Task-Focused Knowledge Transfer from Natural Images for CT Image Quality Assessment

Kazi Ramisa Rifa, Md Atik Ahamed, Jie Zhang, and Abdullah Al Zubaer Imran
University of Kentucky, Lexington, KY, USA

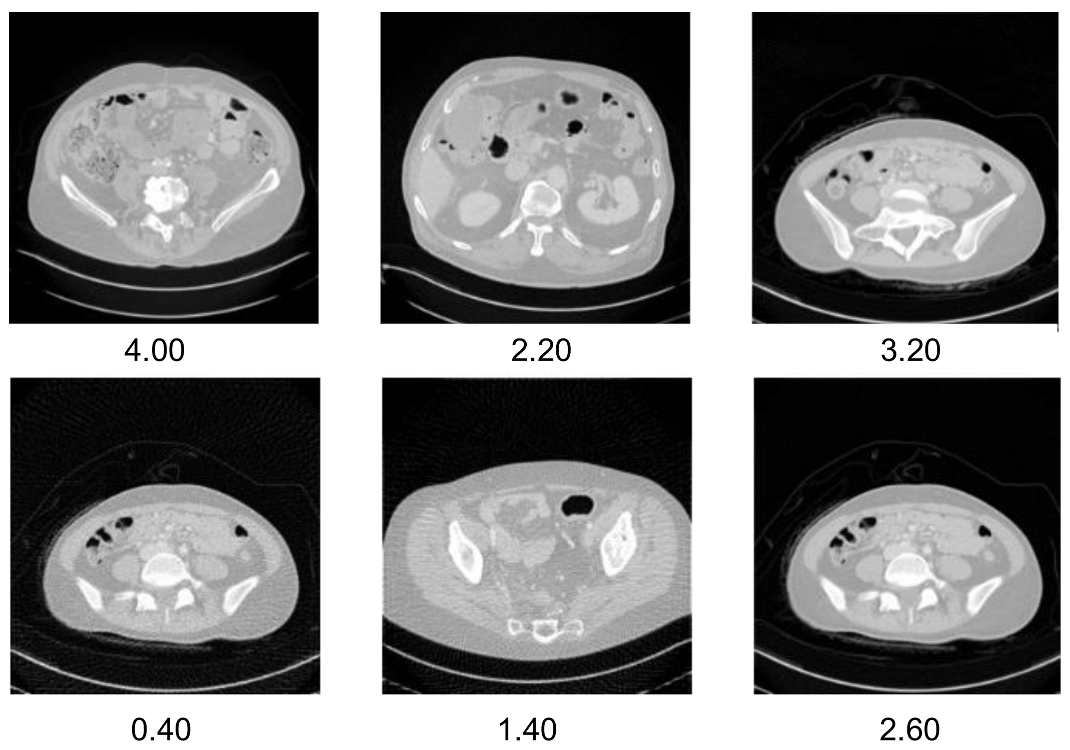
PROBLEM

- ❑ Advanced reconstruction techniques like Iterative Reconstruction and Deep Learning-based Reconstruction have transformed CT image quality assessment (IQA).
- ❑ A standardized metric is urgently needed to objectively assess CT image quality, ensuring diagnostic accuracy while minimizing unnecessary radiation exposure.

CONTRIBUTIONS

- ❑ TFKT: A novel task-specific transfer learning approach with hybrid CNN-Transformer for no-reference assessment of CT image quality leveraging natural images.
- ❑ Extensive experimentation demonstrating the effectiveness of TFKT in predicting radiologists' assigned scores both from in-domain (LDCTIQA) and out-of-domain (in-house) CT images.

KADID[1] images and mean opinion scores (MOS)



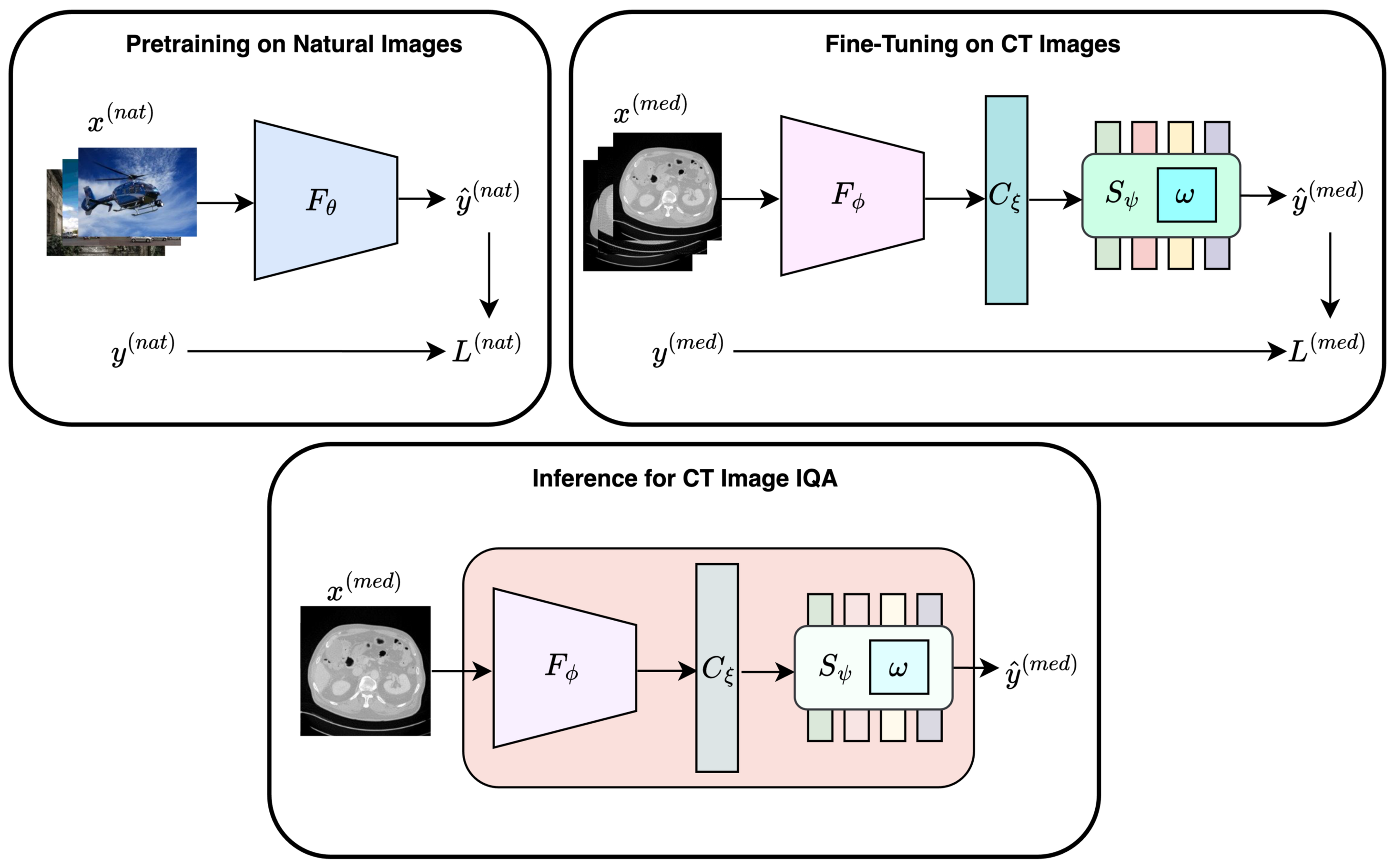
LDCTIQA[2] images and IQA scores

LDCTIQA Scoring Criteria[2]

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

METHODS

- ❑ Pretraining: Leveraging EfficientNet[3] *(F)*, TFKT is trained to predict the MOS scores (1-5) from input natural images.
- ❑ Finetuning:
  - LDCTIQA dataset is used to predict diagnostic quality of CT images.
  - ImageNet pretrained Swin Transformer[4] *(S)* to exploit both local and global features in medical images.
  - A connector module *(C)* to bridge between the *F* and *S*



Schematic diagram of the proposed TFKT-based CT IQA method

EXPERIMENTS

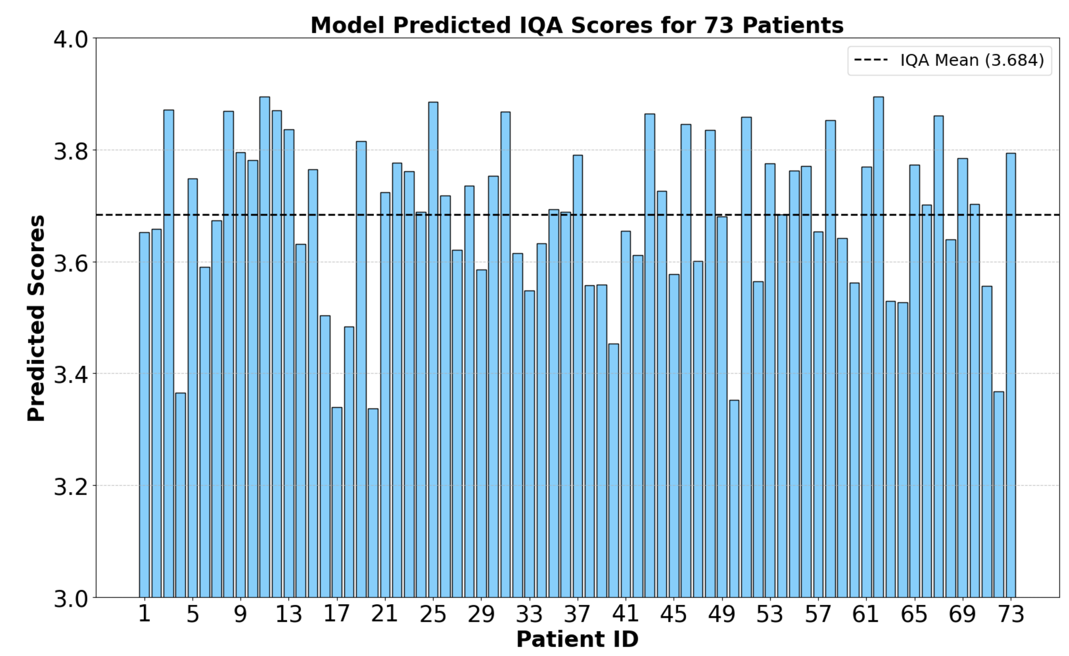
- ❑ Training
  - Phase 1: 10,000 natural distorted images from the KADID dataset
  - Phase 2: 800 images from the LDCTIQA train set
  - Loss: MSE loss is used in both phases
- ❑ Testing
  - Setting A: 200 CT images from the LDCTIQA train set
  - Setting B: LDCTIQA test set of 300 CT images
- ❑ Evaluation
  - Correlation coefficients:
    Pearson's linear correlation coefficient *(r)*,
    Spearman's rank correlation coefficient *(ρ)*, and
    Kendall rank correlation coefficient *(τ)*.
  - Overall model performance *(s)* by aggregating the three correlation coefficients $(r + \rho + \tau)$.

RESULTS

- ❑ TFKT outperforms all the baseline and state-of-the-art methods, demonstrating its effectiveness for CT IQA
- ❑ Ablation study justifies the task-similar knowledge transfer.

Quantitative comparison of TFKT against baseline and state-of-the-art CT IQA methods.

Methods	r	ρ	τ	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
EfficientNetV2L	0.9569	0.9741	0.8772	2.8082
SSIQA	0.9784	0.9767	0.8905	2.8456
TFKT- frozen F	0.9724	0.9757	0.8852	2.8332
TFKT - w/o Pretraining Phase	0.9809	0.9840	0.9097	2.8745
TFKT	**0.9842**	**0.9846**	**0.9126**	**2.8814**
Results on LDCTIQAC Test Set				
TFKT - w/o Pretraining Phase	0.9221	0.9294	0.7746	2.6261
TFKT (LDCTIQAC Test Set)	**0.9434**	**0.9425**	**0.8017**	**2.6876**



- ❑ The TFKT model tested on 73 out-of-domain pediatric abdominal CT scans. Slice-wise predictions are averaged to obtain IQA score in a scan.
- ❑ As expected for clinical images, TFKT predicted scores are also in good agreement (IQA >3).

CONCLUSIONS

- ❑ TFKT model provides a no-reference, fully-automated, and reliable deep learning-based solution for CT image quality assessment.
- ❑ Our ongoing work is focused on large-scale clinical validation with different patient populations across various body parts.

REFERENCES

1. Lin, H., Hosu, V., and Saupe, D., "KADID-10k: a large-scale artificially distorted IQA database," in QoMEX, 1–3, IEEE (2019).
2. Lee, W., Wagner, F., Maier, A., Wang, A., Jongduk, B., Scott, H., and Choi, J.-H., "Low-dose Computed Tomography Perceptual Image Quality Assessment Grand Challenge Dataset," Medical Image Computing and Computer Assisted Intervention, (2023).
3. Tan, M. and Le, Q., "Efficientnetv2: Smaller models and faster training," ICML, 10096–10106, PMLR (2021).
4. Liu, Z., Hu, H., Lin, Y., Yao, Z., Xie, Z., Wei, Y., Ning, J., Cao, Y., Zhang, Z., Dong, L., Wei, F., and Guo, B., "Swin Transformer v2: Scaling up capacity and resolution," CVPR, 11999–12009 (2022).

Acknowledgement

This work was supported by the Igniting Research Collaborations (IRC) program and UNITE Research Priority Area at the University of Kentucky