Electronics Research Institute
Sharif University of Technology

icassp 2022 Singapore

Multi-Head ReLU Implicit Neural Representation Networks

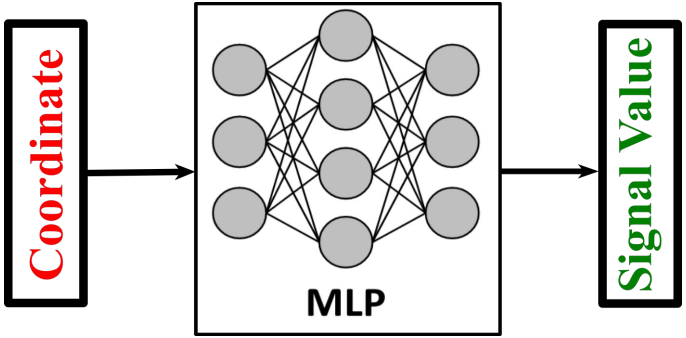
Authors: Arya Aftab, Alireza Morsali, Shahrokh Ghaemmaghami

Presented by: Arya Aftab

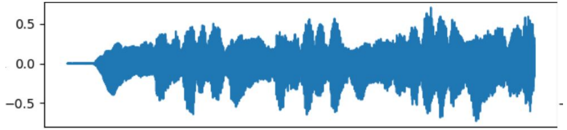
IEEE Signal Processing Society

What is the implicit neural representation?

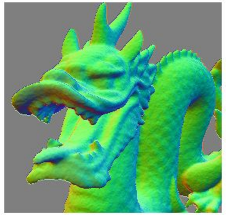
$$S = f(\theta, x)$$

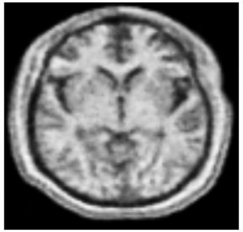


Speech:



Vision

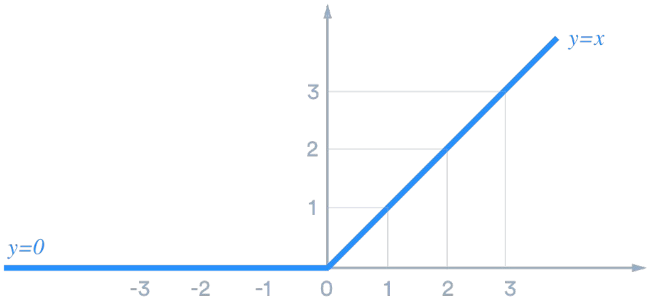


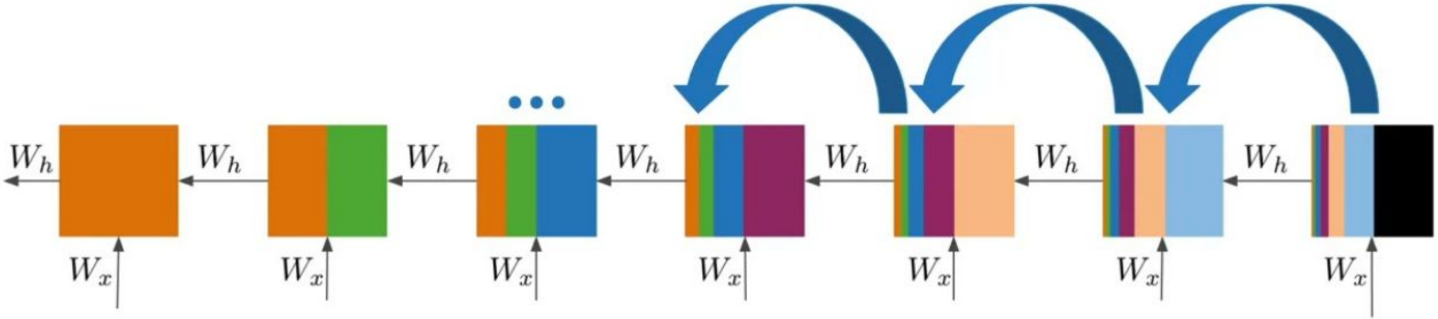


M. Tancik, P. P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J. T. Barron, and R. Ng, "Fourier features let networks learn high frequency functions in low dimensional domains," in Conf. Neural Inf. Process. Syst. (NeurIPS), June 2020.

ReLU Networks

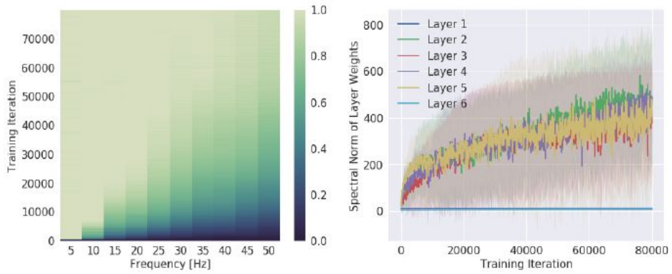
$$f(x) = max(0, x)$$

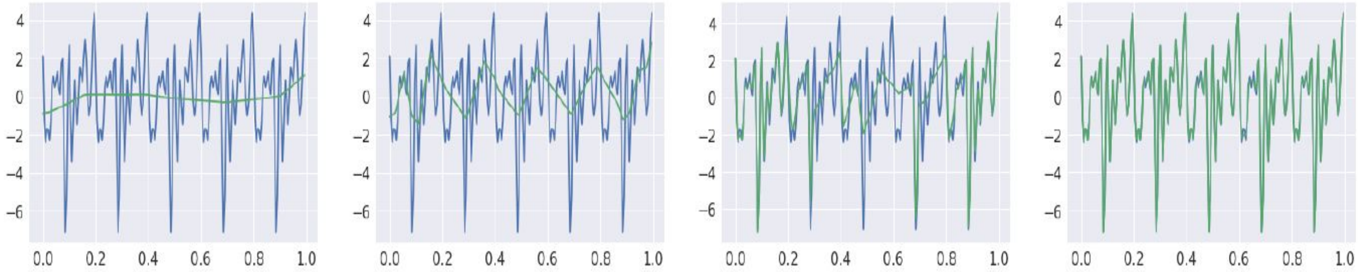




What is the problem with ReLU networks?

Spectral bias





(a) Iteration 100 (b) Iteration 1000 (c) Iteration 10000 (d) Iteration 80000

M. Tancik, P. P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J. T. Barron, and R. Ng, “Fourier features let networks learn high frequency functions in low dimensional domains,” in Conf. Neural Inf. Process. Syst. (NeurIPS), June 2020.

How to Solve this problem?

- **Positional encoding**

$$\gamma(p) = \left(\sin\left(2^0 \pi p\right), \cos\left(2^0 \pi p\right), \cdots, \sin\left(2^{L-1} \pi p\right), \cos\left(2^{L-1} \pi p\right) \right)$$ Eq. 1

- **Fourier Features**

$$\gamma(\mathbf{v}) = \left[a_1 \cos(2\pi \mathbf{b}_1^{\mathrm{T}} \mathbf{v}), a_1 \sin(2\pi \mathbf{b}_1^{\mathrm{T}} \mathbf{v}), \ldots, a_m \cos(2\pi \mathbf{b}_m^{\mathrm{T}} \mathbf{v}), a_m \sin(2\pi \mathbf{b}_m^{\mathrm{T}} \mathbf{v})\right]^{\mathrm{T}}$$ Eq. 2

- **Sine activation function**

$$\gamma(\mathrm{v}) = \sin(W^T \mathrm{v} + b)$$ Eq. 3

Our approach

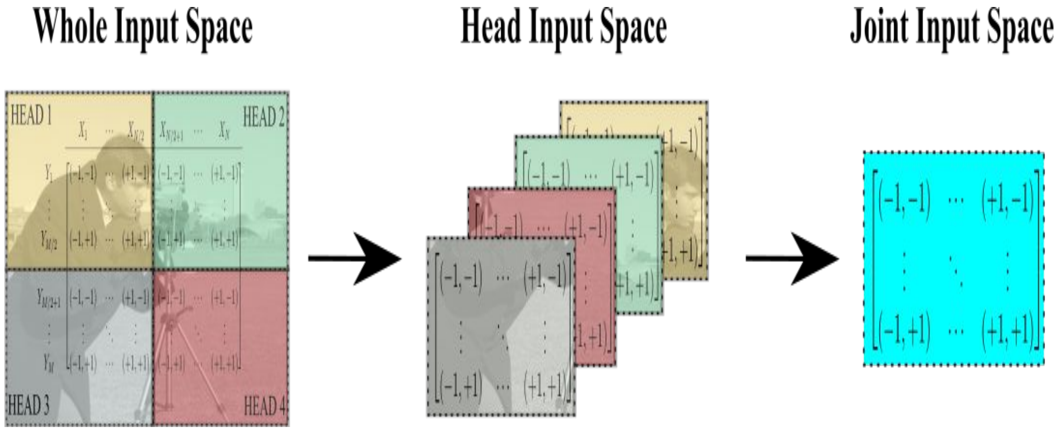
Localization

$$\Phi\big(x(r), y(c)\big) = \mathbf{I}[r, c]$$ (Eq. 1

$$\mathcal{D} = \Big\{\Big(\big(x(r), y(c)\big), \mathbf{I}[r, c]\Big)\Big\}_{r,c=1}^{N_x, N_y}$$ (Eq. 2

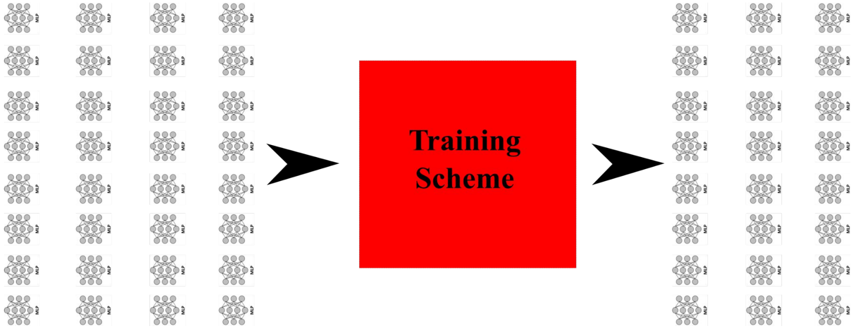
$$\mathbf{I}_{l,k}[\hat{r}, \hat{c}] = \mathbf{I}[\hat{N}_h(l-1) + \hat{r}, \hat{N}_w(k-1) + \hat{c}]$$ (Eq. 3

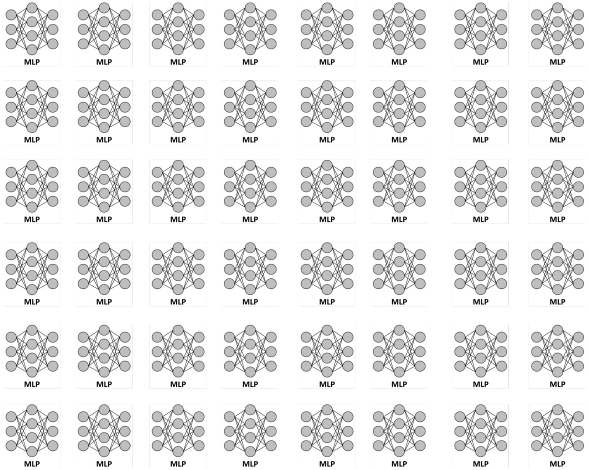
$$\phi_{l,k}\big(\hat{x}(\hat{r}), \hat{y}(\hat{c})\big) = \mathbf{I}_{l,k}[\hat{r}, \hat{c}]$$ (Eq. 4



Challenge in our approach

- **High number of small networks**
- **Difficulty in training the whole structure**



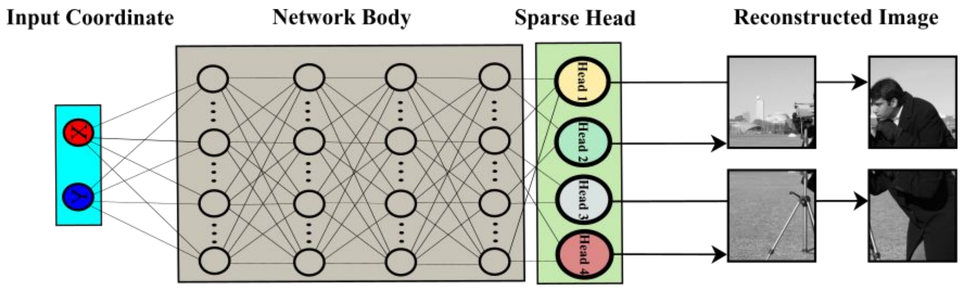


How to solve this Challenge?

- **Share network body between small networks**
- **Consider all heads as a single layer**
- **Use a sparse layer instead of a dense layer for the heads**

$$\phi_{l,k}(\hat{x}, \hat{y}) = \tau_{l,k}(\psi(\hat{x}, \hat{y}))$$ (Eq. 1

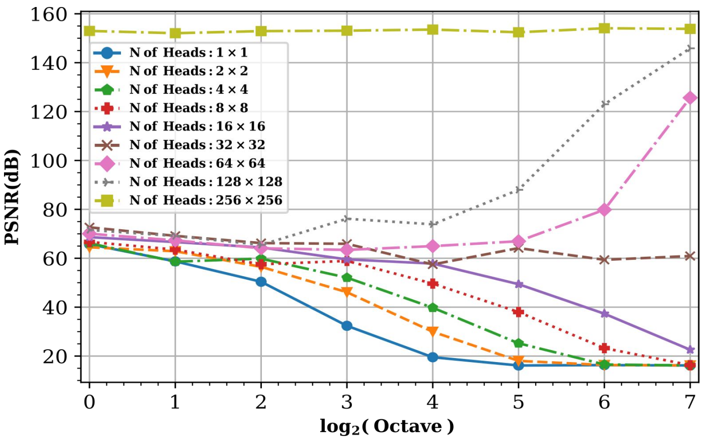
$$\min_{\tau_{l,k}(),\psi()} \sum_{\hat{r}=1}^{\hat{N}_h} \sum_{\hat{c}=1}^{\hat{N}_w} \sum_{l=1}^{H_x} \sum_{k=1}^{H_y} \left(\tau_{l,k}\Big(\psi\big(\hat{x}(\hat{r}), \hat{y}(\hat{c})\big)\Big) - \mathbf{I}_{l,k}[\hat{r}, \hat{c}] \right)^2$$ (Eq. 2

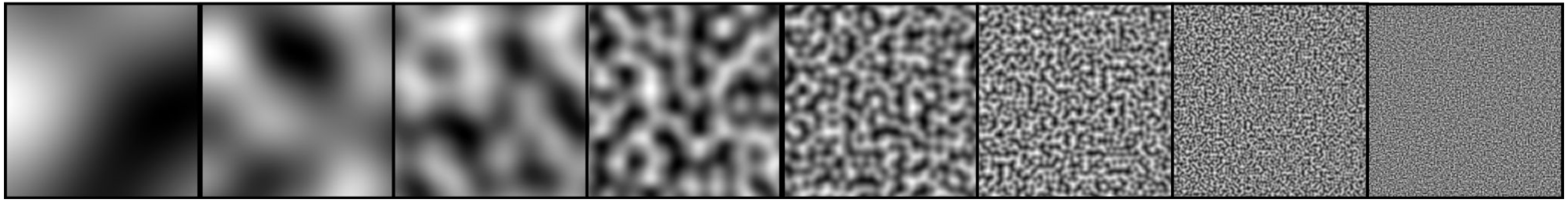


Results

Spectral bias

Train with Perlin noise



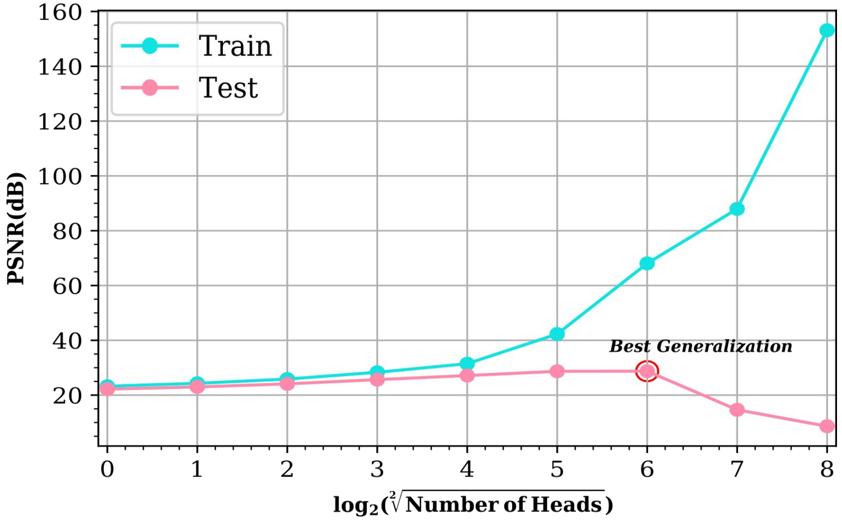


Results

Generalization ability

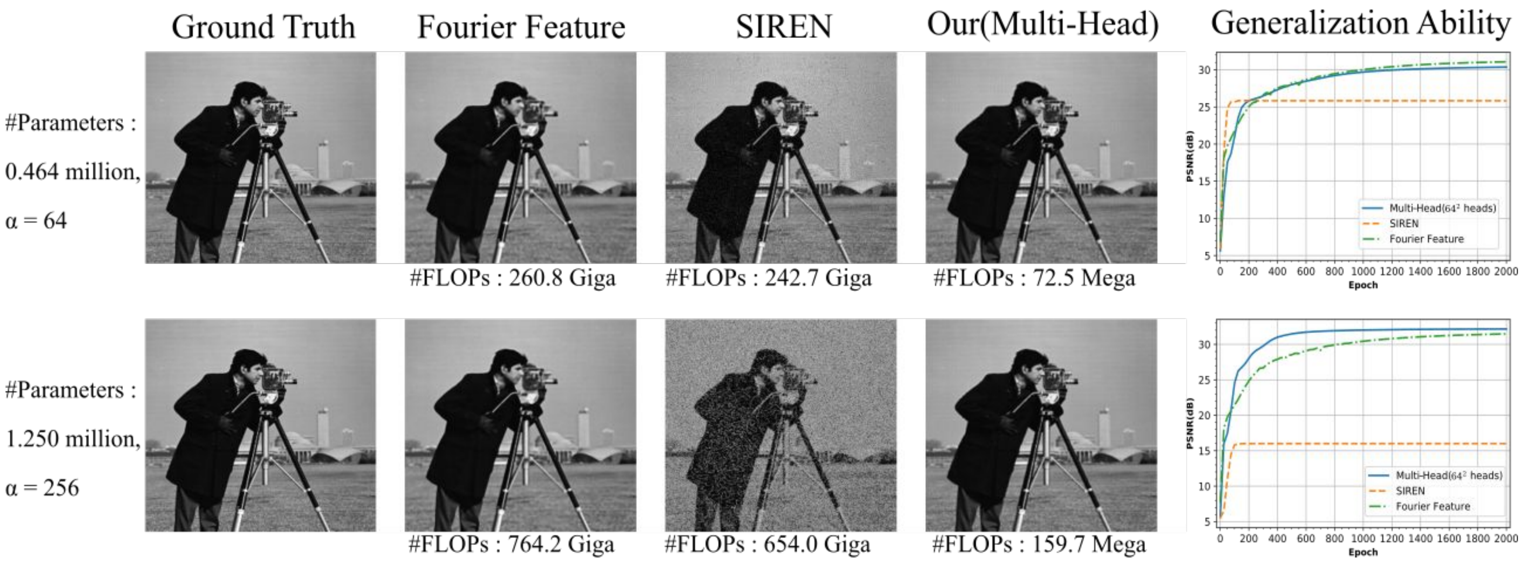
Train with down-sampled image: 256*256

Evaluate with original image: 512*512



Results

Comparison



Conclusions

- We proposed a novel structure that tackle spectral bias.
- The proposed structure has a much lower computational cost, as compared to other methods.
- Due to the shrinking input space with the increasing number of heads, training of the proposed structure takes less time than the time needed in existing methods.
- Experimental results show that the performance of our model is comparable to that of state-of-the-art methods.

References

1-M. Tancik, P. P. Srinivasan, B. Mildenhall, S. Fridovich-Keil, N. Raghavan, U. Singhal, R. Ramamoorthi, J. T. Barron, and R. Ng, "Fourier features let networks learn high frequency functions in low dimensional domains," in Conf. Neural Inf. Process. Syst. (NeurIPS), June 2020.

2-V. Sitzmann, J. Martel, A. Bergman, D. Lindell, and G. Wetzstein, "Implicit neural representations with periodic activation functions," in Conf. Neural Inf. Process. Syst. (NeurIPS), June 2020

3-B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, "NERF: Representing scenes as neural radiance fields for view synthesis," in European Conf. Comput. Vis. (ECCV), Aug. 2020, pp. 405–421.

4-R. Arora, A. Basu, P. Mianjy, and A. Mukherjee, "Understanding deep neural networks with rectified linear units," arXiv preprint arXiv:1611.01491, 2016.

5-N. Rahaman, A. Baratin, D. Arpit, F. Draxler, M. Lin, F. Hamprecht, Y. Bengio, and A. Courville, "On the spectral bias of neural networks," in Int. Conf. Mach. Learn. (ICML). PMLR, July 2019, pp. 5301–5310.