PROJECT REPORT : PRESENT AND IMPLEMENT A SCIENTIFIC PAPER

Chupin Clément
University of Clermont-Ferrand

January 15, 2023

1 Introduction :

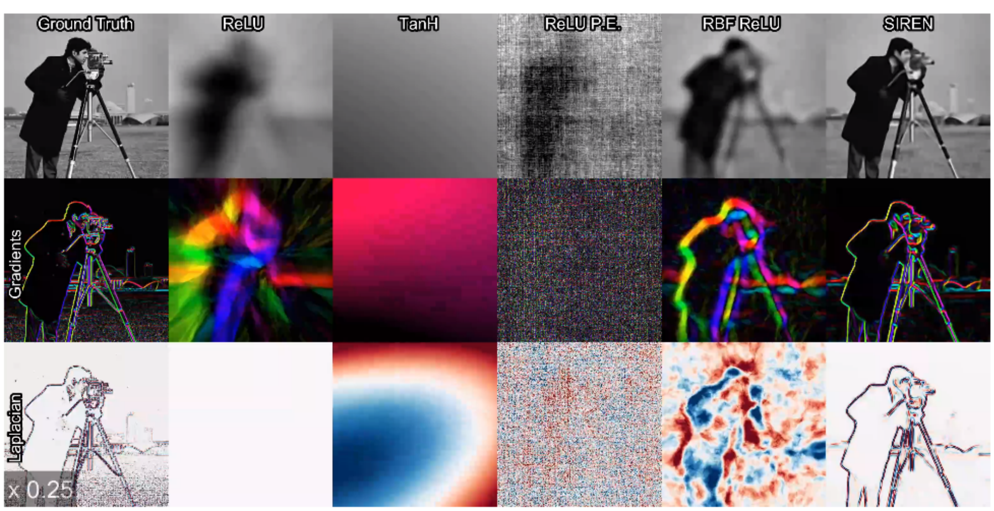
In this project, we will present the paper "Implicit Neural Representations with Periodic Activation Functions" (Sitzmann, Martel et al.,[1]). This paper present a new method to extract input features with a neural network.

This method is design for the signal representation, the neural network task is to predict a specific signal. To evaluate the extraction, the paper try it on data reconstruction (images, audio, 3D shapes ...). This evaluation give the possibility to use the method in different field of research for general and specific task.

Implicit neural representations with periodic activation functions refer to a type of neural network architecture that uses periodic activation functions, such as sine or cosine, as activation functions for the neurons in the network. These activation functions are periodic, meaning that they repeat over a fixed interval, and have the property that the derivative of the function is zero at certain points, which can be beneficial for certain types of problems.

The use of periodic activation functions can help to improve the generalization of the network, as well as reduce the number of parameters that need to be learned. In addition, these networks can be used to learn implicit representations of the data, which can be useful for tasks such as clustering or dimensional reduction.

Finally, implicit neural representations with periodic activation functions can be a powerful tool for modeling and analyzing data with periodic patterns or structures, and have many applications in areas such as signal processing, physics, mathematics and machine learning.



Presentation of different features extraction applied to image reconstruction, SIREN is the method presented by the paper [1]

2 Paper Presentation :

2.1 Paper Motivation

The motivation was to present a new method that leads to a better dimensional understanding by the neural network. In the past, the classical activation layers led to an area-based understanding, where each neuron could represent a different area of its input. With the SIREN method, each new neuron improves the dimensional understanding of the neural network.

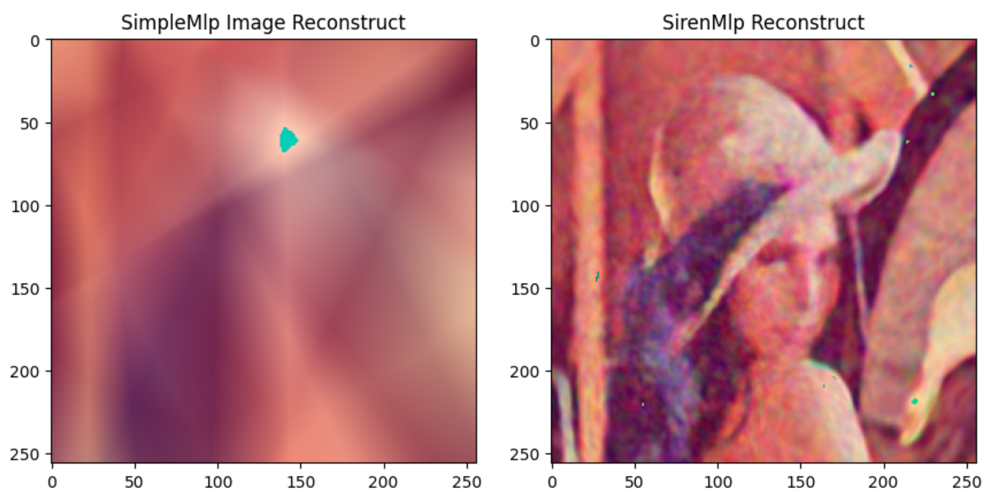
The fact that neural networks are essential to many physical signals implicitly defined as the solution of partial differential equations. The problem is that current network architectures for implicit neural representations are unable to model the signals in detail.

As they show in this paper, a surprising variety of problems in scientific fields fall into this form, such as modeling many types of signals. such as modeling many types of discrete signals in image, video and audio processing using a continuous and differentiable representation, learning 3D shape representations via distance functions signed distance functions and, more generally, the solution of physics's equations, such as Poisson's equations, Helmholtz, or wave equations.

They propose to exploit periodic activation functions for implicit neural representations and show that these networks, are good for representing complex natural signals and their derivatives. SIREN demonstrate the representation of image, wave, video, sound, and their derivatives.

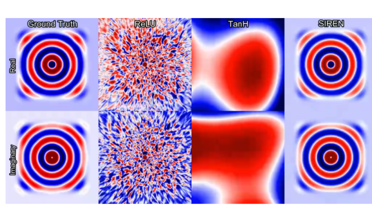
2.2 Paper Result :

Different tasks have been successfully performed, the siren method works on a multitude of tasks. The only limitation is the processing based on the sine activation function, which increases the computation time, compared to the relu activation function for example. This shows that the method can be easily adapted to different problems.

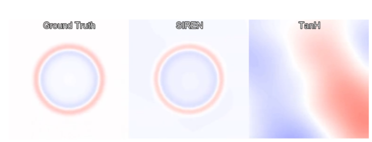


Siren re-implemented on RGB Image Reconstruction

Moreover than the signal representation, the paper bring a method with mathematical basis to extend the signal representation to equation solving with differentials constraints.



Helmholtz equation solving by siren



Wave equation solving by siren

For me, the most impressive contribution of the paper is the mathematical theorizing that allowed the model to succeed to converge. The authors have to deal with a multitude of constraints at the crossroads of mathematics, physics and artificial intelligence to propose a solution, and an analysis of why it's work and how to adjust the fine tuning of the model.

2.3 Paper Conclusion

The paper provides us with a new solution to represent a signal, but, for me, its most important contribution is the theorizing of this technique, and the evaluation of the methods already present in the machine learning field. It provide a solid technical base to apply the SIREN to solve new problems, and improve it or improve other similar methods.

The question of how to represent a signal is covered by several fields of science, Neural implicit representation allows to bring new tools for many fields, whether discrete or continuous, and generalize in several dimensions (sounds, images, 3D objects...).

The technique is also adaptable to solve equations, taking into account differentiable constraints, as well as complex numbers. The paper also brings us several ways to test this technique, that is to say that beyond proposing a new tool, the article also brings us a way to evaluate, in the future, potential new tools, and to compare them to this current method.

The proposed SIREN representation is a powerful tool for deep learning applications involving natural signals, such as images, audio, and video in a deep learning framework. This may be an enabler for downstream tasks involving such signals, such as classification for images or speech-to-text systems for audio. Such applications may be leveraged for both positive and negative ends. SIREN may in the future further enable novel approaches to the generation of such signals.

The periodic aspect of the neural network is useful when it comes to modeling signal problems because many natural signals are periodic and the neural network uses this specificity to have a good understanding and to be able to accomplish the task more precisely than classical neural networks. This can be particularly useful for tasks such as classification or regression, where the goal is to learn a function that can accurately predict the output based on certain input feature, which can be with a high dimensional expression.

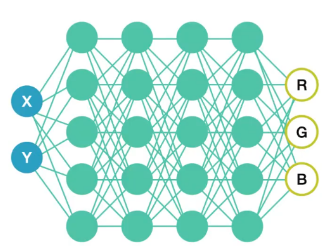
To conclude, in despite of present a new method than is use full, the paper encourages to apply this method to deeper problems, to find new applications for this technique, or to test it for other tasks like reinforcement learning or classification.

3 Implementation :

3.1 The task :

For our implementation, we don't will implement the technical proposed (SIREN is already given). We will focus on the test methods, and the the coherence between the various methods. To go along with the siren function, we will also test a Fourier Features technique ([2]). Finally, we will also study a technique that we propose, which is not created to compete with siren or Fourier features, but which has theoretical benefits in the field of reinforcement learning (triangular input decomposition).

First, we will implement an image reconstruction, it is a good way to have a visual understanding of our neural network along dimensions. The purpose of this implementation is not to feat the image, but to show the advantage and limits of each methods with the use of the minimal neurons and layers possible.



Representation of the neural prediction of the color of a pixel [2]

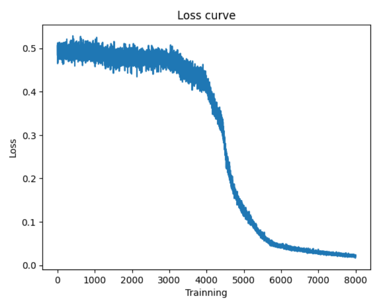
The task is to find a neural network who will, based on the coordinate of a pixel, to find the color of it.

$$F(x, y) => (R, G, B) \tag{1}$$

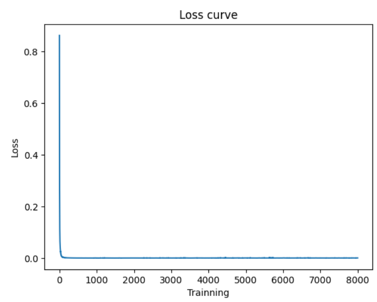
F : The function approximation, in our project, a neural network with a specific feature extraction

3.1.1 Siren :

Siren can be see as a classical MLP, but with Sinus activation function to replace the ReLU activation function, but with a fine tuning bring by the mathematical analysis of gradient to stabilize the training and find a global solution. In my experiment of my internship, it is a critical point, the sinus activation function is difficult to find a solution because, in our training, the sinus activation function will receive a lot of various data, and the gradient will respond with a totally different reaction in despite of the loss to find the good frequency of the sinus function.



Loss of a MLP with Sinus activation function

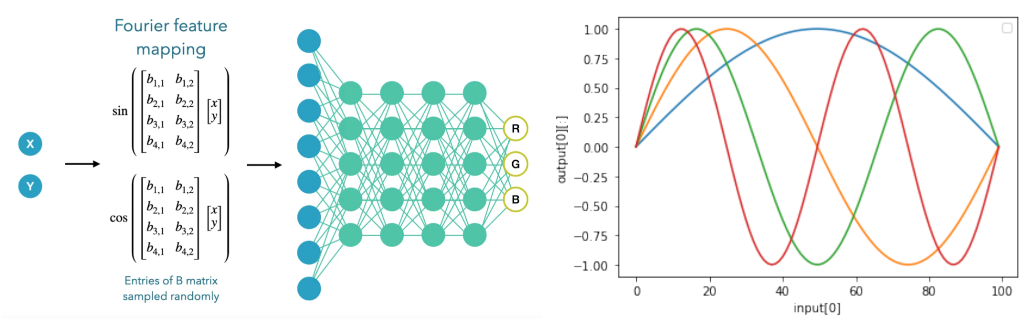


Loss of The SIREN

In these graph, we can see that the simple MLP with Sinus activation function have a lot of difficulties to find a solution, contrary to the siren who learn really fast, on the 2D signal approximation problem. By this experiment, we can clearly see the difficulties to tune a sinus, and the contribution of the SIREN.

3.1.2 Fourier Features extraction :

Fourier Features from (Tancik et al, [2]) :



Representation of the Fourier Features extraction, [2]

The Fourier Features is a transformation of the input to have a better representation of the input for the neural network, the input is convert in a biggest expression space, by multi-frequency representation of each input. We will project the input on different determinate Sinus function to have a better dimensional understanding, and the classical MLP after that will learn in a classical SGD to obtain a solution. The project help to have a better dimensional understanding because a couple of input (a*x1 + b*y2 ...) will be projected with every possibility, the low frequency and the high frequency will work together to give o the MLP the best representation possible.

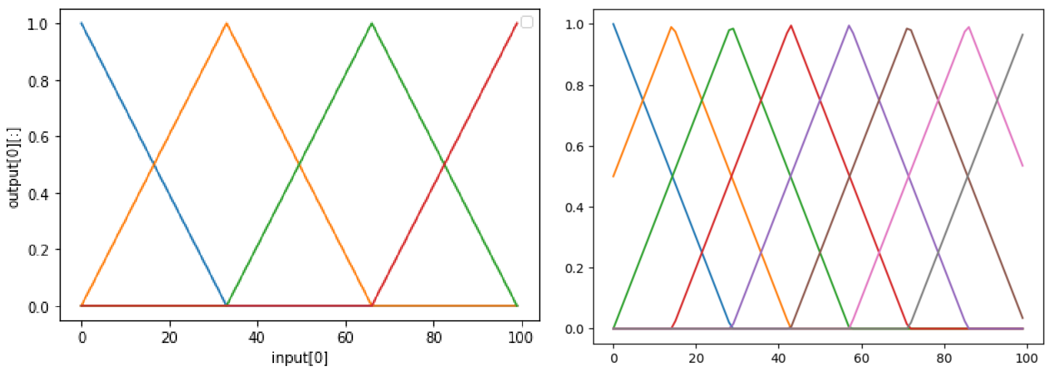
The simplified formula is :

$$FF_i(X) = \omega((X_0 * K_{0,i} + X_1 * K_{1,i} + ... + X_n * K_{n,i}) * 2\pi) \tag{2}$$

3.1.3 The linear input decomposition :

This method will project each input on different activation function who focus on an specific area. This method shown promising improvement in the RL, but don't improve the dimensional encoding.

We present our method, a linear input pre-processing, his purpose is to increase the areas understanding by a pre-layer that can be view as a ReLU pre-trained to focus on each parts of the signal on a independent way.



Input linear feature extraction, by triangular kernel

At the left, we have the general input decomposition, at the right, we have de decomposition used in our neural network, we can increase the variance of the triangular function to have a better ares mixing.

By its design, we can predict that it will be have better result than the simple MLP, but will be worst than the SIREN or the Fourier Features.

3.1.4 Difference between Fourier Features and Siren :

To clarify the difference on this 2 methods :

- Both use sinusoidal to improve their dimensional understanding but there are a lot of little difference.

- Fourier Features use sinusoidal on only its first layer, Siren use sinusoidal on each layer.

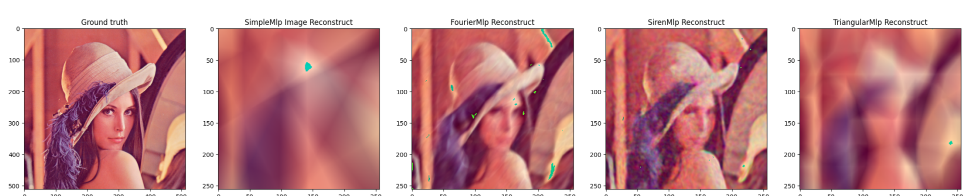
- Fourier Features is for a classical task with a better understanding of the input, siren is oriented to signals processing, with mathematical advanced specifications (differential, high frequency).

- The main difference is the way for the model to learn his task, for the Fourier features, the sinus activation function don't learn, their weights are fixed, for the Siren, the sinus activation function are learned, their weight are find by a training.

To conclude, in our implementation, the Fourier Features may have better result on the image reconstruction, but we can't conclude that it is better than the siren, twice focus on different goals.

3.1.5 Image reconstruction :

In this experiment, we will find the color of a pixel based on his coordinate :



Result of the different methods to

To be able to compare the different models, we will focus on the final loss to reconstruct the image, and the visual aspect of the reconstruction. It's our metrics, the loss is a good value in regression problems, to tune our model and evaluate the method, often, we can adjust the loss function (MSE, MAE or a custom loss function) and analyse the loss curve on the training over the epochs.

Table 1: Methods

Model	Layers	Loss
SimpleMLP	[FC(2,64),FC(64,64),FC(64,3)]	0.018
FourierMLP	[FourierExtract(2,81),FC(81,64),FC(64,3)]	0.004
SirenMLP	[SineLayer(2,64),SineLayer(64,64),SineLayer(64,3)]	0.005
TriangularMLP	[TriangularExtract(2,16),FC(16,64),FC(64,3)]	0.011

Architecture and result of our different methods, on image reconstruction

We can see we have succeed to re-implement this experiment, and we have result coherent with our paper :

- The SimpleMLP have difficulties to predict a pixel, it have a bad dimensional understanding, it can be explained by the simplicity of his architecture, and the ReLU activation function who is not able to correctly predict a signal.

- The Siren succeed to have a good dimensional understanding

- The Fourier feature have very close result to the SIREN

- And finally, the interest of this project, the triangular extraction have better result than the SimpleMLP, but don't have a good dimensional understanding

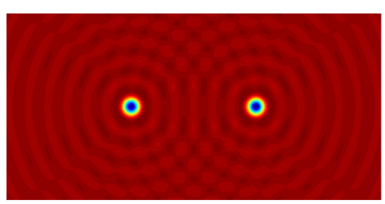
By this simple experiment, we can see the strength of the paper, the SIREN succeed without any difficulties to approximate a 2D signal (here, an image). We also can see the success of the Fourier Features, who is targeted to the input dimensional understanding and not the signal approximation.

3.1.6 Solve Helmholtz equation :

"In mathematics, the eigenvalue problem for the Laplace operator is known as the Helmholtz equation. It corresponds to the linear partial differential equation

$$\nabla^2 f = -k^2 f \tag{3}$$

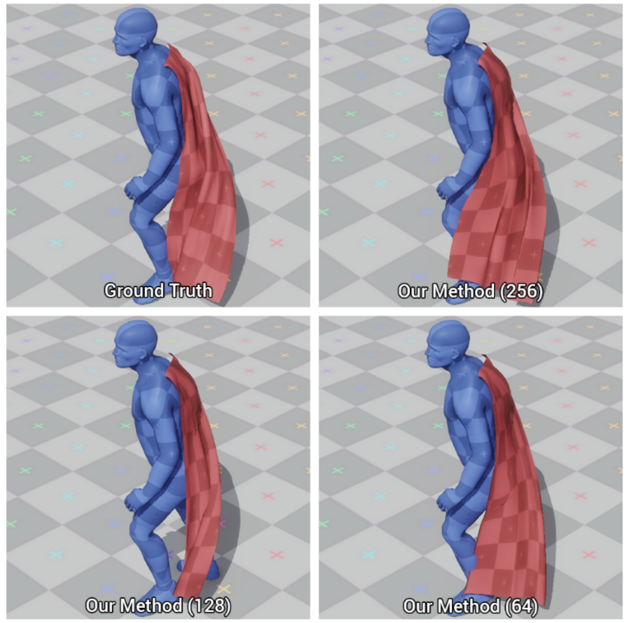
where ∇^2 is the Laplace operator (or "Laplacian"), k^2 is the eigenvalue, and f is the (eigen)function. When the equation is applied to waves, k is known as the wave number. The Helmholtz equation has a variety of applications in physics, including the wave equation and the diffusion equation, and it has uses in other sciences. " [3]



Representation of the Helmholtz equation, applied to physics

As I understand, Helmholtz equation is useful in the physics domain to explain the particle interactions. The mix between physics and machine learning bring new tools to simulate physics faster, contrary to classical physical computation algorithm. Predict a specific point by a neural network is faster tanh simulate the entire environment, machine learning might be a solution to help physician to simulate their research.

For example, Ubisoft (games studios) do a lot of search in this field to be able to simulate physical in real-time for their games (like wind, ocean, cloths...) (Holden et all .,[4]).



Ubisoft method comparison to real physics simulation (Holden et all .,[4])

For our re-implementation, I had difficulties to implement this task, obtain the data set in 3 dimensions with an imaginary part is hard to solve and hard to interprete, I have chosen to feat a Sinus on a 2D space, close to the experiment of Helmholtz, Poison and Wave reconstruction present in the article.

3.2 Result :

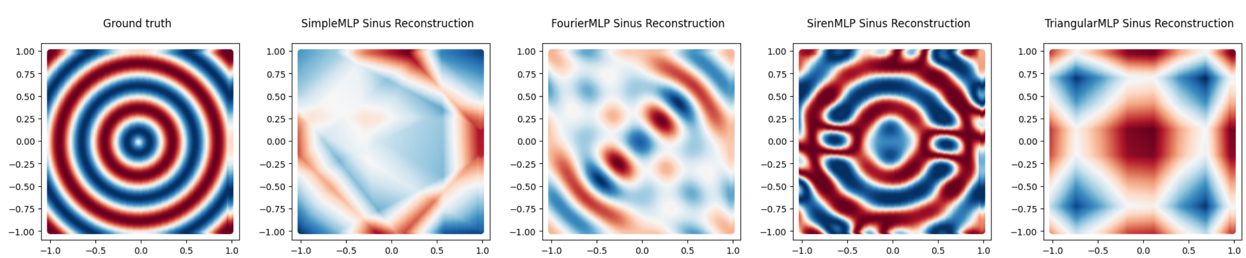
In our first experiment, we will voluntary use small neural network with a bad understanding and prediction of our task. It's to be able to see what each methods bring to the dimensional understanding.

3.2.1 Minimal model

Model	Layers	Loss
SimpleMLP	[FC(2,32),FC(32,1)]	0.39
FourierMLP	[FourierExtract(2,25),FC(25,1)]	0.26
SirenMLP	[SineLayer(2,16),SineLayer(16,1)]	0.10
TriangularMLP	[TriangularExtract(2,16),FC(16,1)]	0.49

Architecture and result of our different methods, on 2D Sinus reconstruction

To be able to compare the different model, we will focus on the final loss to reconstruct the 2D signal, and the visual aspect of the reconstruction.



Result of the different methods on 2D Sinus reconstruction

We can see the difference between each methods, but we can't find the best methods or the limit of each of them

- The MLP try to do a global catching, but don't have enough of areas to do it correctly

- The Fourier features and the SIREN succeed to find a solution across the dimensions, but aren't smart enough to correctly every coordinate.

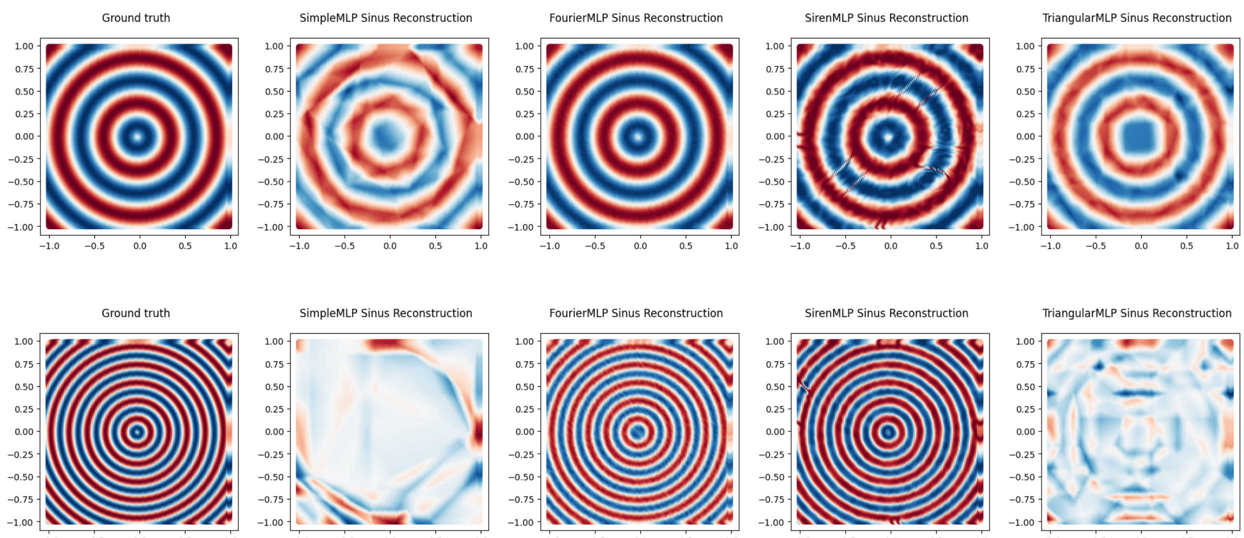
- Finally, we can see the areas decomposition of the triangular decomposition, each parts is process independently, but it's empiric, if the data have more details, we have to proportional adapt our triangular feature extraction, it don't improve the dimensional understanding, it improve the areas understanding.

3.2.2 Maximal model

The second experiment will use neural networks with more complexity, and try to solve an easy task and a hard task, we will try to find the limits of each methods.

Model	Layers	Loss easy	Loss hard
SimpleMLP	[FC(2,32),FC(32,32),FC(32,1)]	0.02	0.45
FourierMLP	[FourierExtract(2,81),FC(81,32),FC(32,3)]	0.00	0.00
SirenMLP	[SineLayer(2,32),SineLayer(32,32),SineLayer(32,1)]	0.00	0.03
TriangularMLP	[TriangularExtract(2,16),FC(16,32),FC(32,1)]	0.01	0.41

Architecture and result of our different methods, on 2D Sinus reconstruction



Result of the different methods on 2D Sinus reconstruction

Similar to the image reconstruction, we have result coherent with our paper :

- The SimpleMLP have difficulties to predict a pixel, it have a bad dimensional understanding, it can be explained by the simplicity of his architecture, and the ReLU activation function who is not able to correctly predict a signal.

- The Siren succeed to have a good dimensional understanding

- The Fourier feature, a methods close to the Siren have very close result

- And finally, the interest of this project, the triangular extraction have better result than the SimpleMLP, but don't have a good dimensional understanding

3.3 Implementation conclusion :

The paper proposed an implementation of the SIREN, and different ways to test and compare this method. In our implementation, we succeeded in implementing some of these tests, but due to the complexity of the problems sometimes treated in the paper, we were not able to treat correctly the Helmholtz equation. By the place of this paper between signal processing, physics and artificial intelligence, it has to propose a large panel of experimentation, with several different techniques, which made our implementation project relevant. We have to keep in mind that some search field close to the machine learning require specific skills.

Indeed, we have tested our triangle decomposition method, which allowed us to highlight its qualities and defaults. This experimentation show than it is worst than a frequency input encoding in the dimensional understanding. For signal representation or improvement of dimensional understanding, it is ineficient, but in the field of reinforcement learning, by slice its input into different areas, it allows a more specific treatment to solve a task.

4 General conclusion

In this project, we studied a paper that presented the Siren method, we have re implemented methods and evaluation from it. This allowed us to better understand the approach of the authors by following their work, and going to the same conclusion. By this project, we can have a better idea of the difficulty and the problematic of proposing a new method in the field of the machine learning.

The re implementation is an essential tool in the research world, when we have a new proposal to make, or we try to compete with something else and we take the same experiments. Or, if we have specific constraints in our field, beyond proposing a method, we have to propose an evaluation that is both relevant and generic.

Moreover, the world of the research in general has the rule of the repeatability of experiments, the new field of machine learning must question itself in relation to this and be able to propose relevant and accessible methods for other researchers.

Finally, the implementation of a scientific paper in the field of machine learning was a success. The proposed methods and techniques were successfully applied to a custom dataset and the results were coherent with those reported in the paper. The project also highlighted the importance of reputability in the field of machine learning and the need for clear and detailed documentation of experimental setups and data to be able to publish and share research project.

References

[1] Vincent Sitzmann, Julien N.P. Martel, Alexander W. Bergman, David B. Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. In *Proc. NeurIPS*, 2020.

[2] Matthew Tancik, Pratul P. Srinivasan, Ben Mildenhall, Sara Fridovich-Keil, Nithin Raghavan, Utkarsh Singhal, Ravi Ramamoorthi, Jonathan T. Barron, and Ren Ng. Fourier features let networks learn high frequency functions in low dimensional domains. *CoRR*, abs/2006.10739, 2020.

[3] Hermann von Helmholtz. Helmholtz equation : https://en.wikipedia.org/wiki/helmholtz_equation.

[4] Holden et al. Subspace neural physics: Fast data-driven interactive simulation. 2019.