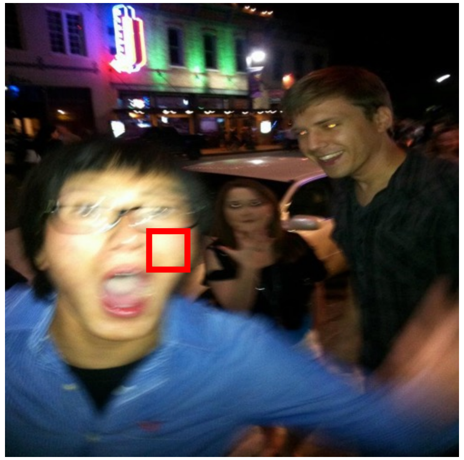
Credit:
[1] Bosse *et al.*, Deep Neural Networks for No-Reference and Full-Reference Image Quality Assessment, In TIP 2018
[2] Liu *et al.*, RankIQA: Learning from Rankings for No-reference Image Quality Assessment, In ICCV 2017
[3] Golestaneh *et al.*, No-Reference Image Quality Assessment via Transformers, Relative Ranking, and Self-Consistency, In WACV 2022

Challenges

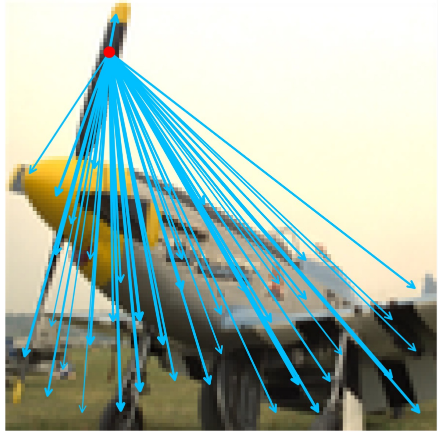


- **Local Modeling** (Convolutional Neural Networks):
  - ✓ Translation Invariance (Pooling)
  - ✓ Translation Equivalence (Convolution)
  - ✓ Sharable Parameters (Weight Sharing)
- **Limitations**:
  - ✓ Small-sized Receptive Field → **Extracted features are too local**
  - ✓ Parameters Fixed across the whole image → **Image content is equally treated**
  - ✓ Lack of Geometric and Relational Modeling → **Missing complex relations and dependencies**

Motivation



Local Feature Extraction



Non-local Dependency

- ✓ HVS is **adaptive to the local content**

  → ***Local feature extraction*** via a pre-trained CNN

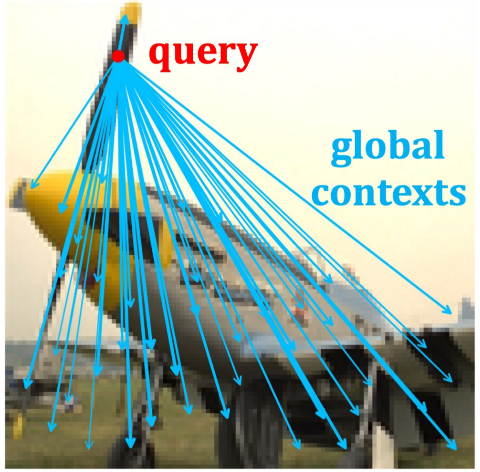
- ✓ HVS perceives image quality with **long-range dependency constructed among different regions**

  → ***Non-local feature extraction*** for long-range dependency and relational modeling

Image Credit: LIVEC and TID2013 Databases

Definition

(a)
Local feature extraction is critical

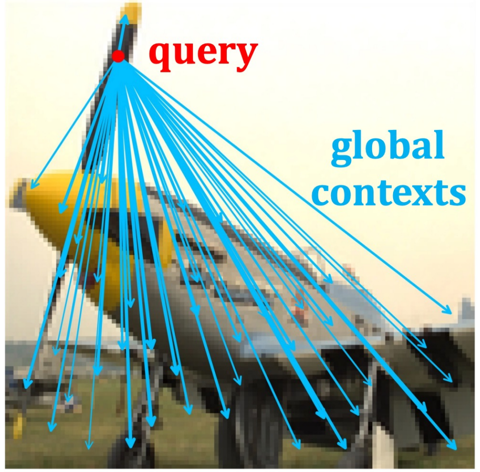


(b)
Non-local dependency
learned by the NLNet

Figure 2: Local region feature extraction and non-local dependency feature extraction

Image Credit: TID2013 and LIVEC Databases

Definition



Convolution:
Pixel-to-Pixel Modeling

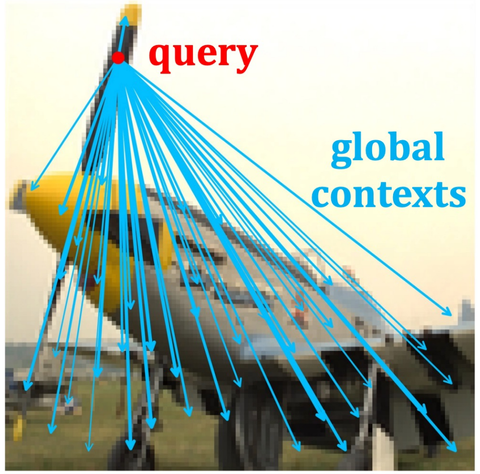
(a)
Local feature extraction is critical

(b)
Non-local dependency learned by the NLNet

Figure 2: Local region feature extraction and non-local dependency feature extraction

Image Credit: TID2013 and LIVEC Databases

Definition



Convolution:
Pixel-to-Pixel Modeling

(a)
Local feature extraction is critical

(b)
Non-local dependency learned by the NLNet

Figure 2: 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

Definition

Non-Local:
Object-to-Pixel
Modeling

Convolution:
Pixel-to-Pixel
Modeling

(a)
Local feature extraction is critical

(b)
Non-local dependency
learned by the NLNet

Figure 2: 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**.

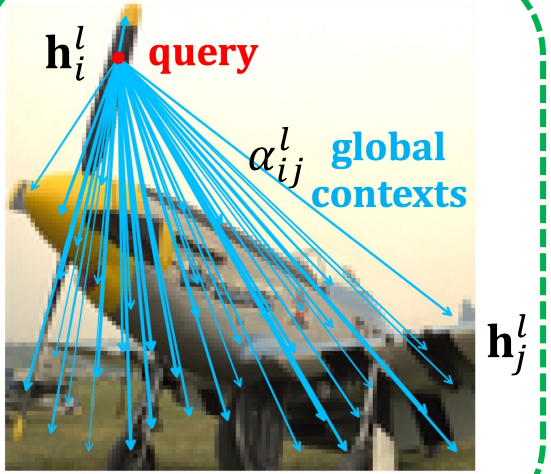
Image Credit: TID2013 and LIVEC Databases

Definition

Non-Local:
Object-to-Pixel Modeling

Convolution:
Pixel-to-Pixel Modeling

(a)
Local feature extraction is critical



(b)
Non-local dependency learned by the NLNet

Figure 2: Local region feature extraction and non-local dependency feature extraction

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

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

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

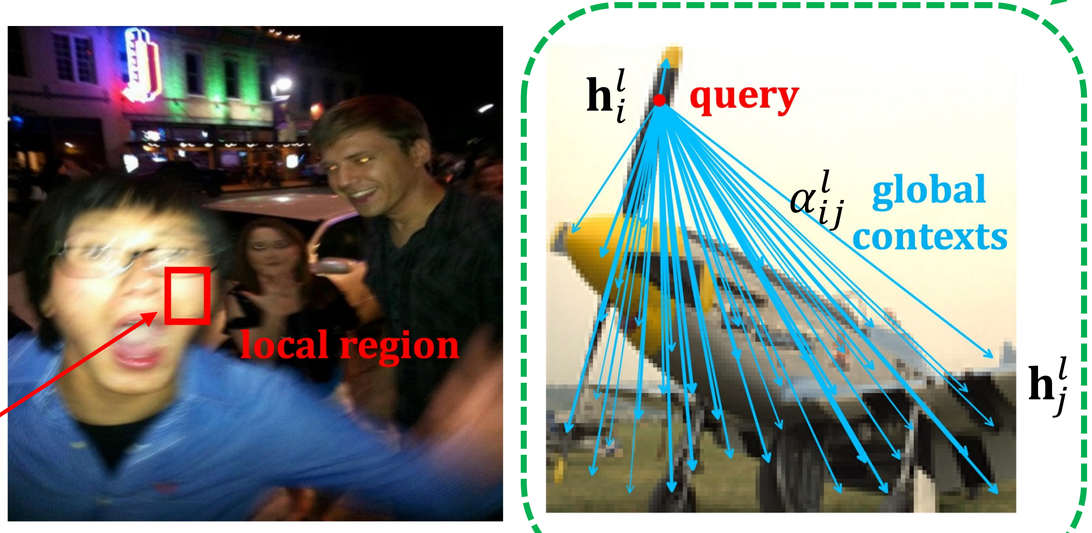
- ✓ **Local Modeling**: encodes spatially proximate **Local Neighborhoods**.
- ✓ **Non-local Modeling**: establishes **Spatial Integration of Information** by **Long- and Short-Range Communications** with **different Spatial Weighting Functions**.

Image Credit: TID2013 and LIVEC Databases

Definition

Non-Local:
Object-to-Pixel Modeling

Beauty is in Simplicity



Convolution:
Pixel-to-Pixel Modeling

(a)
Local feature extraction is critical

(b)
Non-local dependency learned by the NLNet

Figure 2: Local region feature extraction and non-local dependency feature extraction

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

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

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

- ✓ **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

Non-local Behavior

Object-to-Pixel Modeling
Region Feature Extraction

↑

Non-local
Dependency & Relational Modeling

↓

Semantics and Content Understanding

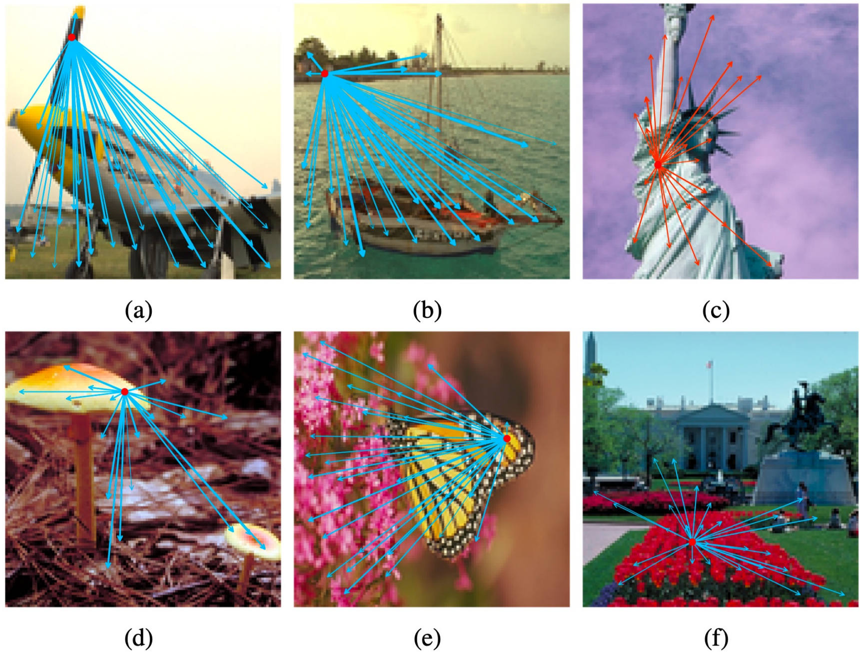


Figure 3.1: The non-local behavior of the long-range dependency and relational modeling. (a) The plane image with a query on wings. (b) The boat image with a query on nearby river bank. (c) The Statue of Liberty image with a query on the lady. (d) The shrooms image with a query on one shroom. (e) The butterfly image with a query on the wing. (f) The Lafayette Square, Washington, D.C. image with a query on flowers.

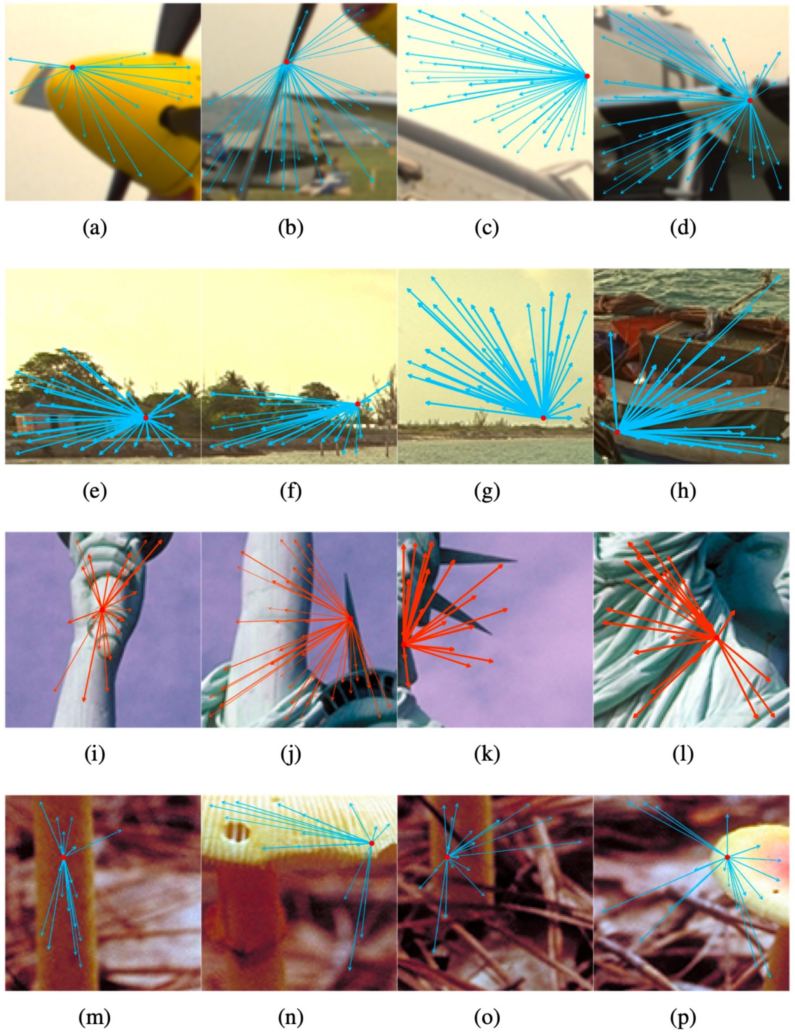


Figure 3.2: Selected demonstrations of the non-local behavior and long-range dependencies with regard to the cropped image patches from the illustrated images. The details of Figure (a) to (p) are described in the thesis.

✓ **Non-local Modeling**: establishes the **Spatial Integration of Information** by **Long- and Short-Range Communications** with **different Spatial Weighting Functions**.

Image Credit: TID2013 and LIVE Databases

Definition

Image Credit: CSIQ and KADID-10k Databases

Definition

Non-Local Recurrence

Figure 4.9: Demonstrations of the global distortions (b/f: GB, c/g: CC, d/h: PN) contaminating the Statue of Liberty and George Rogers Clark Memorial images. Figure (a) and Figure (e) are reference images from the CSIQ database.

Global Distortion

Image Credit: CSIQ and KADID-10k Databases

Definition

Non-Local Recurrence

Figure 4.9: Demonstrations of the global distortions (b/f: GB, c/g: CC, d/h: PN) contaminating the Statue of Liberty and George Rogers Clark Memorial images. Figure (a) and Figure (e) are reference images from the CSIQ database.

Global Distortion

Local Distortion

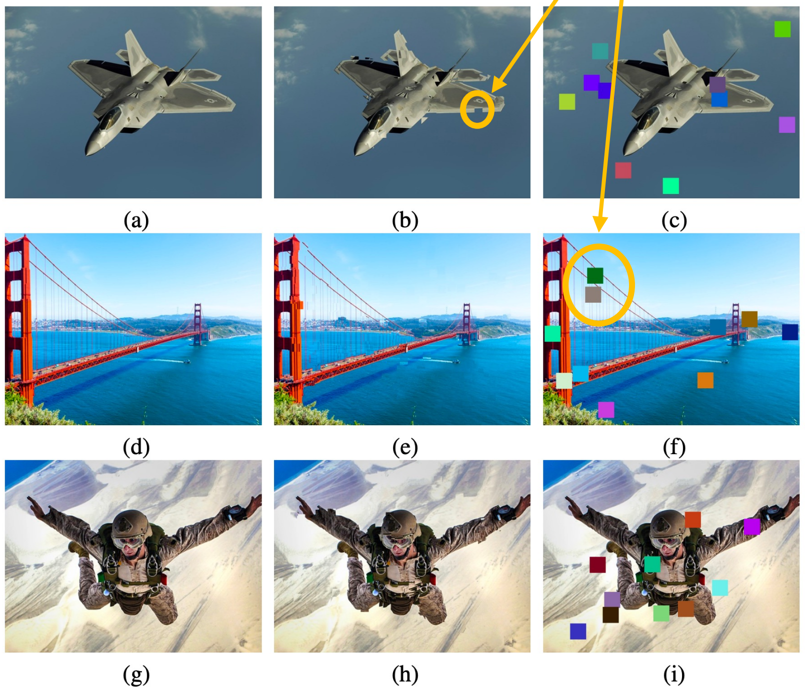


Figure 4.11: Demonstrations of the local distortions (b/e/h: non-eccentricity patch and c/f/i: color block). Figure (a), Figure (d), and Figure (g) are reference images from the KADID-10k database.

Local Distortion

Image Credit: CSIQ and KADID-10k Databases

Definition

Local Distortion

Non-Local Recurrence

Figure 4.9: Demonstrations of the global distortions (b/f: GB, c/g: CC, d/h: PN) contaminating the Statue of Liberty and George Rogers Clark Memorial images. Figure (a) and Figure (e) are reference images from the CSIQ database.

Global Distortion

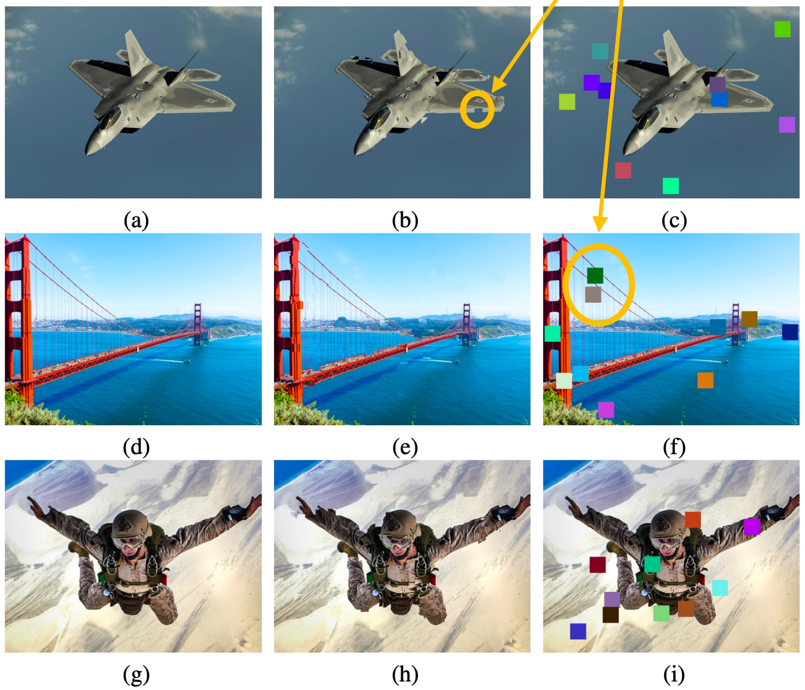


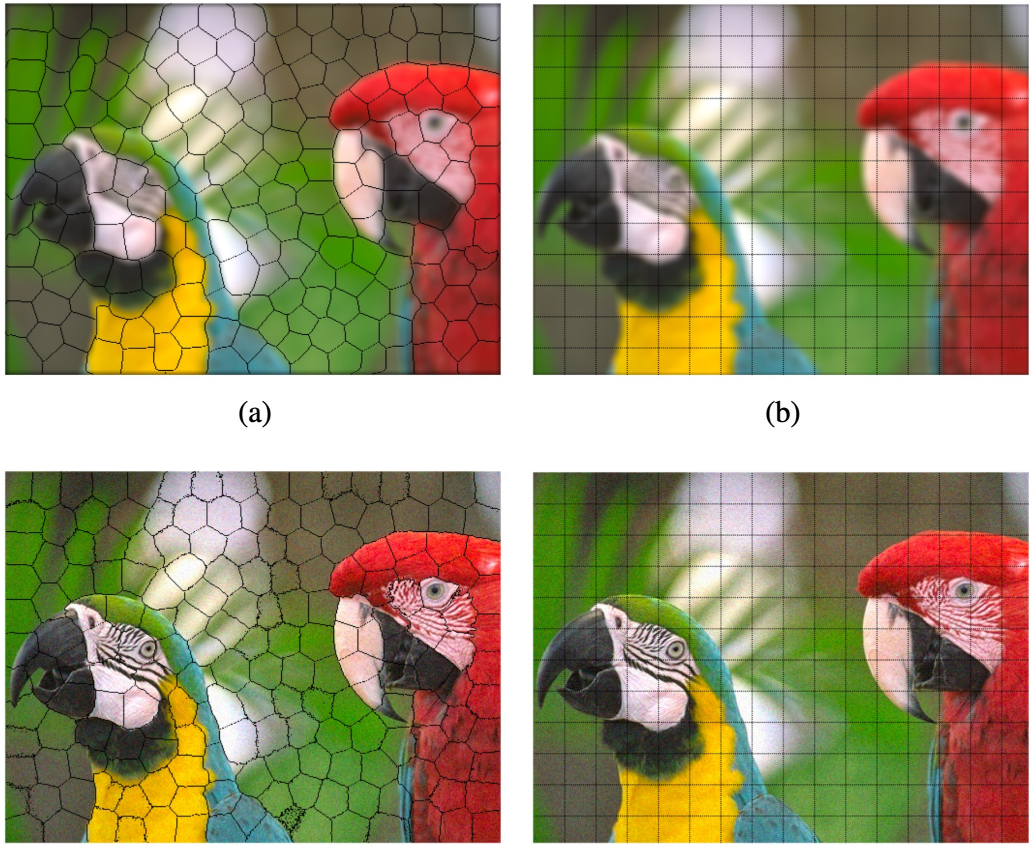
Figure 4.11: Demonstrations of the local distortions (b/e/h: non-eccentricity patch and c/f/i: color block). Figure (a), Figure (d), and Figure (g) are reference images from the KADID-10k database.

Local Distortion

- ✓ **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

Superpixel Segmentation



(a) (b) (c) (d)

Figure 4.2: The superpixel *vs.* square patch representation (with size of $\approx 32 \times 32$) of the plane image from the TID2013 database.

Image Credit: TID2013 Database

Superpixel *vs.* **Square Patch**

- ✓ Adherence to boundaries and **visually meaningful**
- ✓ **Accurate feature extraction**

Superpixel Segmentation

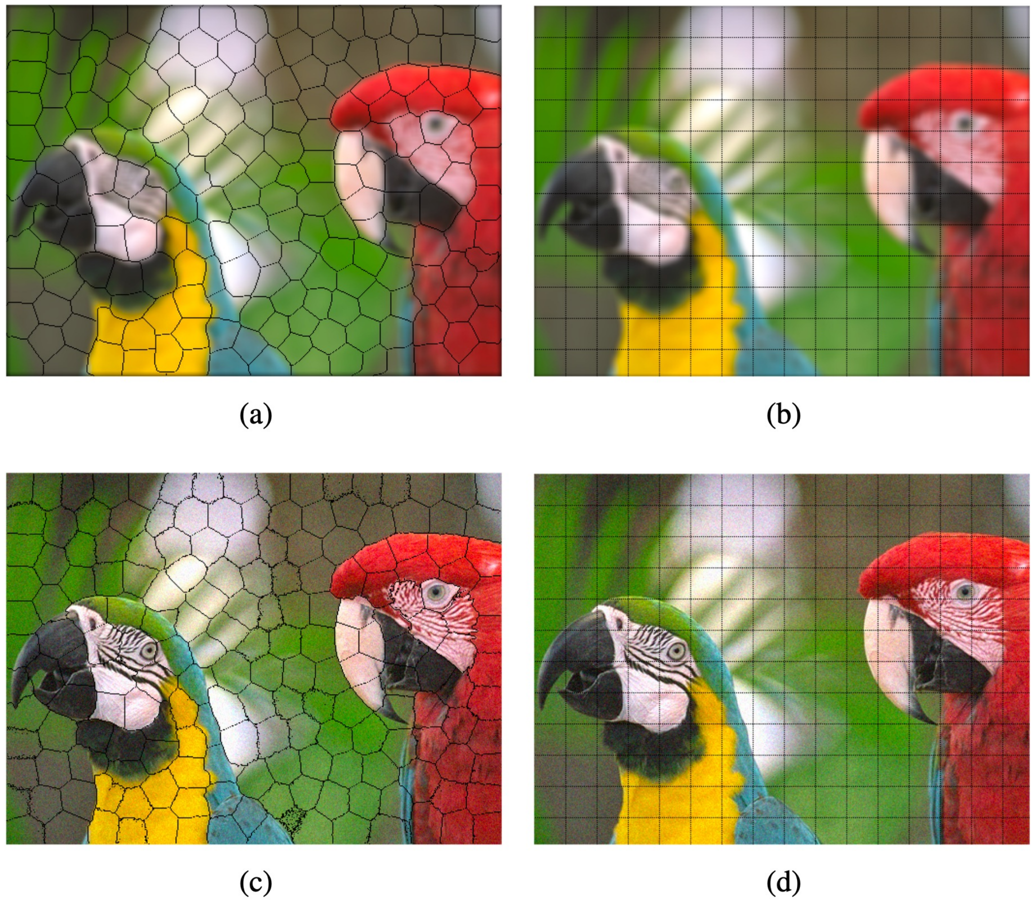


Figure 4.2: The superpixel *vs.* square patch representation (with size of $\approx 32 \times 32$) of the plane image from the TID2013 database.

Image Credit: TID2013 Database

Superpixel Segmentation

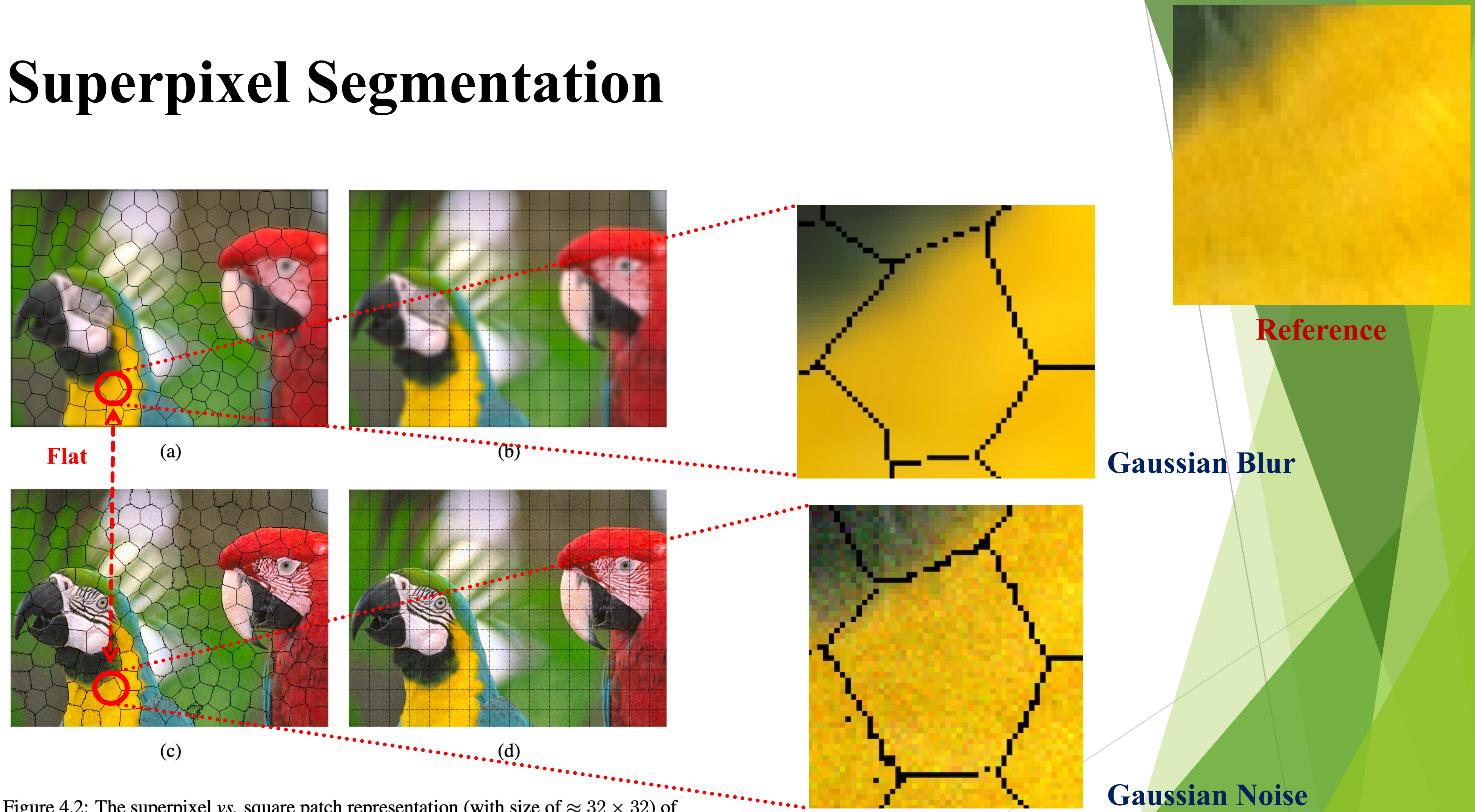


Figure 4.2: The superpixel *vs.* square patch representation (with size of $\approx 32 \times 32$) of the plane image from the TID2013 database.

Image Credit: TID2013 Database

Superpixel Segmentation

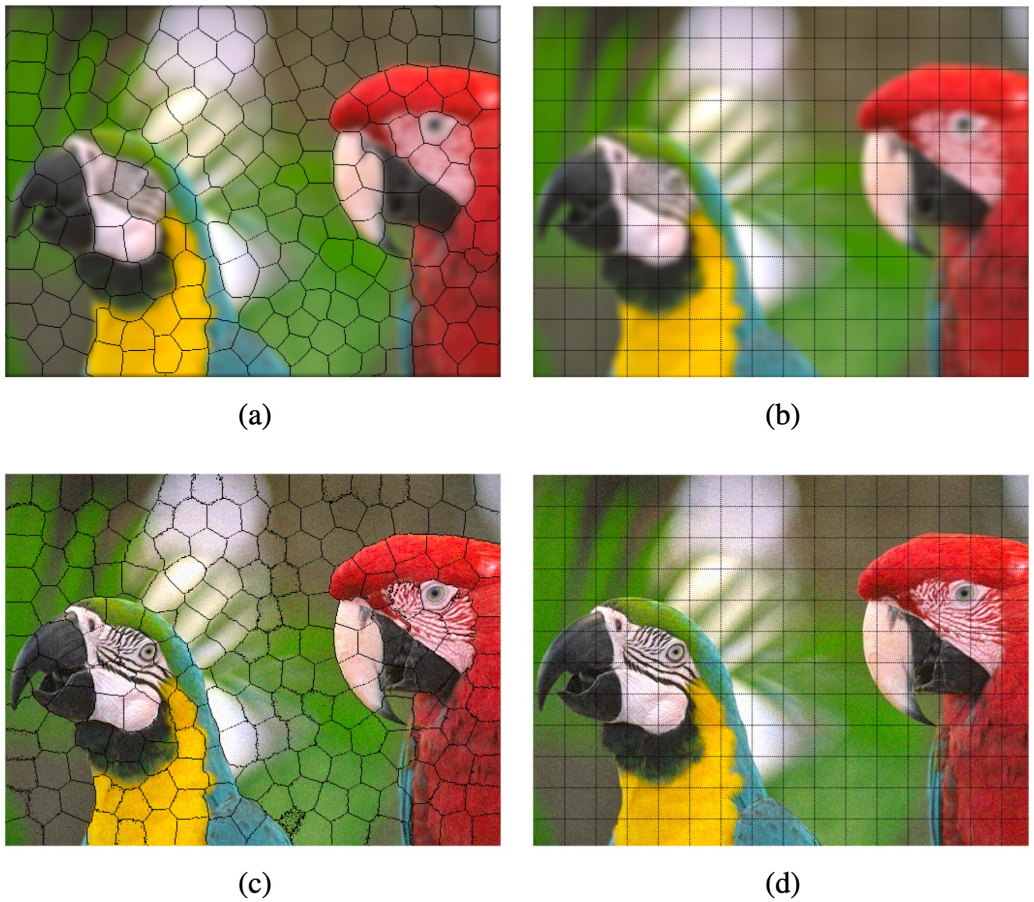


Figure 4.2: The superpixel *vs.* square patch representation (with size of $\approx 32 \times 32$) of the plane image from the TID2013 database.

Image Credit: TID2013 Database

Superpixel Segmentation

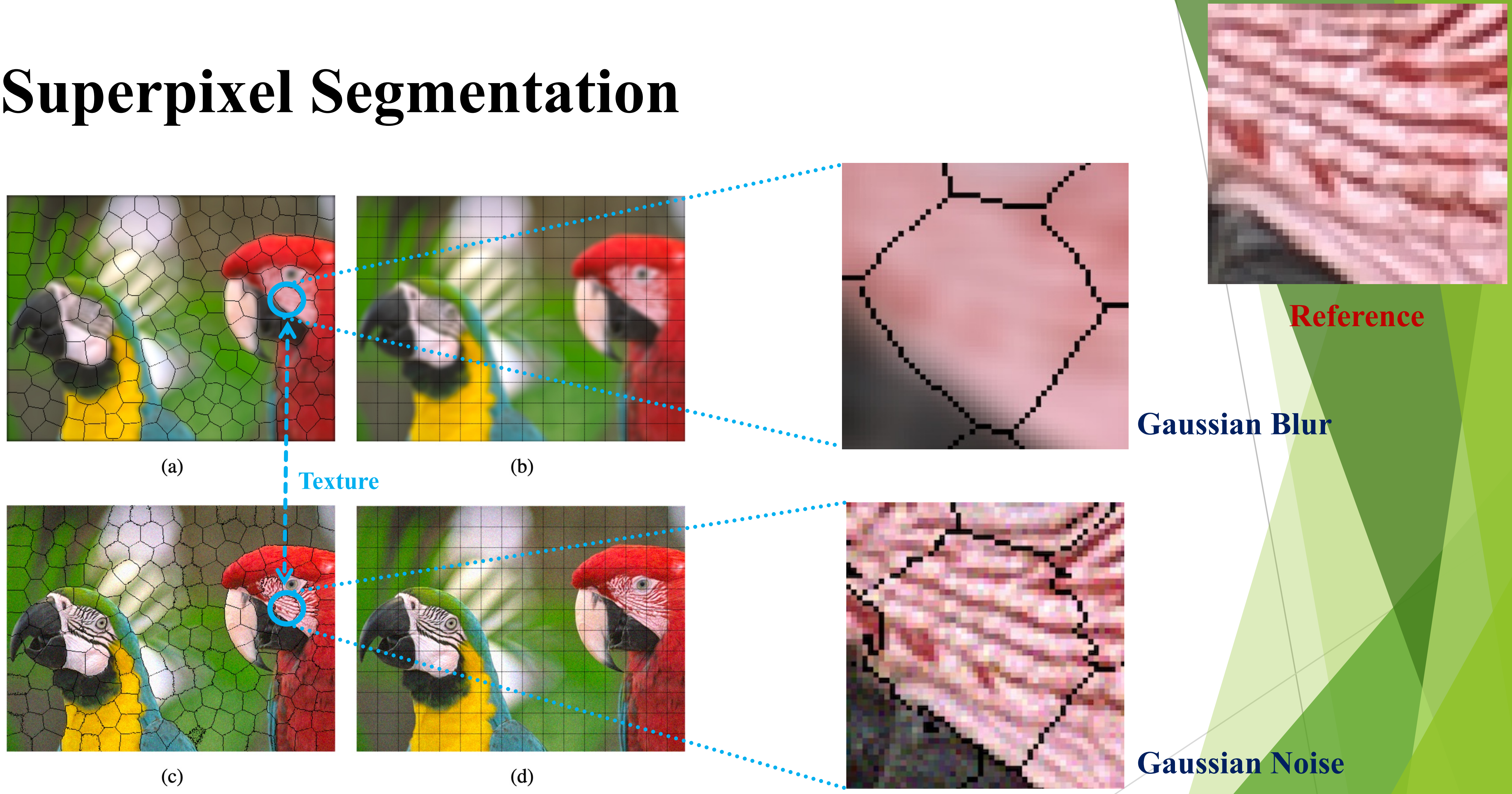


Figure 4.2: The superpixel *vs.* square patch representation (with size of $\approx 32 \times 32$) of the plane image from the TID2013 database.

Image Credit: TID2013 Database

Image Credit: TID2013 Database

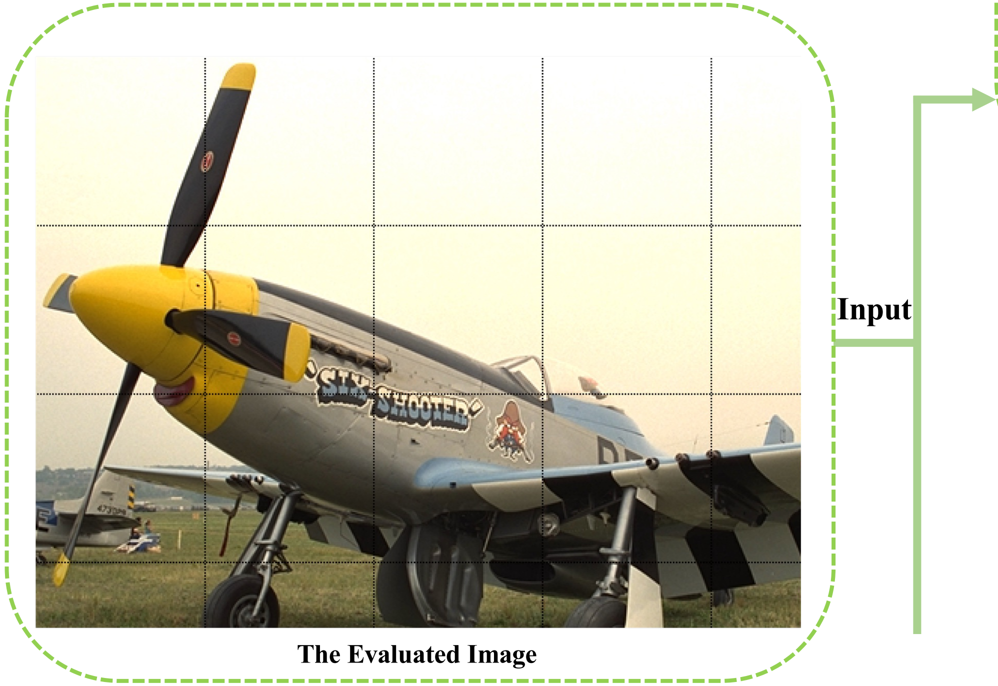
NLNet Architecture

The Evaluated Image

(i) Image Preprocessing

Image Credit: TID2013 Database

NLNet Architecture



(i) Image Preprocessing

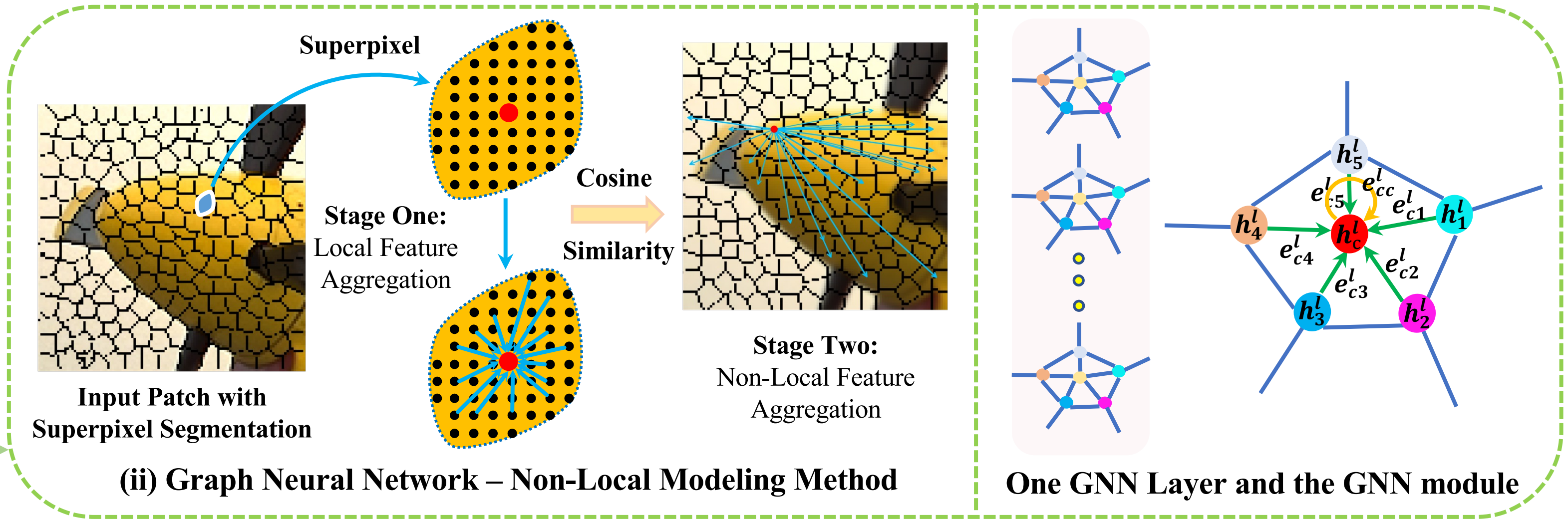
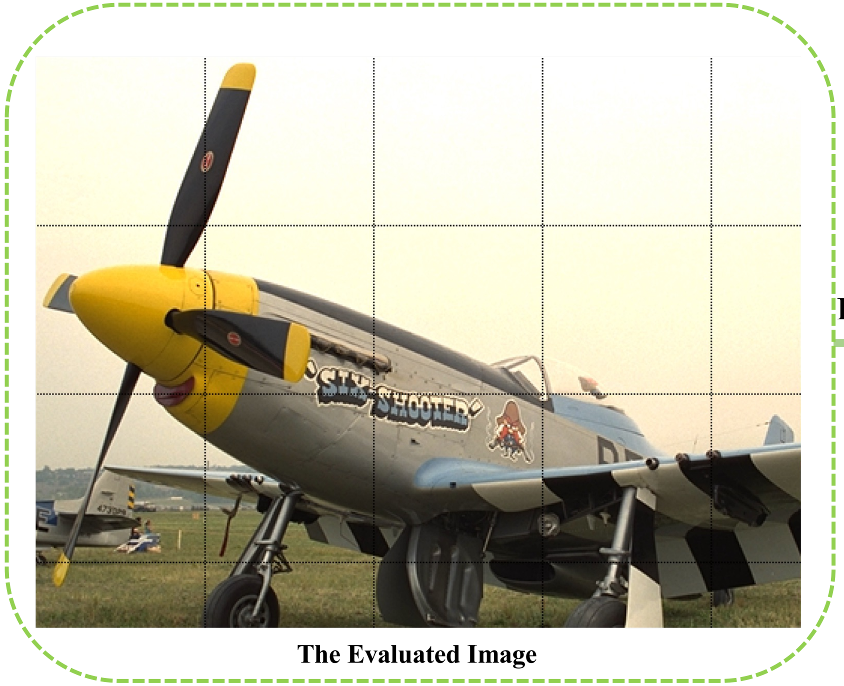


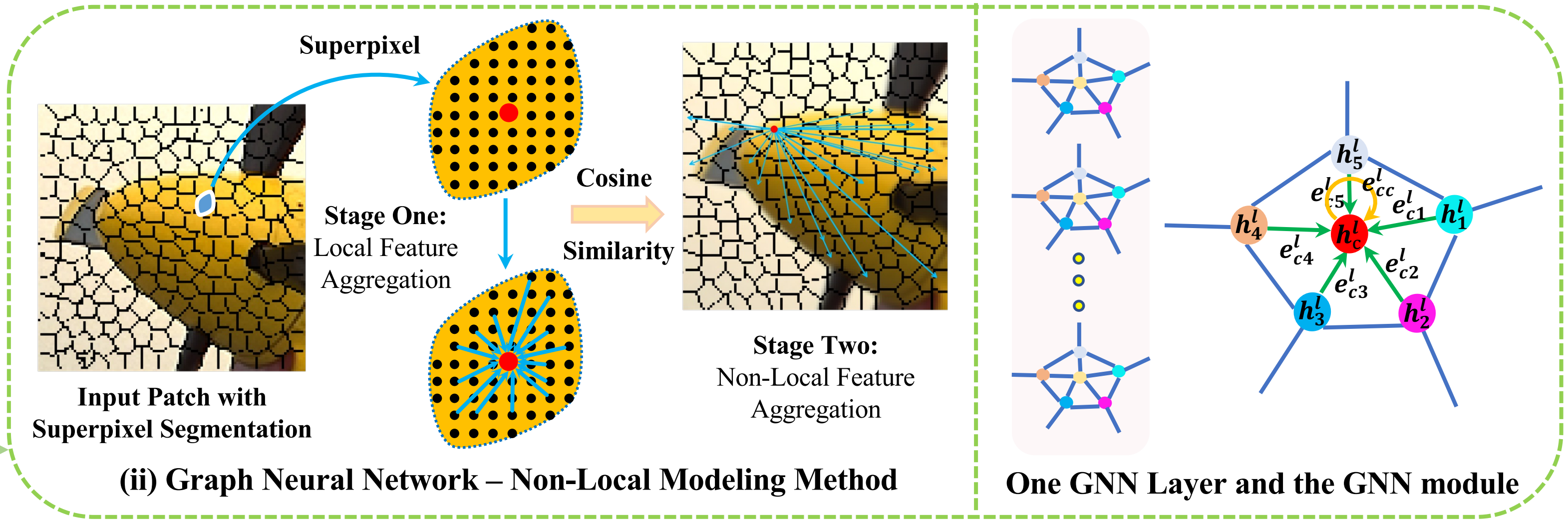
Image Credit: TID2013 Database

NLNet Architecture



(i) Image Preprocessing

Input



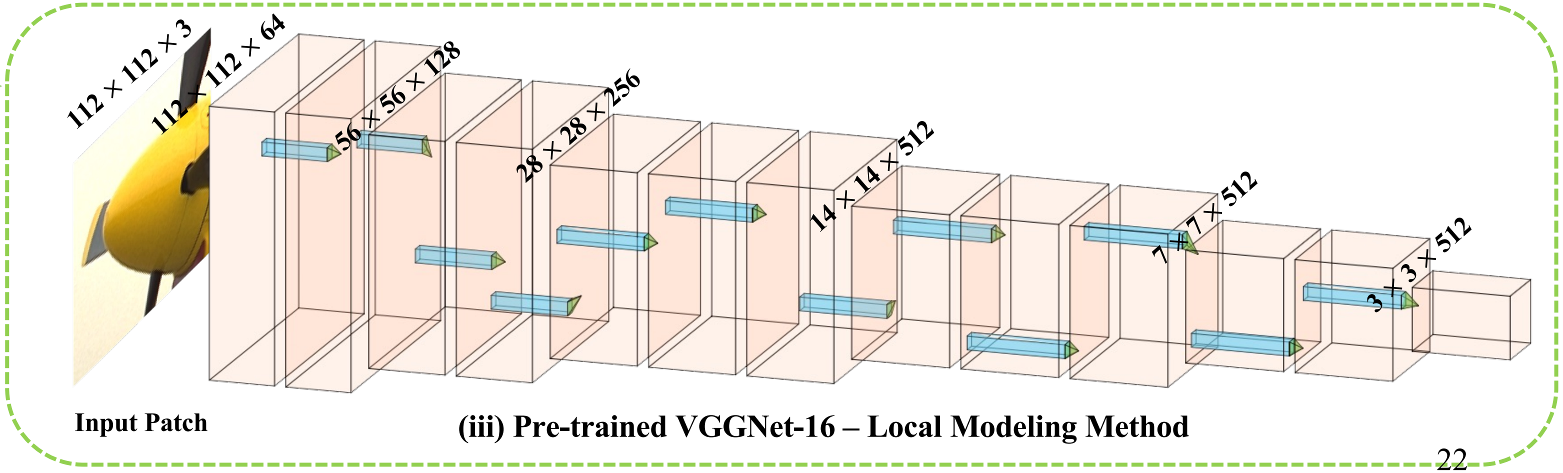


Image Credit: TID2013 Database

NLNet Architecture

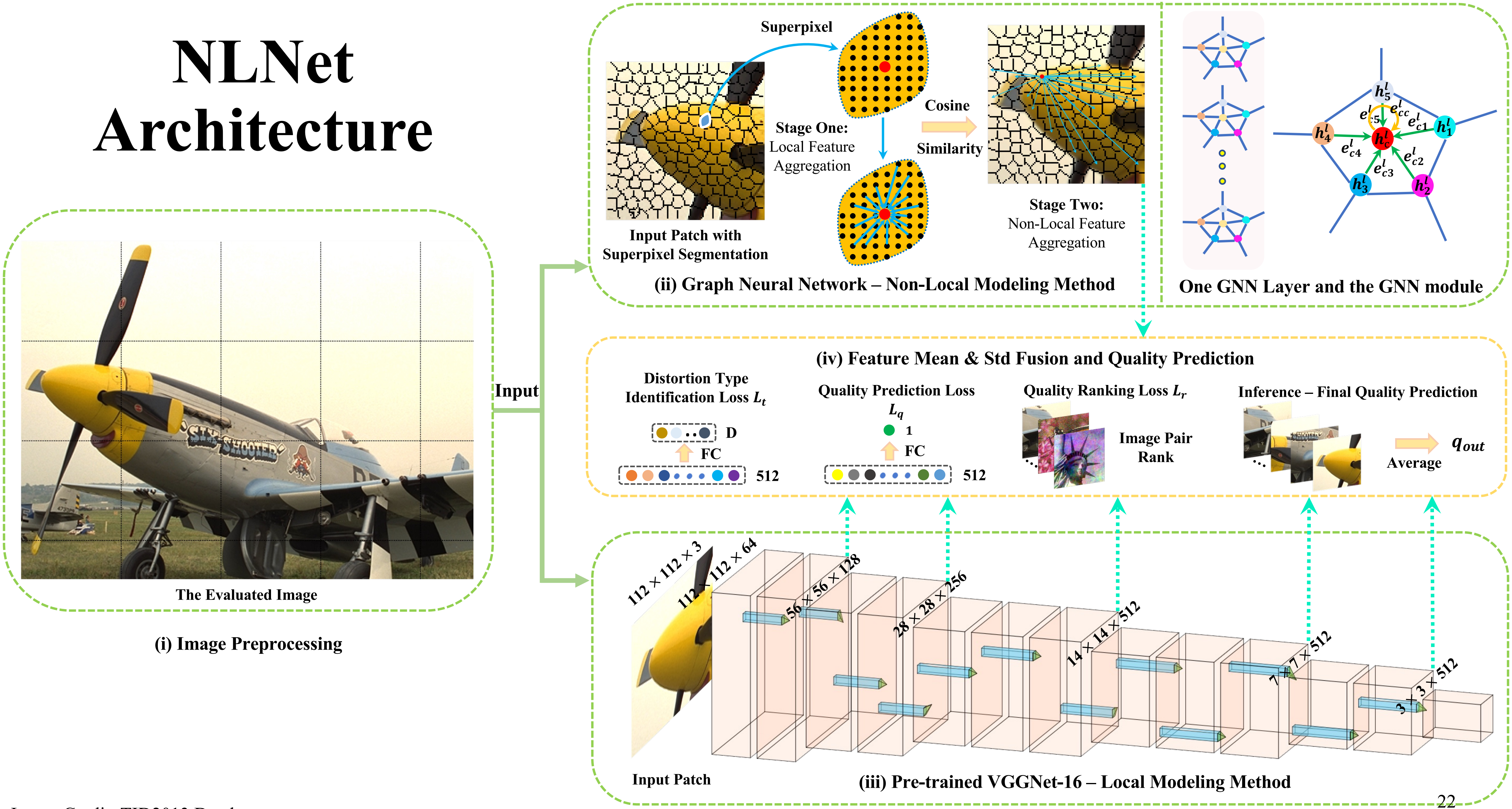


Image Credit: TID2013 Database

Experimental Setup

- **Databases**:
  - ✓ LIVE, CSIQ, TID2013, and KADID-10k
- **Evaluation Metrics**:
  - ✓ SRCC (Spearman Rank-order Correlation Coefficient)
  - ✓ PLCC (Pearson Linear Correlation Coefficient)
- **Experimental Settings**:
  - ✓ Intra-Database Experiments:
    - → 60% training, **20% validation**, and 20% testing, with `random` seeds from 1 to 10
    - → The median SRCC and PLCC are reported.
  - ✓ Cross-Database Evaluations:
    - → One database as the training set, and the other databases as the testing set
    - → Report the last epoch's performance

Figure 1.1: Natural images and a screen content image from the constructed databases. (a) LIVE Database [13] (b) CSIQ Database [14] (c) TID2013 Database [15] (d) KADID-10k Database [16].

Table 4.1: 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	$5 \sim 8$	$3 \sim 5$	5	5
Annotation	DMOS	DMOS	MOS	MOS
Range	[0, 100]	[0, 1]	[0, 9]	[1, 5]

Intra-Database Experiments

Table 4.2: Performance comparisons on the LIVE, CSIQ, and TID2013 databases. 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**

Table 4.3: Performance comparisons on the KADID-10k database. 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

Intra-Database Experiments

Table 4.2: Performance comparisons on the LIVE, CSIQ, and TID2013 databases. 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

Table 4.3: Performance comparisons on the KADID-10k database. 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

Intra-Database Experiments

Table 4.2: Performance comparisons on the LIVE, CSIQ, and TID2013 databases. 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**

Table 4.3: Performance comparisons on the KADID-10k database. 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

Intra-Database Experiments

Table 4.2: Performance comparisons on the LIVE, CSIQ, and TID2013 databases. 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**

Fewer Training Data
↓ 20% Total Data
↑ Highly Competitive Performance

Table 4.3: Performance comparisons on the KADID-10k database. 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

Cross-Database Settings and Evaluations

Table 4.9: 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**

Cross-Database Settings and Evaluations

Table 4.9: 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

Cross-Database Settings and Evaluations

Table 4.9: 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

TID: More Distortion Types & Levels

Single Distortion Type Evaluation

Table 4.4: The average SRCC and PLCC results of the individual distortion type on the LIVE database. 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

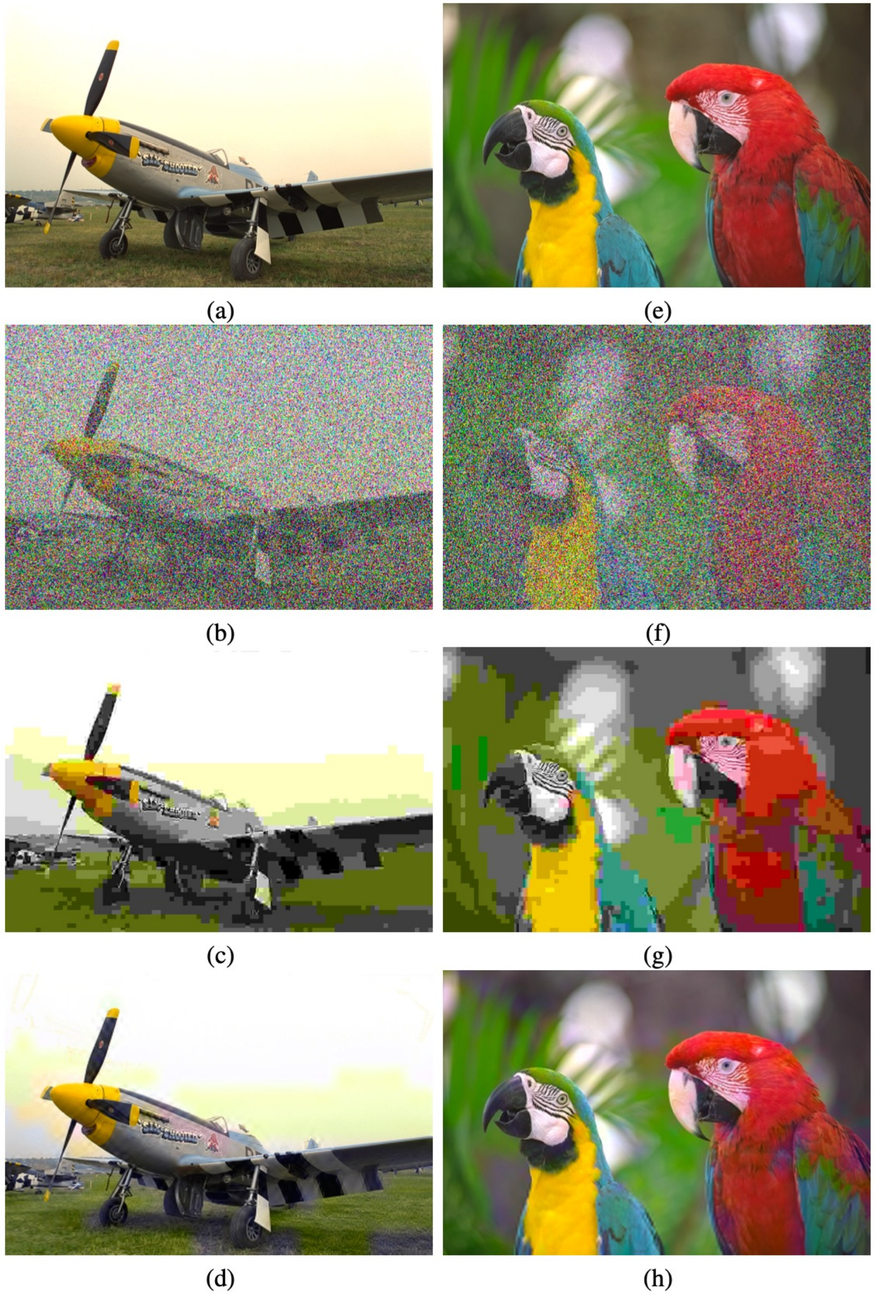


Figure 4.7: Demonstrations of the global distortions (b/f: WN and c/g: JPEG) and local distortions (d/h: FF) contaminating the plane and parrot images. Figure (a) and Figure (e) are reference images from the LIVE database.

Image Credit: LIVE Database

Single Distortion Type Evaluation

Table 4.4: The average SRCC and PLCC results of the individual distortion type on the LIVE database. 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

Non-local Recurrence

Figure 4.7: Demonstrations of the global distortions (b/f: WN and c/g: JPEG) and local distortions (d/h: FF) contaminating the plane and parrot images. Figure (a) and Figure (e) are reference images from the LIVE database.

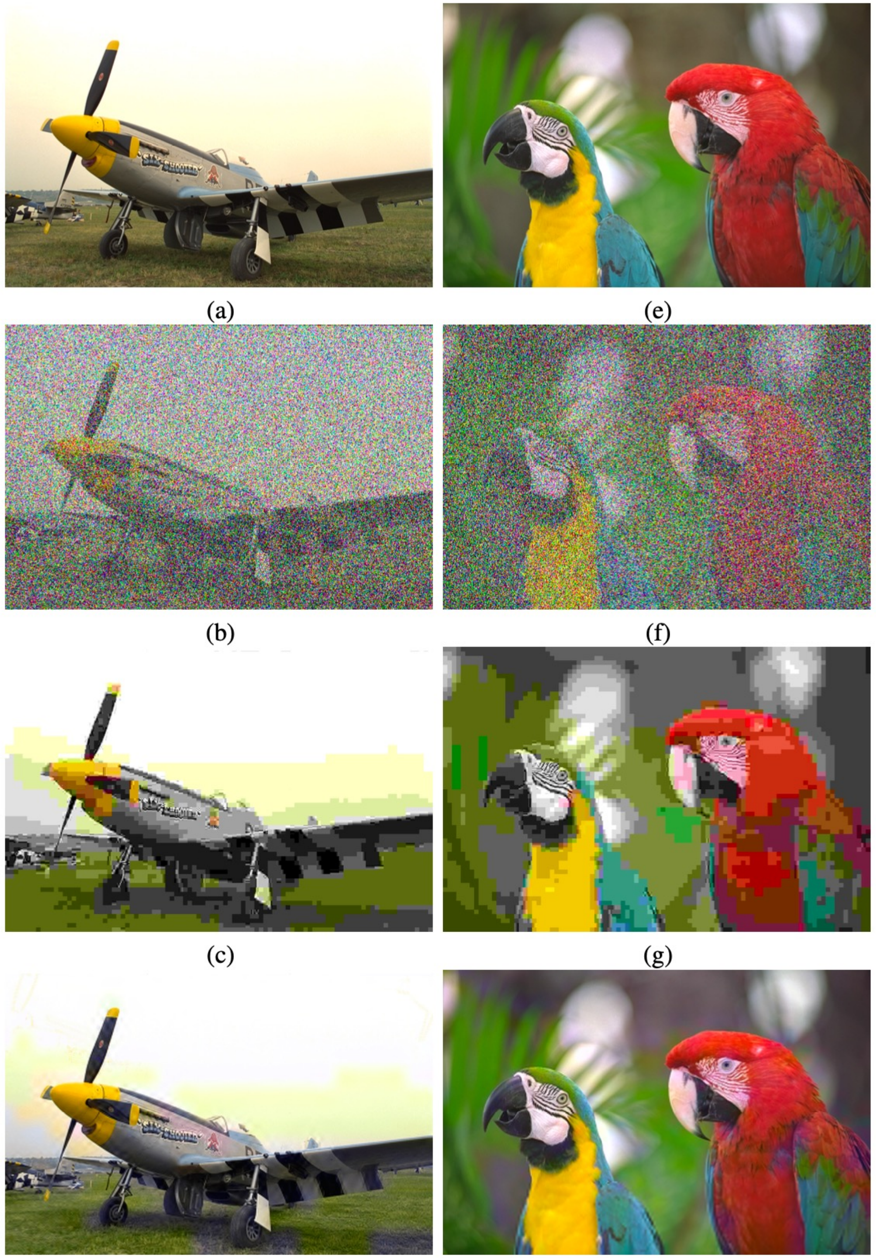
Image Credit: LIVE Database

Single Distortion Type Evaluation

Table 4.4: The average SRCC and PLCC results of the individual distortion type on the LIVE database. 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



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

Figure 4.7: Demonstrations of the global distortions (b/f: WN and c/g: JPEG) and local distortions (d/h: FF) contaminating the plane and parrot images. Figure (a) and Figure (e) are reference images from the LIVE database.

Image Credit: LIVE Database

Single Distortion Type Evaluation

Table 4.4: The average SRCC and PLCC results of the individual distortion type on the LIVE database. 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

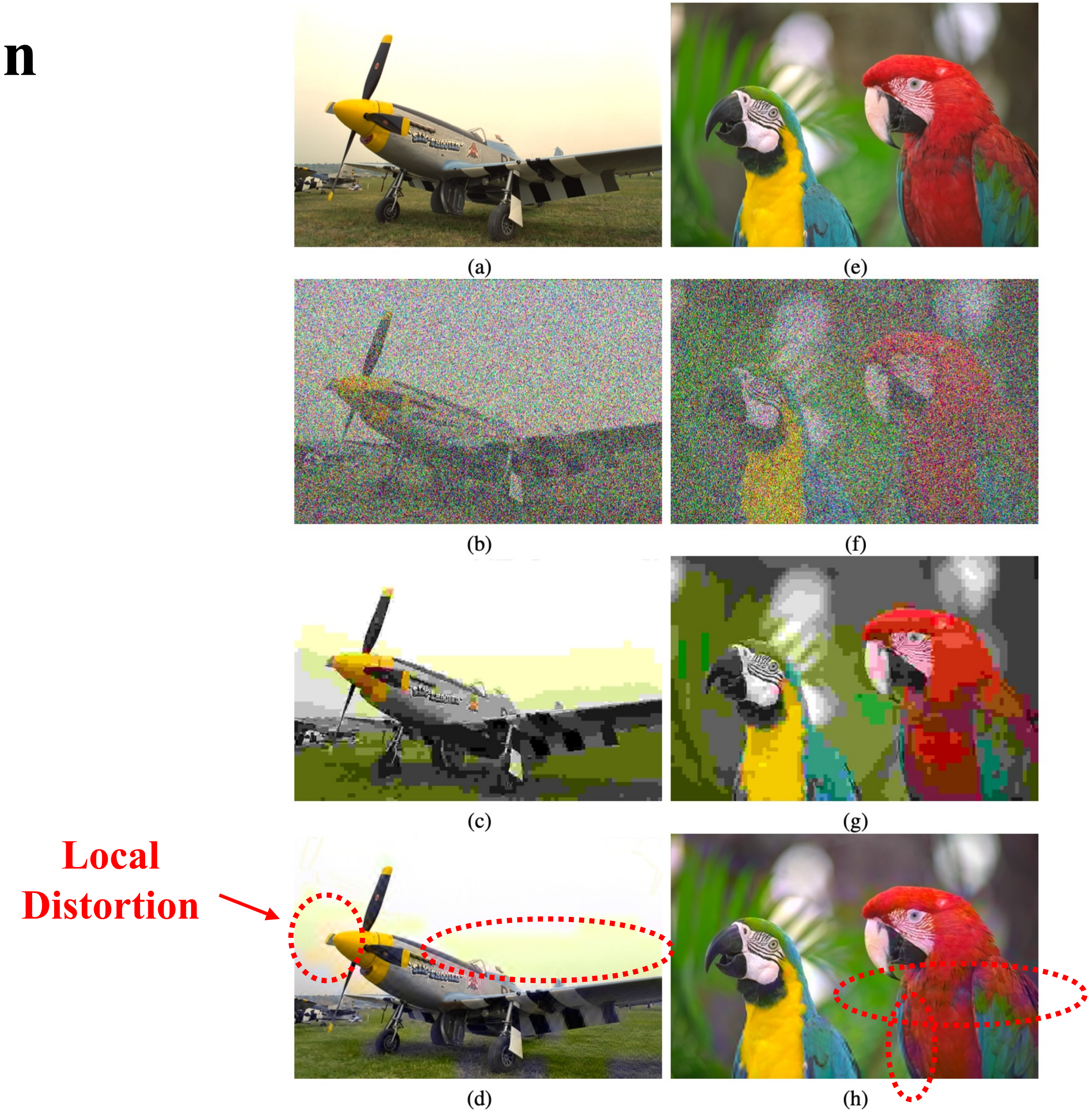


Figure 4.7: Demonstrations of the global distortions (b/f: WN and c/g: JPEG) and local distortions (d/h: FF) contaminating the plane and parrot images. Figure (a) and Figure (e) are reference images from the LIVE database.

Image Credit: LIVE Database

Single Distortion Type Evaluation

Table 4.4: The average SRCC and PLCC results of the individual distortion type on the LIVE database. 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

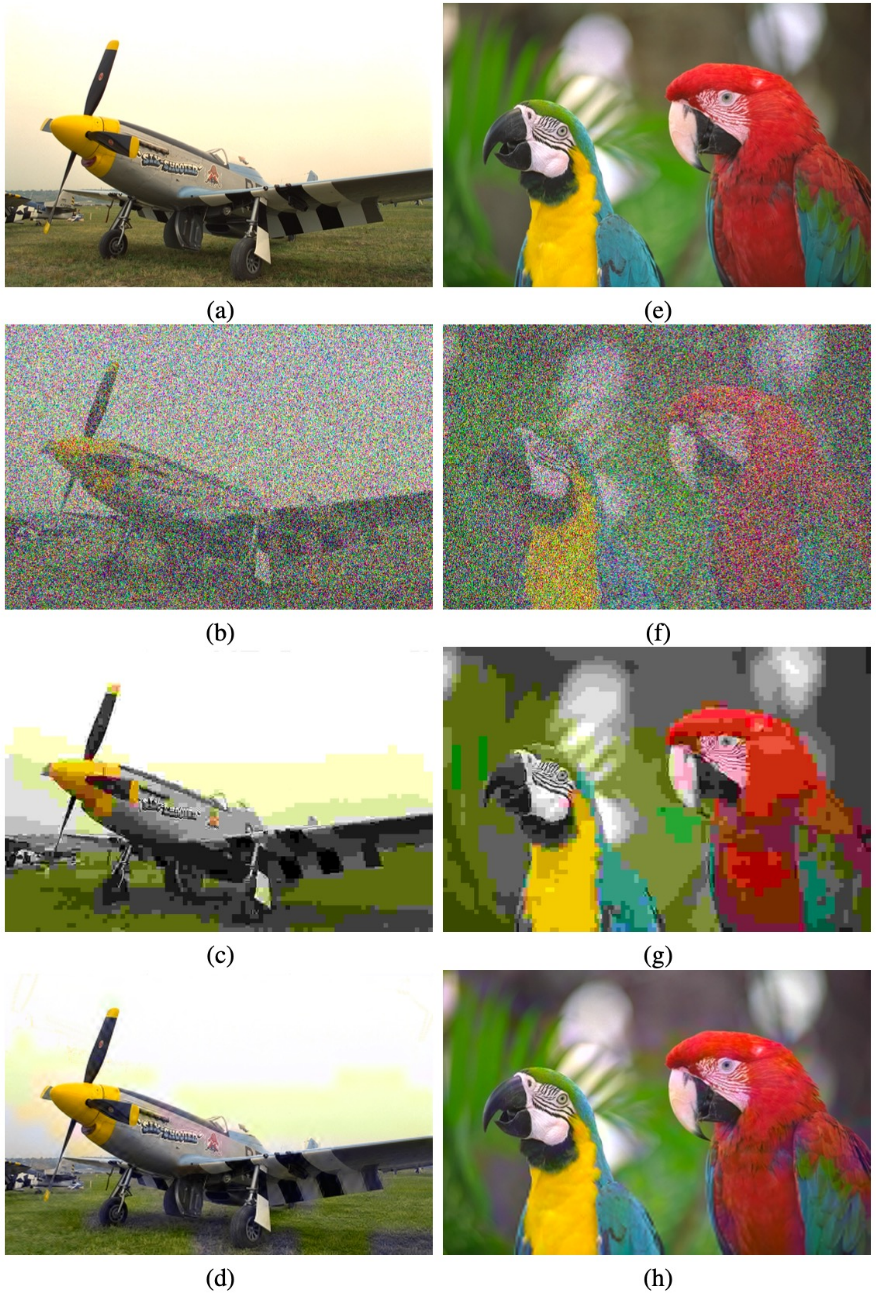


Figure 4.7: Demonstrations of the global distortions (b/f: WN and c/g: JPEG) and local distortions (d/h: FF) contaminating the plane and parrot images. Figure (a) and Figure (e) are reference images from the LIVE database.

Image Credit: LIVE Database

Single Distortion Type Evaluation

Table 4.4: The average SRCC and PLCC results of the individual distortion type on the LIVE database. 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

Global Distortion

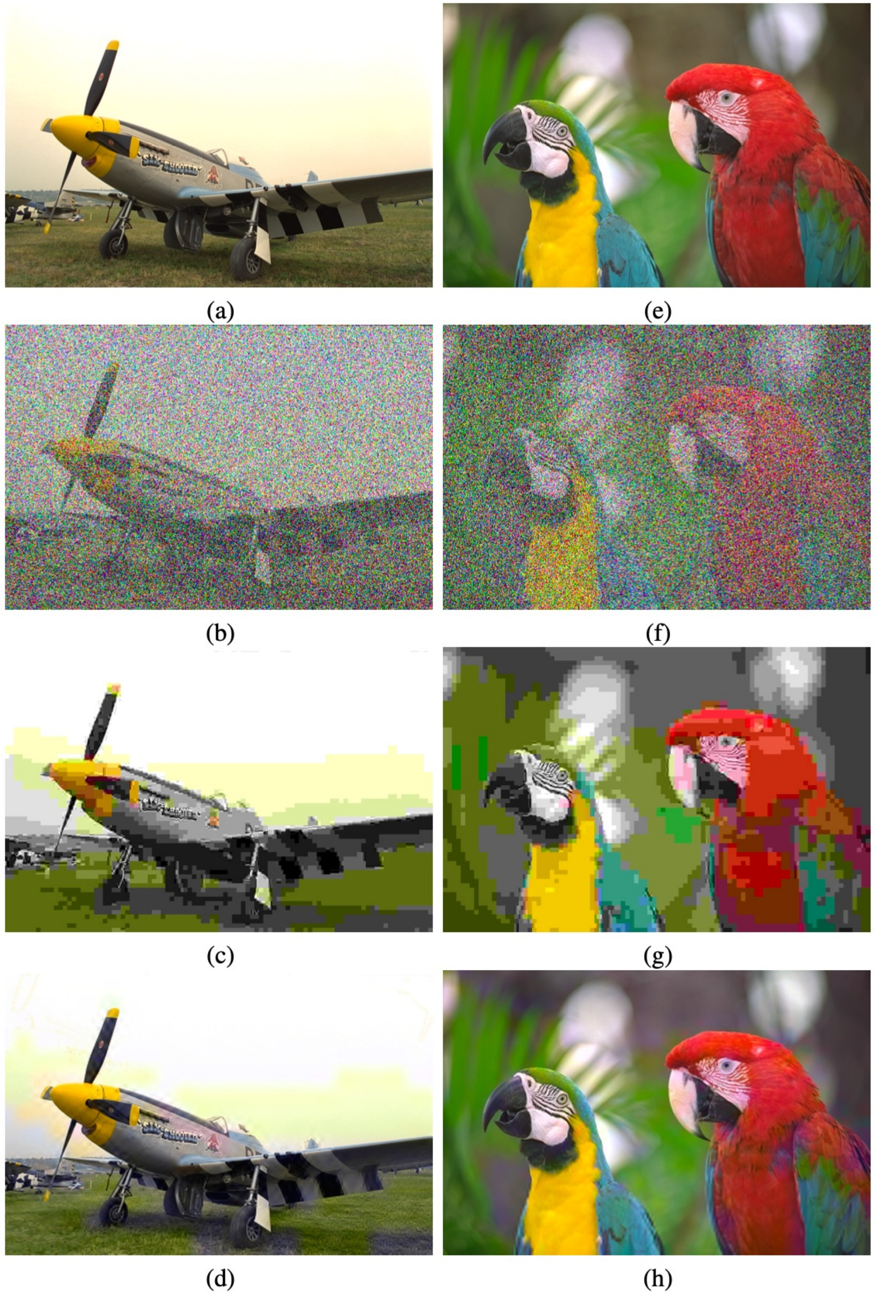


Figure 4.7: Demonstrations of the global distortions (b/f: WN and c/g: JPEG) and local distortions (d/h: FF) contaminating the plane and parrot images. Figure (a) and Figure (e) are reference images from the LIVE database.

Image Credit: LIVE Database

Single Distortion Type Evaluation

Table 4.4: The average SRCC and PLCC results of the individual distortion type on the LIVE database. 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

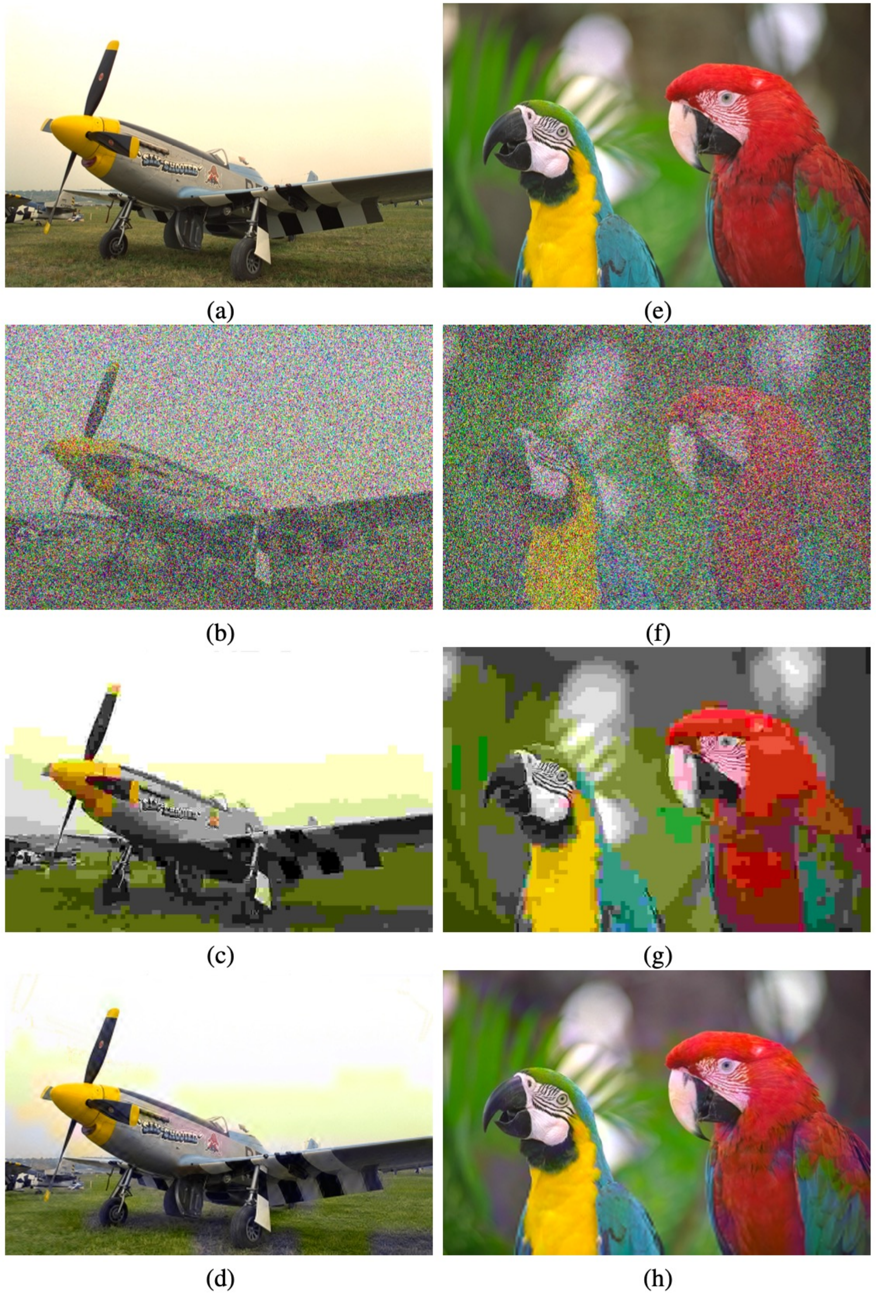


Figure 4.7: Demonstrations of the global distortions (b/f: WN and c/g: JPEG) and local distortions (d/h: FF) contaminating the plane and parrot images. Figure (a) and Figure (e) are reference images from the LIVE database.

Image Credit: LIVE Database

Single Distortion Type Evaluation

Table 4.4: The average SRCC and PLCC results of the individual distortion type on the LIVE database. 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 and Compressed Images

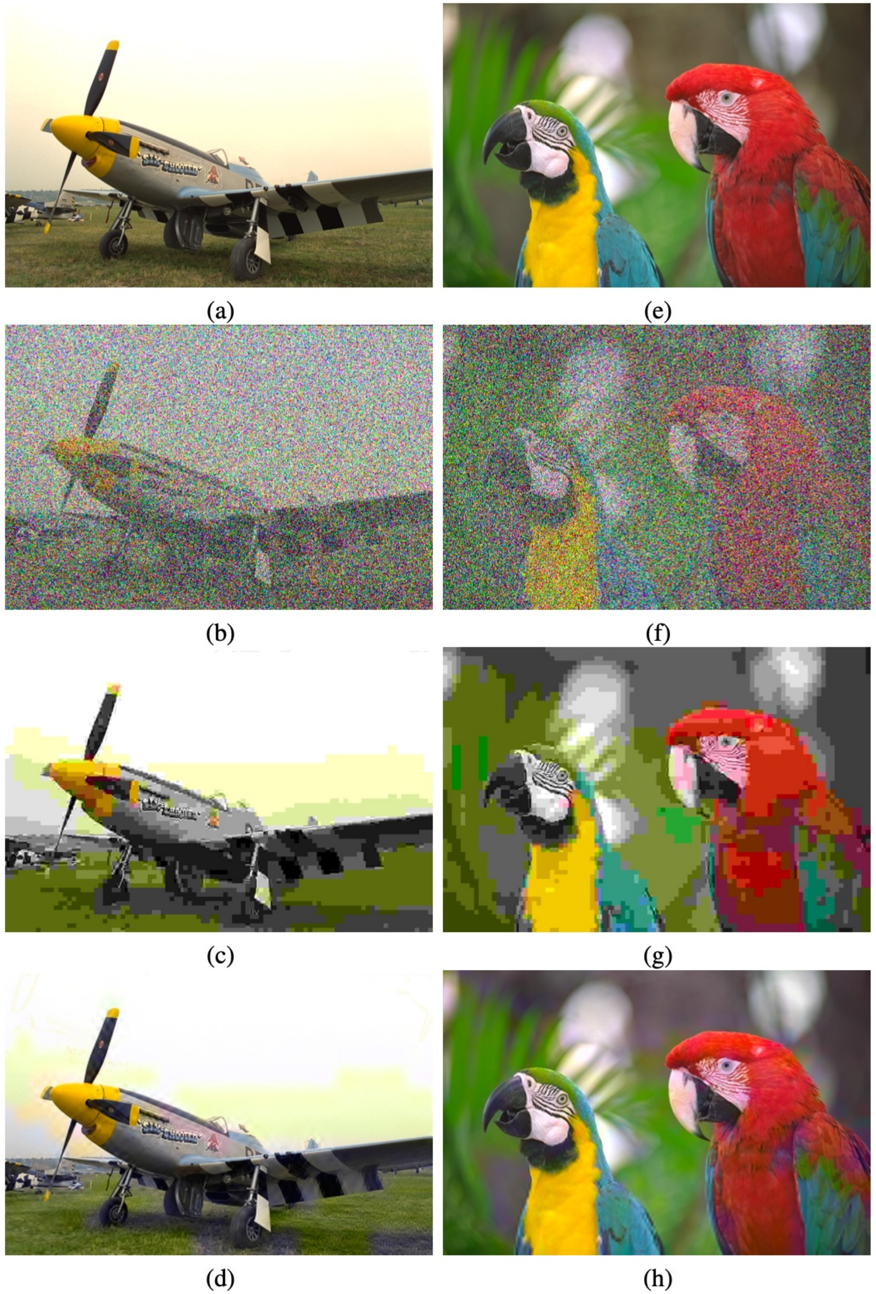


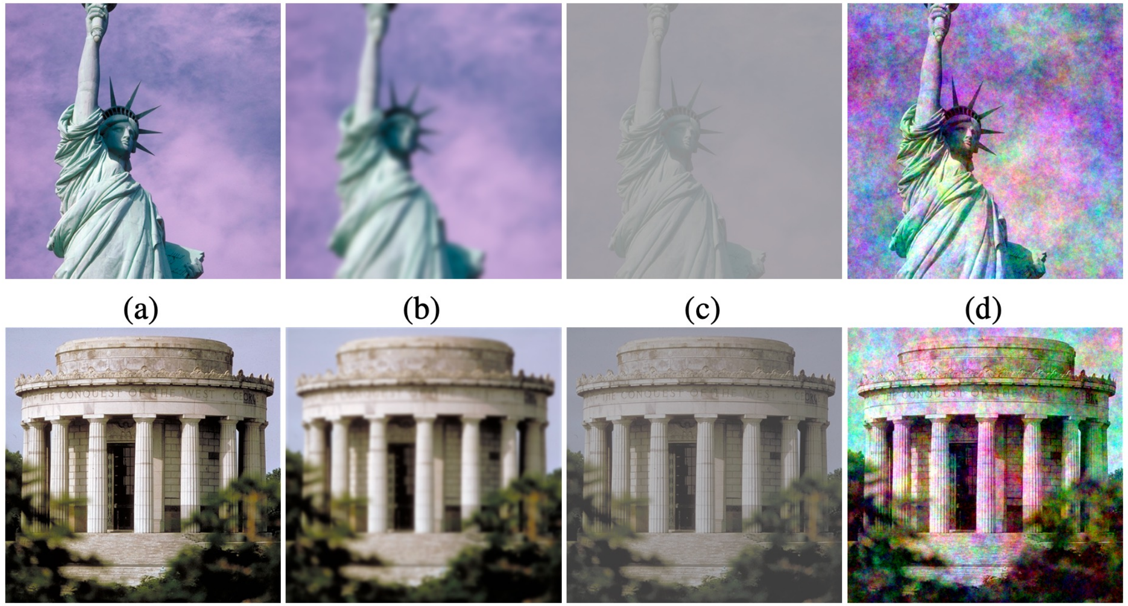
Figure 4.7: Demonstrations of the global distortions (b/f: WN and c/g: JPEG) and local distortions (d/h: FF) contaminating the plane and parrot images. Figure (a) and Figure (e) are reference images from the LIVE database.

Image Credit: LIVE Database

Single Distortion Type Evaluation

Table 4.5: The average SRCC and PLCC results of the individual distortion type on the CSIQ database. 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**



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

Figure 4.9: Demonstrations of the global distortions (b/f: GB, c/g: CC, d/h: PN) contaminating the Statue of Liberty and George Rogers Clark Memorial images. Figure (a) and Figure (e) are reference images from the CSIQ database.

Image Credit: CSIQ Database

Single Distortion Type Evaluation

Table 4.5: The average SRCC and PLCC results of the individual distortion type on the CSIQ database. 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**

Global Distortion

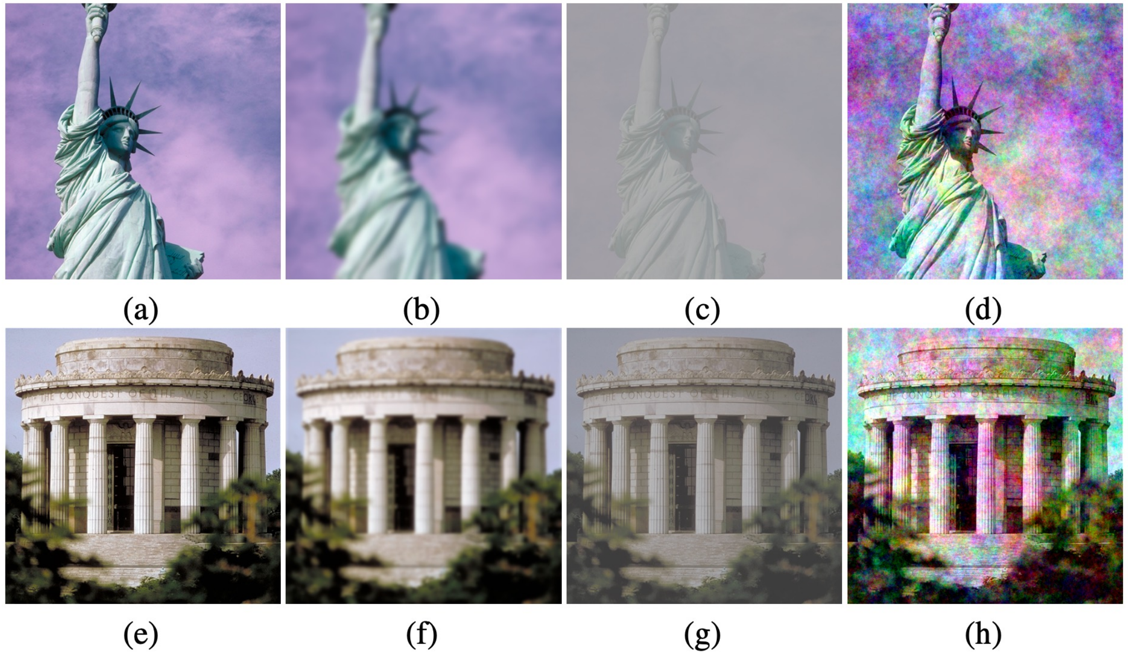


Figure 4.9: Demonstrations of the global distortions (b/f: GB, c/g: CC, d/h: PN) contaminating the Statue of Liberty and George Rogers Clark Memorial images. Figure (a) and Figure (e) are reference images from the CSIQ database.

Image Credit: CSIQ Database

Single Distortion Type Evaluation

Table 4.5: The average SRCC and PLCC results of the individual distortion type on the CSIQ database. 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**

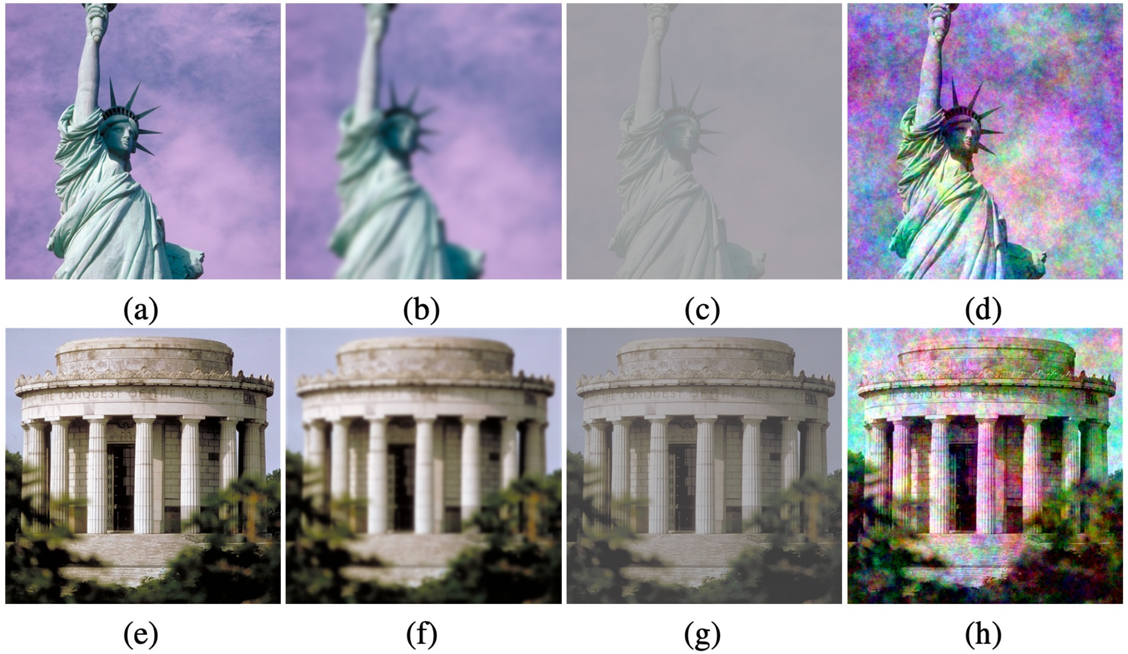


Figure 4.9: Demonstrations of the global distortions (b/f: GB, c/g: CC, d/h: PN) contaminating the Statue of Liberty and George Rogers Clark Memorial images. Figure (a) and Figure (e) are reference images from the CSIQ database.

Image Credit: CSIQ Database

Single Distortion Type Evaluation

Noise-Related Distortions

Table 4.5: The average SRCC and PLCC results of the individual distortion type on the CSIQ database. 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**

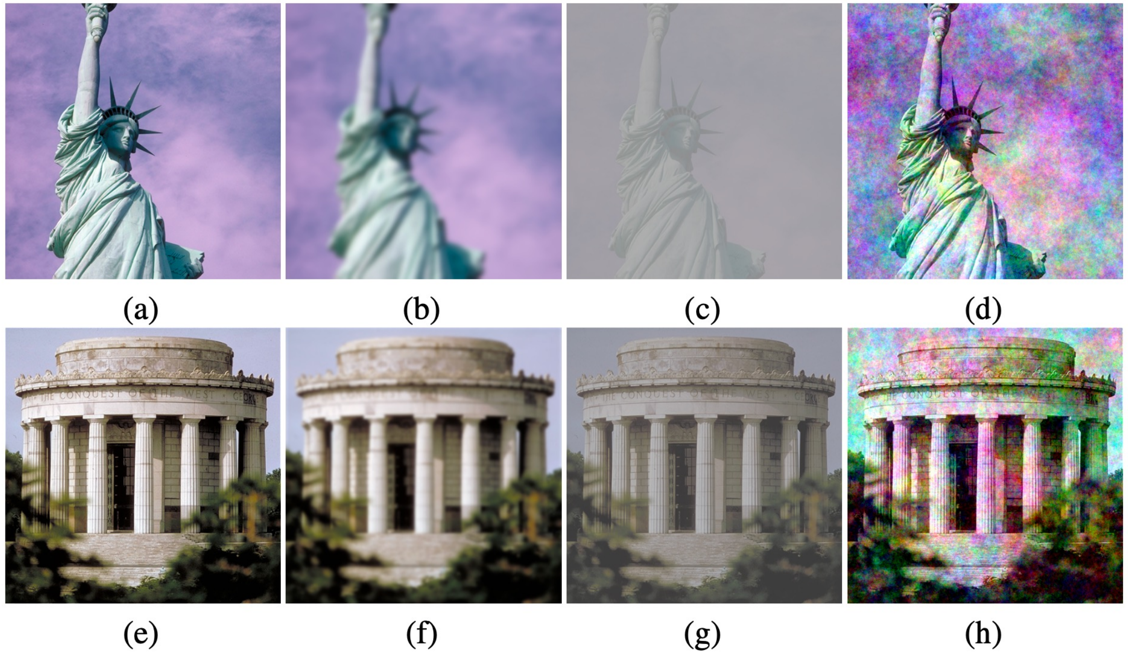


Figure 4.9: Demonstrations of the global distortions (b/f: GB, c/g: CC, d/h: PN) contaminating the Statue of Liberty and George Rogers Clark Memorial images. Figure (a) and Figure (e) are reference images from the CSIQ database.

Image Credit: CSIQ Database

Single Distortion Type Evaluation

Table 4.6: The average SRCC results of the individual distortion type on the TID2013 database. 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**	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**	0.712	**0.935**
	Additive noise in color components	0.432	0.573	0.551	**0.722**	0.560	0.700	0.137	**0.850**
	Comfort noise	0.196	0.318	0.622	0.406	0.353	**0.752**	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**	0.782	0.826	0.741	**0.958**
	High frequency noise	0.842	0.847	0.897	**0.911**	0.900	0.879	0.815	**0.921**
	Impulse noise	0.765	0.730	0.809	**0.901**	0.738	0.708	0.616	**0.913**
	Quantization noise	0.662	0.764	0.815	**0.888**	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**	0.764	**0.882**
	JPEG compression	0.764	0.813	0.891	0.850	0.844	**0.894**	0.797	**0.905**
	JPEG 2000 compression	0.745	0.831	**0.919**	0.891	0.885	0.916	0.846	**0.930**
	Multiplicative Gaussian noise	0.717	0.704	0.768	**0.849**	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**	0.694	**0.805**
	JPEG 2000 transmission errors	0.748	0.660	0.710	0.716	0.718	**0.773**	0.686	**0.875**
	Non eccentricity pattern noise	0.269	0.076	0.242	0.116	0.173	**0.270**	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

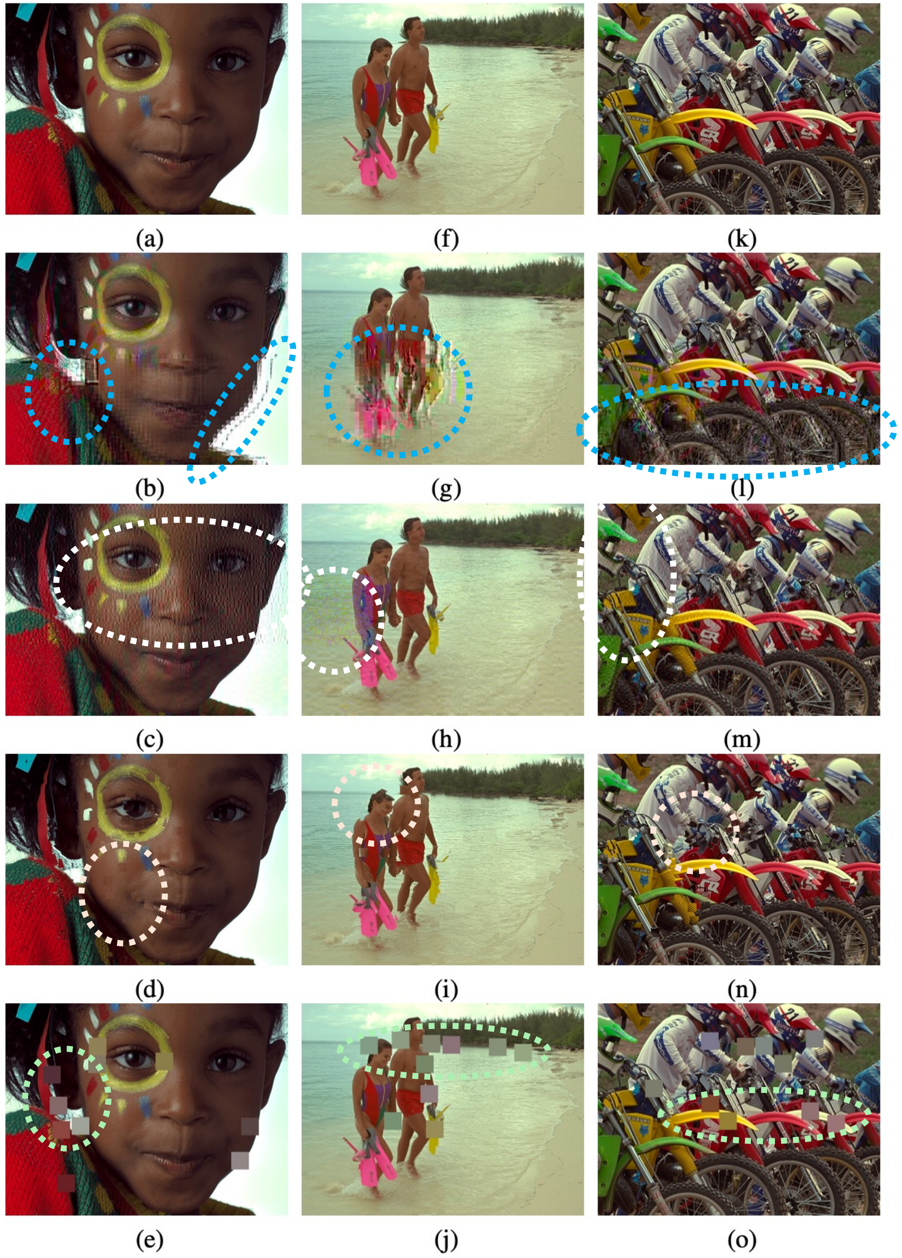


Figure 4.10: Demonstrations of the local distortions (b/g/l: JPEG transmission errors, c/h/m: JPEG2000 transmission errors, d/i/n: non eccentricity pattern noise, e/j/o: local block-wise distortions of different intensity). Figure (a), Figure (f), and Figure (k) are reference images from the TID2013 database.

Image Credit: TID2013 Database

Single Distortion Type Evaluation

Table 4.6: The average SRCC results of the individual distortion type on the TID2013 database. 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%	0.560	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**	0.694	**0.805**
	JPEG 2000 transmission errors	0.748	0.660	0.710	0.716	0.718	**0.773**	0.686	**0.875**
	Non eccentricity pattern noise	0.269	0.076	0.242	0.116	0.173	**0.270**	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

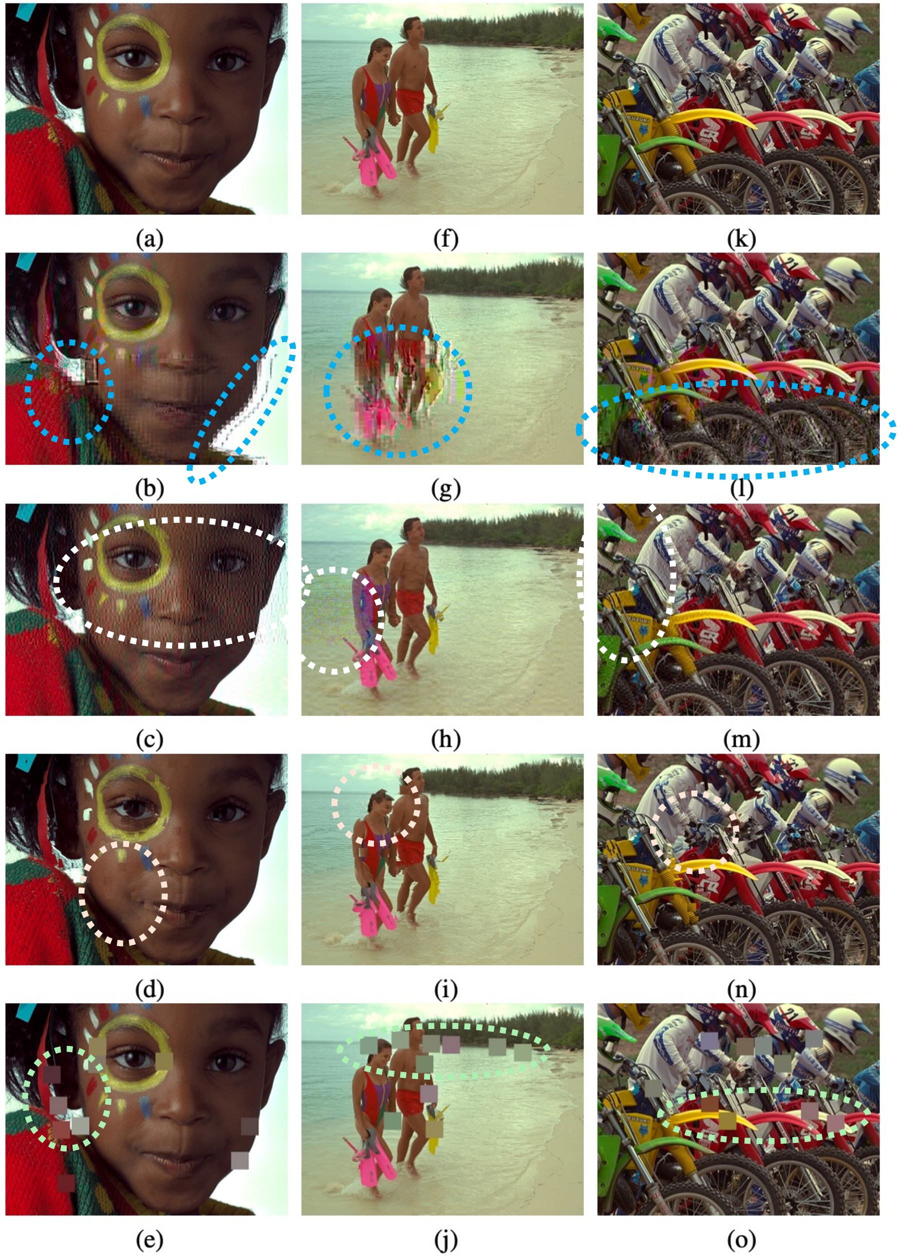


Figure 4.10: Demonstrations of the local distortions (b/g/l: JPEG transmission errors, c/h/m: JPEG2000 transmission errors, d/i/n: non eccentricity pattern noise, e/j/o: local block-wise distortions of different intensity). Figure (a), Figure (f), and Figure (k) are reference images from the TID2013 database.

Image Credit: TID2013 Database

Single Distortion Type Evaluation

Table 4.6: The average SRCC results of the individual distortion type on the TID2013 database. 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%	0.560	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

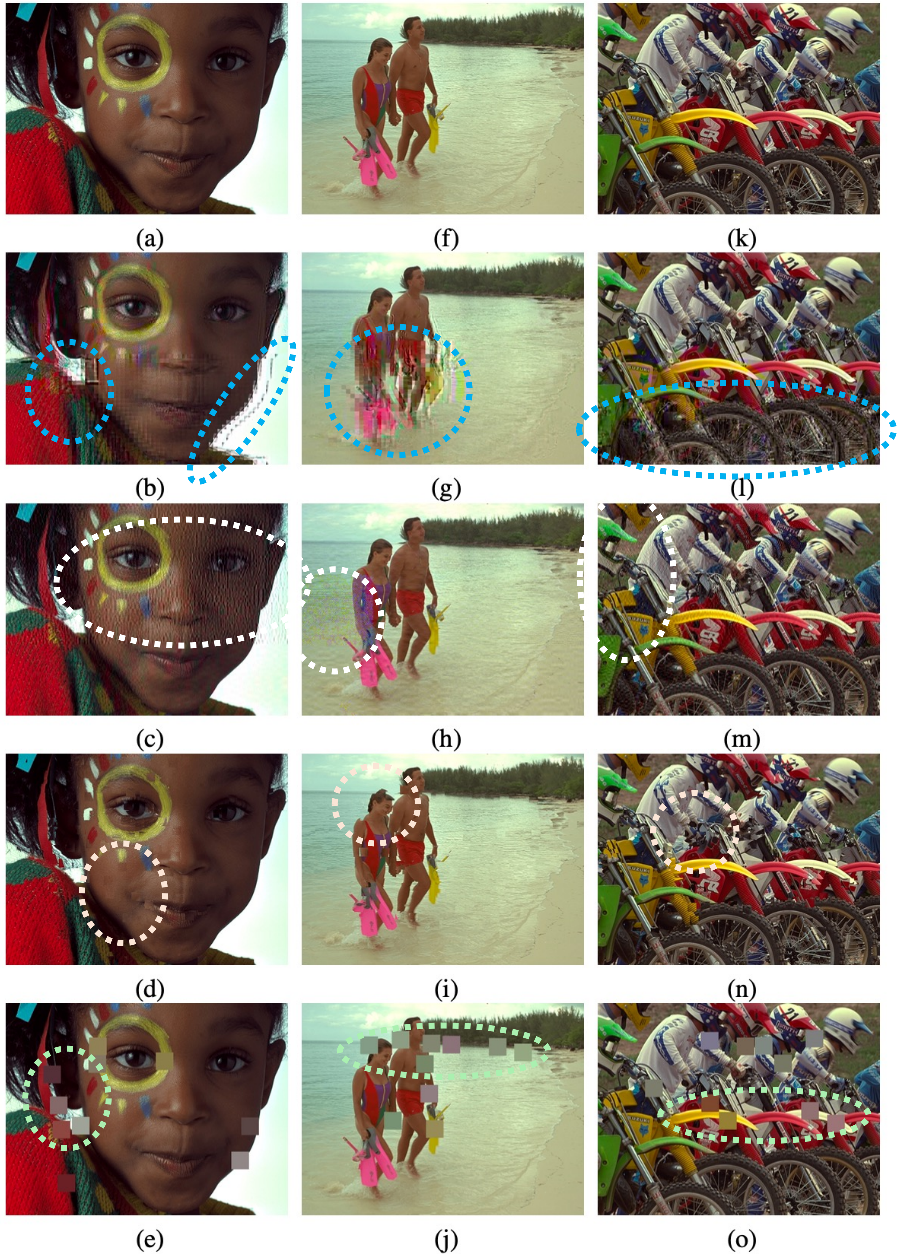


Figure 4.10: Demonstrations of the local distortions (b/g/l: JPEG transmission errors, c/h/m: JPEG2000 transmission errors, d/i/n: non eccentricity pattern noise, e/j/o: local block-wise distortions of different intensity). Figure (a), Figure (f), and Figure (k) are reference images from the TID2013 database.

Image Credit: TID2013 Database

Single Distortion Type Evaluation

Table 4.7: 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**	0.556	0.582	0.530	**0.866**
	JPEG 2000 compression	0.655	0.516	**0.790**	0.342	0.608	0.539	**0.853**
Noise	Denoise	0.457	0.221	**0.856**	0.229	0.247	0.765	**0.953**
	White noise in color component	0.757	0.718	0.841	0.418	0.745	**0.925**	**0.936**
	Multiplicative noise	0.702	0.674	0.682	0.306	0.776	**0.884**	**0.934**
	Impulse noise	0.547	-0.543	0.808	0.219	0.254	**0.814**	**0.916**
	White Gaussian noise	0.628	0.708	0.776	0.357	0.680	**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**

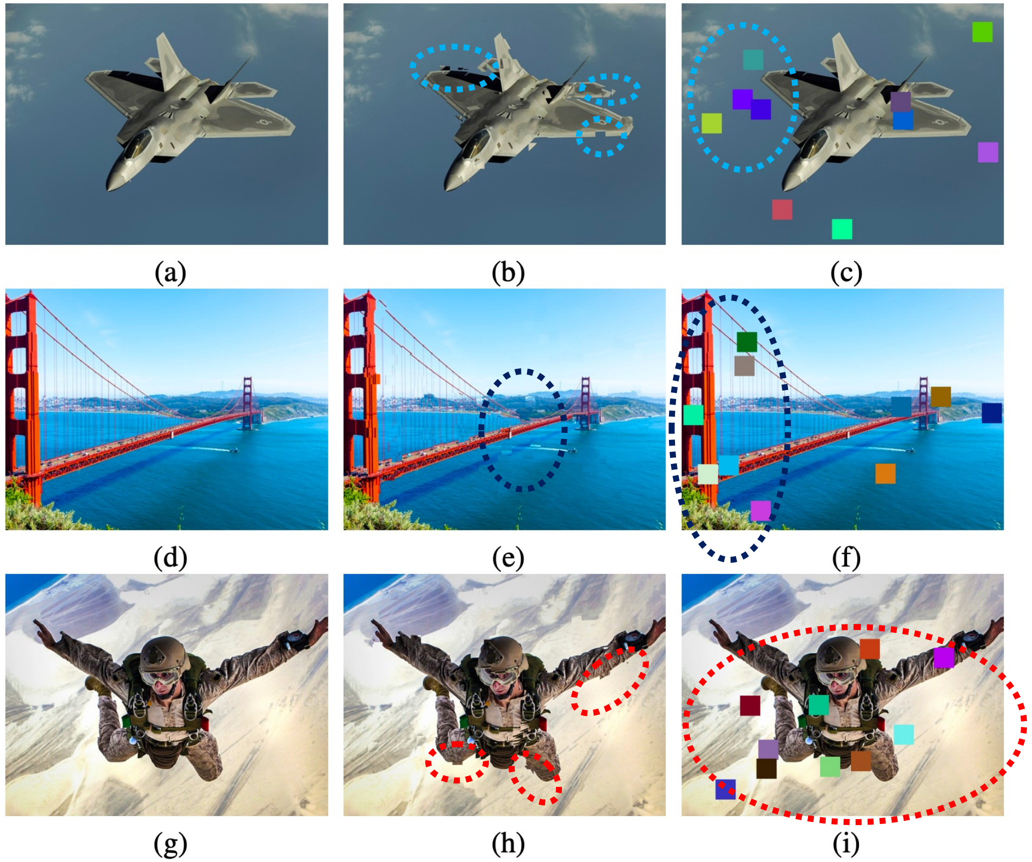


Figure 4.11: Demonstrations of the local distortions (b/e/h: non-eccentricity patch and c/f/i: color block). Figure (a), Figure (d), and Figure (g) are reference images from the KADID-10k database.

Image Credit: KADID-10k Database

Single Distortion Type Evaluation

Table 4.7: 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**

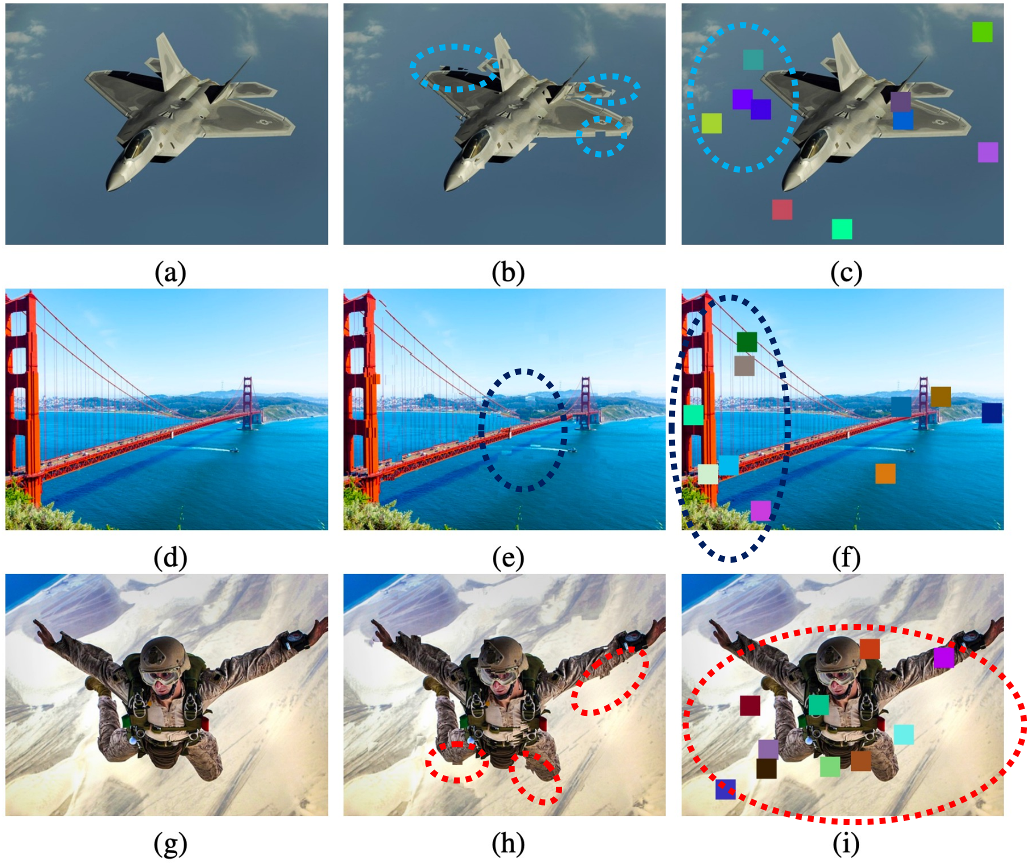


Figure 4.11: Demonstrations of the local distortions (b/e/h: non-eccentricity patch and c/f/i: color block). Figure (a), Figure (d), and Figure (g) are reference images from the KADID-10k database.

Image Credit: KADID-10k Database

Takeaways and Future Work

✓ **Non-local & Local Modeling**

(1) The **Non-local Modeling is complementary** to traditional local methods.

(2) CNN’s **Local Modeling features are effective and robust**.

✓ **Global & Local Distortions**

(1) Handle a wide variety of **Global Distortions**: globally and uniformly distributed with non-local recurrences.

(2) Maintain sensitivity to **Local Distortions**: local nonuniform-distributed distortions in a local region.

(3) Better assess **Noisy and Compressed Images** quality.

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

✓ **Future Work** **Non-local Statistics** [1, 2]

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

EEG Deep Learning

GCNs-Net: A Graph Convolutional Neural Network Approach for Decoding Time-Resolved EEG Motor Imagery Signals

Yimin Hou [1], Shuyue Jia [1, 2], Xiangmin Lun [1], Ziqian Hao [3], Yan Shi [1],

Yang Li [4], Rui Zeng [5], and Jinglei Lv [5] *

[1] School of Automation Engineering, Northeast Electric Power University

[2] Department of Computer Science, City University of Hong Kong

[3] Jinan Vocational College

[4] School of Electrical Engineering, Northeast Electric Power University

[5] School of Biomedical Engineering and Brain and Mind Center, The University of Sydney

EEG Deep Learning Library: https://github.com/SuperBruceJia/EEG-DL

Background

- **BCI**: establish connections between the brain and machines

  (1) **Acquire and analyze brain signals** while conducting **actual** or **imagery** tasks

  (2) **Control machines**

- **Significance**: help the disabled and understand the human brain
- **Types of BCI**:
  - Electroencephalography (EEG)
  - Magnetoencephalography (MEG)
  - Functional Magnetic Resonance Imaging (fMRI)
  - Invasive BCI Technologies (*e.g.*, Neuralink)

- **Reasons for using EEG for this project**:
  - Non-Invasiveness
  - High Temporal Resolution
  - Portability
  - Inexpensive Equipment

  A potential market

- **Specific Task**: EEG Motor Imagery (*e.g.*, control a wheelchair via imagery-based EEG signals)
- **Our Research**: develop EEG-based BCI technologies to improve current stroke rehabilitation strategies

Image Credit: in the public domain.

Key Points in dealing with EEG time series

- **Individual Variability → Lower Classification Accuracy**
  - ✓ Low SNR
  - ✓ Different brain electrical conductivity ← different anatomical structure of brain
  - ✓ Electrodes' positional error
- **Slow Responding → Hard to develop Real-life Applications**
  - ✓ [**most literature**] Trial-level prediction (*e.g.*, 4 s)
  - ✓ Window/Slide-level prediction (*e.g.*, 0.4 s)
  - ✓ Time-resolved prediction (*e.g.*, 6.25 ms) (Our Work)
- **Lower Group-level Accuracy → Hard to develop Applications for a Group of People**
  - ✓ [**most literature**] Subject-level prediction (Our Work)
  - ✓ Group-level prediction (Our Work)

Key Points in dealing with EEG time series

- **Individual Variability → Lower Classification Accuracy**
  - ✓ Low SNR
  - ✓ Different brain electrical conductivity ← different anatomical structure of brain
  - ✓ Electrodes' positional error

Feature Extraction

- **Slow Responding → Hard to develop Real-life Applications**
  - ✓ [**most literature**] Trial-level prediction (*e.g.*, 4 s)
  - ✓ Window/Slide-level prediction (*e.g.*, 0.4 s)
  - ✓ Time-resolved prediction (*e.g.*, 6.25 ms) (Our Work)
- **Lower Group-level Accuracy → Hard to develop Applications for a Group of People**
  - ✓ [**most literature**] Subject-level prediction (Our Work)
  - ✓ Group-level prediction (Our Work)

Key Points in dealing with EEG time series

- **Individual Variability → Lower Classification Accuracy**
  - ✓ Low SNR
  - ✓ Different brain electrical conductivity ← different anatomical structure of brain
  - ✓ Electrodes' positional error

Feature Extraction

EEG Electrodes' Structure Modeling

- **Slow Responding → Hard to develop Real-life Applications**
  - ✓ [**most literature**] Trial-level prediction (*e.g.*, 4 s)
  - ✓ Window/Slide-level prediction (*e.g.*, 0.4 s)
  - ✓ Time-resolved prediction (*e.g.*, 6.25 ms) (Our Work)
- **Lower Group-level Accuracy → Hard to develop Applications for a Group of People**
  - ✓ [**most literature**] Subject-level prediction (Our Work)
  - ✓ Group-level prediction (Our Work)

Key Points in dealing with EEG time series

- **Individual Variability → Lower Classification Accuracy**
  - ✓ Low SNR
  - ✓ Different brain electrical conductivity ← different anatomical structure of brain
  - ✓ Electrodes' positional error

Feature Extraction

EEG Electrodes' Structure Modeling

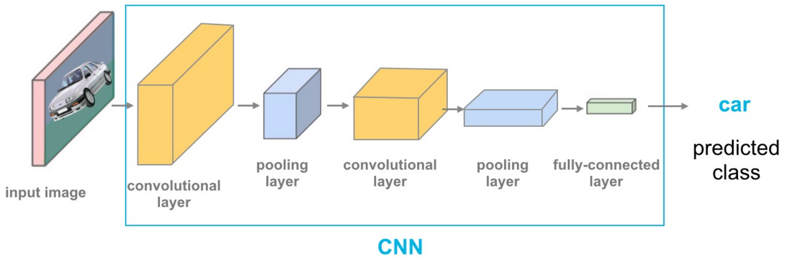
- **Slow Responding → Hard to develop Real-life Applications**
  - ✓ [**most literature**] Trial-level prediction (*e.g.*, 4 s)
  - ✓ Window/Slide-level prediction (*e.g.*, 0.4 s)
  - ✓ Time-resolved prediction (*e.g.*, 6.25 ms) (Our Work)

Time-resolved or Window-based Signal Sampling

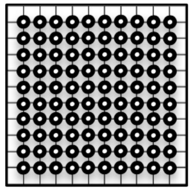
- **Lower Group-level Accuracy → Hard to develop Applications for a Group of People**
  - ✓ [**most literature**] Subject-level prediction (Our Work)
  - ✓ Group-level prediction (Our Work)

Motivation

Convolutional Neural Networks:



- **Module: Convolution → Pooling → Fully-connected**
- **Modeling: Euclidean-Structured Data** (*e.g.*, Image, Speech, Natural Language)

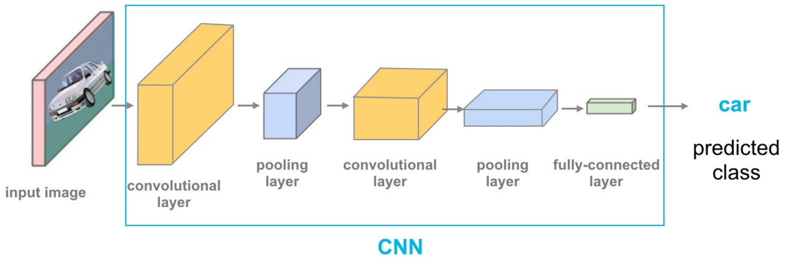


- **Neuroscience** research has increasingly emphasized **Brain Network Dynamics**
  - Model **Functional Topological Connectivity** of EEG Electrodes → **Graph** (Non-Euclidean Structure)

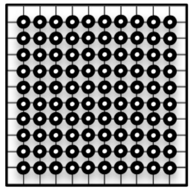
Image Credit: The PhysioNet Dataset and the Functional Network Image is in the public domain.

Motivation

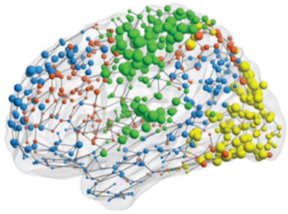
Convolutional Neural Networks:



- **Module**: **Convolution → Pooling → Fully-connected**
- **Modeling**: **Euclidean-Structured Data** (*e.g.*, Image, Speech, Natural Language)

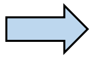


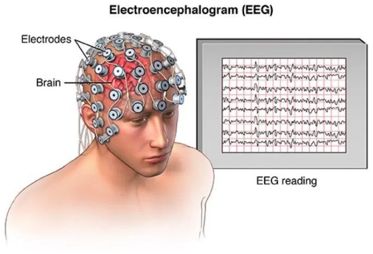
- **Neuroscience** research has increasingly emphasized **Brain Network Dynamics**
  - Model **Functional Topological Connectivity** of EEG Electrodes → **Graph** (Non-Euclidean Structure)

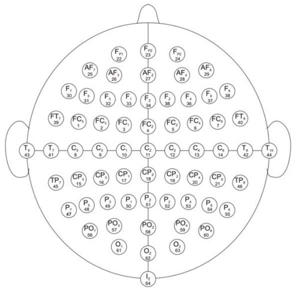


Functional Networks

mapping





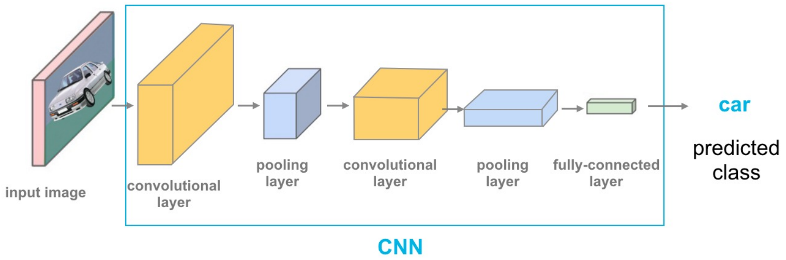


International 10-10 EEG System

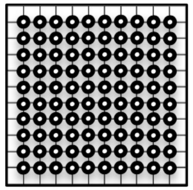
Image Credit: The PhysioNet Dataset and the Functional Network Image is in the public domain.

Motivation

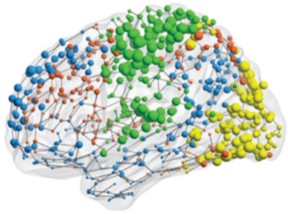
Convolutional Neural Networks:



- **Module: Convolution → Pooling → Fully-connected**
- **Modeling: Euclidean-Structured Data** (*e.g.*, Image, Speech, Natural Language)

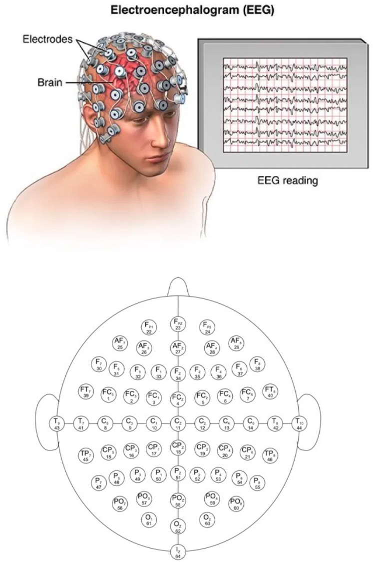


- **Neuroscience** research has increasingly emphasized **Brain Network Dynamics**
  - Model **Functional Topological Connectivity** of EEG Electrodes → **Graph** (Non-Euclidean Structure)



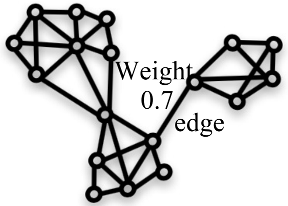
Functional Networks

mapping



International 10-10 EEG System

topology

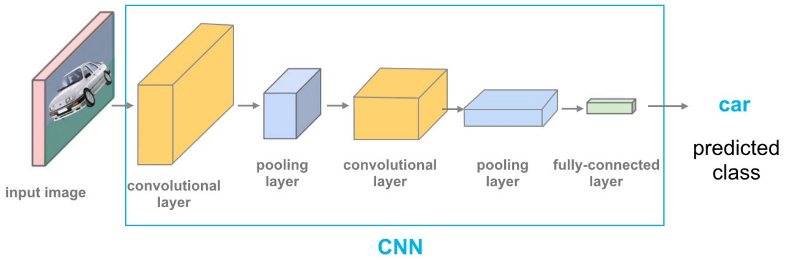


Graph

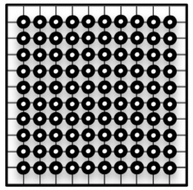
Image Credit: The PhysioNet Dataset and the Functional Network Image is in the public domain.

Motivation

Convolutional Neural Networks:



- **Module**: **Convolution → Pooling → Fully-connected**
- **Modeling**: **Euclidean-Structured Data** (*e.g.*, Image, Speech, Natural Language)



- **Neuroscience** research has increasingly emphasized **Brain Network Dynamics**
  - Model **Functional Topological Connectivity** of EEG Electrodes → **Graph** (Non-Euclidean Structure)

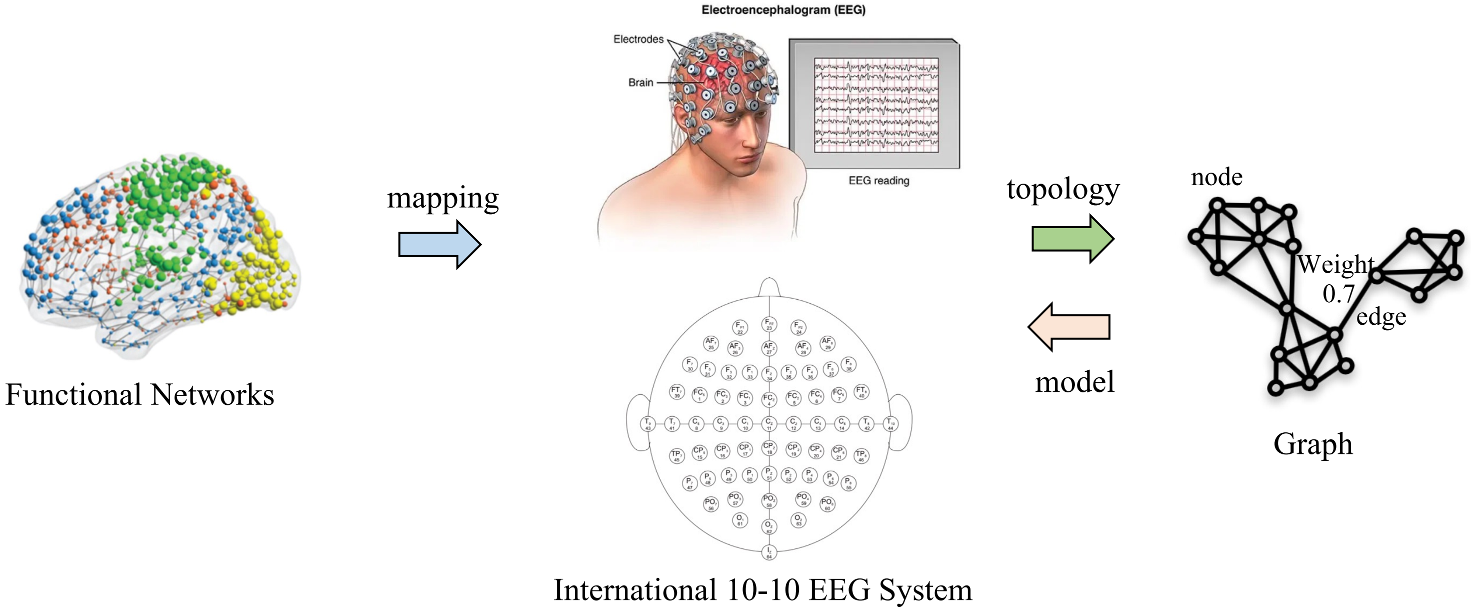
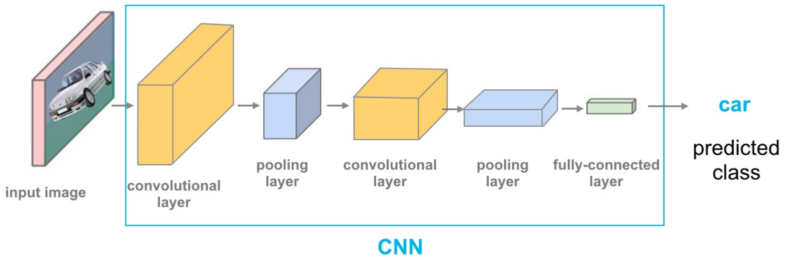
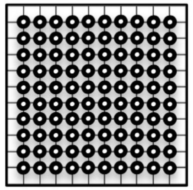


Image Credit: The PhysioNet Dataset and the Functional Network Image is in the public domain.

Motivation

- Convolutional Neural Networks:





  - **Module: Convolution → Pooling → Fully-connected**
  - **Modeling: Euclidean-Structured Data** (*e.g.*, Image, Speech, Natural Language)
- **Neuroscience** research has increasingly emphasized **Brain Network Dynamics**
  - Model **Functional Topological Connectivity** of EEG Electrodes → **Graph** (Non-Euclidean Structure)

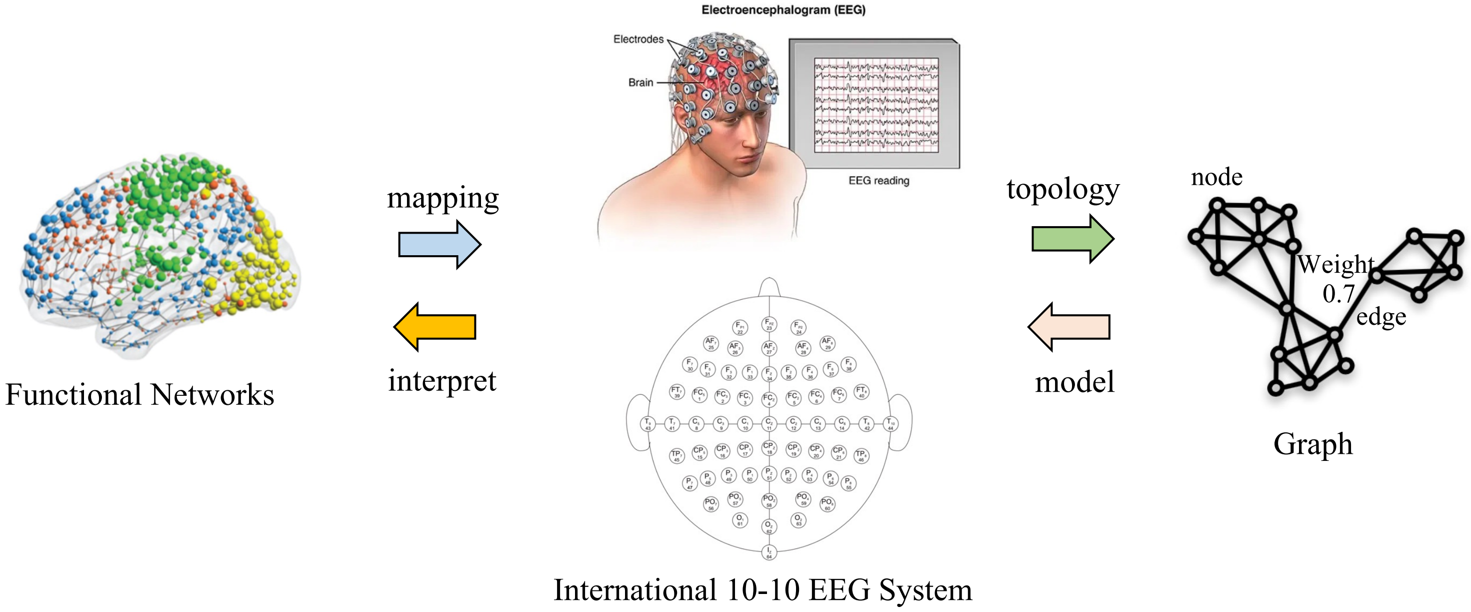
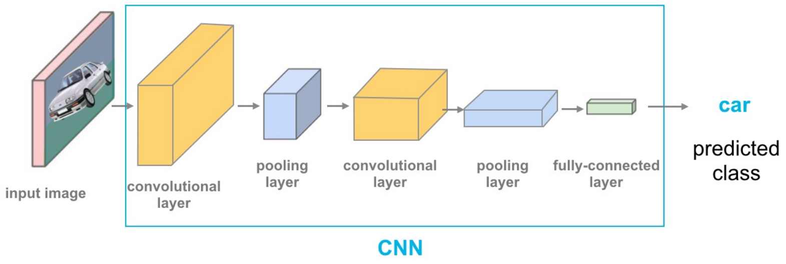


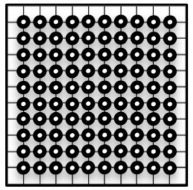
Image Credit: The PhysioNet Dataset and the Functional Network Image is in the public domain.

Motivation

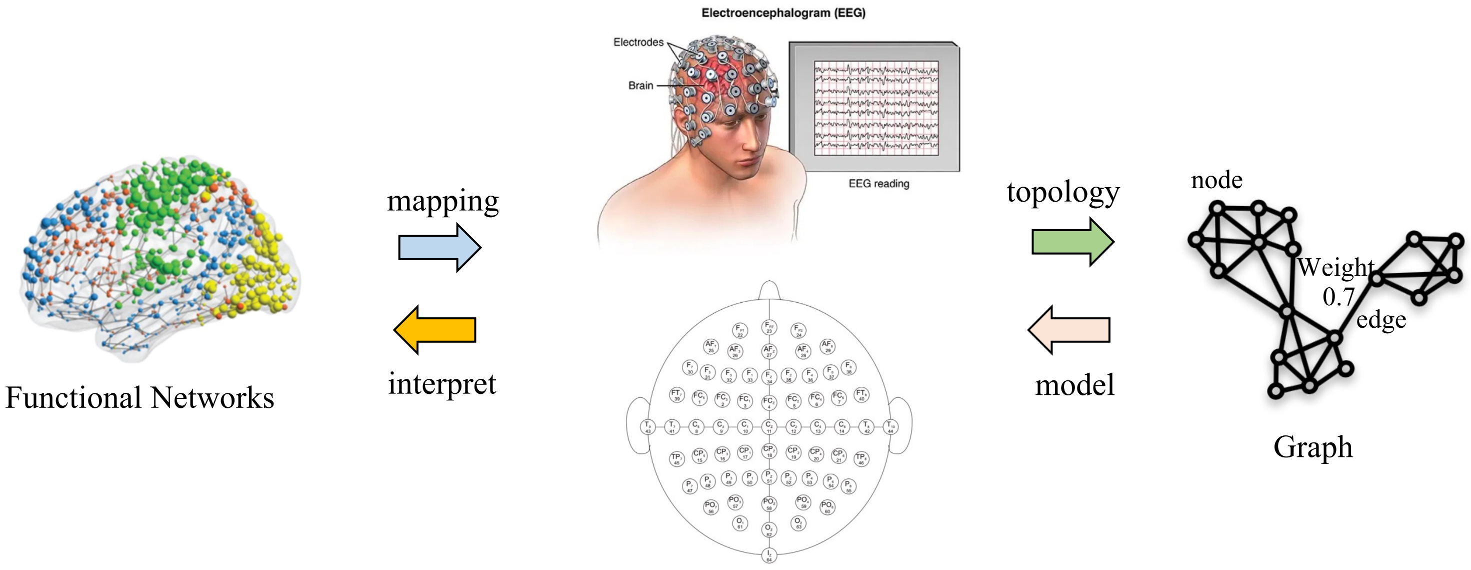
Convolutional Neural Networks:



- **Module: Convolution → Pooling → Fully-connected**
- **Modeling: Euclidean-Structured Data** (*e.g.*, Image, Speech, Natural Language)



- **Neuroscience** research has increasingly emphasized **Brain Network Dynamics**
  - Model **Functional Topological Connectivity** of EEG Electrodes → **Graph** (Non-Euclidean Structure)



International 10-10 EEG System

Our Question

How to model the EEG System as a **Graph**?

How can we process EEG Signals via **Graph Representation Learning**?

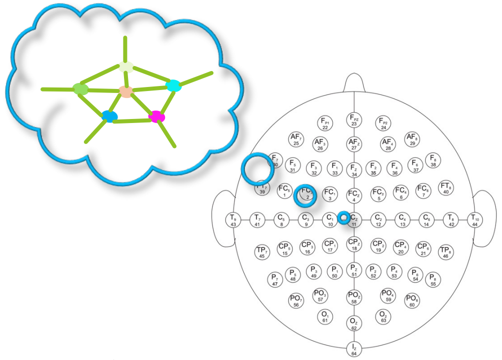
Image Credit: The PhysioNet Dataset and the Functional Network Image is in the public domain.

Can we directly apply convolutions on graphs?

- Traditional CNN **cannot** directly process graph signals
  - **Graph is irregular** (*i.e.*, unordered and vary in size)
  - Convolution **cannot** keep **Translation Invariance** on non-Euclidean signals
- **Graph Convolutional Neural Networks** (GCN)
  - Directly process **non-Euclidean graph-structured signals**
  - Consider **relational properties** (*e.g.*, correlations) between nodes

    → Model **Functional Topological Relationships** among EEG electrodes

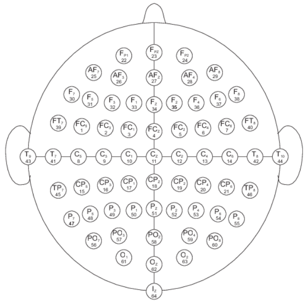
    → Analyze and interpret **Brain Network Dynamics**

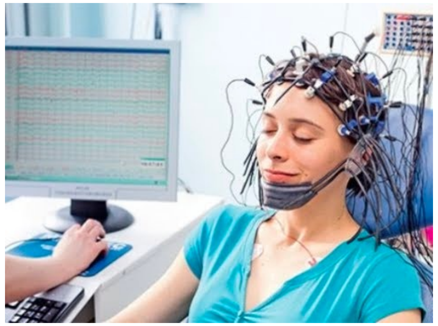


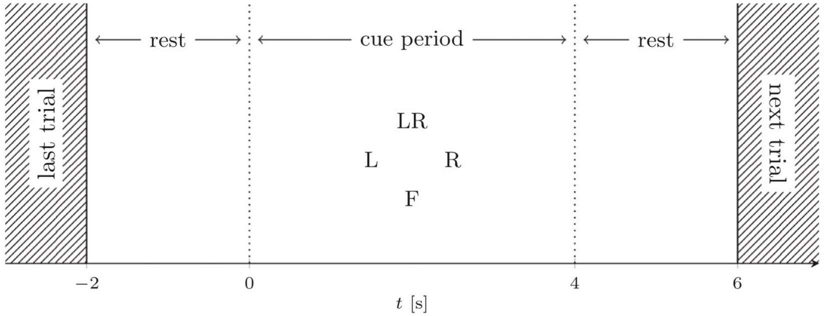
Benchmark Dataset

- The PhysioNet Dataset (EEG Motor Movement/Imagery Dataset)
- International 10-10 EEG System → **64 electrodes**

  (excluding electrodes Nz, F9, F10, FT9, FT10, A1, A2, TP9, TP10, P9, and P10)



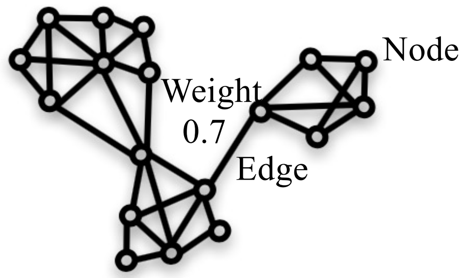




- **109 subjects (the largest number of participants in the field of EEG Motor Imagery)**
- Task: **4-class EEG Motor Imagery Classification**
  - ✓ Imagining (Task 1) left fist, (Task 2) right fist, (Task 3) both fists, (Task 4) both feet
- Each subject → 3 runs, 7 trials, 4 classes → 84 trials in total
- Each trial → 4 seconds experimental duration, 160 Hz Sampling Rate → 640 Time Points
- We apply the **Time-resolved Sampling Method**
  - ✓ Total samples per subject: 3 runs × 7 trials × 4 classes × 4 seconds × 160 Hz = 53,760 samples
  - ✓ Experimental Setting: 90% as the training set and the left 10% as the test set

Image Credit: The PhysioNet Dataset and the middle image is in the public domain.

Preliminary: Graph Representation



Definition: An Undirected and Weighted Graph with N nodes: $\mathbf{G} = \{\mathbf{V}, \mathbf{E}, \mathbf{A}\}$

- **V**: nodes (vertices), $|\mathbf{V}| = N$
- **E**: edges (links) that connect nodes
- **A**: weights (correlations) between nodes

Nodes Correlations: Pearson Matrix $\mathbf{P} \in \mathrm{R}^{N\times N}$ (denotes as PCC matrix)

- Measure the linear correlations between node **x** and node **y**
- μ is the mean, σ is the standard deviation, and $P_{x,y}$ is the Pearson Correlation Coefficient between node **x** and node **y**

$$P_{x,y} = \frac{\mathrm{E}((\mathbf{x} - \mu_x)(\mathbf{y} - \mu_y))}{\sigma_x \sigma_y}$$

- Absolute Pearson Matrix: $|\mathbf{P}| \in \mathrm{R}^{N\times N}$ and $|P_{ij}| \in [0, 1] \rightarrow$ Note: In this work, we only consider **scale**.

Graph Weights: Adjacency Matrix $\mathbf{A} = |\mathbf{P}| - \mathbf{I} \in \mathrm{R}^{N\times N}$, where **I** is an Identity Matrix

Graph Degrees: Degree Matrix $\mathbf{D} \in \mathrm{R}^{N\times N}$

$$D_{ii} = \sum_{j=1}^{N} A_{ij}$$

Graph Representation: Combinatorial Laplacian $\mathbf{L} \in \mathrm{R}^{N\times N}$

$$\mathbf{L} = \mathbf{D} - \mathbf{A}$$

Normalized:

$$\mathbf{L} = \mathbf{I} - \mathbf{D}^{-\frac{1}{2}}\mathbf{A}\mathbf{D}^{\frac{1}{2}}$$

Preliminary:

Spectral Theorem for Graph Laplacian **L**

$$\mathbf{L} = \mathbf{U\Lambda U}^{\mathrm{T}}$$

$$\mathbf{LU} = \mathbf{\Lambda U}$$

– **U**: Fourier basis → **real** and **orthonormal** eigenvectors of **L**

– **Λ**: Fourier modes → the diagonal is the **ordered** and **real nonnegative** eigenvalues of **L**

Graph Fourier Transforms of Signal f

can be seen as the $\mathrm{e}^{-j\omega \mathrm{t}}$ in Fourier Transforms

$$\mathrm{F}[f(\boldsymbol{\lambda})] = \hat{f}(\boldsymbol{\lambda}) = \sum_{i=1}^{\mathrm{n}} f(i) \times U(i)$$

$$\hat{f}(\boldsymbol{\lambda}) = \mathbf{U}^{\mathrm{T}} f \Leftrightarrow f = \mathbf{U}\hat{f}(\boldsymbol{\lambda})$$

$\hat{f}(\boldsymbol{\lambda})$ is the projection value of the Fourier basis **U**

Preliminary: Graph Convolution via Graph Fourier Transform

Notation:

Signal f

Signal h

F: Fourier Transforms

F^{-1}: Inverse Fourier Transforms

$\hat{f}(w)$: F(f)

$\hat{h}(w)$: F(h)

Note: Fourier Transforms of **Convolution in the spatial domain**

⇔

Point-wise Multiplication of two Fourier transformed signals

$$\mathrm{F}\big((f * h)_{\mathbf{G}}\big) = \hat{f}(w)\times\hat{h}(w)$$

Convolution $(f * h)_{\mathbf{G}} = \mathrm{F}^{-1}\big(\hat{f}(w)\times\hat{h}(w)\big)$

$\downarrow$ $\hat{f}(\boldsymbol{\lambda}) = \mathbf{U}^{\mathrm{T}} f$

Hadamard Product
(Element-wise Multiplication)

$$(f * h)_{\mathbf{G}} = \mathrm{F}^{-1}\Big(\big(\mathbf{U}^{\mathrm{T}} f\big) \odot (\mathbf{U}^{\mathrm{T}} h)\Big)$$

$\downarrow$ $f = \mathbf{U}\hat{f}(\boldsymbol{\lambda})$

$$(f * h)_{\mathbf{G}} = \mathbf{U}\Big(\big(\mathbf{U}^{\mathrm{T}} f\big) \odot (\mathbf{U}^{\mathrm{T}} h)\Big)$$

$\downarrow$

$$(f * h)_{\mathbf{G}} = \underbrace{\mathbf{U}}_{[\mathrm{n}\times\mathrm{n}]}\ \underbrace{\mathrm{diag}[\hat{h}(\lambda_1), \hat{h}(\lambda_2), \ldots, \hat{h}(\lambda_{\mathrm{n}})]}_{[\mathrm{n}\times\mathrm{n}]}\ \underbrace{\mathbf{U}^{\mathrm{T}}}_{[\mathrm{n}\times\mathrm{n}]}\ \underbrace{f}_{[\mathrm{n}\times\mathrm{d}]}$$

Source: https://en.wikipedia.org/wiki/Convolution_theorem

Graph Convolution

$$(f * h)_{\mathbf{G}} = \mathbf{U}\ \boxed{\mathrm{diag}[\hat{h}(\lambda_1), \hat{h}(\lambda_2), \ldots, \hat{h}(\lambda_{\mathrm{n}})]}\ \mathbf{U}^{\mathrm{T}} f$$

Activation Function

$$\mathcal{Y} = \sigma(\mathbf{U}\mathbf{g}_{\boldsymbol{\theta}}\mathbf{U}^{\mathrm{T}}\boldsymbol{\chi})$$

K^{th} Polynomial Function

$$\mathbf{g}_{\boldsymbol{\theta}}(\boldsymbol{\Lambda}) = \sum_{k=0}^{K} \theta_k \boldsymbol{\Lambda}^k$$

Approximate

$$\mathcal{Y} = \sigma(\mathbf{U}\mathbf{g}_{\boldsymbol{\theta}}(\boldsymbol{\Lambda})\mathbf{U}^{\mathrm{T}}\boldsymbol{\chi})$$
$$\boldsymbol{\Lambda} = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_{\mathrm{n}})$$

$$\mathcal{Y} = \sigma\left(\mathbf{U}\sum_{k=0}^{K} \theta_k \boldsymbol{\Lambda}^k\ \mathbf{U}^{\mathrm{T}}\boldsymbol{\chi}\right) = \sigma\left(\sum_{k=0}^{K} \theta_k\ (\mathbf{U}\boldsymbol{\Lambda}^k\mathbf{U}^{\mathrm{T}})\boldsymbol{\chi}\right) = \sigma\left(\sum_{k=0}^{K} \theta_k\ (\mathbf{U}\boldsymbol{\Lambda}\mathbf{U}^{\mathrm{T}})^k\boldsymbol{\chi}\right) = \sigma\left(\sum_{k=0}^{K} \theta_k\ \mathbf{L}^k\boldsymbol{\chi}\right)$$

$$\boxed{\mathcal{Y} = \sigma\left(\sum_{k=0}^{K} \theta_k\ \mathbf{L}^k\boldsymbol{\chi}\right)}$$

Credit: Defferrard *et al.*, Convolutional Neural Networks on Graphs with Fast Localized Spectral Filtering, In NeurIPS 2016.

Graph Convolution

Node Aggregation
K is Filter Size

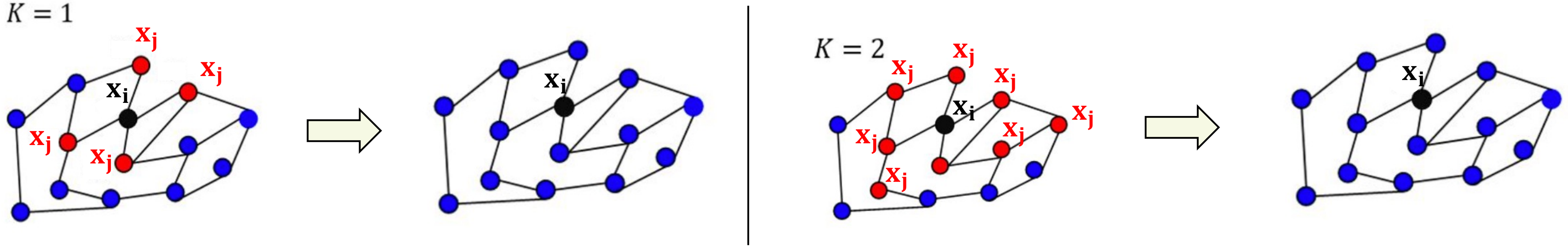
$$\mathcal{Y} = \sigma\left(\sum_{k=0}^{K} \theta_k \, \mathbf{L}^k \boldsymbol{\chi}\right)$$

GCN Key Idea: Use "**edge information**" to **aggregate** "**node information**" to generate a new "**node representation**"

Localize in Space

Laplace Operator
Local connectivity

$$\mathbf{x_{new}} \leftarrow \mathbf{Lx_i} = \sum_j A_{ij}(\mathbf{x_i} - \mathbf{x_j})$$



Pros:

1. No need for Spectral Decomposition of $\mathbf{L}$
2. Less number of parameters (decrease model complexity) $\rightarrow K \ll N$

Cons: Need to compute $\mathbf{L}^k$

Graph Convolution

Convolution:
Weighted Sum

Node Aggregation
K is Filter Size

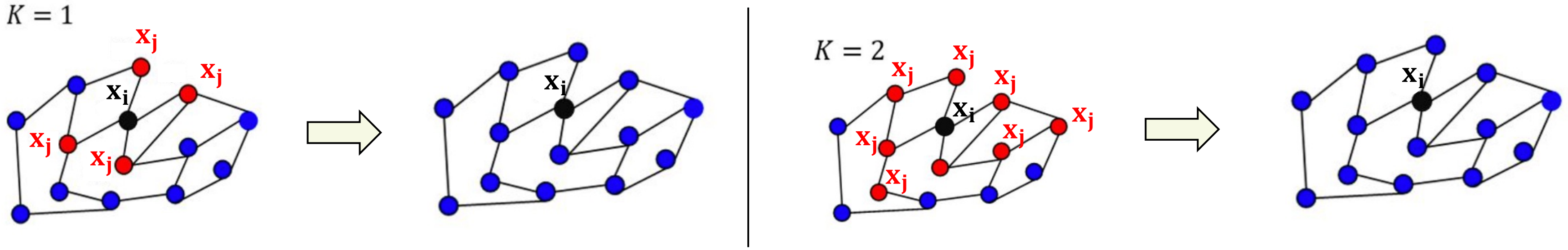
$$\mathcal{Y} = \sigma\left(\sum_{k=0}^{K} \boxed{\theta_k\, \mathbf{L}^k \boldsymbol{\chi}}\right)$$

GCN Key Idea: Use "**edge information**" to **aggregate** "**node information**" to generate a new "**node representation**"

Laplace Operator
Local connectivity

$$\mathbf{x}_{\mathbf{new}} \leftarrow \mathbf{L}\mathbf{x}_{\mathbf{i}} = \sum_{j} A_{ij}(\mathbf{x}_{\mathbf{i}} - \mathbf{x}_{\mathbf{j}})$$

Localize in Space



Pros:

1. No need for Spectral Decomposition of $\mathbf{L}$
2. Less number of parameters (decrease model complexity) $\rightarrow K \ll N$

Cons: Need to compute $\mathbf{L}^k$

Graph Convolution

Node Aggregation
K is Filter Size

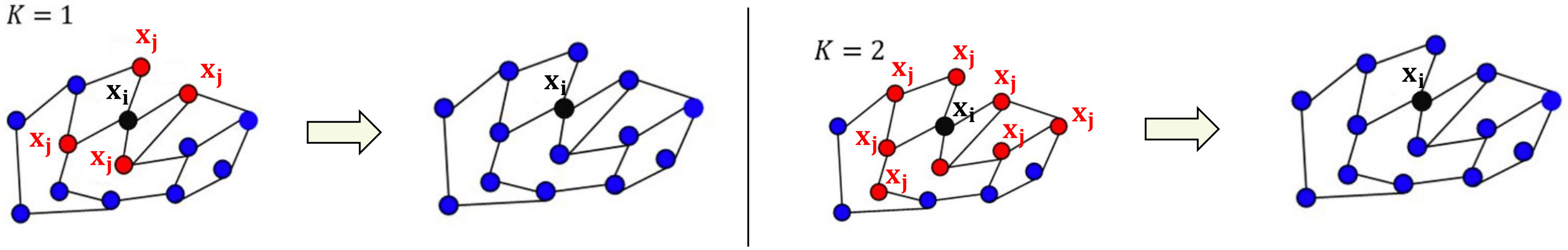
$$\mathcal{Y} = \sigma\left(\sum_{k=0}^{K} \theta_k \, \mathbf{L}^k \boldsymbol{\chi}\right)$$

GCN Key Idea: Use "**edge information**" to **aggregate** "**node information**" to generate a new "**node representation**"

Laplace Operator
Local connectivity

$$\mathbf{x_{new}} \leftarrow \mathbf{L}\mathbf{x_i} = \sum_j A_{ij}(\mathbf{x_i} - \mathbf{x_j})$$

Localize in Space



Pros:

1. No need for Spectral Decomposition of $\mathbf{L}$
2. Less number of parameters (decrease model complexity) $\rightarrow K \ll N$

Cons: Need to compute $\mathbf{L}^k$

Graph Convolution

Node Aggregation

K is Filter Size

$$\boldsymbol{y} = \sigma\left(\sum_{k=0}^{K} \theta_k \boxed{\mathbf{L}^k \boldsymbol{\chi}}\right)$$

No need for Fourier Transform

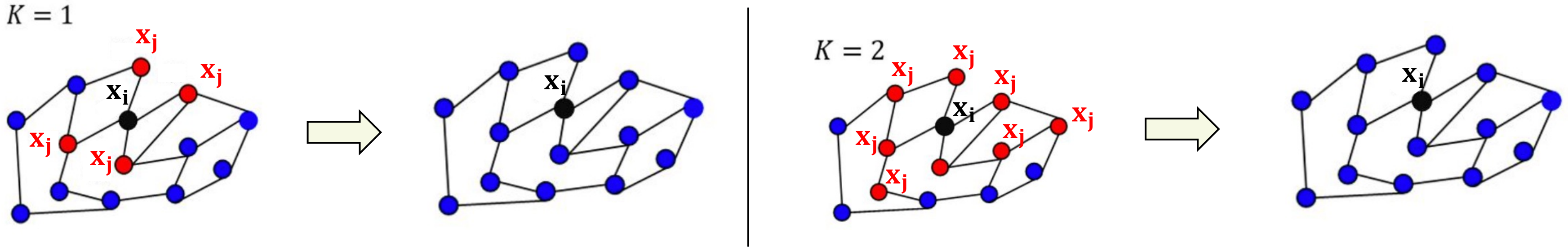
GCN Key Idea: Use "**edge information**" to **aggregate** "**node information**" to generate a new "**node representation**"

Laplace Operator

Local connectivity

$$\mathbf{x_{new}} \leftarrow \mathbf{L}\mathbf{x_i} = \sum_j A_{ij}(\mathbf{x_i} - \mathbf{x_j})$$

Localize in Space



Pros:

1. No need for Spectral Decomposition of $\mathbf{L}$
2. Less number of parameters (decrease model complexity) $\rightarrow K \ll N$

Cons: Need to compute $\mathbf{L}^k$

Graph Convolution

Node Aggregation
K is Filter Size

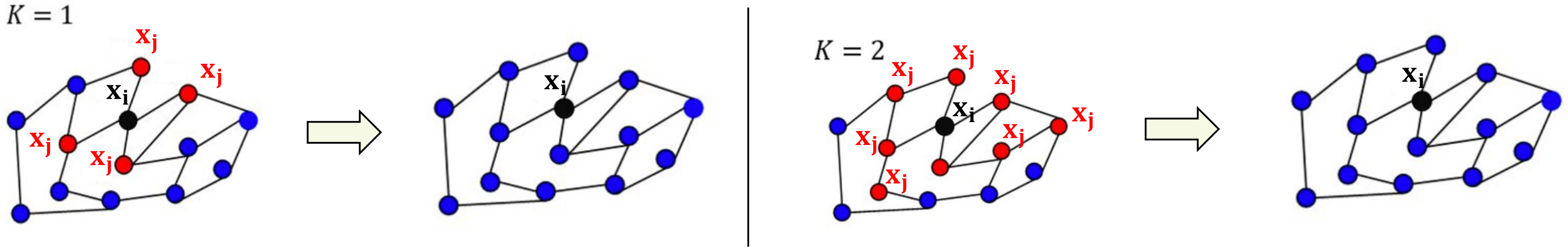
$$\boldsymbol{\mathcal{Y}} = \sigma\left(\sum_{k=0}^{K} \theta_k \, \mathbf{L}^k \boldsymbol{\chi}\right)$$

GCN Key Idea: Use "**edge information**" to **aggregate** "**node information**" to generate a new "**node representation**"

Laplace Operator
Local connectivity

$$\mathbf{x}_{\mathbf{new}} \leftarrow \mathbf{L}\mathbf{x}_{\mathbf{i}} = \sum_{j} A_{ij}(\mathbf{x}_{\mathbf{i}} - \mathbf{x}_{\mathbf{j}})$$

Localize in Space



Pros:

1. No need for Spectral Decomposition of $\mathbf{L}$
2. Less number of parameters (decrease model complexity) $\rightarrow K \ll N$

Cons: Need to compute $\mathbf{L}^k$

Graph Convolution

Node Aggregation

K is Filter Size

Weight Sharing

$$\mathcal{Y} = \sigma\left(\sum_{k=0}^{K} \boxed{\theta_k}\, \mathbf{L}^k \boldsymbol{\chi}\right)$$

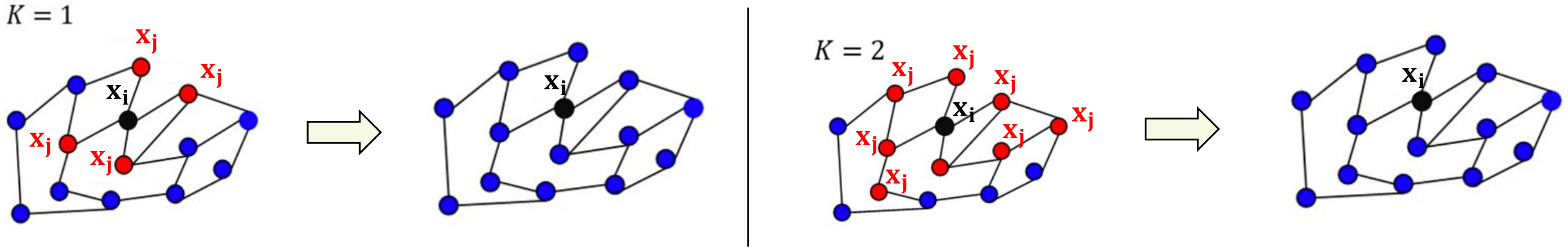
GCN Key Idea: Use "**edge information**" to **aggregate** "**node information**" to generate a new "**node representation**"

Laplace Operator

Local connectivity

$$\mathbf{x_{new}} \leftarrow \mathbf{L}\mathbf{x_i} = \sum_j A_{ij}(\mathbf{x_i} - \mathbf{x_j})$$

Localize in Space



Pros:

1. No need for Spectral Decomposition of $\mathbf{L}$
2. Less number of parameters (decrease model complexity) $\rightarrow K \ll N$

Cons: Need to compute $\mathbf{L}^k$

Graph Convolution

Node Aggregation

K is Filter Size

$$\mathcal{Y} = \sigma\left(\sum_{k=0}^{K} \theta_k \, \mathbf{L}^k \boldsymbol{\chi}\right)$$

Beauty is in Simplicity

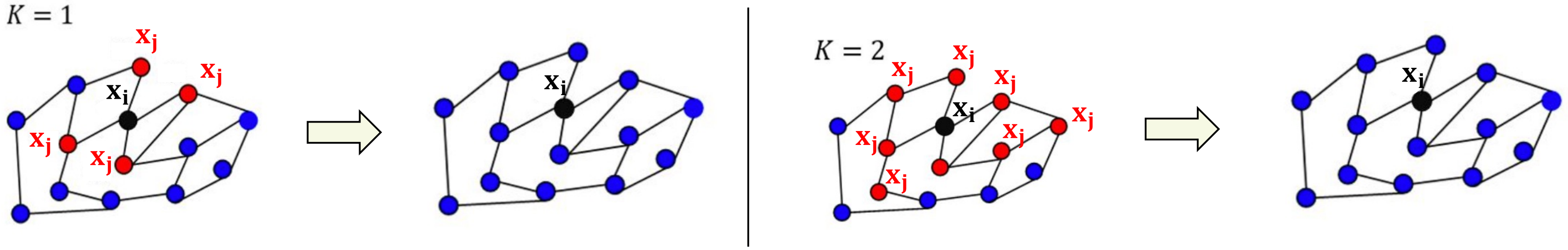
GCN Key Idea: Use "**edge information**" to **aggregate** "**node information**" to generate a new "**node representation**"

Laplace Operator

Local connectivity

$$\mathbf{x_{new}} \leftarrow \mathbf{Lx_i} = \sum_j A_{ij}(\mathbf{x_i} - \mathbf{x_j})$$

Localize in Space



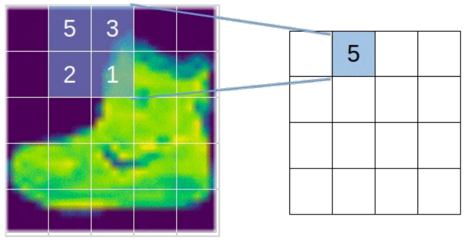
Pros:

1. No need for Spectral Decomposition of $\mathbf{L}$

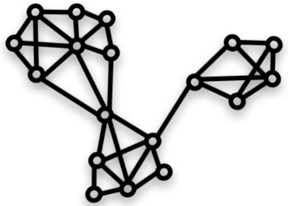
2. Less number of parameters (decrease model complexity) $\rightarrow K \ll N$

Cons: Need to compute $\mathbf{L}^k$

Pooling on Graphs (Graph Coarsening)



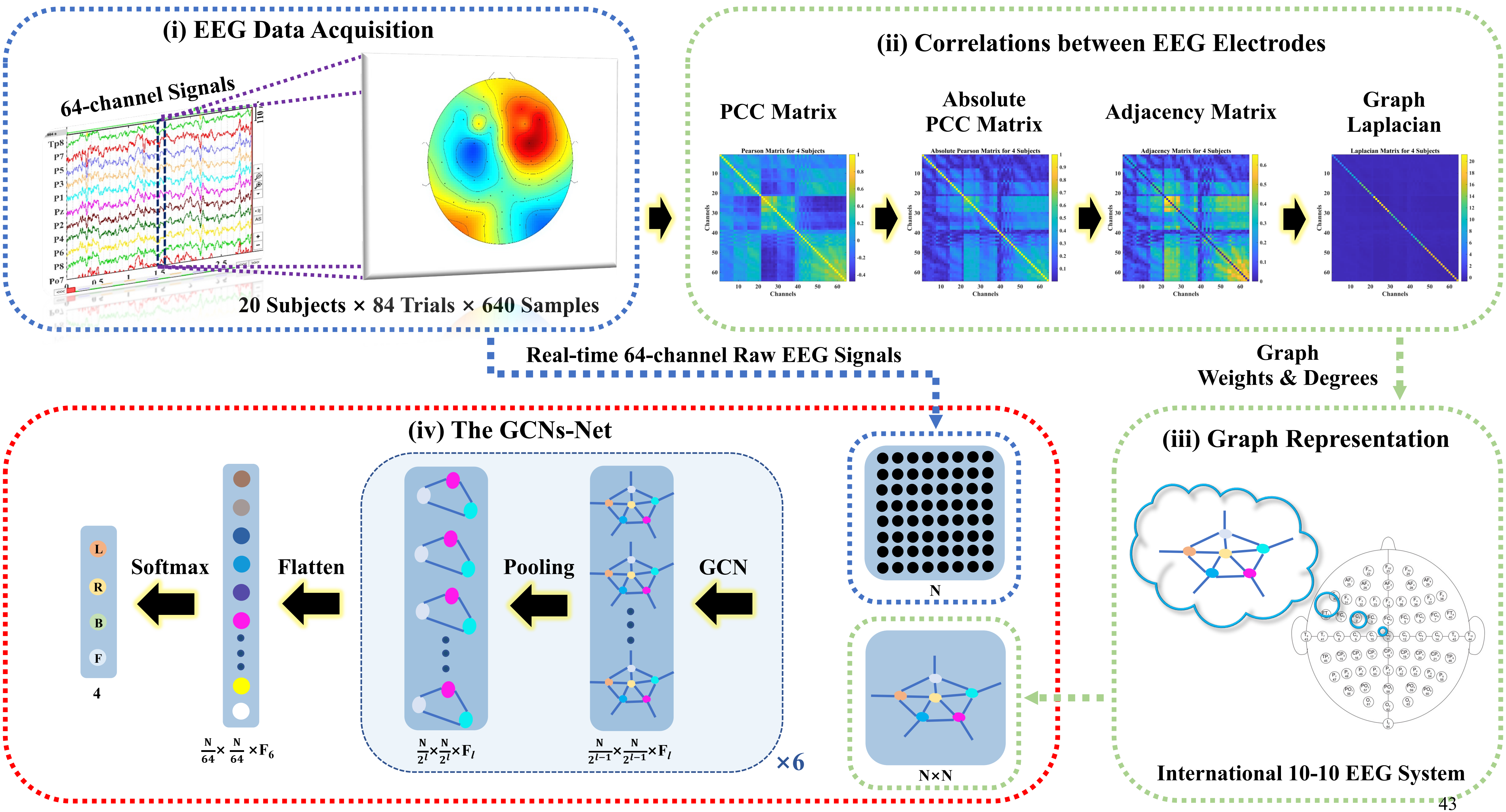
- Traditional CNN doesn't need to consider neighbors after convolutions
  - [**Euclidean Structure**] The output Feature Maps are "regular"
  - The neighbor is "meaningful"
- GCNs need to consider neighbors after convolutions
  - [**Non-Euclidean Structure**] The output **graphs' nodes are not arranged in any meaningful way**
  - Use **Graclus Multilevel Clustering Algorithm** to find "meaningful" neighbors
  - Minimize the ***Local Normalized Cut*** (a cluster grouping method)



$$-W_{ij}\left(\frac{1}{d_i}+\frac{1}{d_j}\right)$$

  - i and j denote node i and node j
  - W_{ij} is the **learned weight** between node i and node j

Image Credit: in the public domain.



(i) EEG Data Acquisition
64-channel Signals
20 Subjects × 84 Trials × 640 Samples
(ii) Correlations between EEG Electrodes
PCC Matrix
Absolute PCC Matrix
Adjacency Matrix
Graph Laplacian
Pearson Matrix for 4 Subjects
Absolute Pearson Matrix for 4 Subjects
Adjacency Matrix for 4 Subjects
Laplacian Matrix for 4 Subjects
Channels
Real-time 64-channel Raw EEG Signals
Graph Weights & Degrees
(iii) Graph Representation
International 10-10 EEG System
(iv) The GCNs-Net
Softmax
Flatten
Pooling
GCN
L
R
B
F
4
$\frac{N}{64}\times\frac{N}{64}\times F_6$
$\frac{N}{2^l}\times\frac{N}{2^l}\times F_l$
$\frac{N}{2^{l-1}}\times\frac{N}{2^{l-1}}\times F_l$
×6
N
N×N

Correlation among EEG electrodes
Two Subjects: Subject 10 and 5

Problem: **Individual Variability**

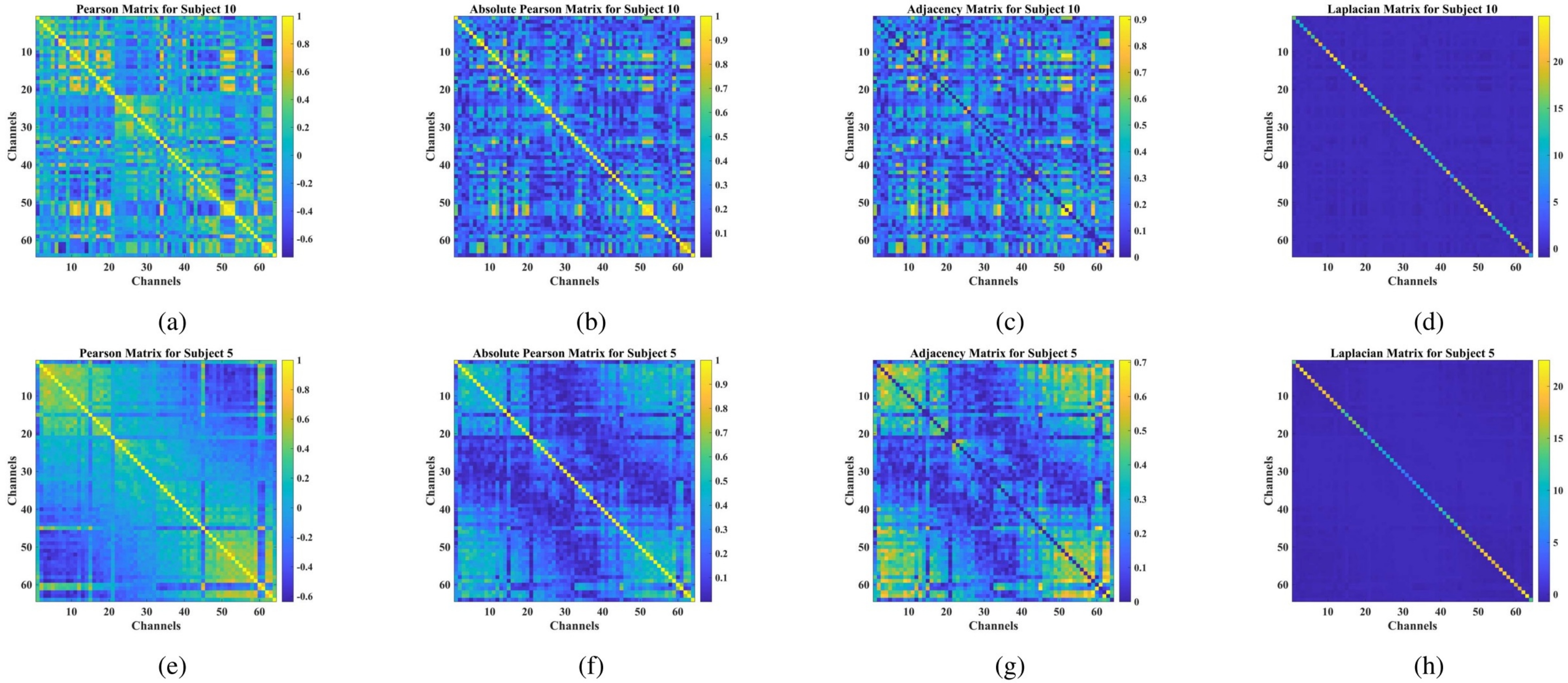


Fig. 6. PCC matrix, absolute PCC matrix, adjacency matrix, and graph Laplacian for Subjects 10 and 5 from the PhysioNet dataset. (a) PCC matrix for Subject 10. (b) Absolute PCC matrix for Subject 10. (c) Adjacency matrix for Subject 10. (d) Graph Laplacian for Subject 10. (e) PCC matrix for Subject 5. (f) Absolute PCC matrix for Subject 5. (g) Adjacency matrix for Subject 5. (h) Graph Laplacian for Subject 5.

Correlation among EEG electrodes
20 Subjects and 100 Subjects

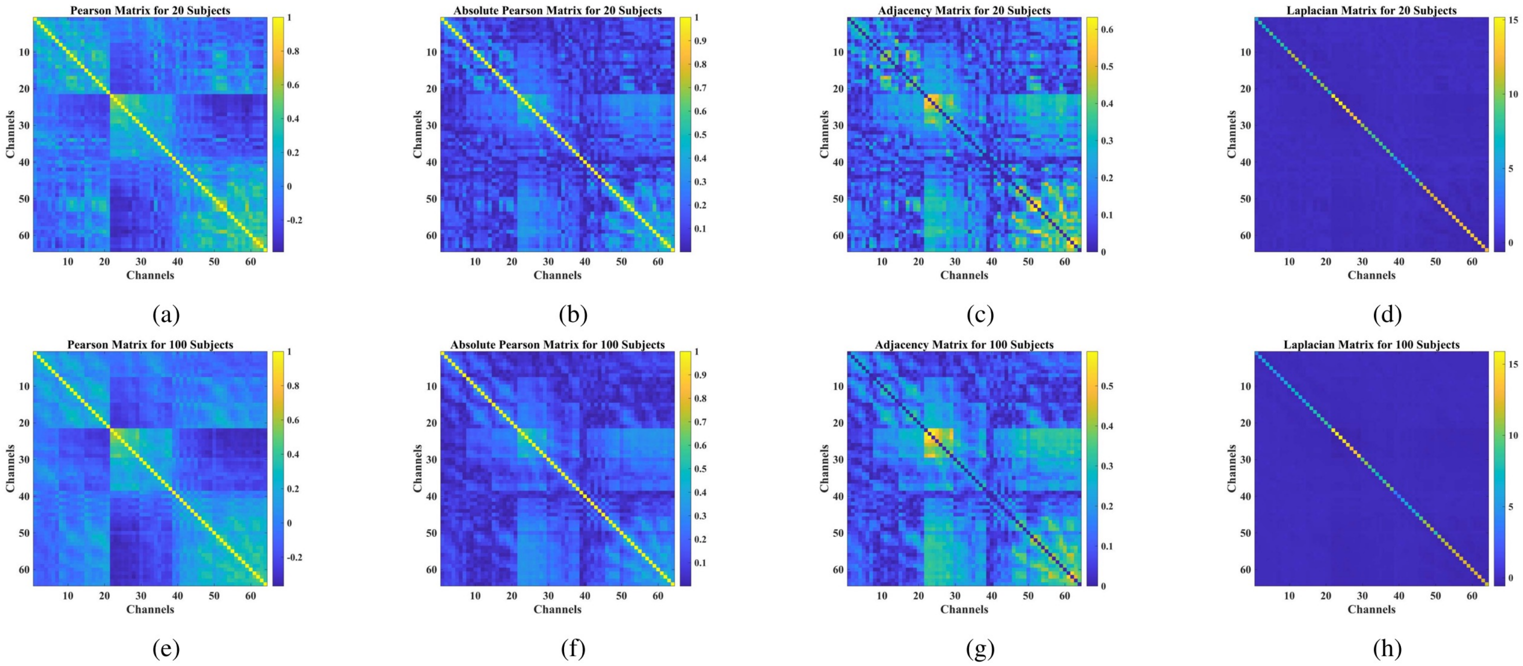


Fig. 2. PCC matrix, absolute PCC matrix, adjacency matrix, and graph Laplacian for 20 and 100 subjects, respectively, from the PhysioNet dataset. (a) PCC matrix for 20 subjects. (b) Absolute PCC matrix for 20 subjects. (c) Adjacency matrix for 20 subjects. (d) Graph Laplacian for 20 subjects. (e) PCC matrix for 100 subjects. (f) Absolute PCC matrix for 100 subjects. (g) Adjacency matrix for 100 subjects. (h) Graph Laplacian for 100 subjects.

Increasing ***the number of subjects*** alleviates ***individual variability***

Model Design for 64-electrode EEG System

TABLE I
IMPLEMENTATION DETAILS OF THE PROPOSED GCNS-NET ON THE PHYSIONET DATASET

Layer	Type	Maps	Size	Edges	Polynomial Order	Pooling Size	Activation	Weights	Bias
Softmax	Fully-connected	–	O	–	–	–	Softmax	$\frac{N}{64} \times \frac{N}{64} \times F_6 \times O$	O
Flatten	Flatten	–	$\frac{N}{64} \times \frac{N}{64} \times F_6$	–	–	–	–	–	–
P6	Max-pooling	F_6	$\frac{N}{32}$	$\sum_{i=1}^{\frac{N}{32}-1} i$	–	2	–	–	–
C6	Convolution	F_6	$\frac{N}{32}$	$\sum_{i=1}^{\frac{N}{32}-1} i$	K	–	Softplus	$F_5 \times F_6 \times K$	$\frac{N}{32} \times F_6$
P5	Max-pooling	F_5	$\frac{N}{16}$	$\sum_{i=1}^{\frac{N}{16}-1} i$	–	2	–	–	–
C5	Convolution	F_5	$\frac{N}{16}$	$\sum_{i=1}^{\frac{N}{16}-1} i$	K	–	Softplus	$F_4 \times F_5 \times K$	$\frac{N}{16} \times F_5$
P4	Max-pooling	F_4	$\frac{N}{8}$	$\sum_{i=1}^{\frac{N}{8}-1} i$	–	2	–	–	–
C4	Convolution	F_4	$\frac{N}{8}$	$\sum_{i=1}^{\frac{N}{8}-1} i$	K	–	Softplus	$F_3 \times F_4 \times K$	$\frac{N}{8} \times F_4$
P3	Max-pooling	F_3	$\frac{N}{4}$	$\sum_{i=1}^{\frac{N}{4}-1} i$	–	2	–	–	–
C3	Convolution	F_3	$\frac{N}{4}$	$\sum_{i=1}^{\frac{N}{4}-1} i$	K	–	Softplus	$F_2 \times F_3 \times K$	$\frac{N}{4} \times F_3$
P2	Max-pooling	F_2	$\frac{N}{2}$	$\sum_{i=1}^{\frac{N}{2}-1} i$	–	2	–	–	–
C2	Convolution	F_2	$\frac{N}{2}$	$\sum_{i=1}^{\frac{N}{2}-1} i$	K	–	Softplus	$F_1 \times F_2 \times K$	$\frac{N}{2} \times F_2$
P1	Max-pooling	F_1	N	$\sum_{i=1}^{N-1} i$	–	2	–	–	–
C1	Convolution	F_1	N	$\sum_{i=1}^{N-1} i$	K	–	Softplus	$1 \times F_1 \times K$	$N \times F_1$
Input	Input	1	N	$\sum_{i=1}^{N-1} i$	–	–	–	–	–

Model Optimization

- **Ablation Study**: **Optimal Model Structure** (64-electrode EEG System)
  - C6-P6-K2 with [16, 32, 64, 128, 256, 512] filters
- **Gradient Iterative Solver**: **Adam Optimizer** with **Stochastic Gradient Descent** (SGD) algorithm
  - Learning Rate: 0.01
  - Batch Size: 1,024
- **Activation Function**: **Softplus** (Smooth Rectified Linear Unit)

$$\mathrm{F}(\mathbf{x}) = \log(1 + \mathrm{e}^{\mathbf{x}})$$

- **Model Output**: **Softmax**: **y** are labels, $\hat{\mathbf{y}}$ are the final output tasks

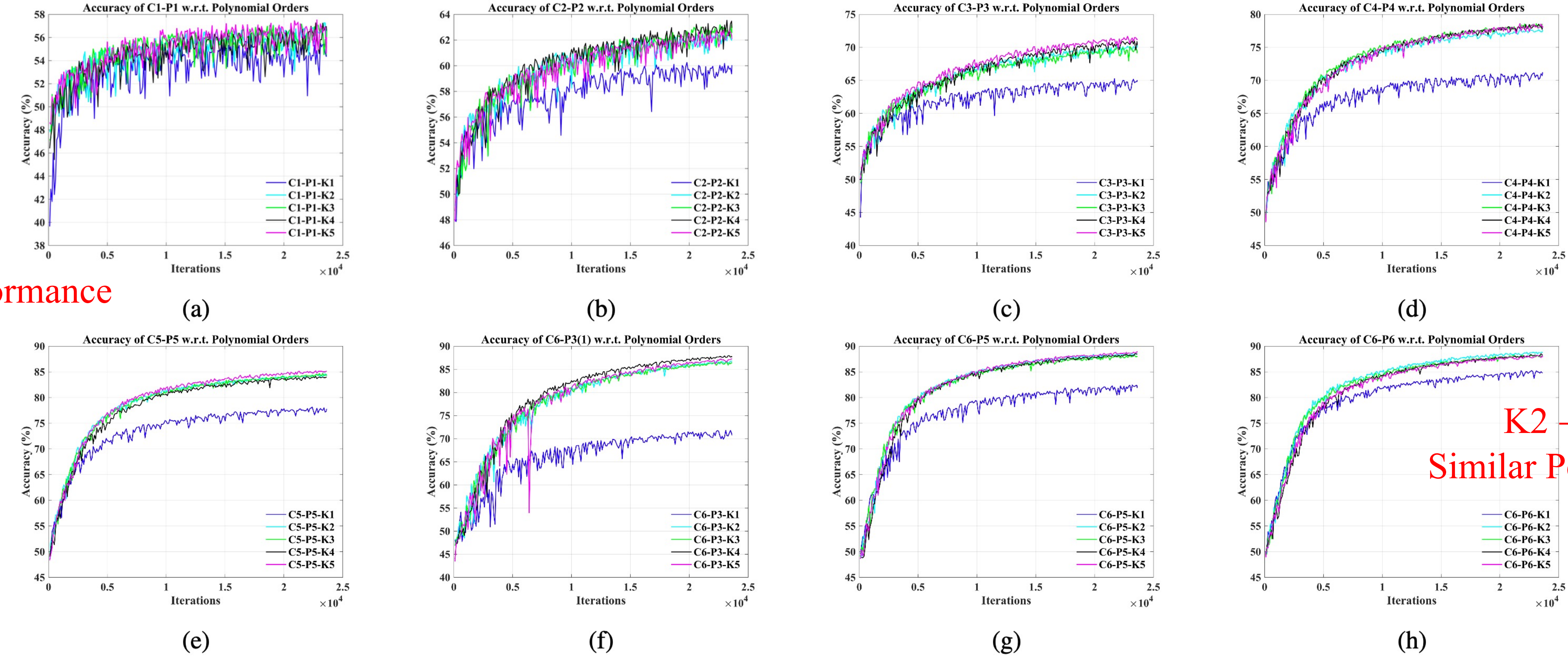
$$\hat{y}_i = \operatorname{argmax}\left(\frac{\mathrm{e}^{y_\mathrm{i}}}{\sum_{\mathrm{i}=1}^{4} \mathrm{e}^{y_\mathrm{i}}}\right)$$

- **Loss Function**: **Cross-entropy Loss** with **L2 regularization**

$$\mathrm{Loss} = -\sum_{\mathrm{i}=1}^{4} y_i \log(\hat{y}_i) + \lambda\left(\sum_{\mathrm{j}=1}^{\mathrm{n}} w_\mathrm{j}^2 + b_\mathrm{j}^2\right)$$

$\lambda = 1 \times 10^{-6}$ is the coefficient of the L2 regularization.

Ablation Study



K1
Poor Performance

K2 → K5
Similar Performance

Fig. 3. Accuracy of some selected models regarding different polynomial approximation order. The models are selected from Table II. (a) Accuracy of the model C1-P1 (model 1). (b) Accuracy of the model C2-P2 (model 3). (c) Accuracy of the model C3-P3 (model 6). (d) Accuracy of the model C4-P4 (model 10). (e) Accuracy of the model C5-P5 (model 14). (f) Accuracy of the model C6-P3 (model 16). (g) Accuracy of the model C6-P5 (model 19). (h) Accuracy of the model C6-P6 (model 20).

Experimental Results
Groupwise Prediction and Subject-specific Adaptation

TABLE IV

PERFORMANCE COMPARISONS ON THE PHYSIONET DATASET

Related Work	Max. Accuracy	Avg. Accuracy	p-value	Level	Approach	Num. of Subjects
Dose *et al*. (2018) [22]	–	58.58%	–	Group	CNNs	105
	80.38%	68.51%	< 0.05	Subject		1
Ma *et al*. (2018) [60]	82.65%	68.20%	–	Group	RNNs	12
Hou *et al*. (2020) [20]	94.50%	–	–	Group	ESI-CNNs	10
	96.00%	–	> 0.05	Subject		1
Hou *et al*. (2022) [34]	94.64%	–	–	Group	BiLSTM-GCN	20
	98.81%	95.48%	> 0.05	Subject		1
Jia *et al*. (2022) [40]	94.16%	93.78%	–	Group	Graph ResNet	20
	98.08%	94.18%	> 0.05	Subject		1
Author	**89.39%**	**88.57%**	–	**Group**	**GCNs-Net**	**20**
	88.14%	–		**Group**		**100**
	98.72%	**93.06%**		**Subject**		**1**

Note: ***p*-value** < 0.05 → Statistically Significant Difference

Takeaways and Future Work

✓ **Graph Representation**

Graph Representation Learning to deeply extract **Network Patterns of Brain Dynamics** for EEG classification.

✓ **Model Converge**

Converge for both **Personalized and Groupwise Predictions**, indicating that the GCNs-Net is able to build a generalized representation of EEG time-series **against both Personalized and Groupwise Variations**.

✓ **Future Work**

Model EEG signals as **Dynamic Graphs** and process them via **Dynamic Graph Representation Learning**.

Deep Feature Mining via Attention-based BiLSTM-GCN for Human Motor Imagery Recognition

Yimin Hou [1], Shuyue Jia [1, 2 *], Xiangmin Lun [1], Shu Zhang [3], Tao Chen [1], Fang Wang [1], and Jinglei Lv [4]

[1] School of Automation Engineering, Northeast Electric Power University

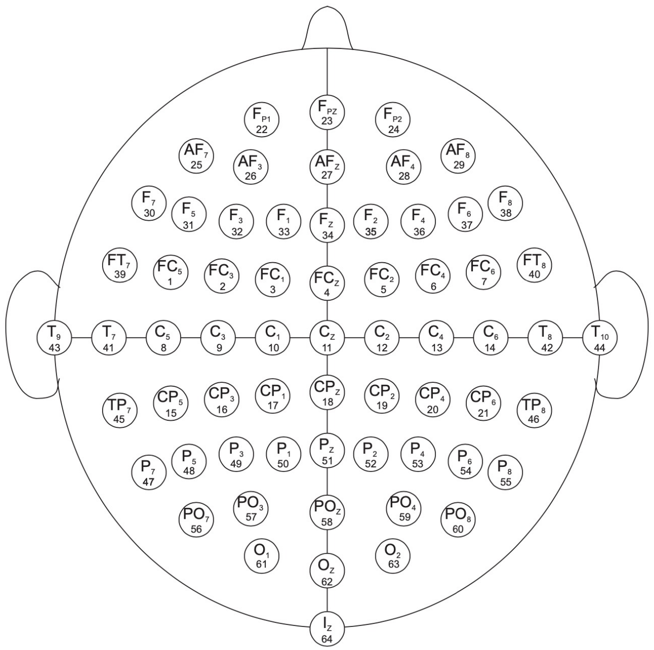
[2] Department of Computer Science, City University of Hong Kong

[3] School of Computer Science, Northwestern Polytechnical University

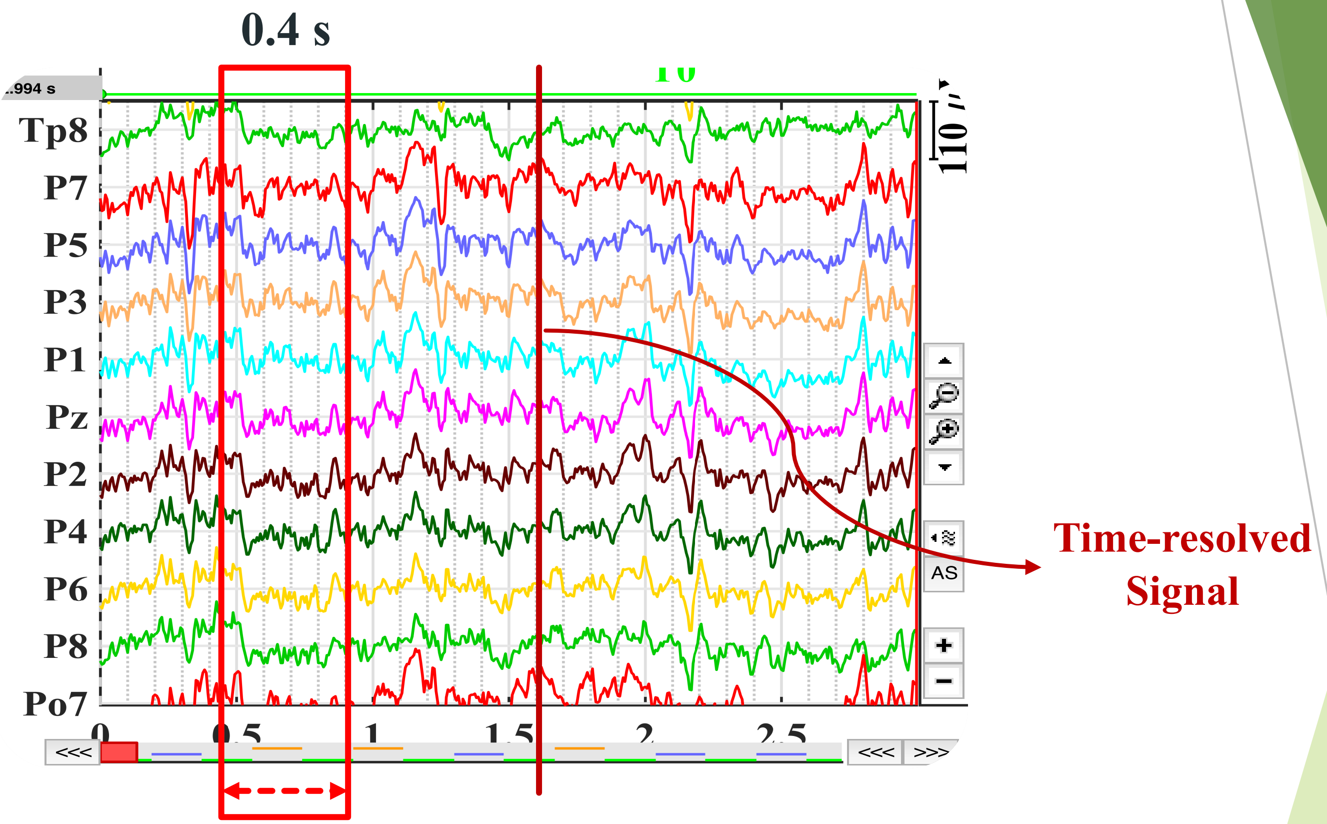
[4] School of Biomedical Engineering and Brain and Mind Center, The University of Sydney

EEG Deep Learning Library: https://github.com/SuperBruceJia/EEG-DL

One Problem of the GCNs-Net



Spatial information



Temporal information

- ✓ GCNs-Net is based on **Time-resolved Signal** → doesn't consider **Temporal Information**

Motivation:

- ✓ **[Spatial-Temporal Analysis] Consider Temporal and Spatial Information from EEG signals**
- ✓ **[Responsive] Maintain High Responding Time**

Image Credit: The PhysioNet Dataset.

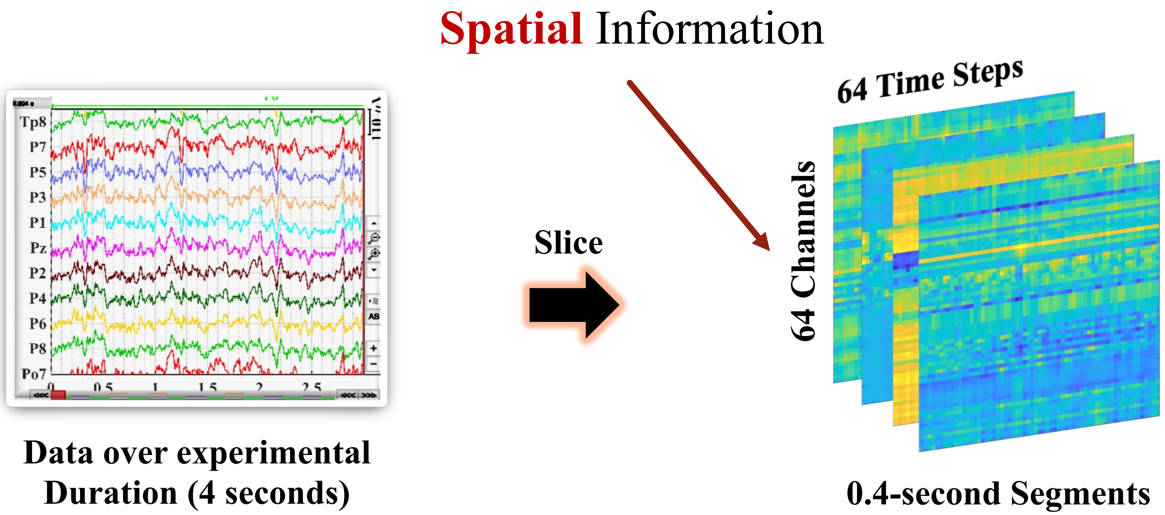
64-channel Raw EEG Signals Acquisition

Data over experimental Duration (4 seconds)

0.4-second Segments

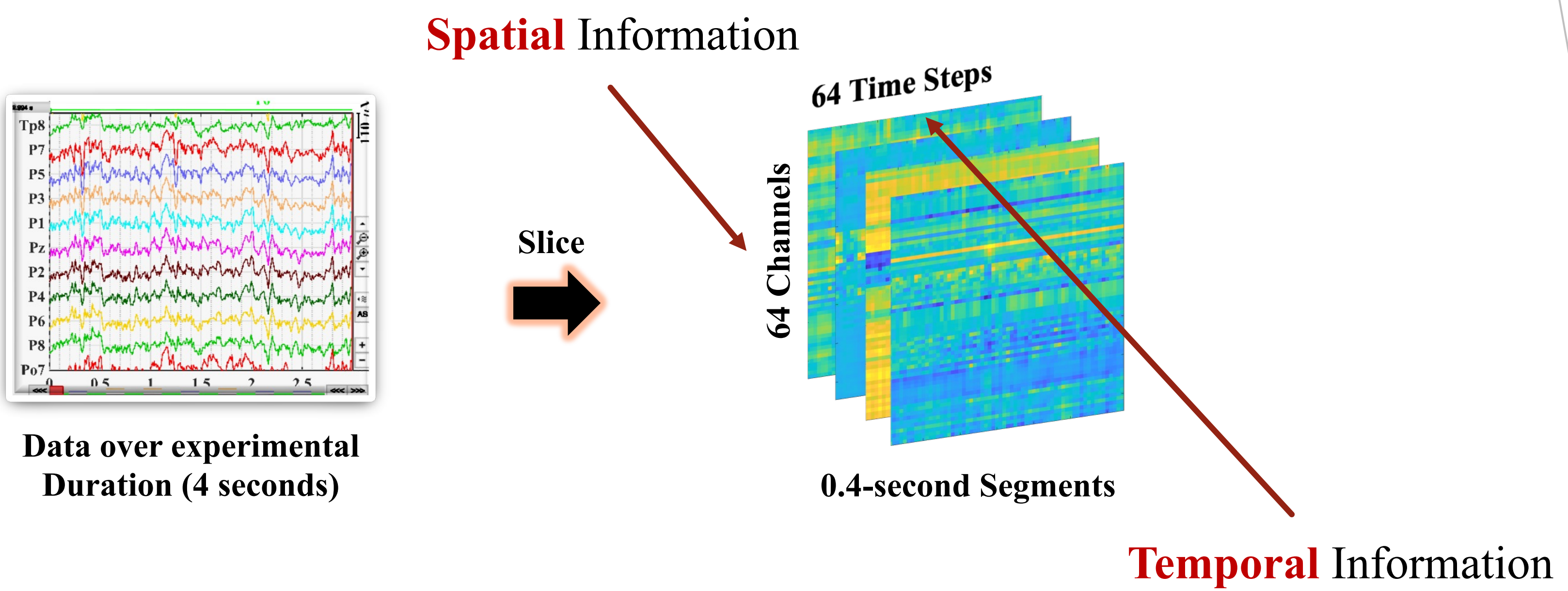
- ✓ 4-s Signals (experimental duration): **0.4-s segments** over time
- ✓ Each Segment: **64 channels × 64 time steps**
- ✓ Pre-processed Data: **Temporal** Information + **Spatial** Information

64-channel Raw EEG Signals Acquisition



- ✓ 4-s Signals (experimental duration): **0.4-s segments** over time
- ✓ Each Segment: **64 channels × 64 time steps**
- ✓ Pre-processed Data: **Temporal** Information + **Spatial** Information

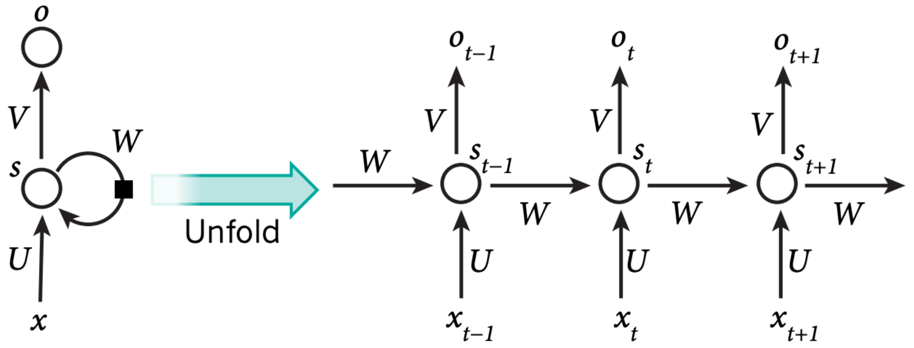
64-channel Raw EEG Signals Acquisition



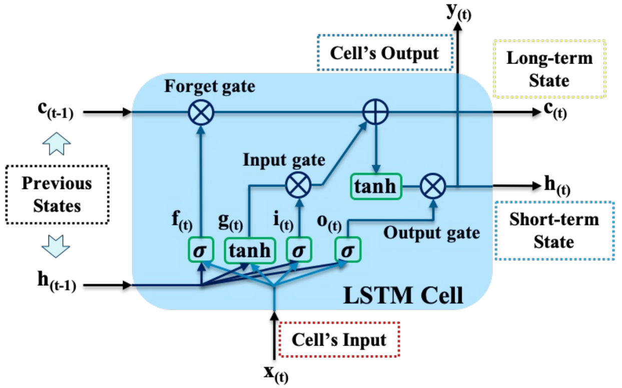
- ✓ 4-s Signals (experimental duration): **0.4-s segments** over time
- ✓ Each Segment: **64 channels × 64 time steps**
- ✓ Pre-processed Data: **Temporal** Information + **Spatial** Information

Temporal Information Extraction

unrolling the network through time



Recurrent Neural Network (RNN)

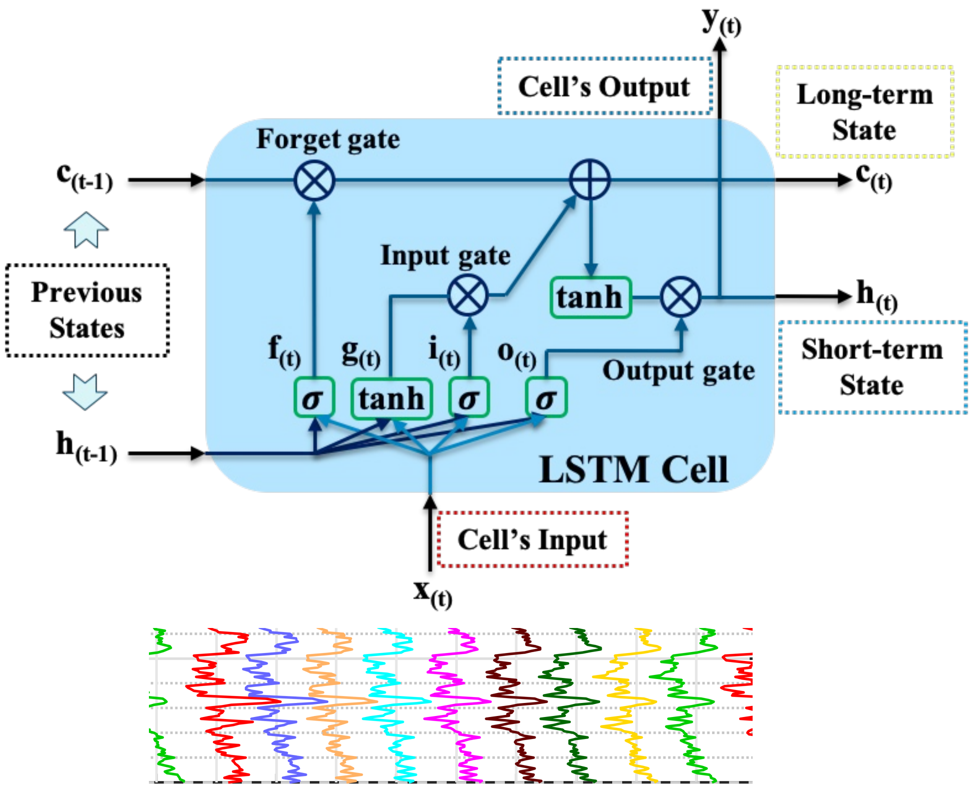


Long Short-Term Memory (LSTM)

- ✓ Designed for **order-mattered sequential data**, *e.g.*, time series
- ✓ The learned features at *time step t* are affected by $\mathbf{x}_t$ and $\mathbf{x}_{t-1}$ → **continuly learn** from time series
- ✓ LSTM: better capture **long-range sequence dependencies**
- ✓ Gated Recurrent Units (GRU): **lightweight** architecture with comparable performance

Image Credit: LeCun *et al.*, Deep Learning, In Nature.

Long Short-term Memory (LSTM)



64-channel Raw EEG Signals at Time Step $x_{(t)}$

- ✓ Capture **Long-range Dependencies** by the long-term state path $\mathbf{c}_{t-1} \rightarrow \mathbf{c}_t$
- ✓ **Input Gate**: store $\mathbf{x}_t$ and control $\mathbf{c}_t$'s input
- ✓ **Forget Gate**: control $\mathbf{c}_{t-1}$
- ✓ **Output Gate**: control $\mathbf{c}_t$'s output
  → short-term state $\mathbf{h}_t$ (**Cell's Output**)
- ✓ More parameters to store information
- ✓ Bidirectional:
  (1) $\mathbf{x}_1 \rightarrow \mathbf{x}_t$
  (2) $\mathbf{x}_t \rightarrow \mathbf{x}_1$

Attention Mechanism

- ✓ Signals or Outputs

Equally treated/contributed

vs.

Differently treated/contributed with preference/importance

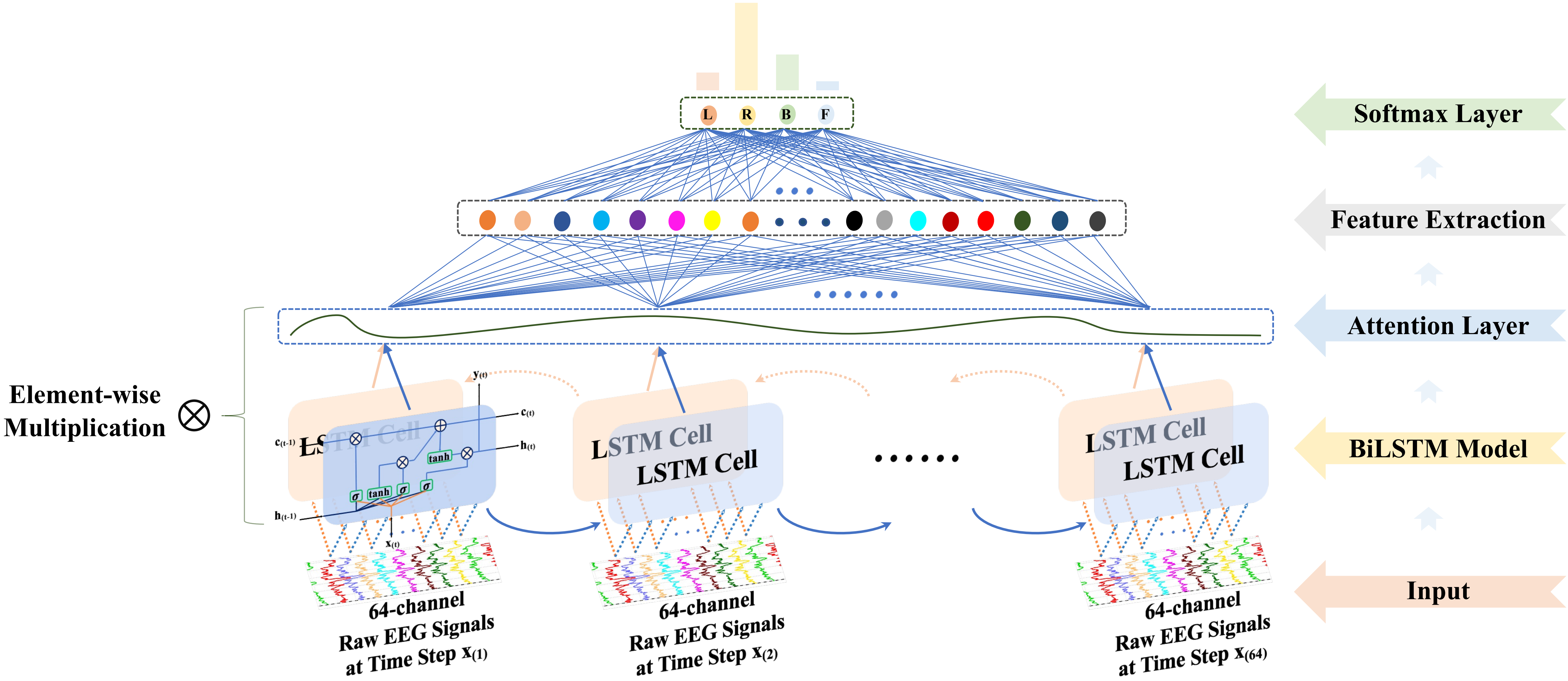
FC Layer $$\mathbf{U_t} = \tanh(\mathbf{W_w y_t} + \mathbf{b_w})$$

Attentional Weights $$\boldsymbol{\alpha}_\mathbf{t} = \frac{\exp(\mathbf{U_t^T W_U})}{\sum_t \exp(\mathbf{U_t^T W_U})}$$

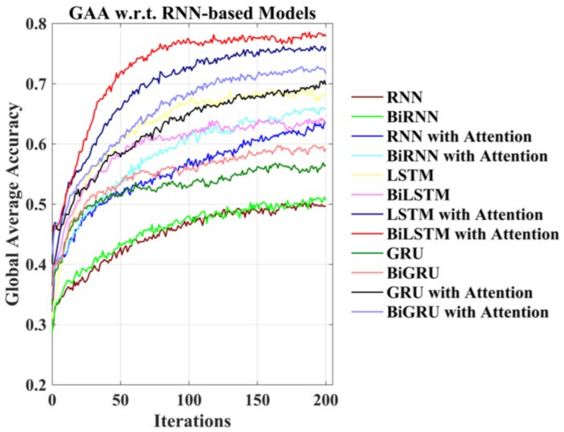
Weighted Sum $$\hat{\mathbf{U}}_\mathrm{t} = \sum_\mathrm{t} \boldsymbol{\alpha}_\mathrm{t} \mathbf{y}_\mathrm{t}$$

Image Credit: Prof. Bolei Zhou.

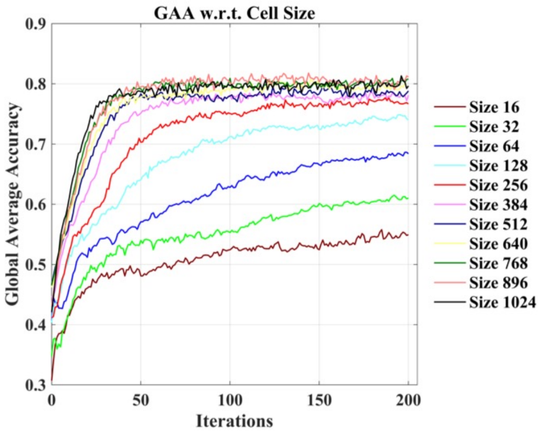
Attention-based Bidirectional Long Short-term Memory (Bi-LSTM)



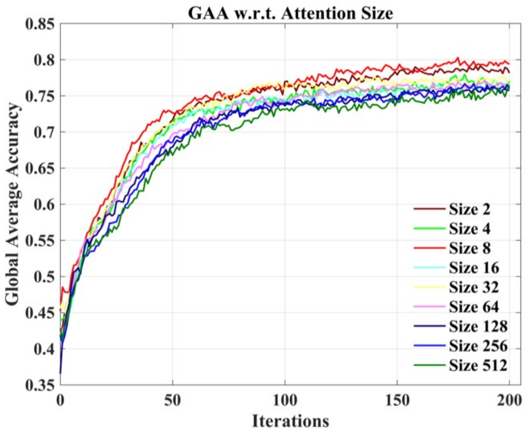
Model Design Ablation Study



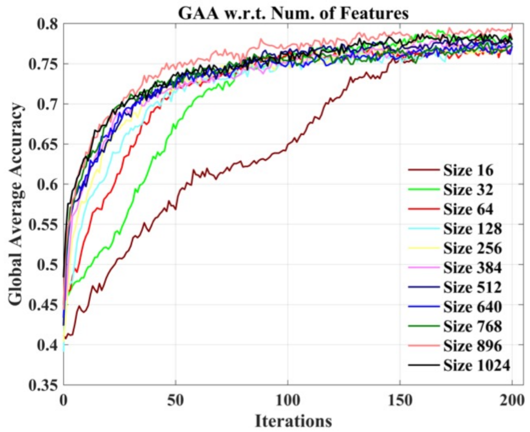
(a) GAA w.r.t. RNN-based Models



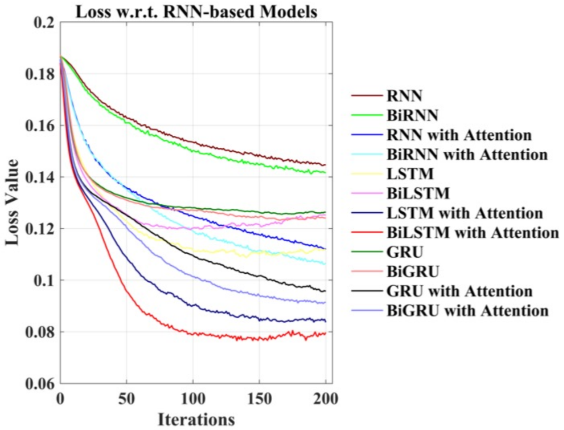
(b) GAA w.r.t. the BiLSTM Cell Size



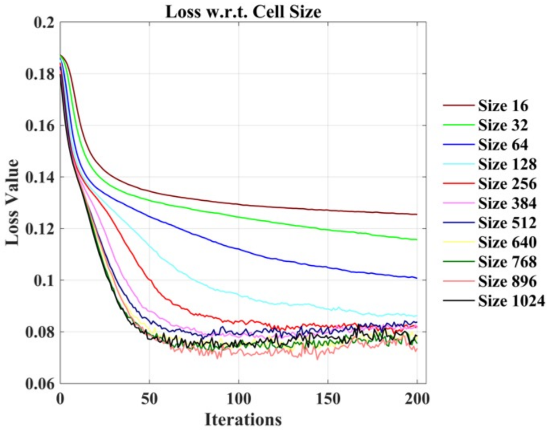
(c) GAA w.r.t. the Attention Size of the BiLSTM



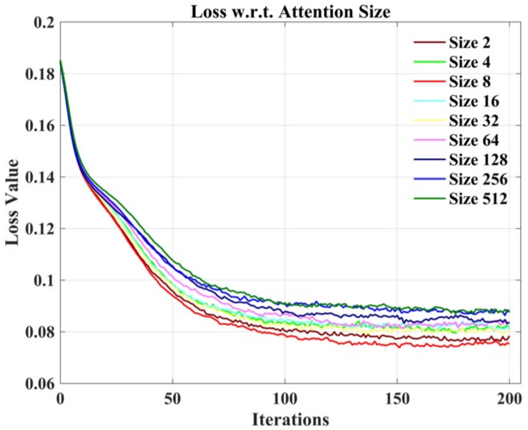
(d) GAA w.r.t. Num. of Extracted Features



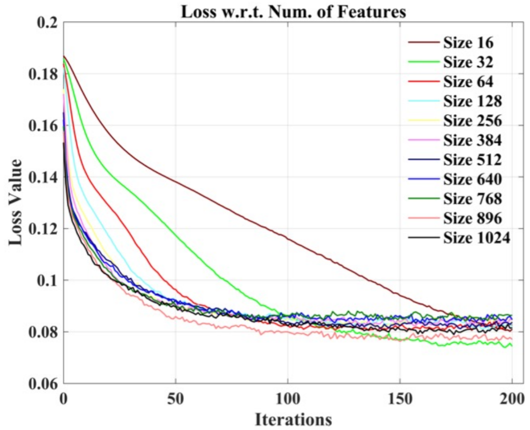
(e) Loss w.r.t. RNN-based Models



(f) Loss w.r.t. the BiLSTM Cell Size



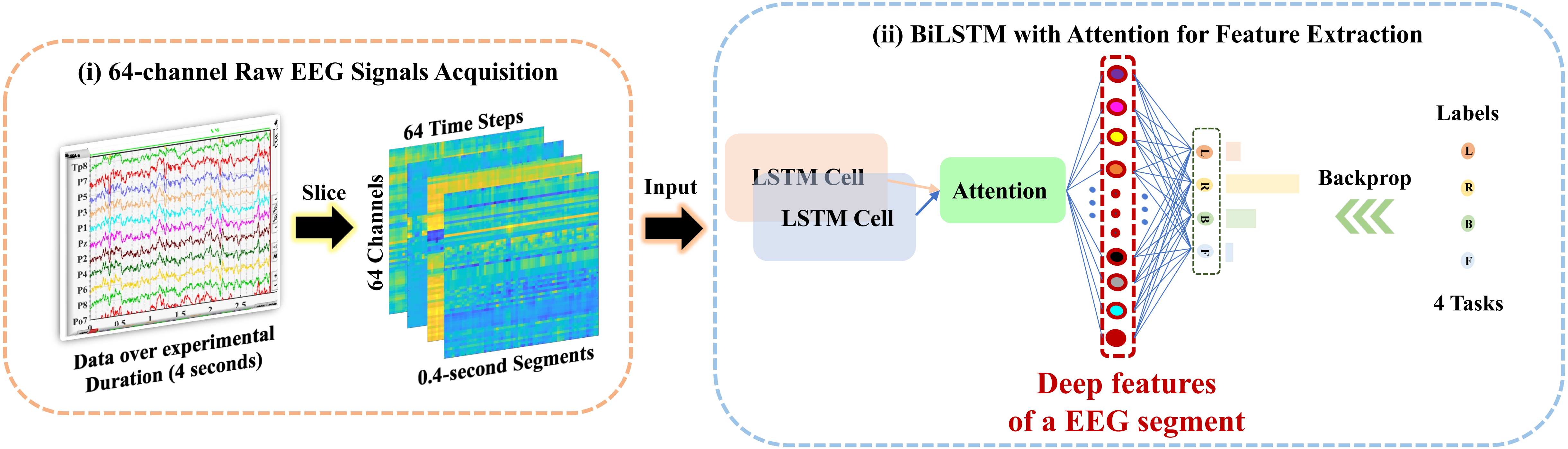
(g) Loss w.r.t. the Attention Size of the BiLSTM



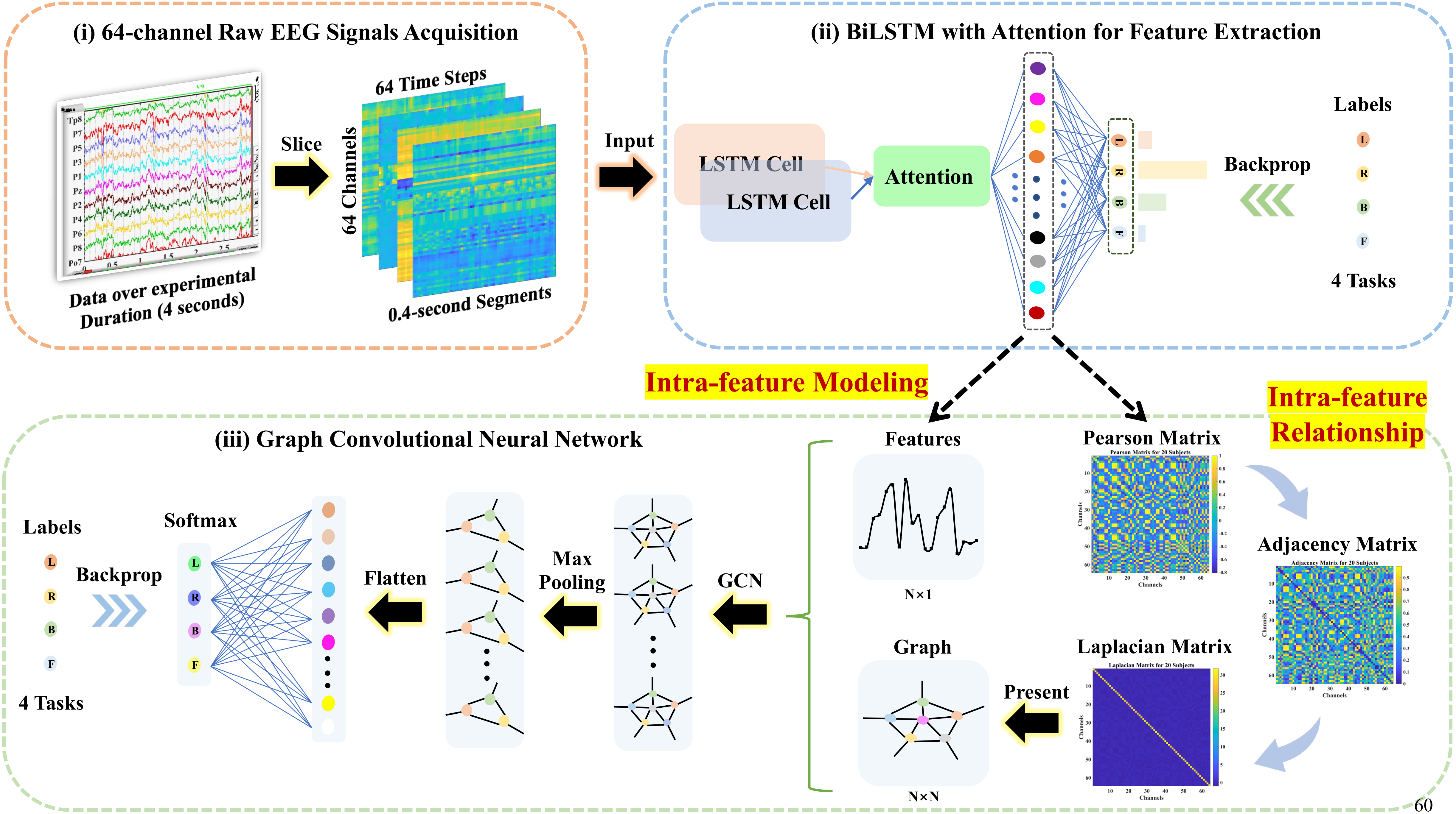
(h) Loss w.r.t. Num. of Extracted Features

Fig. 3: Models and Hyperparameters Comparison w.r.t. the RNN-based Methods for Feature Extraction

Topological Structure of Features

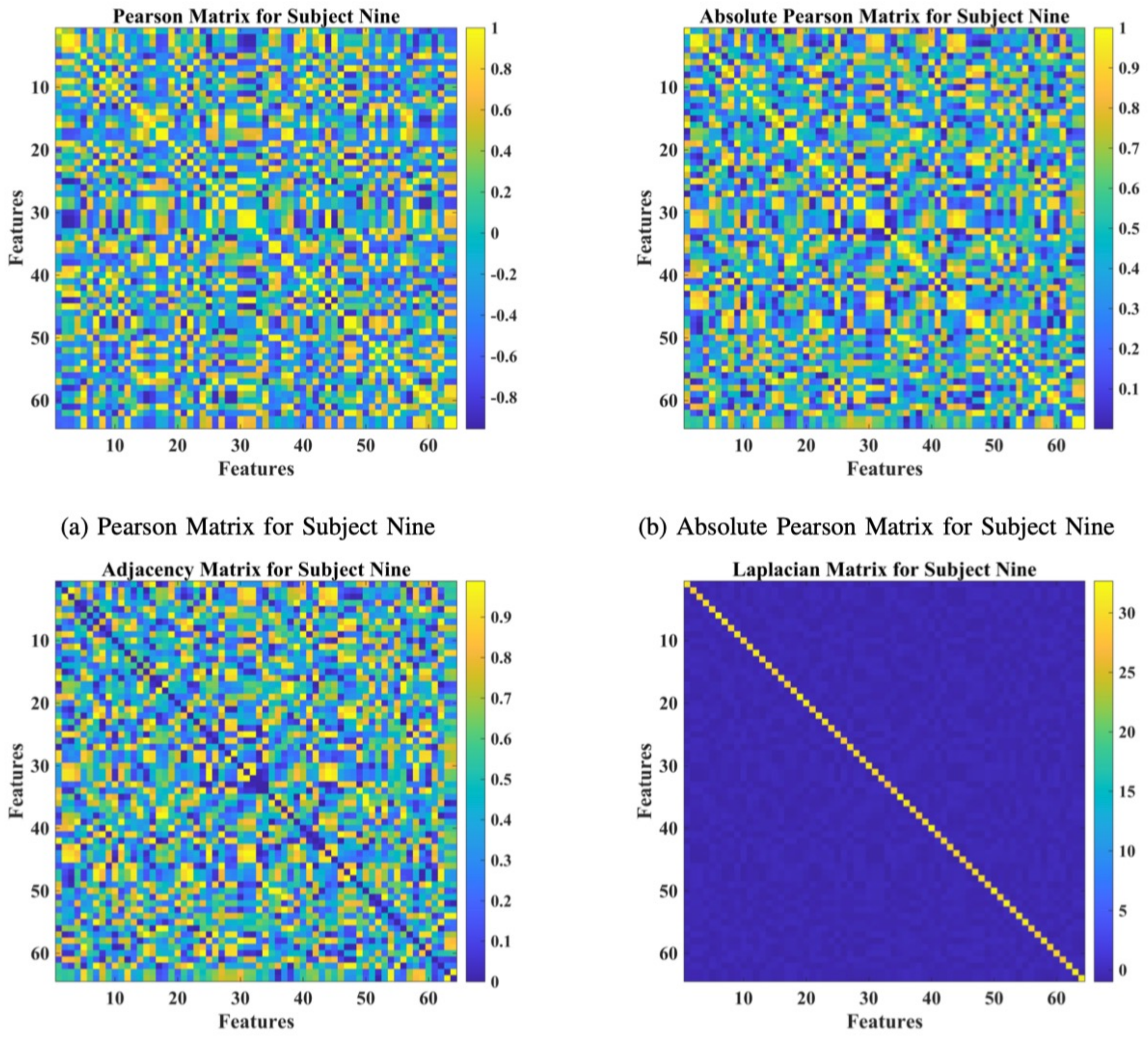


✓ **Deep Feature Mining → Intra-feature Relationship → Intra-feature Modeling**



(i) 64-channel Raw EEG Signals Acquisition
Data over experimental Duration (4 seconds)
Slice
64 Time Steps
64 Channels
0.4-second Segments
Input
(ii) BiLSTM with Attention for Feature Extraction
LSTM Cell
LSTM Cell
Attention
Labels
Backprop
4 Tasks
Intra-feature Modeling
Intra-feature Relationship
(iii) Graph Convolutional Neural Network
Features
N×1
Pearson Matrix
Adjacency Matrix
Graph
N×N
Laplacian Matrix
Present
GCN
Max Pooling
Flatten
Softmax
Backprop
Labels
4 Tasks

Topological Structure of Features



(a) Pearson Matrix for Subject Nine

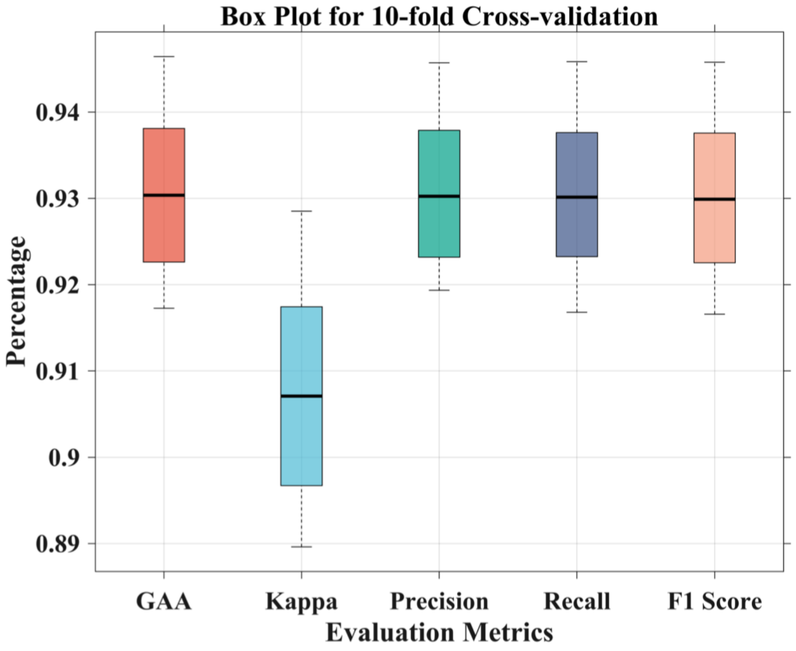
(b) Absolute Pearson Matrix for Subject Nine

(c) Adjacency Matrix for Subject Nine

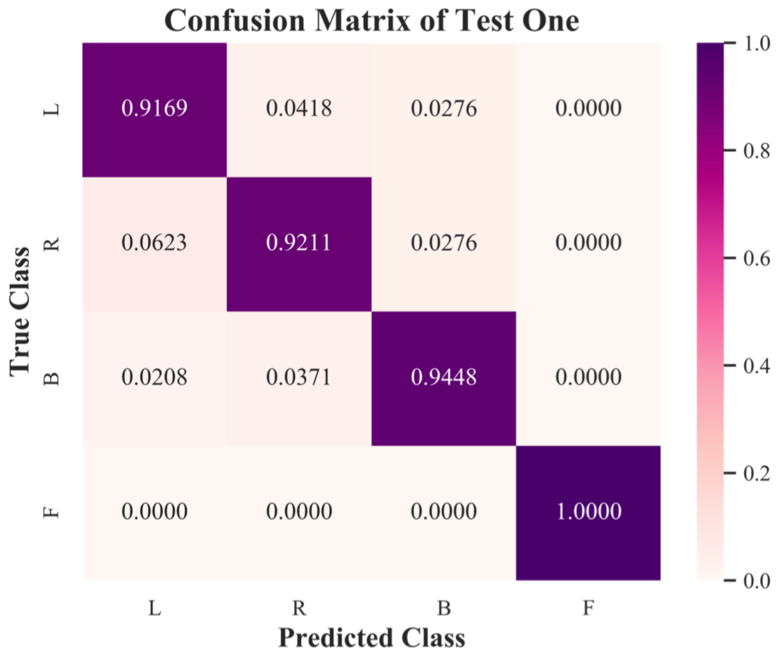
(d) Laplacian Matrix for Subject Nine

Fig. 4: The Pearson, Absolute Pearson, Adjacency, and Laplacian Matrices for Subject Nine.

Experimental Results - Groupwise Prediction



(a) Box Plot for 10-fold cross validation



(b) Confusion Matrix of Test One

Fig. 5: Box plot and confusion matrix for 10-fold cross validation.

Note:

(1) Box Plot (Maximum Score, Upper Quartile, Median, Lower Quartile, and Minimum Score)

(2) Confusion Matrix: TP, TN, FP, and FN

Experimental Results - Subject-Specific Adaptation

TABLE II: Subject-level Evaluation

No. of Subject	GAA	Kappa	Precision	Recall	F1 Score
1	94.05%	92.06%	94.20%	94.32%	94.16%
2	96.43%	95.19%	96.06%	96.06%	96.06%
3	97.62%	96.79%	97.33%	97.08%	97.18%
4	90.48%	87.34%	91.30%	91.11%	90.42%
5	95.24%	93.61%	95.96%	95.06%	95.38%
6	94.05%	92.02%	93.40%	94.96%	93.66%
7	98.81%	98.40%	98.81%	99.07%	98.92%
8	95.24%	93.60%	95.39%	95.04%	95.19%
9	98.81%	98.39%	99.11%	98.68%	98.87%
10	94.05%	91.98%	93.39%	94.70%	93.61%
Average	**95.48%**	**93.94%**	**95.50%**	**95.61%**	**95.35%**

TABLE III: Current studies comparison on subject-level prediction

Related Work	Max. GAA	Approach	Database
Ortiz-Echeverri *et al.* (2019)	94.66%	Sorted-fast ICA-CWT + CNNs	
Sadiq *et al.* (2019)	95.20%	EWT + LS-SVM	BCI Competition IV-a Dataset
Taran *et al.* (2018)	96.89%	TQWT + LS-SVM	
Zhang *et al.* (2019)	83.00%	CNNs-LSTM	
Ji *et al.* (2019)	95.10%	SVM	BCI Competition IV-2a Dataset
Amin *et al.* (2019)	95.40%	MCNNs	
Dose *et al.* (2018)	68.51%	CNNs	
Hou *et al.* (2019)	96.00%	ESI + CNNs	Physionet Database
This work	**98.81%**	**Attention-based BiLSTM-GCN**	

Takeaways and Future Work

✓ Spatial-Temporal Analysis

(1) Converge to both **Subject-level and Groupwise Predictions** and handle **Individual Variability**.

(2) The 0.4-s sample size **Time-Resolved Solution** toward fast response.

✓ Deep Feature Mining

(1) ↑ **Highest Accuracy**

(2) Advance **Clinical Translation** of EEG-based BCI technology to meet diverse demands, such as those of paralyzed patients.

✓ Future Work

Long-range Dependencies among intra-subject or inter-subject EEG signals can be modeled via **Non-local Modeling**, **Self-attention Mechanism**, **Transformer**, and **AI foundation Models**.

Thank you!

Any question?

Acknowledgment: Dr. Zhijian Hou and Dr. Dingquan Li.