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

*Kazi Ramisa Rifa, Jie Zhang, Abdullah-Al-Zubaer Imran*

**University of Kentucky, Lexington, KY, 40506, USA**

ISBI 2025
2025 IEEE International Symposium on Biomedical Imaging
April 14-17, 2025 | Houston, TX, USA

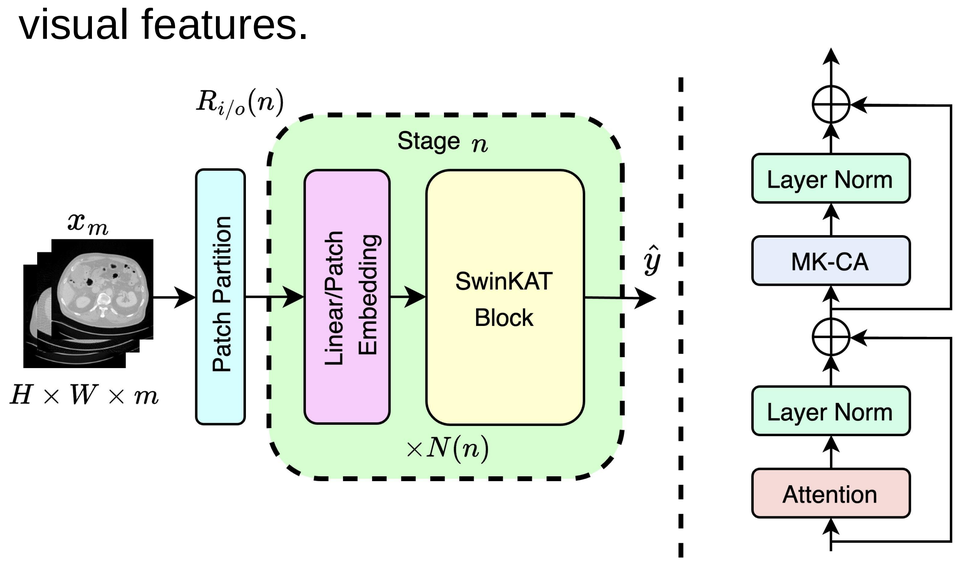
**Poster No.** 1571090127

## INTRODUCTION

- **Goal:** CT Image quality assessment (IQA) aligned with radiologists' perceptions
- **Contribution:** Novel transformer-based architecture (Swin-KAT) integrating KAN into Swin Transformer, with an Innovative cross-attention approach combining MLP and KAN
- **Outcome:** Reference-free and efficient CT IQA

## METHOD

- Modifies Swin Transformer to help capture hierarchical visual features.



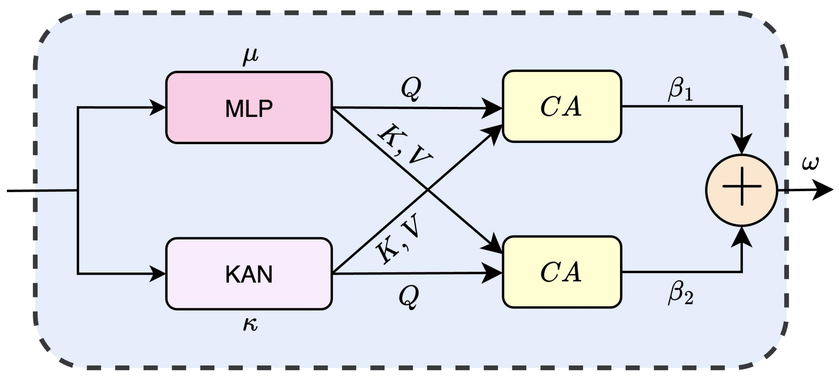

Proposed Swin-KAT architecture

- MLP-KAN Cross Attention:

$$\beta_1 = CrossAttention(Q_\mu, K_{\kappa_\eta}, V_{\kappa_\eta})$$
$$\beta_2 = CrossAttention(Q_{\kappa_\eta}, K_\mu, V_\mu)$$

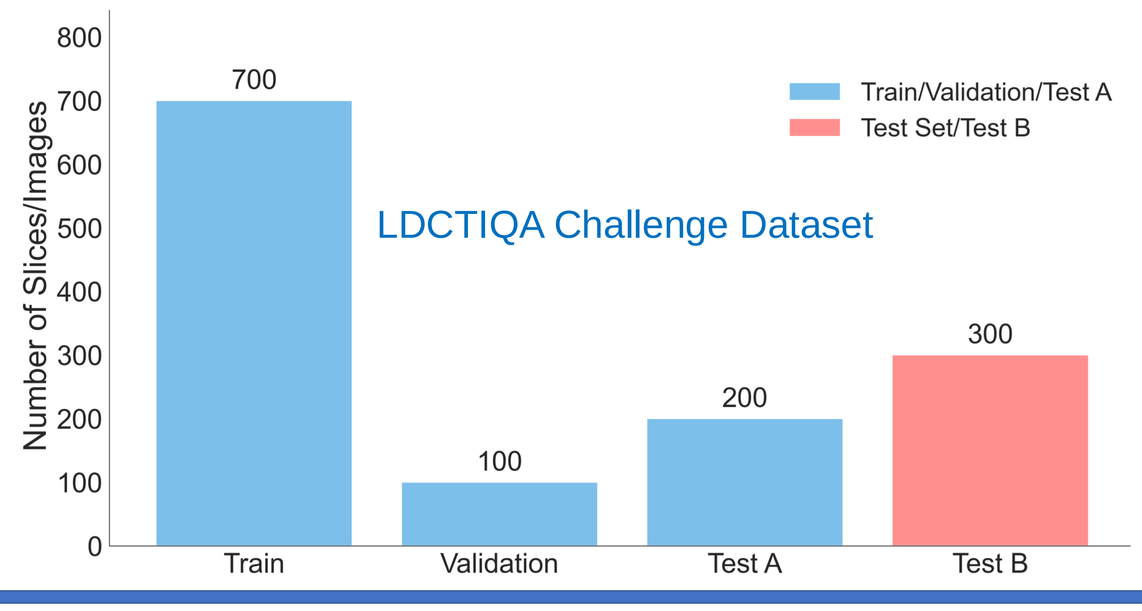
The resulting outputs $\beta_1$ and $\beta_2$ are then combined using element-wise summation as:

$$\omega = \beta_1 + \beta_2$$




Dual cross-attention paths (MK-CA) combining MLP and KAN

## DATA




LDCTIQA Challenge Dataset

## RESULTS

Swin-KAT **predictions** are in good agreement with [reference].

IQA: **0.61** [0.50] IQA: **1.80** [1.87] IQA: **2.39** [2.33] IQA: **3.00** [3.10] IQA: **3.89** [4.00]

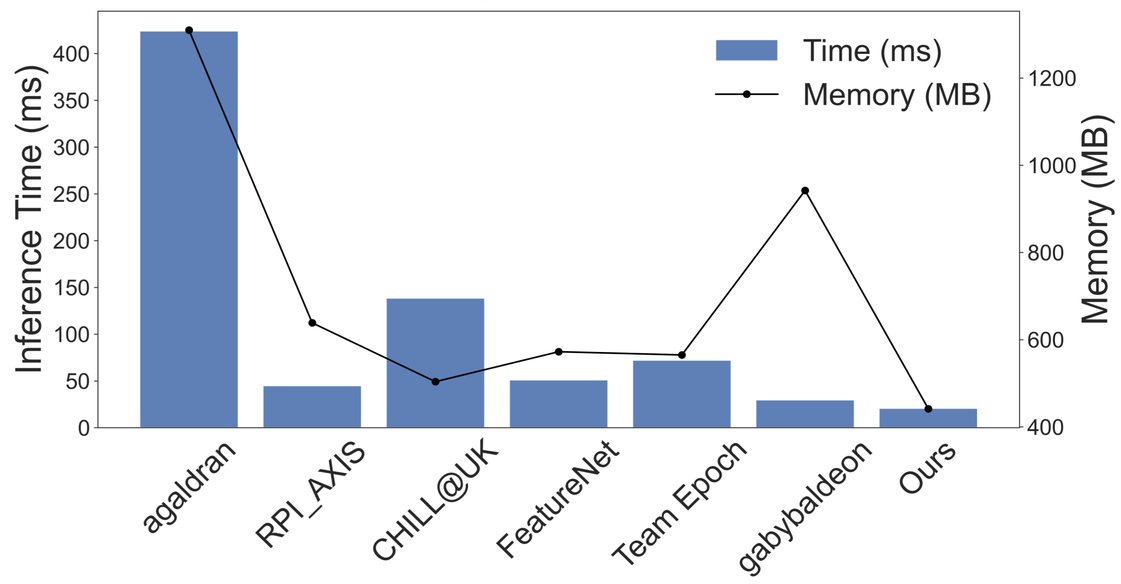
TestA: Quantitative comparison with state-of-the-art methods where Swin-KAT achieves the best overall performance.

| Method | $r$ | $\rho$ | $\tau$ | $s$ |
|---|---|---|---|---|
| DBCNN | 0.9714 | 0.9734 | 0.8808 | 2.8255 |
| MD-IQA | 0.9771 | 0.9793 | **0.9106** | 2.8670 |
| MANIQA | 0.9768 | 0.9786 | 0.8891 | 2.8445 |
| AHIQ | 0.9762 | 0.9746 | 0.8810 | 2.8317 |
| QPT | 0.9743 | 0.9732 | 0.8797 | 2.8272 |
| SSIQA | 0.9784 | 0.9767 | 0.8905 | 2.8456 |
| **Swin-KAT** | **0.9831** | **0.9825** | 0.9031 | **2.8687** |

TestB: Evaluation with different operations demonstrating the superiority of the proposed MK-CA block which is $\omega$.

| Method | Operation | $r$ | $\rho$ | $\tau$ | $s$ | MAE ↓ |
|---|---|---|---|---|---|---|
| $\mu$ only | – | 0.9331 | 0.9331 | 0.7854 | 2.6516 | 0.3279 |
| $\kappa$ only | – | 0.9373 | 0.9382 | 0.7954 | 2.6709 | 0.2952 |
| $\mu - \kappa$ | Average | 0.9405 | 0.9375 | 0.7938 | 2.6718 | 0.2919 |
| $\mu - \kappa$ | Concat | 0.9391 | 0.9367 | 0.7902 | 2.6659 | 0.2923 |
| $\mu - \kappa$ | Sum | 0.9436 | 0.9380 | 0.7921 | 2.6738 | 0.2906 |
| $\mu - \kappa$ | CA | 0.9352 | 0.9321 | 0.7855 | 2.6529 | 0.4022 |
| $\mu - \kappa$ | CA | 0.9368 | 0.9298 | 0.7793 | 2.6460 | 0.3034 |
| $\omega$ | CA | **0.9454** | **0.9389** | **0.7967** | **2.6811** | **0.2826** |

- Swin-KAT is faster and more memory-efficient, processing *approximately* 50 CT images per second.




Comparison to top LDCTIQA Challenge Algorithms

## CONCLUSION

- No-reference and reliable deep learning-based IQA solution
- Future research centers on localized IQA across various body regions

## REFERENCE

[1] Liu, Z. et al. KAN: Kolmogorov–Arnold Networks, ICLR, 2025.
[2] Liu, Z. et al. Swin Transformer: Hierarchical vision transformer using shifted windows, ICCV, 2021
[3] Lee, W. et al. Low-dose computed tomography perceptual image quality assessment, Medical Image A2025.

## Acknowledgements

This work was supported by UNITE Research Priority Area at the University of Kentucky

**Project Site**

University of Kentucky