Resolution-Invariant Image Classification via Implicit Neural Representations: An Application to Medical Whole-Slide Imaging

Woojeh Chung, Computer Science.
Mentor: Kookjin Lee, Assistant Professor.
School of Computing and Augmented Intelligence

Introduction and Motivation

- Classification task Limitations: Traditional Deep Neural Networks(DNNs) are confined to specific resolutions for classification tasks, limiting their applicability.
- High-Resolution Medical Imaging: Prevailing methods struggle with giga-pixel images such as whole-slide images (WSIs), crucial in medical diagnostics.
- Implicit Neural Representation (INR): We leverage these INR representations to handle varying resolutions, optimizing image classification.
- Resolution Invariant Classification: Our approach aims to break free from the constraints of fixed resolutions in image classification tasks.
- Medical Image Application: The focus is on significantly improving the classification efficiency of vast medical images, facilitating advanced medical diagnostics.

Encoding images with INR

Implicit Neural Networks

- INRs are a way to encode images as continuous functions rather than discrete arrays, allowing for more flexible and detailed image descriptions.
- Unlike traditional array representations, INRs can represent images at arbitrary resolutions and scales, providing a more versatile framework for image processing.
- INRs are particularly suited for tasks requiring high-resolution detail or images that vary significantly in size, such as medical imaging.

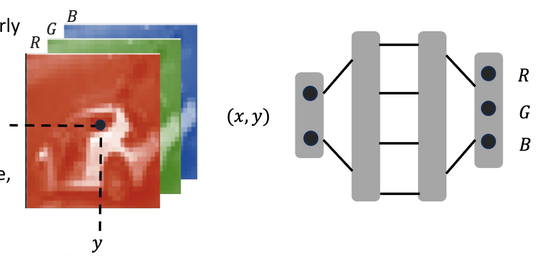

Figure 1: Schematic of the Implicit Neural Representation

Meta-Learning and Modulations

- The base neural network is trained using meta-learning techniques, which allow it to learn general features across various tasks and datasets.
- Shift modulations are applied to the meta-learned base network to encode image-specific variations
- Meta-learning optimizes the base network, enabling it to rapidly update data-specific parameters for each new image, bypassing the need for extensive retraining.

$$x \rightarrow \sin(\omega_0 W^{(i)} x + b^{(i)} + s^{(i)})$$

$$s = W' z + b'$$

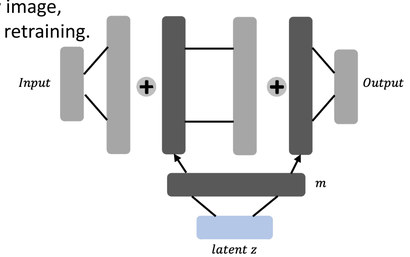

Figure 2: Latent Modulated SIREN

Results

- Successfully trained our model on the CIFAR-10 dataset, integrating INRs with shift modulations for image classification.
- Attained a top-1 classification accuracy of 59%, showcasing the potential of our method in handling standard resolution images.

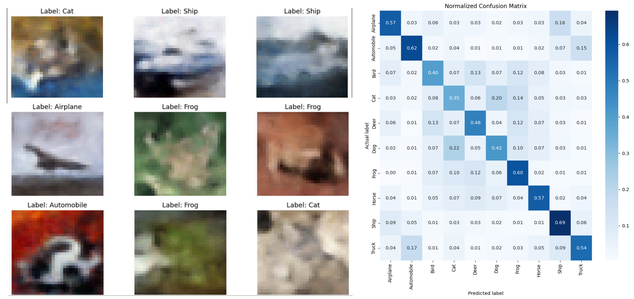

Figure 3: Reconstructed image and Normalized Confusion Matrix from CIFAR-10 Classification

Conclusions & Future work

- We were able to validate the approach as a viable method for classifying images using a base network and INR technique.
- Dataset Expansion: We are planning to augment our dataset to include images of various resolutions, broadening the scope of our model's applicability.
- Constructing Additional Functa-Sets: We will construct more functa-sets to facilitate the classification of these diverse-resolution images.
- Maintaining Accuracy: We aim to maintain sufficient accuracy levels post-expansion, ensuring the model's robustness and scalability.
- Whole-Slide Imaging (WSI) Application: Applying our trained model to WSI datasets, taking steps towards real-world medical diagnostic applications.

Acknowledgement

Special thanks to Dr. Kookjin Lee, and John Perkins (Masters student in School of Computintg and Augmented Intelligence) for their unparalleled support and contribution to the project

References

[1] Sitzmann, V., Martel, J. N. P., Bergman, A. W., Lindell, D. B., & Wetzstein, G. (2020). *Implicit neural representations with periodic activation functions. arXiv. Retrieved from https://ar5iv.org/abs/2006.09661&#8203*

[2] Dupont, E., Kim, H., Eslami, S. M. A., Rezende, D., & Rosenbaum, D. (2022). *From data to functa: Your data point is a function and you can treat it like one. Proceedings of Machine Learning Research, 162, 5694-5725.*

FURI

Ira A. Fulton Schools of Engineering
Arizona State University