

# Investigating the Implicit Bias of Activations in Coordinate-MLPs

Bruce Balfour, Ruben Schenk, Alexandra Trofimova

ETH Zürich

Institute for Visual Computing  
Computer Vision and Learning Group  
Universitätstrasse 6  
8092 Zürich

{balfourb, rschenk, atrofimo}@ethz.ch

[github.com/rubenwgs/ibac](https://github.com/rubenwgs/ibac)

## Abstract

The accurate and efficient representation of multidimensional continuous signals is essential in computer vision and graphics. Coordinate-based Multi-Layer Perceptrons (MLPs), or implicit neural representations (INRs), offer a promising alternative to traditional grid-based methods by encoding signals continuously and compactly. This project explores the implicit biases of non-periodic activation functions – Gaussian, Laplacian, and Quadratic – in the context of 2D video approximation and sparse 3D reconstruction. We conducted extensive experiments to evaluate these activation functions, focusing on their performance and generalisation capabilities. Our findings reveal that non-periodic activations can provide smoother interpolations and improved robustness to sparse inputs compared to periodic activations like SIREN. Additionally, we propose a geometric initialisation scheme for Gaussian activation that enhances its stability and convergence in 3D approximation tasks. This work extends the understanding of activation functions in coordinate MLPs and highlights their potential for diverse signal representations.

## 1. Introduction

**Goals.** The accurate and efficient representation of multidimensional continuous signals is a crucial task in computer vision and graphics, particularly for applications such as 2D video approximation and sparse 3D reconstruction [3]. Traditional grid-based methods for representing these signals often suffer from limitations in resolution and memory efficiency. Coordinate-based MLPs, also known as INRs, offer a promising alternative by encoding signals continuously and compactly [4]. However, the choice of activation functions in these networks significantly impacts

their performance and generalisation capabilities. This work aims to explore the implicit biases of non-periodic activation functions, specifically Gaussian, Laplacian, and Quadratic, in the context of 2D video approximation and sparse 3D reconstruction.

**Problems.** Despite the advancements brought by periodic activations such as those proposed in [6], existing methods still face challenges. Sinusoidal activations can effectively handle high-frequency components without positional embeddings, but their performance is highly sensitive to initialisation schemes and often lacks robustness [4]. The reliance on periodic activations limits the flexibility and applicability of these models to a broader range of signals, particularly those with non-periodic characteristics. Furthermore, traditional activations like ReLU in coordinate MLPs fail to adequately capture high-frequency details, necessitating the use of positional embeddings, which adds complexity and computational overhead, as highlighted in [4].

**Solution.** Our proposed solution involves investigating the performance and the implicit bias of non-periodic activation functions – Gaussian, Laplacian, and Quadratic – in coordinate MLPs. By moving beyond the periodic activations, we aim to develop an understanding of models that are more robust to sparse inputs and capable of capturing a wider variety of signal characteristics. Non-periodic activations can potentially offer smoother interpolations and better generalisation properties, ensuring that nearby coordinate points in the input space produce similar output behaviours.

**Contributions.** Our work makes the following contributions:

- **Extensive Empirical Evaluation on 2D Video Approximation:** We conducted a comprehensive set of experiments to assess the performance of Gaussian, Laplacian, and Quadratic activation functions in the

context of 2D video approximation. This included varying the percentages for the train/test split and evaluating the impact of ResFields.

- **Experiments on Sparse 3D Reconstruction:** Building on the insights from our 2D video approximation experiments, we carried out a subset of experiments on sparse 3D reconstruction. This allowed us to assess the applicability of non-periodic activations in a different but related domain, providing a broader perspective on their performance.
- **Improvement in 3D Approximation for Gaussian Activation:** We specifically enhanced the performance of the Gaussian activation function in 3D approximation tasks by deriving a geometric initialisation scheme. This initialisation approach improved the stability and convergence of models using Gaussian activations, demonstrating the practical benefits of tailored initialisation methods.

By conducting these experiments, and by developing an initialisation method for the Gaussian activation, our work extends the current understanding of activation functions in coordinate MLPs.

## 2. Related Work

The field of INRs has seen significant advancements, with various methods proposed to enhance the efficiency and accuracy of representing multidimensional continuous signals.

**Periodic activation functions** in INRs have been notably advanced by [6], who introduced SIREN, a method that uses sinusoidal activations to model high-frequency details without positional embeddings. This approach has proven effective for high-frequency components, but its sensitivity to initialisation schemes can limit robustness and applicability across diverse signal types. Similar to this approach, [5] proposed the Hyperbolic Oscillation (HOSC) activation function, which is designed to better preserve high-frequency details compared to traditional periodic activations like SIREN. HOSC has been shown to outperform both ReLU and SIREN in convergence speed and loss reduction in signal encoding tasks, making it a strong candidate for INRs involving detailed signal representations.

**Non-periodic activation functions** in INRs have also been explored to address the limitations of periodic activations. [4] proposed a unifying framework for activations in coordinate MLPs, advocating for the use of non-periodic activation functions such as Gaussian and Laplacian. These functions exhibit flexibility in encoding signals with varying local Lipschitz smoothness, broadening the range of applicable signals.

**Architectural enhancements** in INRs have further advanced the field. [3] introduce ResFields, which use residual connections within neural fields to improve the learning of spatiotemporal patterns. This approach demonstrates enhanced performance in complex signal domains such as video approximation and 3D reconstruction, emphasising the importance of architectural modifications in complementing activation function improvements to achieve state-of-the-art results. Similarly, the HOIN framework proposed by [1] addresses the spectral bias in INRs by introducing high-order interaction blocks. These architectural enhancements effectively capture high-frequency components while mitigating spectral bias, showing significant improvements in model performance when compared to traditional methods.

## 3. Method

**Preliminary Knowledge.** INRs use neural networks to encode continuous multidimensional signals. A common approach involves coordinate MLPs, where the input is a coordinate (e.g., spatial coordinates for images) and the output is the signal value at that coordinate (e.g., pixel intensity). The performance of these MLPs is significantly influenced by the choice of activation functions.

**Activation Functions.** We explore both periodic and non-periodic activation functions in our experiments. The activations used are as follows:

- Gaussian Activation:  $\sigma(x) = e^{\frac{-0.5x^2}{\alpha^2}}$
- Laplacian Activation:  $\sigma(x) = e^{\frac{-|x|}{\alpha}}$
- Quadratic Activation:  $\sigma(x) = \frac{1}{1+\alpha x^2}$
- ReLU (Rectified Linear Unit):  $\sigma(x) = \max(0, x)$
- SIREN (Sinusoidal Activation):  $\sigma(x) = \sin(30x)$

**Model.** The coordinate MLP can be described by the following equations:

$$f(\mathbf{x}) = \mathbf{W}_L \sigma(\mathbf{W}_{L-1} \sigma(\dots \sigma(\mathbf{W}_1 \mathbf{x} + \mathbf{b}_1) \dots + \mathbf{b}_{L-1}) + \mathbf{b}_L), \quad (1)$$

where  $\mathbf{x}$  is the input coordinate,  $\mathbf{W}_i$  and  $\mathbf{b}_i$  are the weights and biases of the  $i$ -th layer, and  $\sigma$  is the activation function.

**Experimental Setup.** We conducted extensive experiments to evaluate the performance of different activation functions in both 2D video approximation and sparse 3D reconstruction.

For the 2D video approximation, we used two videos, `cat.mp4` and `bikes.mp4`. We experimented with five different activation functions: Gaussian, Quadratic, Laplacian, ReLU, and SIREN. For each activation function, we tested configurations with 256 and 512 hidden features, and

varied the percentage of data used for testing (90%, 50%, 30%, 10%). The rationale for varying the test data percentage is to assess model performance under different levels of data sparsity, mimicking conditions of sparse reconstruction where less training data is available. Each experiment was conducted with and without ResField, resulting in a total of 80 experiments. We recorded the Loss, Train PSNR, and Test PSNR at every 20,000 steps up to 100,000 steps. All experiments were run on ETH’s Euler computer, equipped with an RTX 3090 GPU, 6 cores, and either 4GB of RAM per core for the 256 hidden feature cases or 6GB per core for the 512 hidden feature cases.

For sparse 3D reconstruction, we evaluated the performance of Gaussian, Laplacian, and Quadratic activation functions on both DyNeRF (Dynamic Neural Radiance Fields) [2] and TNeRF (Temporal Neural Radiance Fields) [2] networks. DyNeRF and TNeRF are networks designed to model dynamic and temporal aspects of 3D scenes. We conducted experiments with no ResField layers and with ResField layers at  $i = 1, 2, 3$ . We evaluated the performance using PSNR and SSIM metrics. Similar to the video approximation experiments, these were conducted on the Euler computer with an RTX 3090 GPU, 6 cores, and 6GB of RAM per core.

**ResFields Model.** To address the capacity bottleneck in modelling complex spatiotemporal signals, we introduce residual field layers (ResFields [3]) within the MLP architecture. A ResField layer can be defined as:

$$\phi_i(t, \mathbf{x}_i) = \sigma_i((\mathbf{W}_i + \mathcal{W}_i(t))\mathbf{x}_i + \mathbf{b}_i), \quad (2)$$

where  $\mathcal{W}$  is a time-dependent weight matrix that models the residuals of the network weights. This formulation increases the model capacity via additional trainable parameters without modifying the overall network architecture.

The residual of network weights is defined as:

$$\mathcal{W}_i(t) = \sum_{r=1}^{R_i} \mathbf{v}_i(t)[r] \cdot \mathbf{M}_i[r], \quad (3)$$

where the coefficients  $\mathbf{v}_i(t) \in \mathbb{R}^{R_i}$  and the spanning set  $\mathbf{M}_i \in \mathbb{R}^{R_i \times N_i \times M_i}$  are trainable parameters. This low-rank factorisation reduces the total number of trainable parameters, helping to prevent overfitting.

## 4. Experiments

### 4.1. Overview

**Datasets.** We utilised two datasets for our experiments: one for 2D video approximation and another for sparse 3D reconstruction. For the 2D video approximation tasks, we used the following videos:

- Cat Video from [6]: A 512x512 pixel video, 12 seconds long at 25 frames per second (fps), resulting in 300 frames.
- Bikes Video from [6]: A 640x272 pixel video, 10 seconds long at 25 fps, resulting in 250 frames.

For the sparse 3D reconstruction tasks, we used the basketball and dancer sequences from the Owlii dataset [7]. Each sequence is captured at 30 fps over a 20-second period. Depth maps were not used for training. These sequences provide complex, dynamic 3D scenes ideal for evaluating the performance of our models.

**Evaluation Metrics.** To assess the performance of our models, we used the following metrics:

- PSNR (Peak Signal-to-Noise Ratio): Defined as

$$\text{PSNR} = 10 \log_{10} \left( \frac{\text{MAX}^2}{\text{MSE}} \right), \quad (4)$$

where MAX is the maximum possible pixel value of the image. Higher PSNR indicates better reconstruction quality.

- SSIM (Structural Similarity Index Measure): Measures the similarity between two images, considering luminance, contrast, and structure. It is defined as:

$$\text{SSIM}(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}, \quad (5)$$

where  $\mu_x$  and  $\mu_y$  are the means of  $x$  and  $y$ ,  $\sigma_x^2$  and  $\sigma_y^2$  are the variances,  $\sigma_{xy}$  is the covariance of  $x$  and  $y$ , and  $C_1$  and  $C_2$  are constants to stabilise the division. Higher SSIM values indicate better structural similarity.

**Baselines and Comparisons.** For the 2D video approximation task, we compared our proposed non-periodic activation functions (Gaussian, Laplacian, Quadratic) against traditional activation functions (ReLU) and periodic activation functions (SIREN). The results of the 3D sparse reconstruction task were compared to the baselines established in the ResFields paper [3].

## 4.2. Results and Discussion

### 4.2.1 2D Video Approximation

**Results.** We evaluated the performance of different activation functions on the `cat.mp4` and `bikes.mp4` videos using two configurations: 256 and 512 hidden neurons, and with test data percentages of 90% and 30%.

For the 256 neuron configuration with 90% test data (Tab. 1), Gaussian and Quadratic activations achieved similar Train and Test PSNR values, with Gaussian showing

a minor drop ( $\Delta = -0.43$ ) and Quadratic ( $\Delta = -0.41$ ). Laplacian had a slightly smaller drop ( $\Delta = -0.33$ ), while ReLU showed the smallest difference ( $\Delta = -0.13$ ). SIREN exhibited the highest Train PSNR (44.19) but a significant drop in Test PSNR ( $\Delta = -14.39$ ), indicating overfitting. When using 30% test data (Tab. 2), all activations performed better with smaller differences between Train and Test PSNR values. ReLU again showed the smallest drop ( $\Delta = -0.02$ ), and SIREN’s performance improved but still had notable overfitting ( $\Delta = -1.04$ ).

For the 512 neuron configuration, Gaussian activation had a Train PSNR of 32.88 and a Test PSNR of 30.85 ( $\Delta = -2.04$ ) with 90% test data (Tab. 1). Quadratic and Laplacian had Train/Test PSNR differences of  $-0.68$  and  $-1.82$ , respectively. ReLU continued to perform stably with a  $\Delta$  of  $-0.22$ , while SIREN had the highest Train PSNR (52.04) but the largest drop ( $\Delta = -22.21$ ). With 30% test data (Tab. 2), Gaussian, Quadratic, and Laplacian activations maintained small differences between Train and Test PSNR values ( $\Delta$  of  $-0.10$ ,  $-0.07$ , and  $-0.16$ , respectively). ReLU showed the smallest drop ( $\Delta = -0.03$ ), and SIREN, although improved, still showed notable overfitting ( $\Delta = -2.10$ ). Some example frames for the results of the 512 neuron configuration are shown in Fig. 1 for the `cat.mp4` sequence, and in Fig. 2 for the `bikes.mp4` sequence.

**Discussion.** The results indicate that non-periodic activation functions (Gaussian, Quadratic, Laplacian) generally achieve consistent performance across different test data percentages, with smaller differences between Train and Test PSNR values. ReLU showed the smallest difference, indicating good generalisation capability, while SIREN, despite achieving high Train PSNR, consistently showed significant drops in Test PSNR, highlighting issues with overfitting.

The data suggests that non-periodic activations may have stronger implicit biases that allow them to generalise better in sparse data conditions. This is particularly evident in the smaller drops in PSNR values across different test scenarios. These activations potentially offer smoother interpolations and better robustness compared to periodic activations like SIREN, which tend to overfit, especially with a higher number of neurons and a lower percentage of training data.

#### 4.2.2 Sparse 3D Reconstruction

**3D Sparse Reconstruction.** We evaluated the performance of different activation functions on both DyNeRF and TNeRF networks using two sequences from the Owlii dataset, namely "Basketball" and "Dancer." The results are summarised in Tab. 3 and Tab. 4.

For the DyNeRF baseline without ResFields, the network achieved an SSIM of 92.05 and a PSNR of 23.41 on

		90% Test Data		
	$\sigma(\mathbf{x})$	$t \downarrow$	Train PSNR	Test PSNR
256 Neurons	Gaussian	1.6h	28.80	28.37
	Quadratic	1.6h	28.26	27.85
	Laplacian	1.5h	26.15	25.82
	ReLU	0.9h	23.96	23.84
	SIREN	1.0h	44.19	29.80
512 Neurons	Gaussian	4.0h	32.88	30.85
	Quadratic	3.9h	29.79	29.11
	Laplacian	3.9h	31.52	29.70
	ReLU	2.6h	25.32	25.09
	SIREN	2.9h	52.04	29.83

Table 1. Mean PSNR values of different activation functions on the `cat.mp4` and `bikes.mp4` video with ResField and 90% test data.

		30% Test Data		
	$\sigma(\mathbf{x})$	$t \downarrow$	Train PSNR	Test PSNR
256 Neurons	Gaussian	1.6h	29.04	28.98
	Quadratic	1.6h	28.09	28.04
	Laplacian	1.5h	27.20	27.13
	ReLU	0.9h	23.98	23.97
	SIREN	1.0h	39.49	38.45
512 Neurons	Gaussian	4.0h	30.41	30.31
	Quadratic	3.9h	29.18	29.11
	Laplacian	3.9h	30.26	30.11
	ReLU	2.6h	25.18	25.15
	SIREN	2.9h	43.75	41.65

Table 2. Mean PSNR values of different activation functions on the `cat.mp4` and `bikes.mp4` video with ResField and 30% test data.

average. Adding ResFields improved performance to an SSIM of 93.59 and a PSNR of 24.99. This improvement was consistent across both sequences, indicating that ResFields help capture more details in the reconstruction.

With Laplacian activation, the DyNeRF network achieved an SSIM of 89.30 and a PSNR of 22.14 without ResFields. Adding ResFields slightly improved the results to an SSIM of 89.87 and a PSNR of 21.52 on average. Similar patterns were observed for both sequences.

Gaussian and Quadratic activations did not learn effectively with the current initialisation scheme, resulting in poor performance with an SSIM of 86.51 and a PSNR of 12.39 for both. These results remained the same with and without ResFields, suggesting that the network did not learn anything meaningful.

For the TNeRF baseline without ResFields, the network achieved an SSIM of 94.18 and a PSNR of 26.18 on average. Adding ResFields further improved performance to an SSIM of 95.21 and a PSNR of 27.44. This trend was consistent across both sequences, similar to the DyNeRF network.

Using Laplacian activation, the TNeRF network achieved an SSIM of 89.31 and a PSNR of 22.15 without ResFields. Adding ResFields resulted in a slight improvement, with an SSIM of 89.88 and a PSNR of 21.52. The patterns observed were consistent across both sequences.

Similar to the DyNeRF network, Gaussian and Quadratic activations in the TNeRF network also failed to learn ef-

Network	$t \downarrow$	Mean		Basketball		Dancer	
		SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑
DyNeRF + ResFields ( $i = 1, 2, 3$ )	12h	92.05 93.59	23.41 24.99	92.56 93.49	23.49 24.77	91.54 93.69	23.33 25.22
DyNeRF Laplacian + ResFields ( $i = 1, 2, 3$ )	18h	89.30 89.87	22.14 21.52	90.01 89.46	22.38 20.82	88.60 90.29	21.91 22.22
DyNeRF Gaussian + ResFields ( $i = 1, 2, 3$ )	18h	86.51 86.51	12.39 12.39	86.01 86.01	12.09 12.09	87.01 87.01	12.69 12.69
DyNeRF Quadratic + ResFields ( $i = 1, 2, 3$ )	21h	86.51 86.51	12.39 12.39	86.01 86.01	12.09 12.09	87.01 87.01	12.69 12.69

Table 3. Results for DyNeRF networks with different activation functions. The table shows the SSIM and PSNR values for the Basketball and Dancer sequences, as well as the mean of both.

Network	$t \downarrow$	Mean		Basketball		Dancer	
		SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑
TNeRF + ResFields ( $i = 1, 2, 3$ )	12h	94.18 95.21	26.18 27.44	94.57 95.84	26.33 27.98	93.53 94.87	25.09 26.55
TNeRF Laplacian + ResFields ( $i = 1, 2, 3$ )	17h	89.31 89.88	22.15 21.52	90.01 89.46	22.38 20.82	88.60 90.29	21.91 22.22
TNeRF Gaussian + ResFields ( $i = 1, 2, 3$ )	18.5h	86.51 86.51	12.39 12.39	86.01 86.01	12.09 12.09	87.01 87.01	12.69 12.69
TNeRF Quadratic + ResFields ( $i = 1, 2, 3$ )	20h	86.51 86.51	12.39 12.39	86.01 86.01	12.09 12.09	87.01 87.01	12.69 12.69

Table 4. Results for TNeRF networks with different activation functions. The table shows the SSIM and PSNR values for the Basketball and Dancer sequences, as well as the mean of both.

fectively with the current initialisation scheme, resulting in poor performance with an SSIM of 86.51 and a PSNR of 12.39 for both. These results were unchanged with or without ResFields.

**Discussion.** The results highlight the importance of proper initialisation for different activation functions. While the SIREN initialisation works well for periodic activations, it appears inadequate for non-periodic ones like Gaussian and Quadratic, leading to poor learning outcomes. This indicates a bias in the initialisation scheme, which needs to be tailored to the specific activation function to ensure effective learning. The Laplacian activation, although better than Gaussian and Quadratic, still underperformed compared to the baseline, further emphasising the need for suitable initialisation strategies for non-periodic activations.

Insights from the 2D video approximation task suggest that non-periodic activations have potential for better generalisation under data sparsity conditions. However, this potential did not fully transfer to the 3D reconstruction task. The poor performance of Gaussian and Quadratic activations in the 3D task highlights the necessity of appropriate initialisation, as the same activations performed adequately in the 2D task.

### 4.3. Geometric Initialisation

**Introduction.** In the context of INRs, the initialisation of network parameters significantly influences the ability of the network to learn complex signals. Standard initialisation

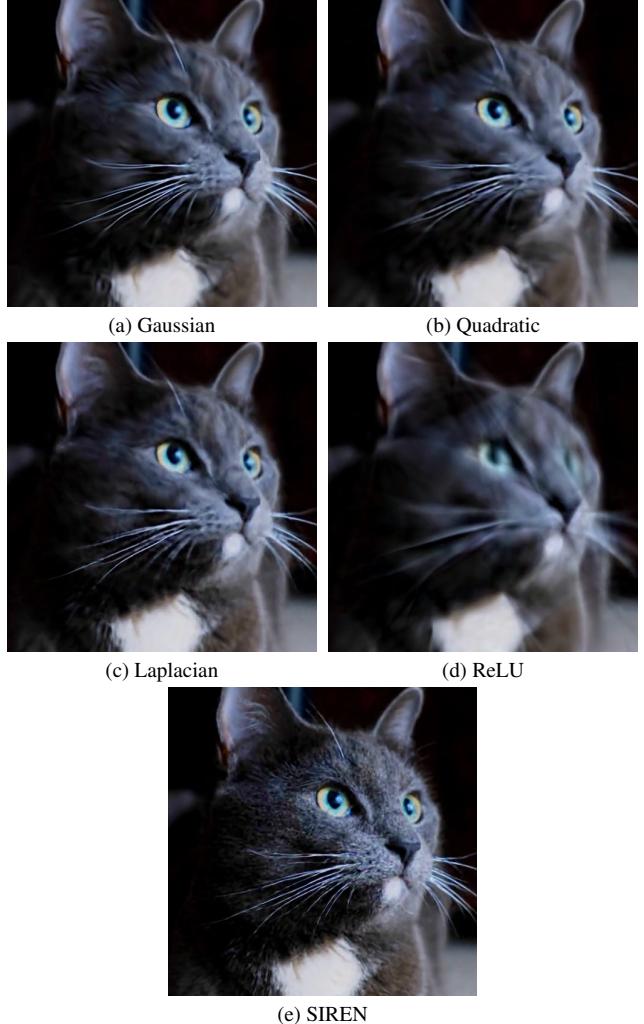


Figure 1. Frames from the 2D video approximation tasks using the `cat.mp4` video, with 512 hidden features, ResFields, and 90% test data.

techniques, such as those used for periodic activations like in SIREN networks, might not be suitable for non-periodic activations like Gaussian functions. This mismatch can lead to suboptimal performance, as observed in our experiments with Gaussian and Quadratic activations. To address this issue, we derived a tailored geometric initialisation for the Gaussian activation function, ensuring the network starts with parameters that allow for effective learning.

**Derivation of the Geometric Initialisation for Gaussian Activation.** The Gaussian function used in our experiments is defined as

$$g(x) = e^{-\frac{0.5x^2}{a^2}}$$

with  $a = 1$ . For an  $n$ -dimensional vector  $z$  where each entry is  $\frac{1}{n}$ , and an  $n \times m$  matrix  $W$  with i.i.d. entries  $W \sim$



(a) Gaussian



(b) Quadratic



(c) Laplacian



(d) ReLU



(e) SIREN

Figure 2. Frames from the 2D video approximation tasks using the `bikes.mp4` video, with 512 hidden features, ResFields, and 90% test data.

$\mathcal{N}(0, 1)$ , we aim to compute the expectation of

$$f(x) = z^\top \sum_{i=1}^n g(w_i x),$$

where  $w_i$  are the rows of  $W$ . First, we simplify the expression:

$$f(x) = \frac{1}{n} \sum_{i=1}^n g(w_i x).$$

Given that  $w_i x \sim \mathcal{N}(0, \|x\|^2)$ , we can use the result for a normally distributed variable  $Y \sim \mathcal{N}(0, \sigma^2)$ , which states:

$$\mathbb{E}\left[e^{-\frac{Y^2}{2}}\right] = \left(1 + \frac{\sigma^2}{2}\right)^{-1/2}.$$

Substituting  $\sigma^2 = \|x\|^2$ , we obtain:

$$\mathbb{E}\left[e^{-0.5\|x\|^2}\right] = (1 + \|x\|^2)^{-1/2}.$$

Thus, the expectation for the Gaussian activation is:

$$\mathbb{E}[f(x)] = (1 + \|x\|^2)^{-1/2}.$$

More generally, this can be expressed as:

$$\frac{1}{\sqrt{1 + \frac{a^2\|x\|^2}{2}}}.$$

This derivation provides a geometric initialisation specifically tailored for Gaussian activations, ensuring the network parameters are initialised in a manner that provides effective learning.

**3D Sparse Reconstruction with Gaussian Initialisation.** To evaluate the effectiveness of the new initialisation, we conducted additional experiments using the TNeRF network on the "Basketball" and "Dancer" sequences. The results, summarised in Tab. 5, demonstrate significant improvements compared to the standard initialisation.

Before the geometric initialisation, the TNeRF network with Gaussian activation failed to learn effectively, yielding poor performance with an average SSIM of 86.51 and PSNR of 12.39 for both sequences (see Tab. 4).

After applying the geometric initialisation specifically derived for Gaussian activations, the performance improved substantially. The TNeRF network with the new initialisation achieved an average SSIM of 91.27 and a PSNR of 23.08, indicating a more effective learning process and better reconstruction quality.

Comparatively, the Gaussian initialisation performed better than the Laplacian activation, which achieved an average SSIM of 89.31 and a PSNR of 22.15 (see Tab. 5). The baseline TNeRF without ResFields showed the best results with an SSIM of 94.18 and a PSNR of 26.18. Although the Gaussian initialisation did not outperform the baseline, it significantly closed the gap, indicating its potential for improving non-periodic activations in 3D reconstruction tasks. Visual results of these experiments can be seen in Fig. 3.

Network	$t \downarrow$	Mean		Basketball		Dancer	
		SSIM↑	PSNR↑	SSIM↑	PSNR↑	SSIM↑	PSNR↑
TNeRF	12h	94.18	26.18	94.57	26.33	93.53	25.09
TNeRF Laplacian	17h	89.31	22.15	90.01	22.38	88.60	21.91
TNeRF Gaussian Init.	19h	91.27	23.08	91.15	23.01	91.39	23.16

Table 5. Results for TNeRF networks with different activation functions after applying geometric Gaussian initialisation. The table shows the SSIM and PSNR values for the Basketball and Dancer sequences, as well as the mean of both.

## 5. Conclusion

In this work, we have investigated the implicit biases of non-periodic activation functions—Gaussian, Laplacian, and Quadratic—with coordinate-based MLPs for 2D video approximation and sparse 3D reconstruction. Our contributions include:

- Comprehensive empirical evaluation of Gaussian, Laplacian, and Quadratic activations in 2D video approximation, demonstrating their potential for better generalisation and smoother interpolations compared to periodic activations.
- Evaluation of these non-periodic activations in sparse 3D reconstruction tasks, providing insights into their applicability across different signal domains.
- Development of a geometric initialisation scheme for Gaussian activation, which significantly improves its stability and convergence in 3D approximation tasks.

While our results indicate that non-periodic activation functions can offer robust performance and better generalisation in certain scenarios, there are limitations to our approach. The initialisation schemes for Gaussian and Quadratic activations need further refinement to enhance their effectiveness across all tasks. Additionally, the current evaluation is limited to specific datasets and network configurations, which may not generalise to all possible use cases.

Future work could explore the development of tailored initialisation methods for other non-periodic activations and extend the evaluation to a broader range of datasets and network architectures. Furthermore, integrating architectural enhancements such as residual connections with non-periodic activations could provide additional performance improvements. These directions hold promise for advancing the capabilities of implicit neural representations in diverse applications.

## 6. Contributions of team members

Each team member made the following contributions:

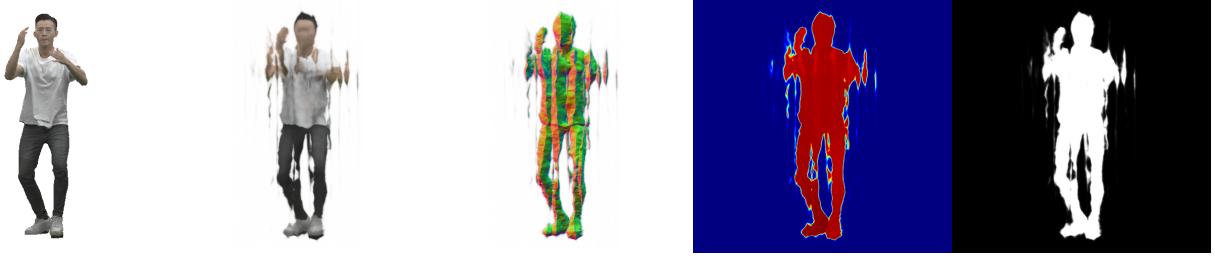
- Bruce Balfour: Rewriting of code base, coding of new activation functions, derivation of geometric initialisation for Gaussian activation, creation of slides.
- Alexandra Trofimova: Creation of slides, writing of mid-term report.
- Ruben Schenk: Running of experiments, writing of final report, writing of mid-term report, creation of slides.

## Acknowledgements

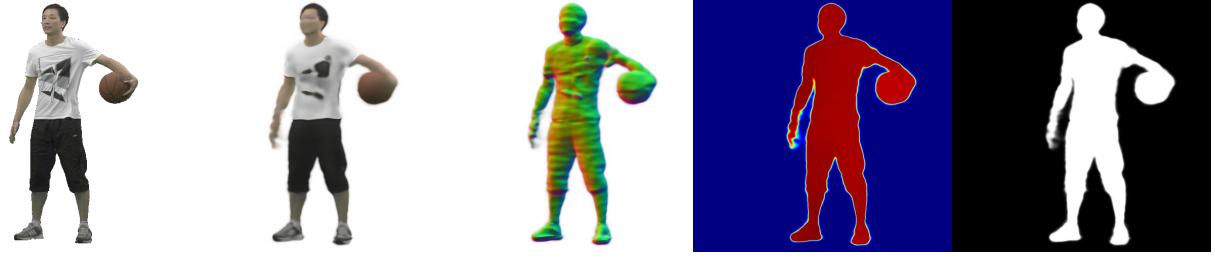
We would like to thank our supervisor, Marko Miha-jlovic, for his continuous support and guidance throughout this project. His explanations and ideas helped us solve many problems we encountered. We also thank Prof. Siyu Tang for giving us the opportunity to work on these interesting topics through her course. Her support and the learning environment she provided were a big part in our project’s success.



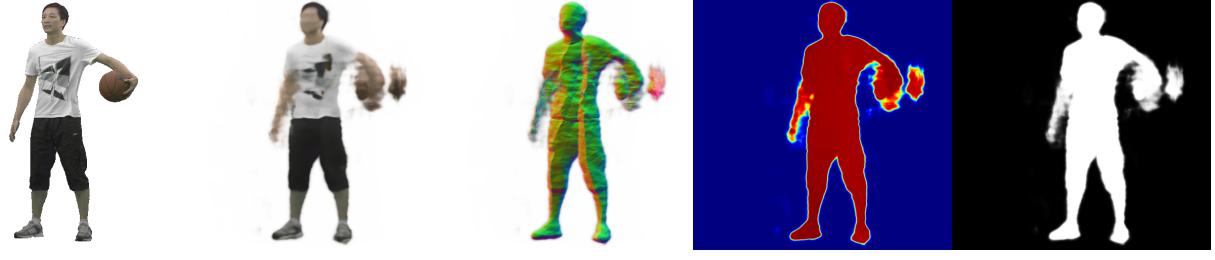
(a) Gaussian activation with geometric initialisation on the "Dancer" sequence



(b) Laplacian activation on the "Dancer" sequence



(c) Gaussian activation with geometric initialisation on the "Basketball" sequence



(d) Laplacian activation on the "Basketball" sequence

Figure 3. Frames from the 3D sparse reconstruction tasks using the TNeRF network without ResFields. The Gaussian activations were trained with our derived geometric initialisation.

## References

- [1] Yang Chen, Ruituo Wu, Yipeng Liu, and Ce Zhu. Hoin: High-order implicit neural representations, 2024. [2](#)
- [2] Tianye Li, Mira Slavcheva, Michael Zollhoefer, Simon Green, Christoph Lassner, Changil Kim, Tanner Schmidt, Steven Lovegrove, Michael Goesele, Richard Newcombe, and Zhaoyang Lv. Neural 3d video synthesis from multi-view video, 2022. [3](#)
- [3] Marko Mihajlovic, Sergey Prokudin, Marc Pollefeys, and Siyu Tang. Resfields: Residual neural fields for spatiotemporal signals, 2024. [1](#), [2](#), [3](#)
- [4] Sameera Ramasinghe and Simon Lucey. Beyond periodicity: Towards a unifying framework for activations in coordinate-mlps, 2022. [1](#), [2](#)
- [5] Danzel Serrano, Jakub Szymkowiak, and Przemyslaw Musialski. Hosc: A periodic activation function for preserving sharp features in implicit neural representations, 2024. [2](#)
- [6] Vincent Sitzmann, Julien N. P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions, 2020. [1](#), [2](#), [3](#)
- [7] Yi Xu, Yao Lu, and Ziyu Wen. Owlii dynamic human mesh sequence dataset. In *ISO/IEC JTC1/SC29/WG11 m41658, 120th MPEG Meeting*, 2017. [3](#)