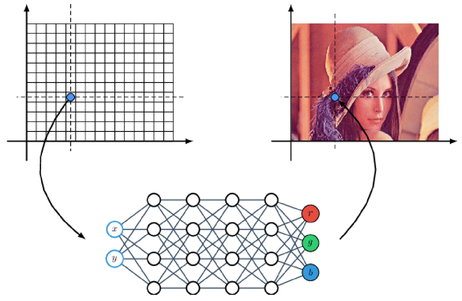
Meta Learning For Feature Grid Based INRs

Andrew Tianbao Zhong, Darshit Miteshkumar Desai, Hitesh Kyatham, Jayasuriya K Suresh, Vinay Krishna Bukka

Introduction of the Implicit Neural Representation

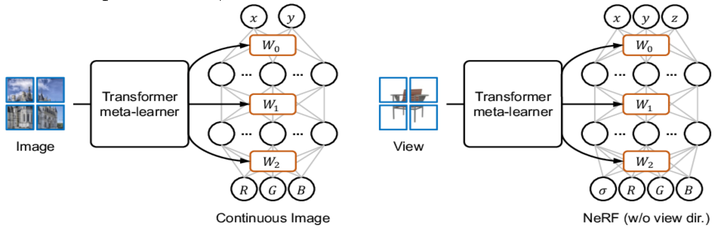
- INRs are typically multilayer perceptrons organized to learn a scalar or vector field. They take a coordinate as input and predict a continuous signal value as output(such as pixel).



Picture credit: Adversarial Generation of Continuous Images.

Introduction of the Meta Learning Model

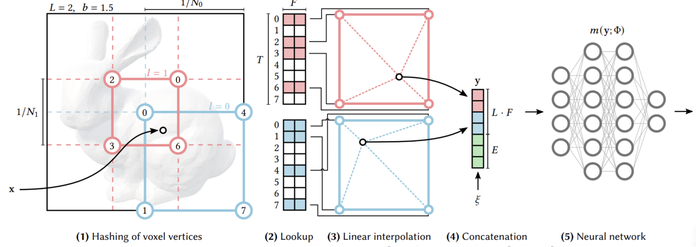
- A Transformer hypernetwork to infer the whole weights in an INR (Implicit Neural Representation)



Picture credit: Transformers as Meta Learners for Neural Representations

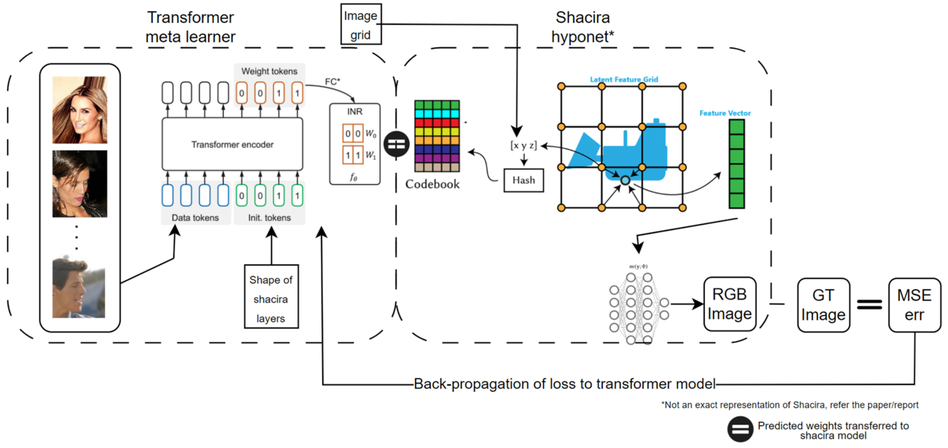
Introduction of the Feature Grid Based INR

- Feature grids are a type of input encoding for neural networks that store trainable features in a grid structure, which are passed as input to INR.



Picture credit: Instant Neural Graphic Primitives

Pipeline (for each epoch)



Model Architecture Description

Transformer with GELU + Feature Grid INR
(Trans INR) (SHACIRA)

Experimental Details and Figures

Learning Rate & Batch Size	Latent Quantization	Color Decoder MLP	Loss	PSNR
Lr: 0.05 Batch Size: 9	Enabled	Enabled	0.0708	11.8
Lr: 0.1 Batch Size: 6	Disabled	Disabled	0.0668	12.078

Table : Hyperparameter variation

Codebook Size	Loss	PSNR
47737	0.0708	11.806
26704	0.0668	12.07
4096	0.0714	11.737

Table : Codebook Size Variation

Hyperparameters	Network Model	LOSS	PSNR
Learning rate: 0.01 Number of Epochs:60k Batch size: 1	SHACIRA Standalone ~ 10 mins	0.001	40.67
Learning rate: 0.01 Number of Epochs:30 Batch size: 16	HypoShacira ~ 8 hrs (TransINR+ Shacira)	0.1	10

Table : Network Model Variations

Results / Findings so far

- Implementation and study of existing works trans inr and SHACIRA.
- Current results of feature grid based inr with meta learned

Fig: Standalone Shacira result

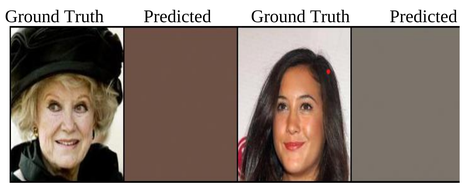
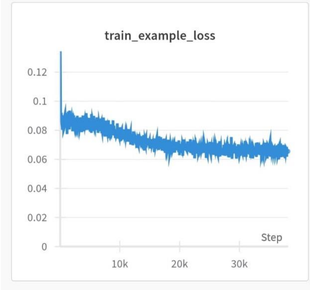
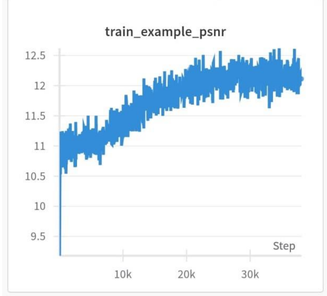


Fig: TransShacira results of Reconstructed Images with Color Decoder+Codebook enabled

- Training loss and PSNR Graphs





Why the loss saturates after some steps?

We haven't found the cause, but while debugging it was observed that the transformer predicts the weight values which are almost the same (at step 20 and at step 20000), we have debugged end to end to check why this is happening but haven't found a concrete cause.

Conclusions

- With implementation of the proposed methodology, the reconstruction has not been successful so far.
- To improve the results through hyperparameter tuning of transformer network is planned and explore alternate methods.

Applications

Higher quality image and video compression with faster training times