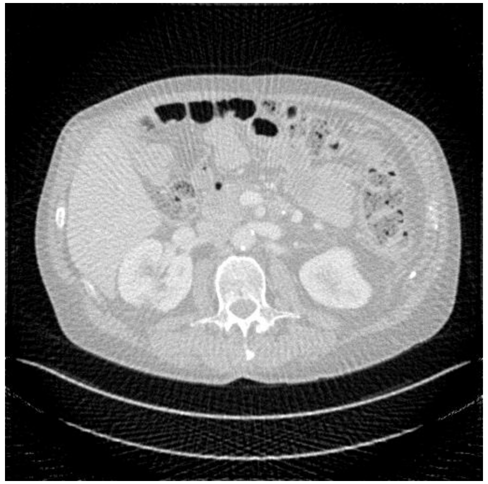
Swin-KAT: Advancing Swin Transformer with Kolmogorov-Arnold Network for CT Image Quality Assessment

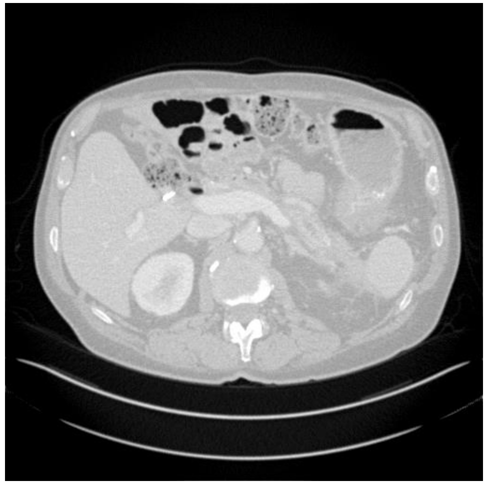
Kazi Ramisa Rifa[1], Jie Zhang[2], Abdullah-Al-Zubaer Imran[1]
[1]Department of Computer Science, [2]Department of Radiology

CT Image Quality Assessment

- ❑ Goal: Non-reference assessment of CT image quality



Poor Quality



Good Quality

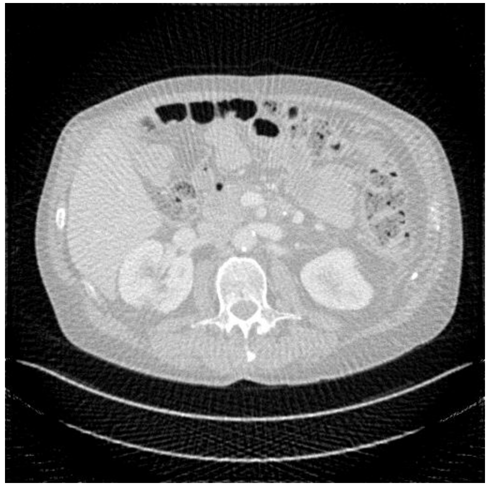
Image Source: LDCTIQAC Dataset

IQA Scoring Criteria

Score	Quality	Diagnostic Quality Criteria
0	Bad	Desired features are not shown
1	Poor	Diagnostic interpretation is impossible
2	Fair	Suitable for compromised interpretation
3	Good	Good for diagnostic interpretation
4	Excellent	Anatomical features are clearly visible

IQA Scoring Criteria

Score	Quality	Diagnostic Quality Criteria
0	Bad	Desired features are not shown
1	Poor	Diagnostic interpretation is impossible
2	Fair	Suitable for compromised interpretation
3	Good	Good for diagnostic interpretation
4	Excellent	Anatomical features are clearly visible

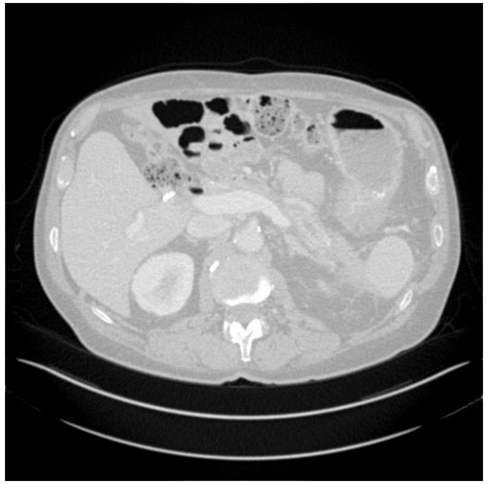


IQA: 0.80

❑ Radiologists' assigned scores are averaged

IQA Scoring Criteria

Score	Quality	Diagnostic Quality Criteria
0	Bad	Desired features are not shown
1	Poor	Diagnostic interpretation is impossible
2	Fair	Suitable for compromised interpretation
3	Good	Good for diagnostic interpretation
4	Excellent	Anatomical features are clearly visible



IQA: 3.60

❑ Radiologists' assigned scores are averaged

Clinical Motivations

- **Low-dose challenge:** Dose reduction compromises image quality in clinical CT scans
- **IQA role:** Image Quality Assessment ensures diagnostic reliability in low-dose settings

Technical Motivations

- ❑ **Transformer efficiency:** Enables scalable feature extraction with lower computational cost
- ❑ **CT assessment:** Well-suited for efficient analysis of CT image slices

Contributions

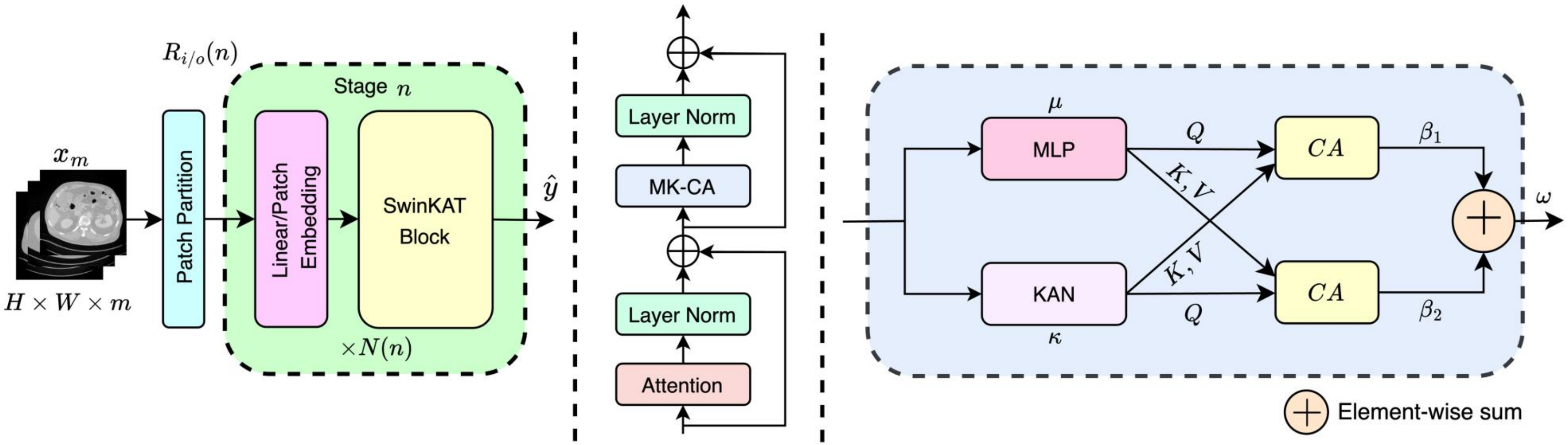
- ❑ Novel transformer-based architecture (Swin-KAT) integrating KAN into Swin Transformer
- ❑ An innovative attention-based approach combining MLP and KAN
- ❑ Generalized performance of Swin-KAT in predicting IQA from abdominal CT images

Existing Models

Model	Method	Architecture	Details
MD-IQA	Unsupervised IQA (no labels needed)	Vision Transformer and ConvNeXt	IQA using many images without labeled quality scores.
D-BIQA	Generated-reference IQA	Vision Transformer, Swin Transformer and Transposed Attention Blocks	IQA by comparing to algorithmically generated references.

Proposed Swin-KAT

- Modifies Swin Transformer to help capture hierarchical visual features
- Dual cross-attention paths (MK-CA) combining multilayer perceptron and Kolmogrov-Arnold network

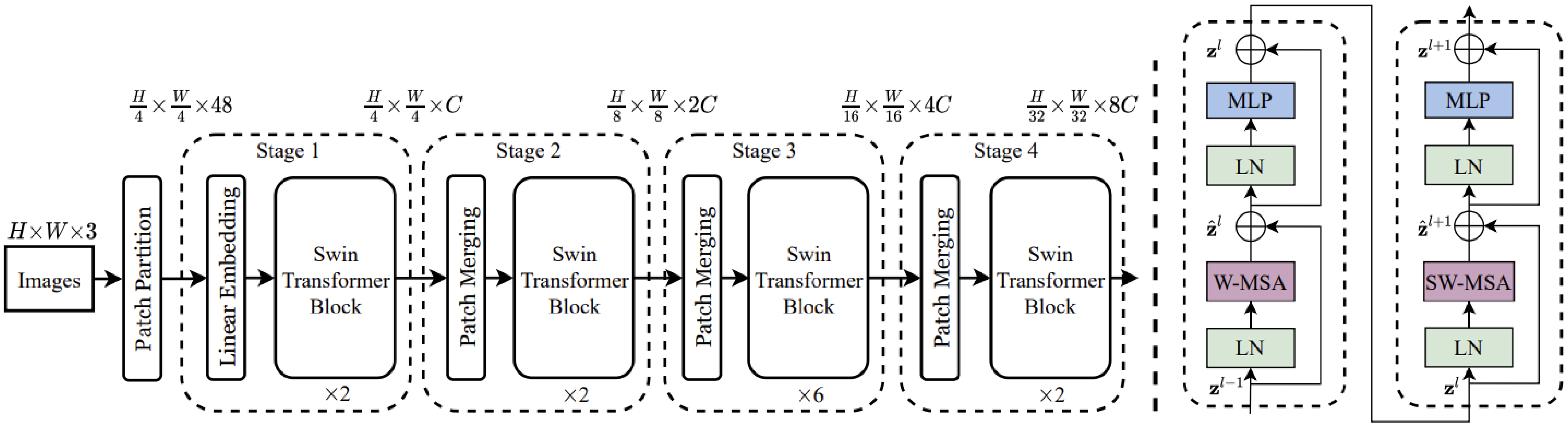


(a) The architecture of Swin-KAT

(b) MK-CA embedded in the Swin-KAT block

(c) MK-CA

Swin Transformer



(a) Architecture

(b) Two Successive Swin Transformer Blocks

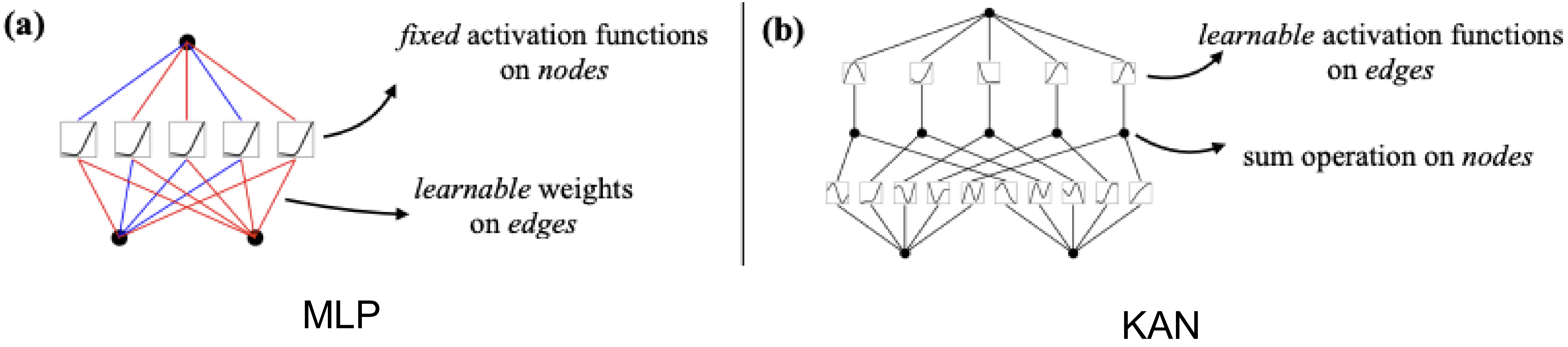
W-MSA: Multi-head self attention w/ regular windowing

SW-MSA: Multi-head self attention w/ shifted windowing

Liu et al. ICCV, 2021

Kolmogorov-Arnold Network (KAN)

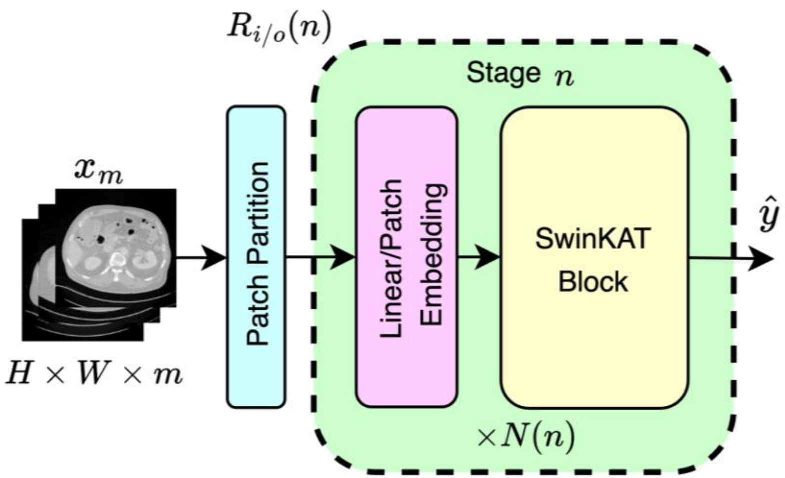
- KANs have learnable activation functions on edges which improves scalability
- This design allows KANs to capture intricate data structures better than traditional MLPs



Liu et al. ICLR 2025

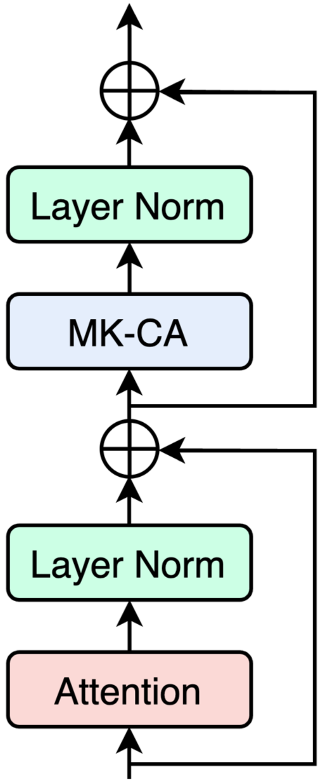
Swin-KAT

- ❑ N denotes the number of repetitions of the procedure.
- ❑ n is the number of stages.
- ❑ $R_{i/o}(n) = (\frac{H}{d_n}, \frac{W}{d_n}, C_n)$ and $n \in \{1,2,3,4\}$.
- ❑ d_n is the downsampling factor.
- ❑ C_n is the channel multiplier.



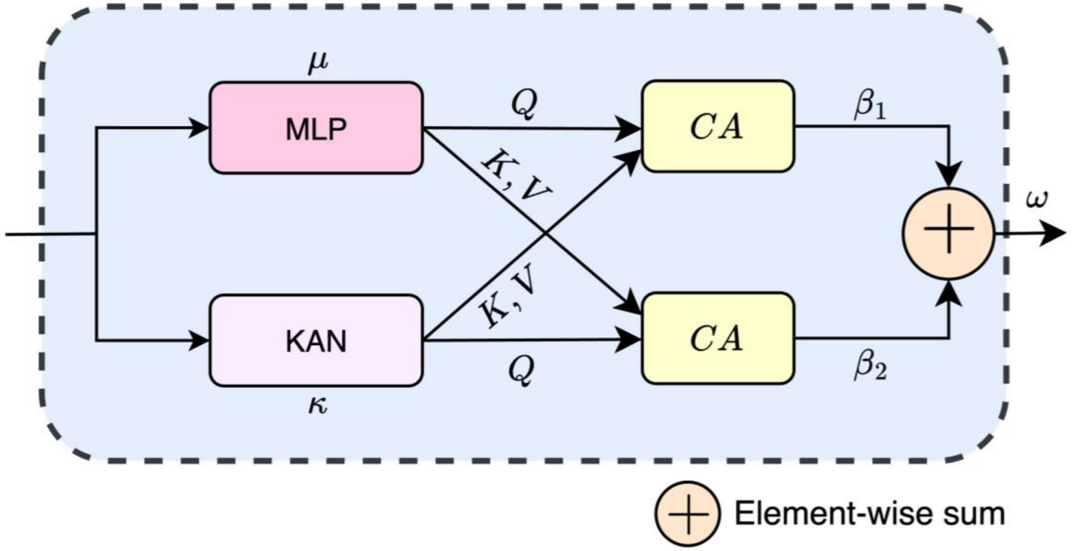
Swin-KAT Block

MK-CA path is embedded in each of the Swin-KAT blocks



MK-CA Path

A cross-attention based approach combining MLP and KAN



Datasets

- Low-Dose CT Image Quality Assessment (LDCTIQA) Challenge dataset

Set	Count	Details
Training Set	1,000	To compare against the existing methods: train (700), val (100), and test (200)
Test Set	300	Additional evaluation to ensure comprehensive assessment

Quantitative Results

No.	Methods	r	ρ	τ	s
1	DBCNN	0.9714	0.9734	0.8808	2.8255
2	MD-IQA	0.9771	<u>0.9793</u>	**0.9106**	<u>2.8670</u>
3	MANIQA	0.9768	0.9786	0.8891	2.8445
4	AHIQ	0.9762	0.9746	0.8810	2.8317
5	QPT	0.9743	0.9732	0.8797	2.8272
6	SSIQA	<u>0.9784</u>	0.9767	0.8905	2.8456
7	Swin-KAT	**0.9831**	**0.9825**	<u>0.9031</u>	**2.8687**

We use Pearson's (r), Spearman's (ρ), and Kendall's (τ) correlation coefficients, with the overall score (s) as their aggregate.

MK-CA Variants

No.	Methods	Operation	r	ρ	τ	s
1	μ only	-	0.9331	0.9331	0.7854	2.6516
2	κ only	-	0.9373	0.9382	0.7954	2.6709
3	μ-κ	Average	0.9405	0.9375	0.7938	2.6718
4	μ-κ	Concat	0.9391	0.9367	0.7902	2.6659
5	μ-κ	Sum	0.9436	0.9380	0.7921	2.6738
6	μ-κ	CA	0.9352	0.9321	0.7855	2.6529
7	μ-κ	CA	0.9368	0.9298	0.7793	2.6460
8	ω	CA	**0.9454**	**0.9389**	**0.7967**	**2.6811**

μ refers to MLP and κ refers to KAN

Time and Memory Comparisons

Swin-KAT is faster and memory efficient than the LDCTIQA challenge's top algorithms

No.	Team	Model	Time (ms)	Memory
1	agaldran	Swin & BiTResNeXt50	424.07	1309.89
2	RPI_AXIS	MANIQA	44.87	638.51
3	CHILL@UK	EfficientNet-V2L	138.46	503.63
4	FeatureNet	ViT & GLCM	51.08	572.31
5	Team Epoch	EDCNN	72.11	564.94
6	gabybaldeon	CNN-ViT	29.69	942.13
7	**Ours**	**Swin-KAT**	**20.69**	**441.50**

Conclusions

- ❑ No-reference and reliable deep learning-based IQA solution
- ❑ Transformer model with fixed and learnable activation functions using cross-attention
- ❑ Swin-KAT reliably quantifies noisy and artifact-affected CT images
- ❑ The model achieves a notable reduction in both memory usage and runtime
- ❑ Future research centers on localized IQA across various body regions

University of Kentucky

Acknowledgement

Funding:

High Impact General Pilot,
UNITE Research Priority Area,
University of Kentucky

UNITE
RESEARCH PRIORITY AREA

TRANSFORMING TOMORROW

By Bringing Together Many People in a Global Community

Swin-KAT: Advancing Swin Transformer with Kolmogorov-Arnold Network for CT Image Quality Assessment

Kazi Ramisa Rifa, Jie Zhang, Abdullah Al Zubaer Imran

Questions?

Contact: ramisa.rifa@uky.edu

Project Site

Personal Site