Efficient Deep Learning for 3D and 4D Shape Analysis

From Medical Applications to Generative AI

Hamid Laga

Murdoch University, Australia

H.Laga@murdoch.edu.au

AI and Computer Vision Research Group

Fundamental and strategic research, combining a range of computational and mathematical domains to solve long-standing problems in artificial intelligence, machine learning and computer vision

Artificial Intelligence

- Foundations of AI and ML
- Edge AI and embedded AI
- Human-AI Teaming

Computer Vision & Machine Learning

- 3D and 4D Computer Vision
- Image and video analysis and understanding
- Image and video captioning
- Natural Language Understanding

Cyber Security and AI

- AI for Cyber Security
- Cyber Security in AI

AI and Computer Vision Research Group

Translational research, working closely with various industries to drive innovation in these areas of strategic importance

Artificial Intelligence

Computer Vision & Machine Learning

Cyber Security and AI

Health (AI & Digital Pathology, Aboriginal Health, Training and Rehabilitation)

Digital Agriculture (Plant / soil health, plant yield, food security, automation)

Environment and Sustainability (biosecurity, environment conservation, etc.)

Mining

AI and Computer Vision Research Group

Ethical, Responsible and Explainable AI

Artificial Intelligence

Computer Vision & Machine Learning

Cyber Security and AI

Health (AI & Digital Pathology, Aboriginal Health, Training and Rehabilitation)

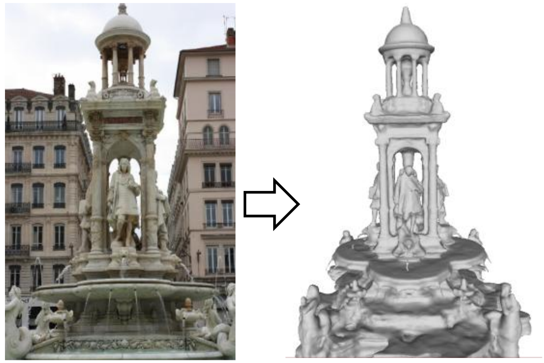
Digital Agriculture (Plant / soil health, plant yield, food security, automation)

Environment and Sustainability (biosecurity, environment conservation, etc.)

Mining

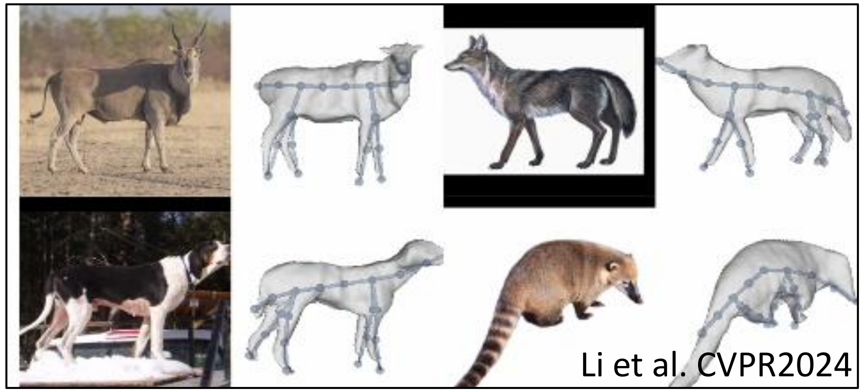
My Current Research Focus

Mathematical models that reproduce the physical world and the underlying processes that control Geometry, Appearance, Behaviors, Interactions, and Growth

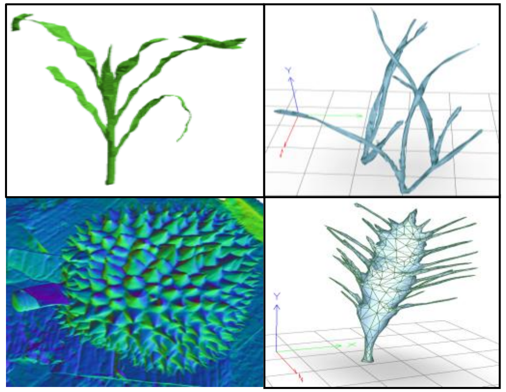


Digital twins

Virtual Humans



Virtual Animals (and insects)



Virtual Plants

Why is it Important? - Applications

Engineering

Health/Medicine

Agriculture

Livestock management

Urban planning

Sports
(Amateur / high performance)

Wildlife preservation

Disaster simulation and management

Animal welfare

Biology and evolution

Cultural heritage preservation

Post injury Rehabilitation

What do We Need?

- An efficient representation of the world
  - How can we mathematically represent the geometry, appearance, motion and dynamics of real-world objects?
- Learning this representation from observations
  - How can we recover this representation from observations of the world (e.g., images and videos captured at different scales and times)?
- Generative modelling

What do We Need?

- An efficient representation of the world
  - How can we mathematically represent the geometry, appearance, motion and dynamics of real-world objects?
- Learning this representation from observations
  - How can we recover this representation from observations o the world (e.g., images and videos captured at different scales and times)?
- Generative modelling

Digital Representation of Signals

- The world is a continuous function,

1D Signals

$f: \mathbb{R} \to \mathbb{R}$

Images (RGB, hyperspectral)

$f: \mathbb{R}^2 \to [0, 1]^3$

3D geometry

$f: \mathbb{R}^3 \to \mathbb{R}$

3D geometry and appearance

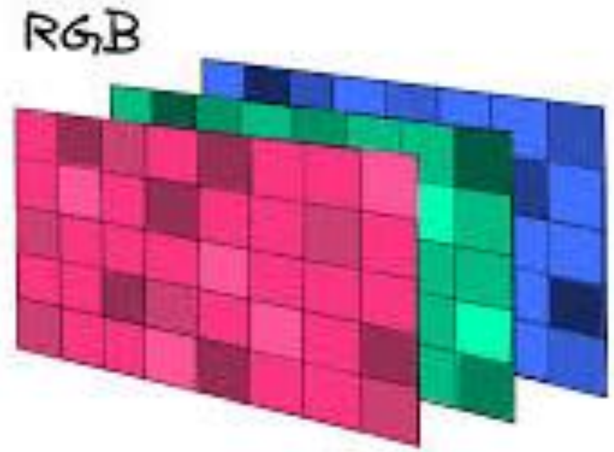
$f: \mathbb{R}^3 \times S^2 \to \mathbb{R} \times [0, 1]^3$

Digital Representation of Signals

- The world is a continuous function,
- But is discretized during the acquisition, storage, and visualization
  - Require high storage and cannot represent every level of detail

Images (RGB, hyperspectral)

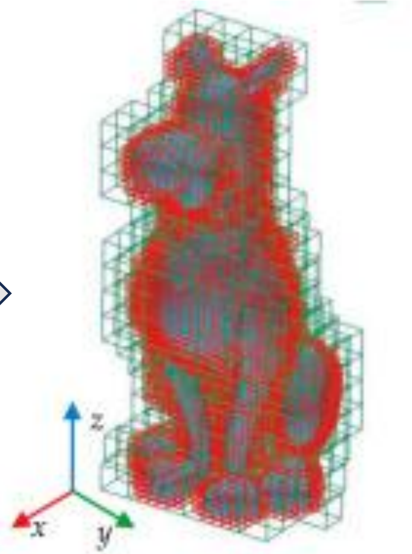
$f\colon \mathbb{R}^3 \rightarrow [0,1]^3$



On computer

3D Geometry

$f\colon \mathbb{R}^3 \rightarrow \mathbb{R}$



The World as a Continuous Function

[OFFICIAL]

The World as a Continuous Function

Point $p \in \mathbb{R}^3$ → **F** →

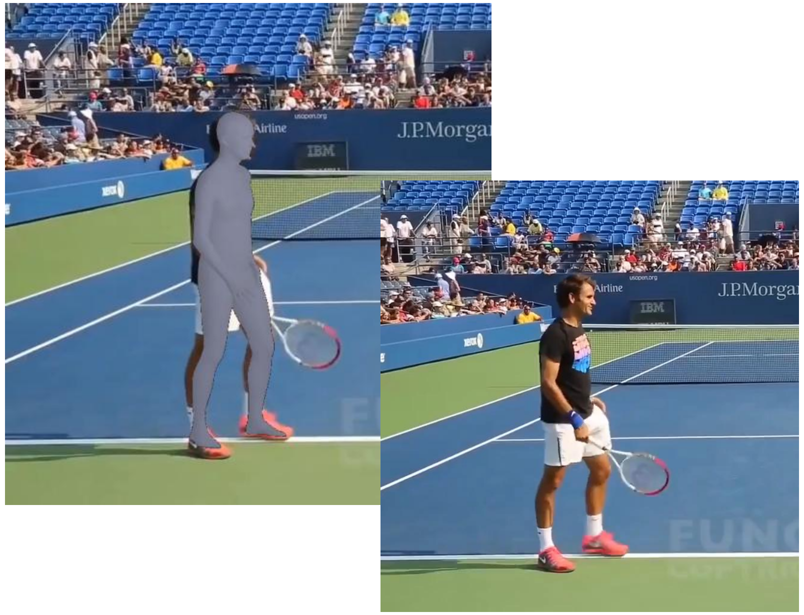
3D Geometry

The World as a Continuous Function

Point $p \in \mathbb{R}^3$ →
Viewing direction →

F

→ 3D Geometry + Color

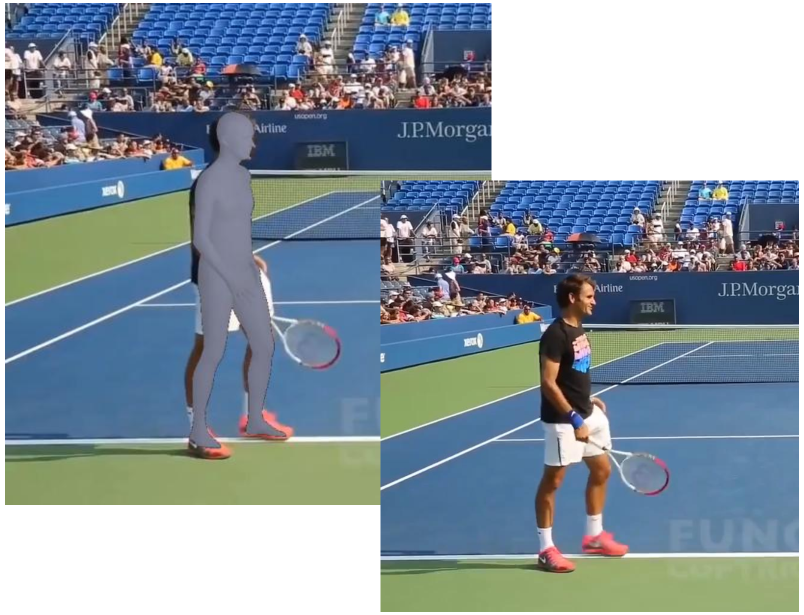


The World as a Continuous Function

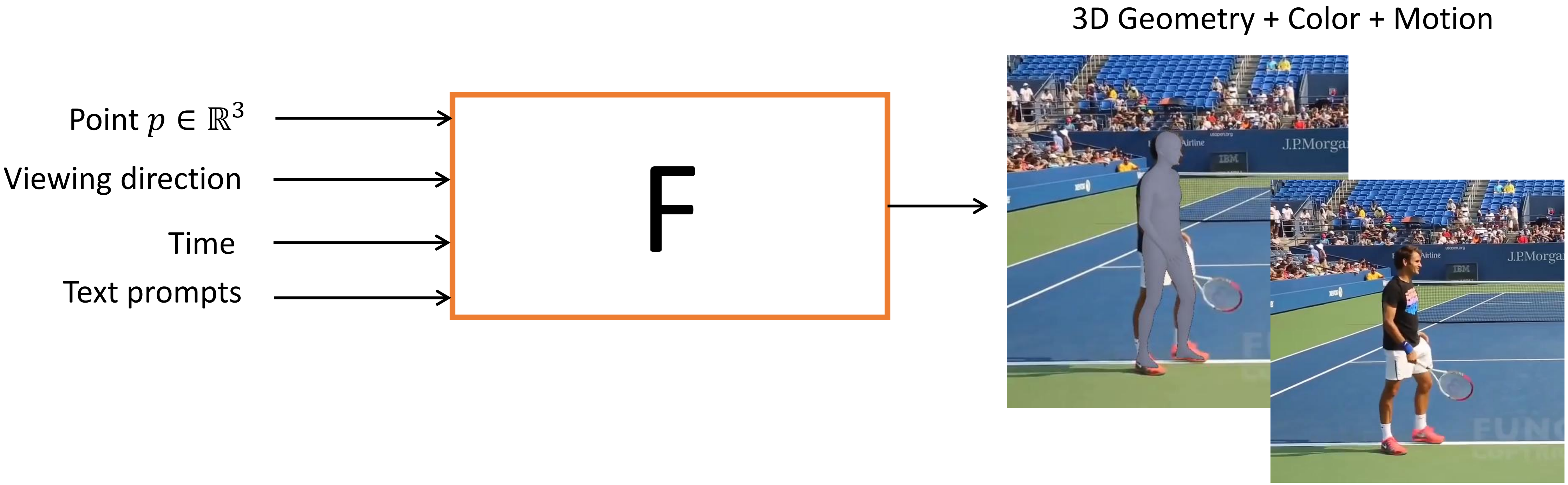
Point $p \in \mathbb{R}^3$ →
Viewing direction →
Time →

F →

3D Geometry + Color + Motion

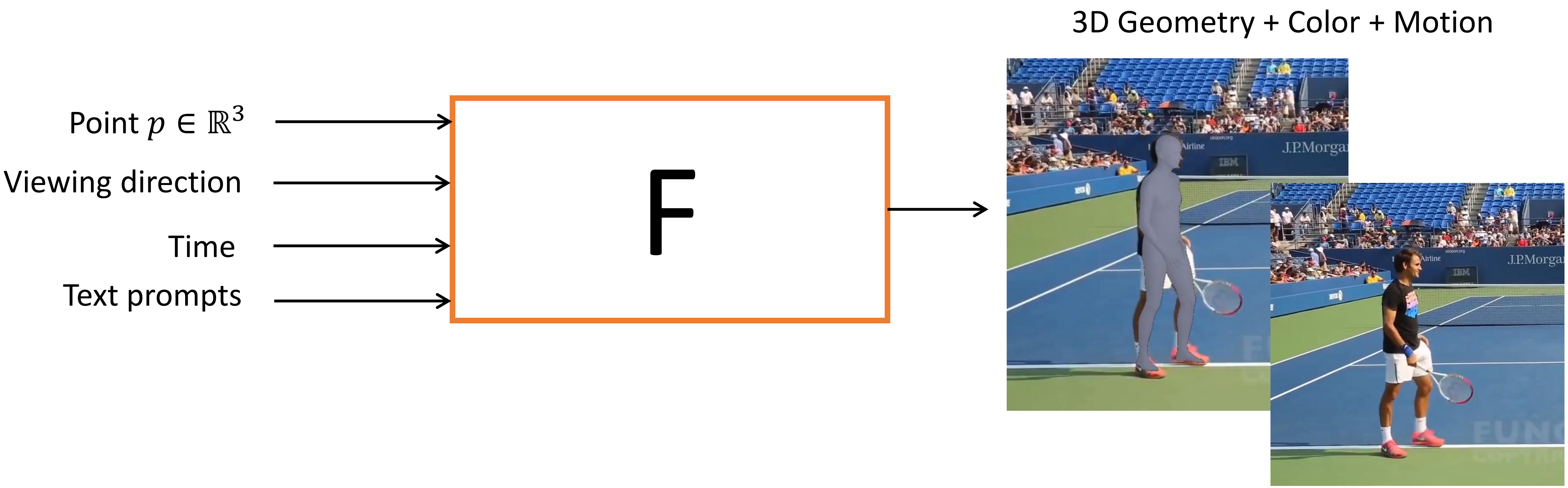


The World as a Continuous Function



The World as a Continuous Function

- F is a complex function - Has no analytical form



The World as a Continuous Function

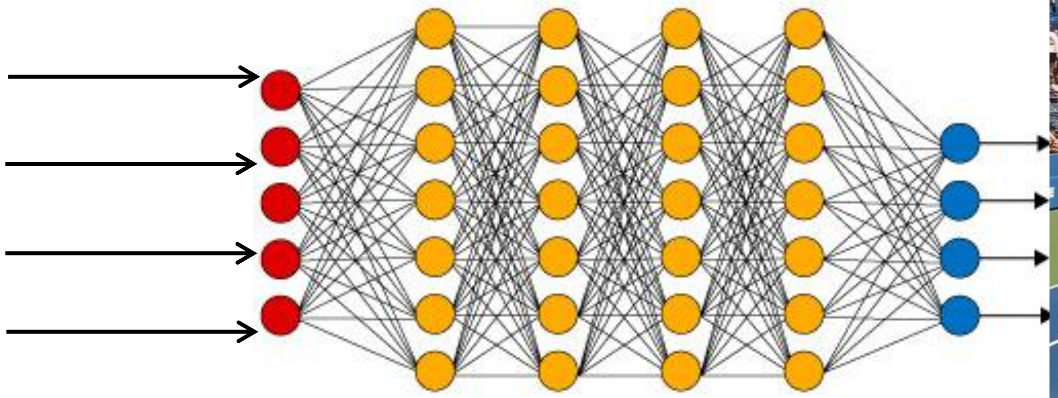
- F is a complex function - Has no analytical form
  - Represent it using Neural Networks, which are universal approximators
  - Learn its parameters from (training) data

Point $p \in \mathbb{R}^3$

Viewing direction

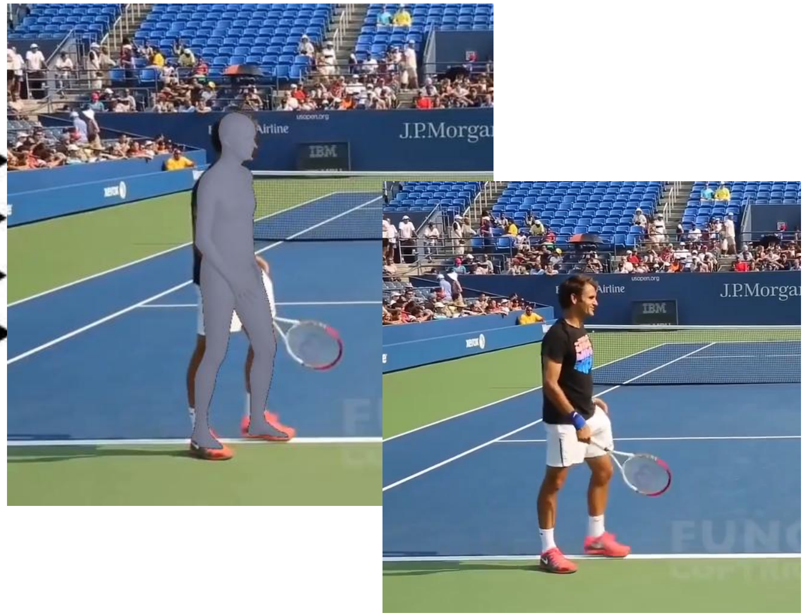
Time

Text prompts



Approximate the function F using Deep Neural Networks

3D Geometry + Color + Motion



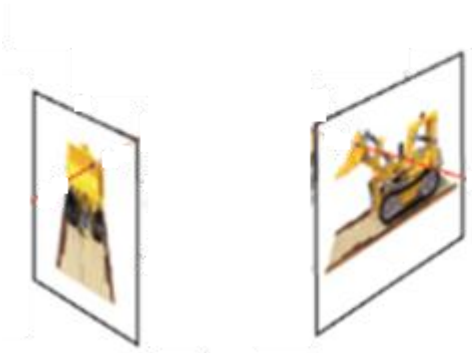
Problem 1 – 3D Reconstruction from Images

- A 3D scene is a function f that takes
  - A point p in the 3D space and returns its geometry value (either density or SDF)
  - A point p and viewing direction v and returns its color as seen from v

$$
\begin{aligned}
f:\ & \mathbb{R}^3 \times S^2 \rightarrow \mathbb{R} \times [0,1]^3 \\
& p, v \rightarrow \mathrm{geometry}(p), \mathrm{color}(p, v)
\end{aligned}
$$

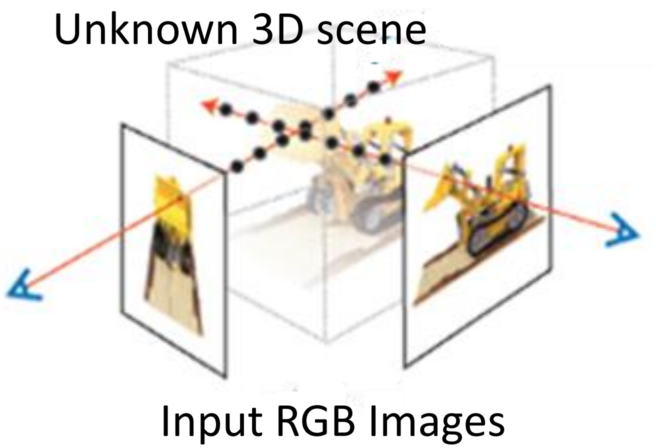
- 3D reconstruction is the problem of learning this function from images

Problem 1 – 3D Reconstruction from Image

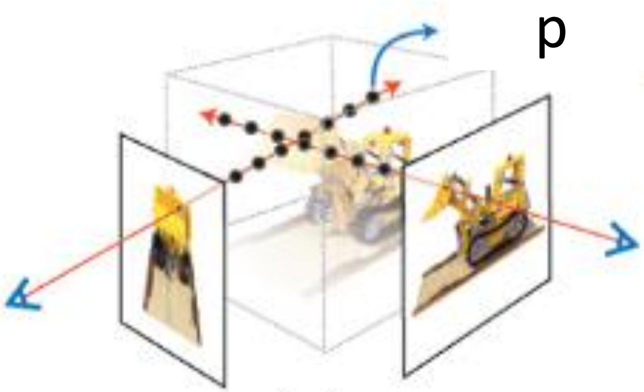


RGB Images

Problem 1 – 3D Reconstruction from Image



Problem 1 – 3D Reconstruction from Image

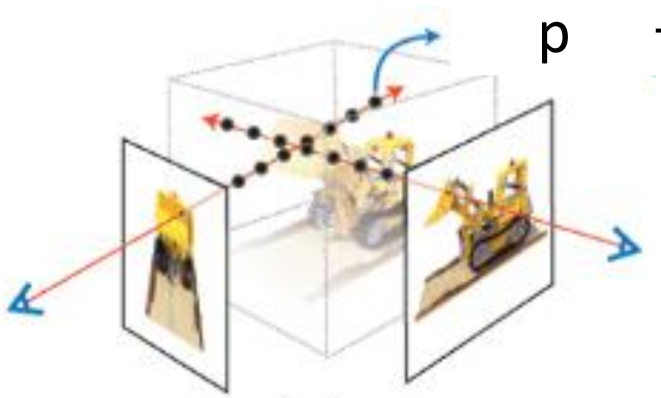


(a)

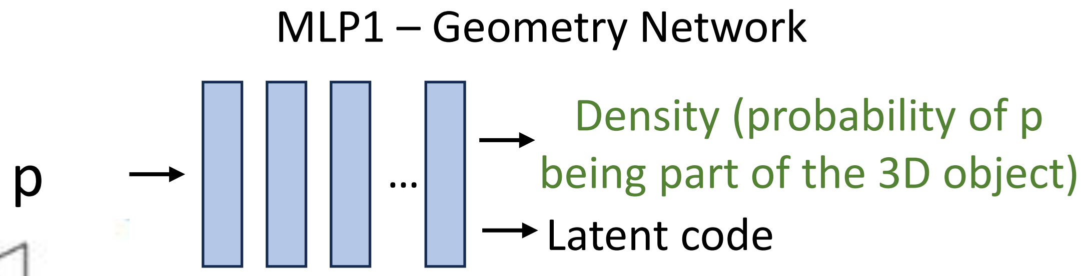
MLP1 – Geometry Network

p → [| | | … |] → Density (probability of p being part of the 3D object)

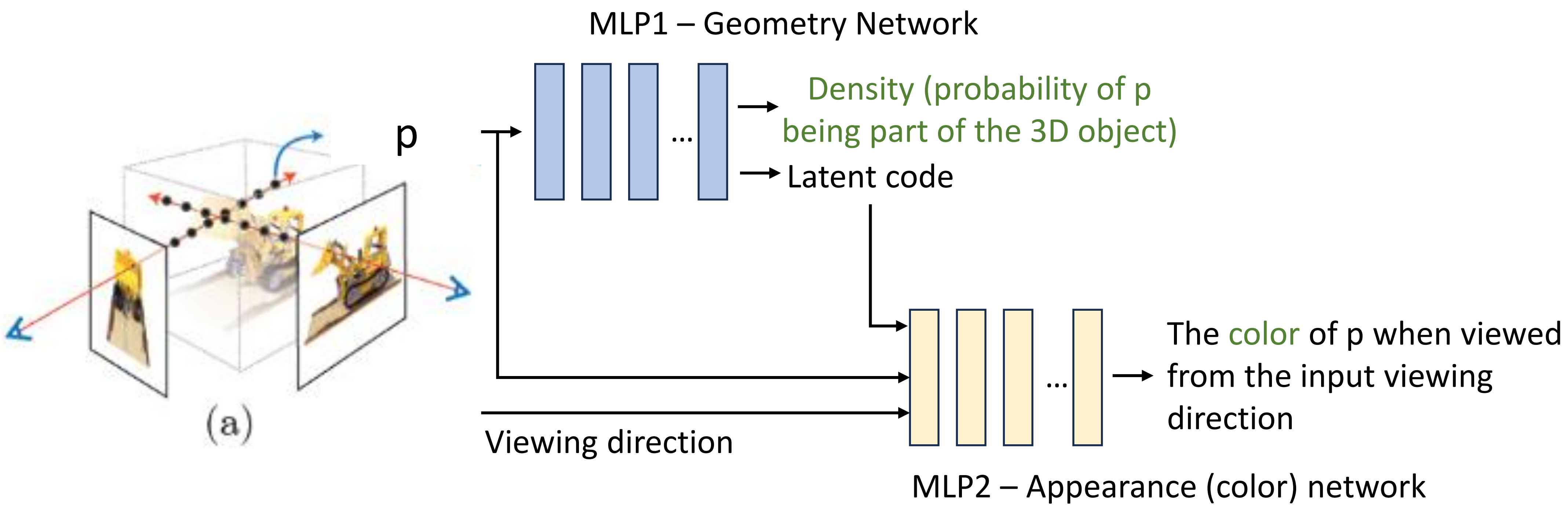
Problem 1 – 3D Reconstruction from Image



(a)

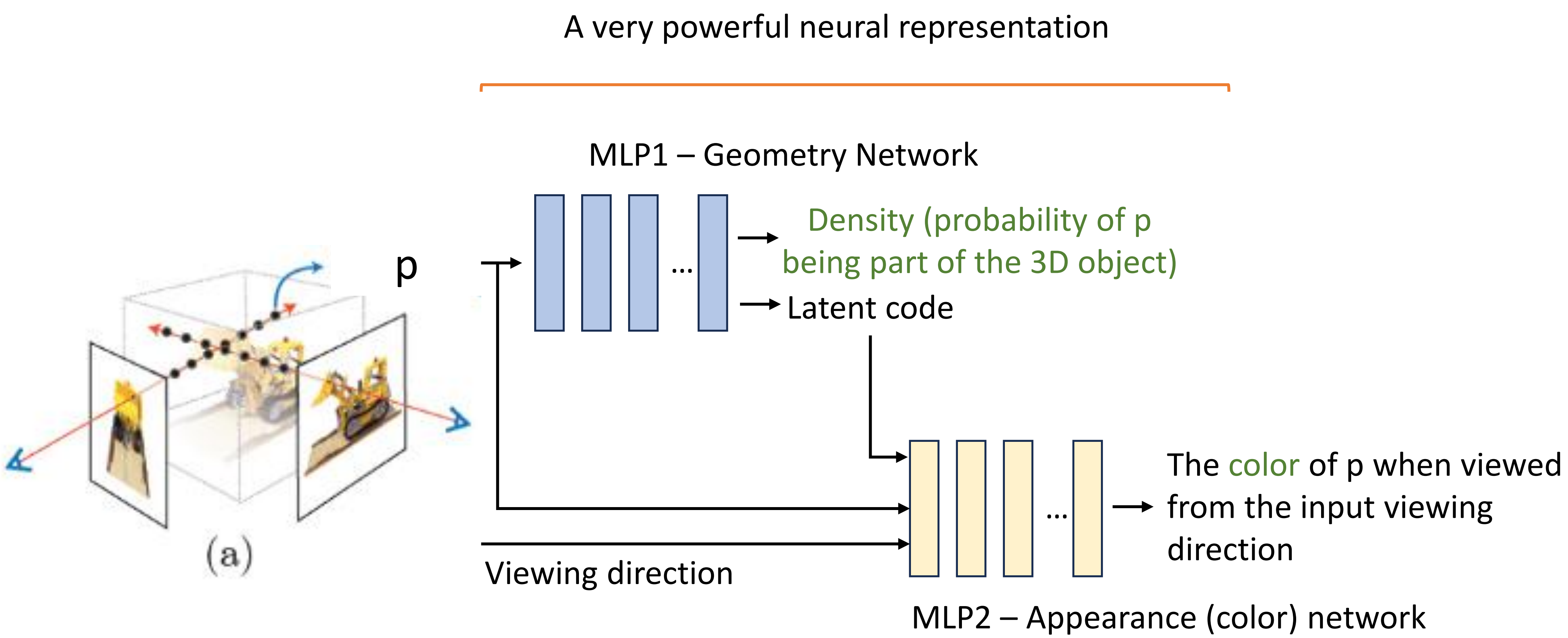


Problem 1 – 3D Reconstruction from Image



Problem 1 – 3D Reconstruction from Image

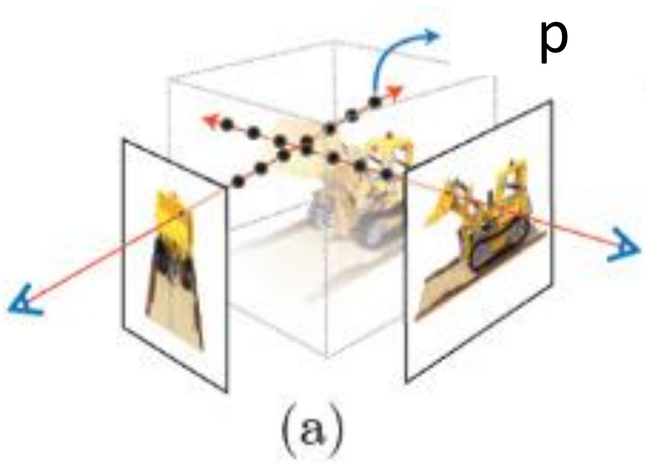
- Neural Radiance Fields (NeRF)

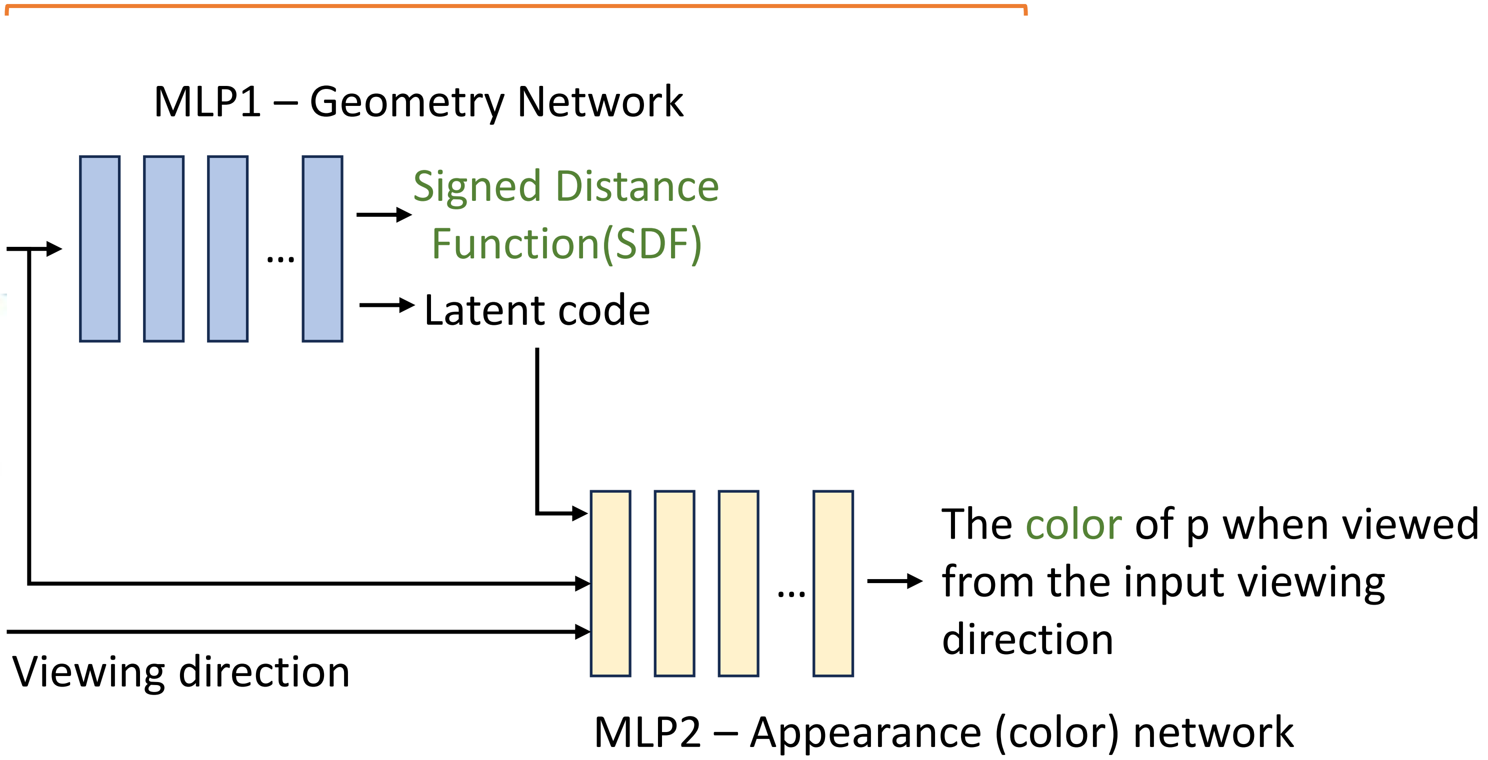


Problem 1 – 3D Reconstruction from Image

- Neural Implicit Surfaces (NeuS)

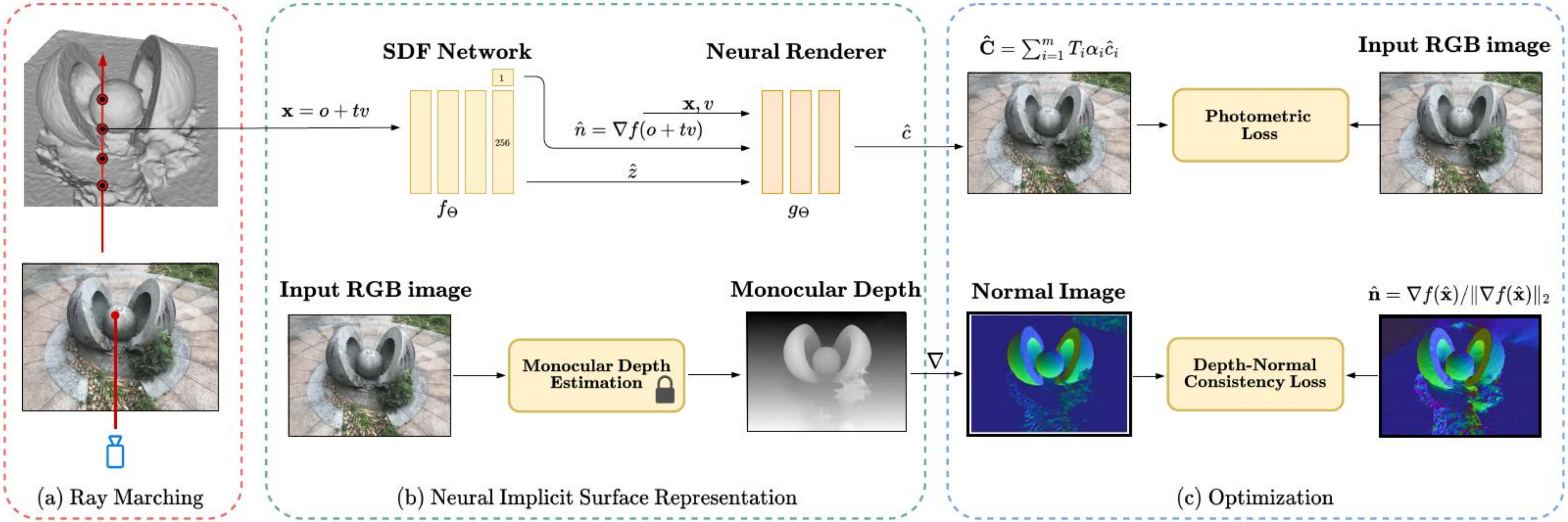
A very powerful neural representation





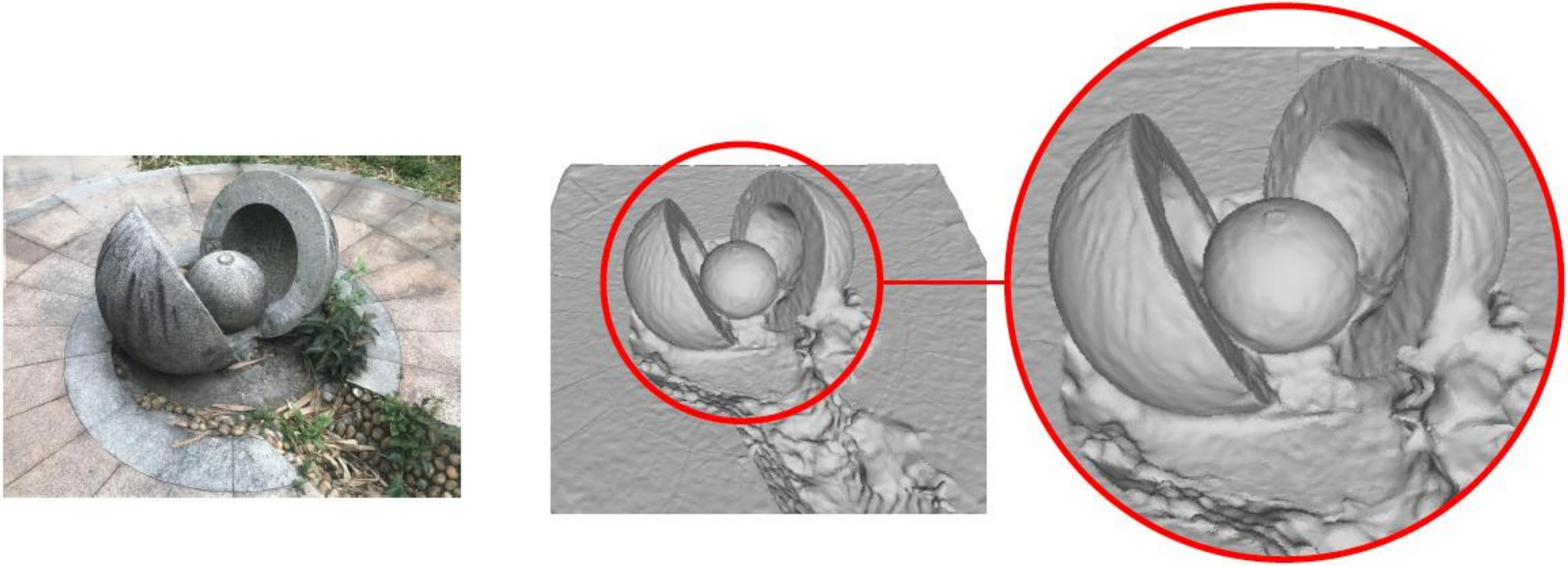
Problem 1 – 3D Reconstruction from Image

- Normal-guided Neural Implicit Representation of 3D shapes
  - Normals is all what you need to train the neural implicit representation
  - Highly detailed 3D reconstruction from as few as 2 RGB images (with minimum overlap)



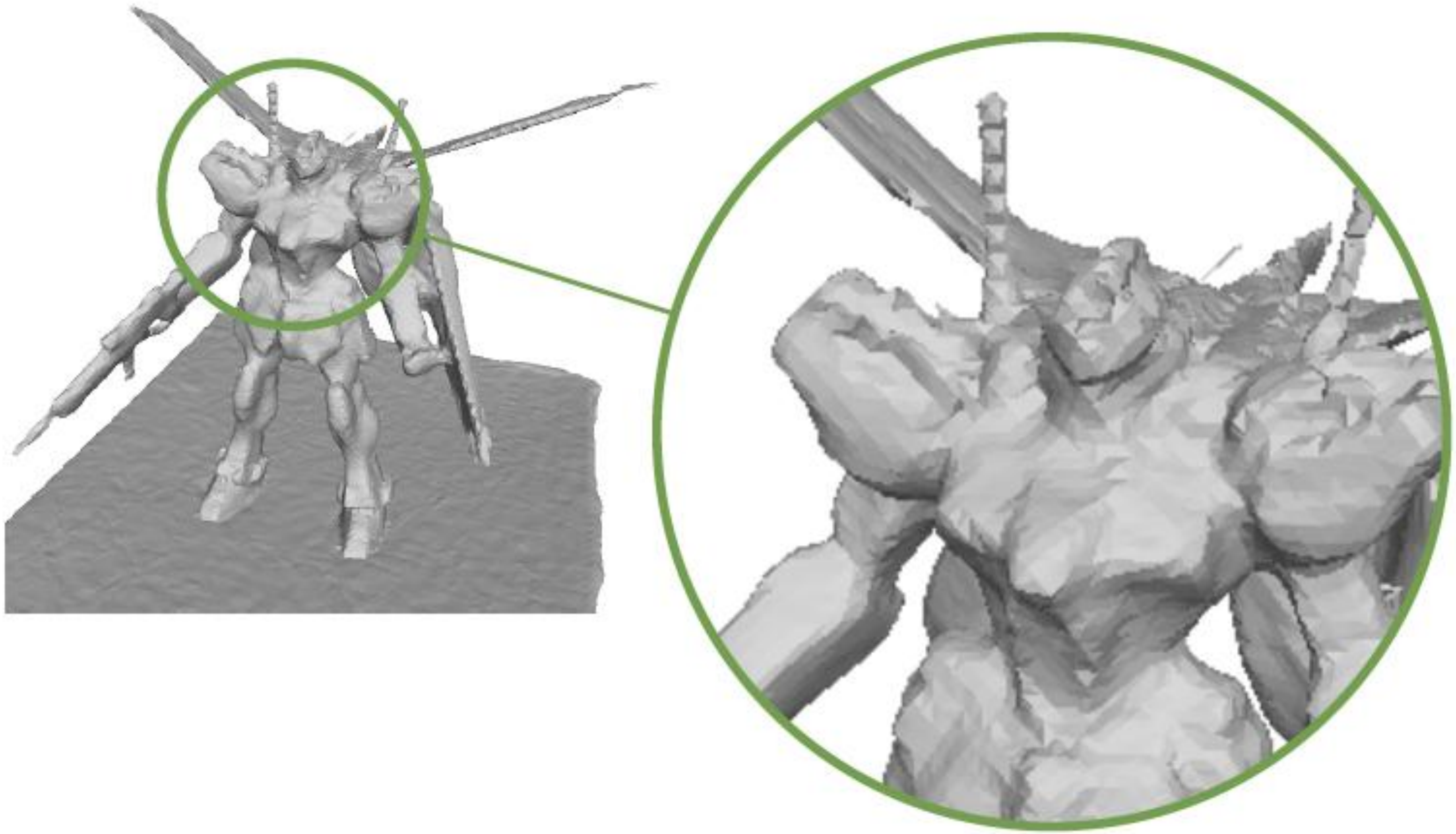
3D Reconstruction Results

Recovering the 3D geometry of real objects from as few as two images



3D Reconstruction Results

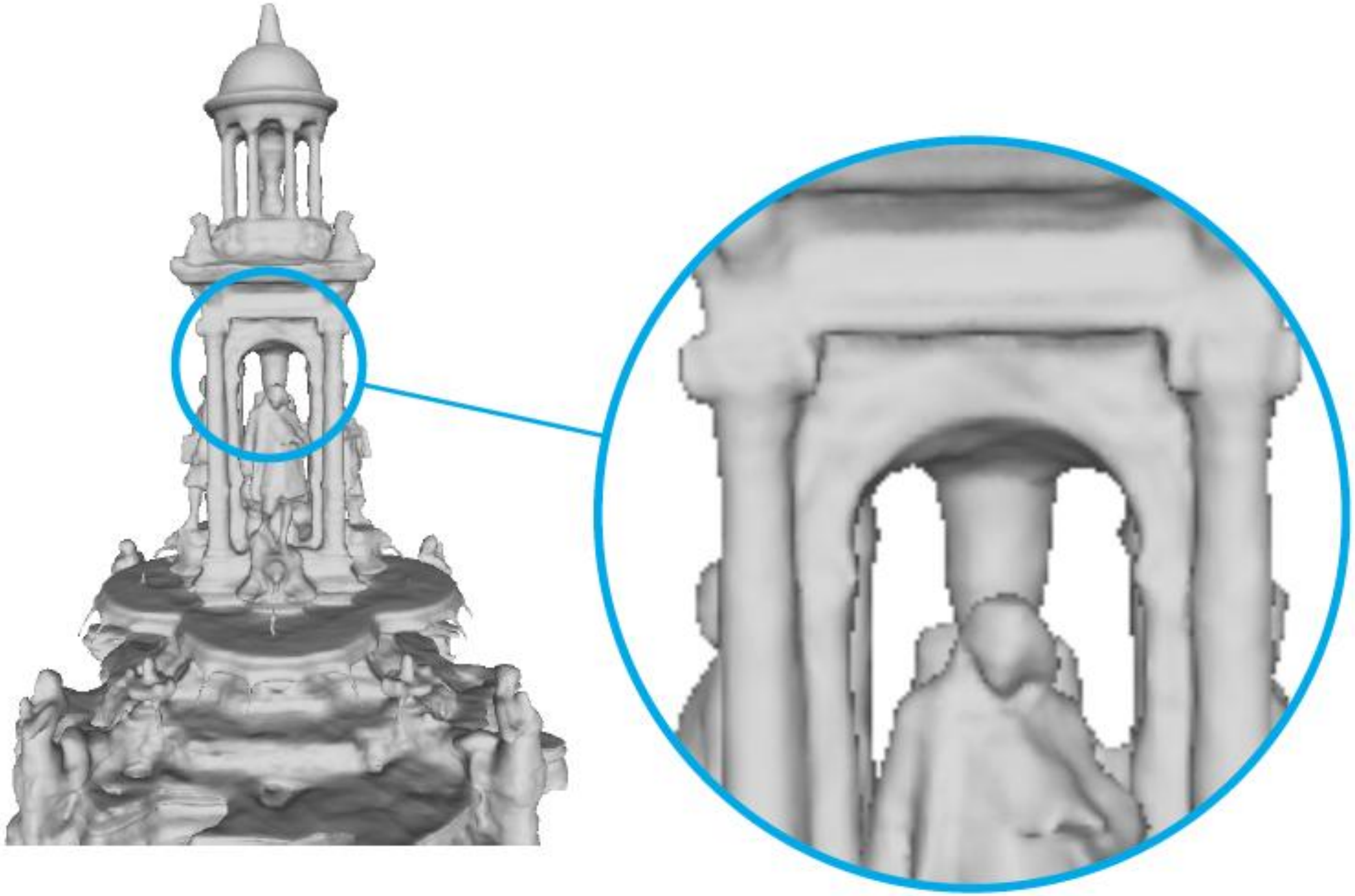
Recovering the 3D geometry of real objects from as few as two images



[C

3D Reconstruction Results

Recovering the 3D geometry of real objects from as few as two images



More Results

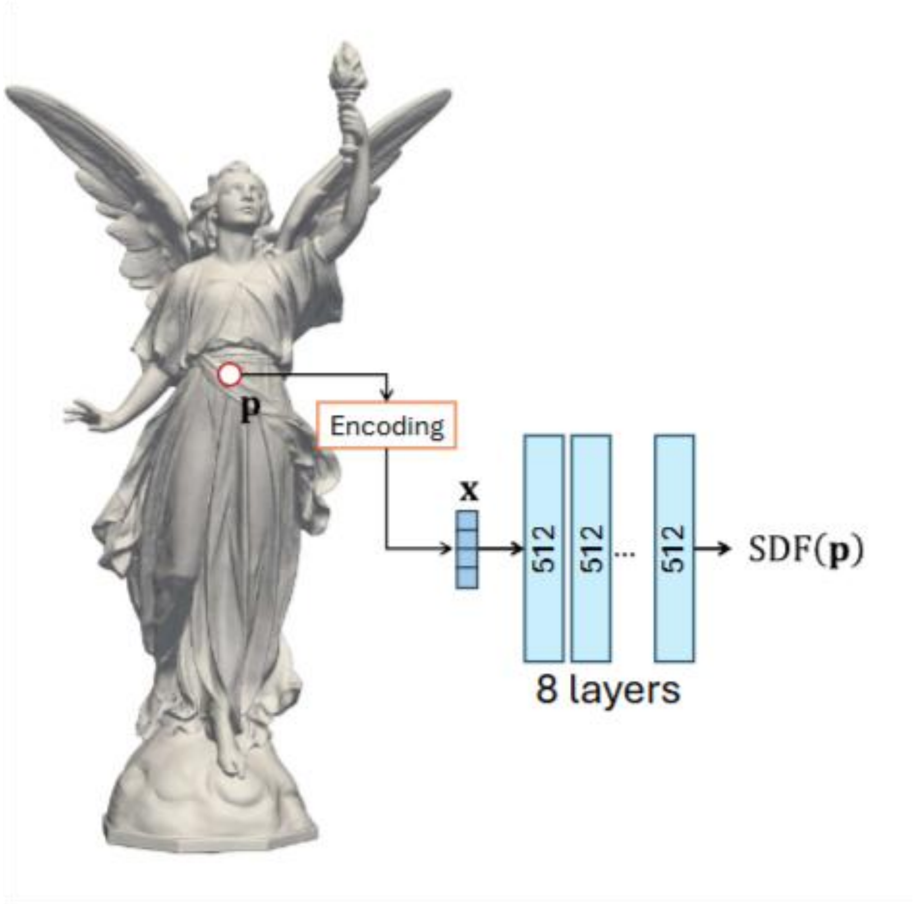
https://sn-nir.github.io/

Issues with MLP-based Representations

- Very slow to train
  - Can take up to 6 hrs per scene
- Memory
  - Need to store parameters of the (deep and wide) neural networks, which can be large
  - Can also be slow at inference time
  - Many acceleration techniques focused on speeding the inference time

Are MLP-based Neural Networks the Right Ones for representing the world?

Are MLP-based Representations the Right Ones?

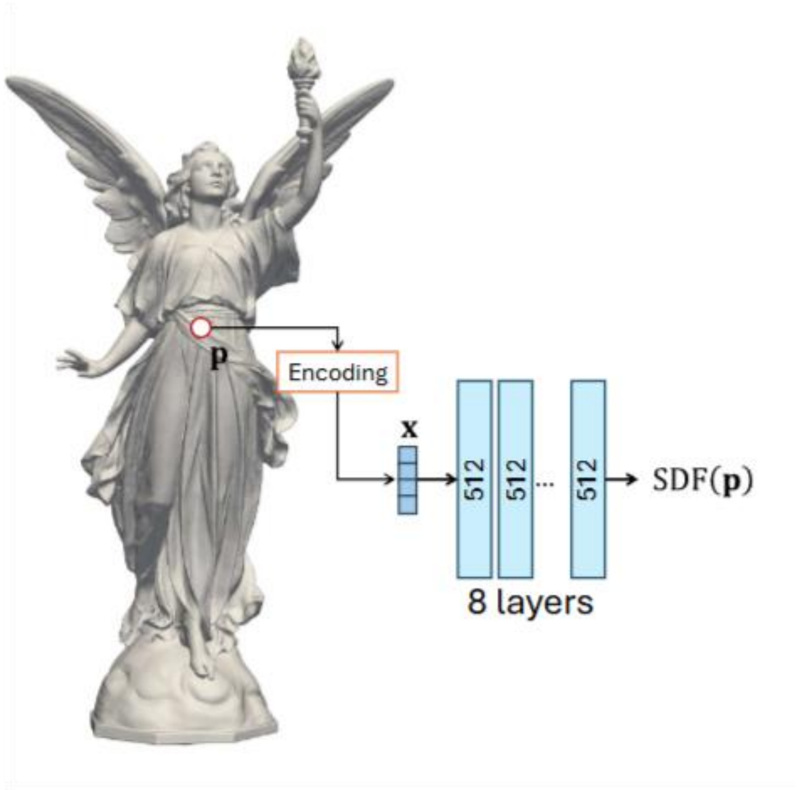


Traditional Deep SDF

- Most neural representations use MLPs with ReLU activation functions
- ReLU is a piecewise linear function
  - It represents complex functions using piecewise linear functions
  - We need a large network to capture complex nonlinearities
- How about if we let the Neurons do a bit more than just a dot product followed by ReLU?

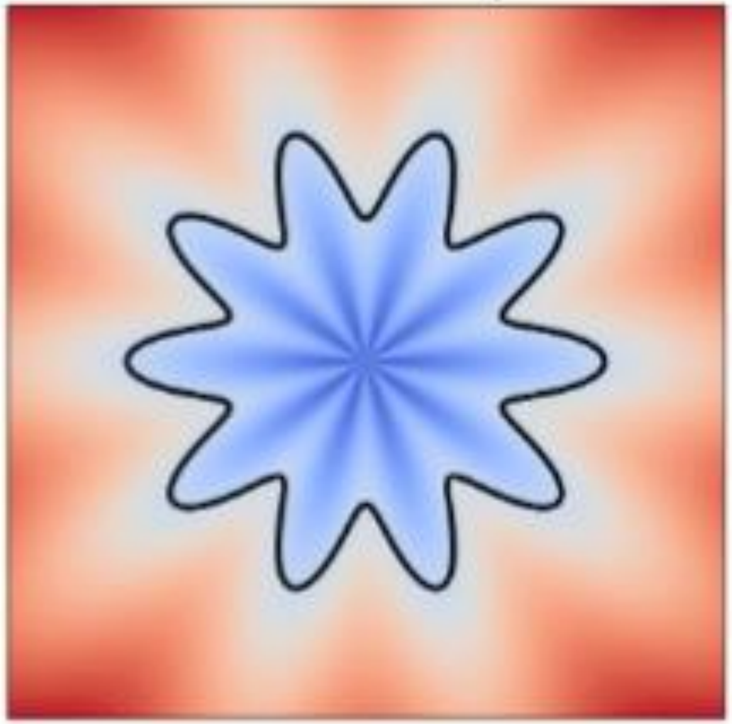
MLPs with Quadratic Layers

- Replace the last layer in the MLP by a layer of neurons with a Quadratic Activation Function (instead of ReLU)



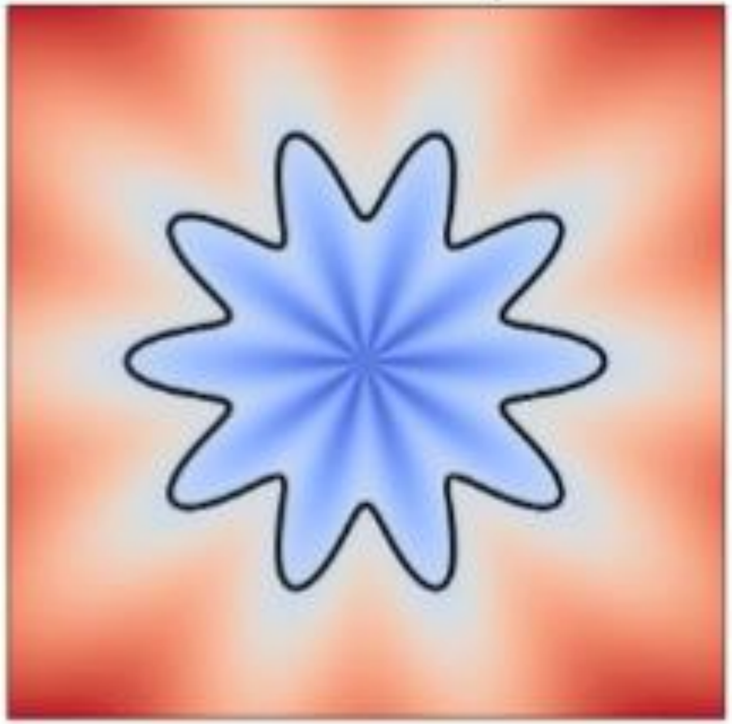
Traditional Deep SDF

Are MLP-based Representations the Right Ones?

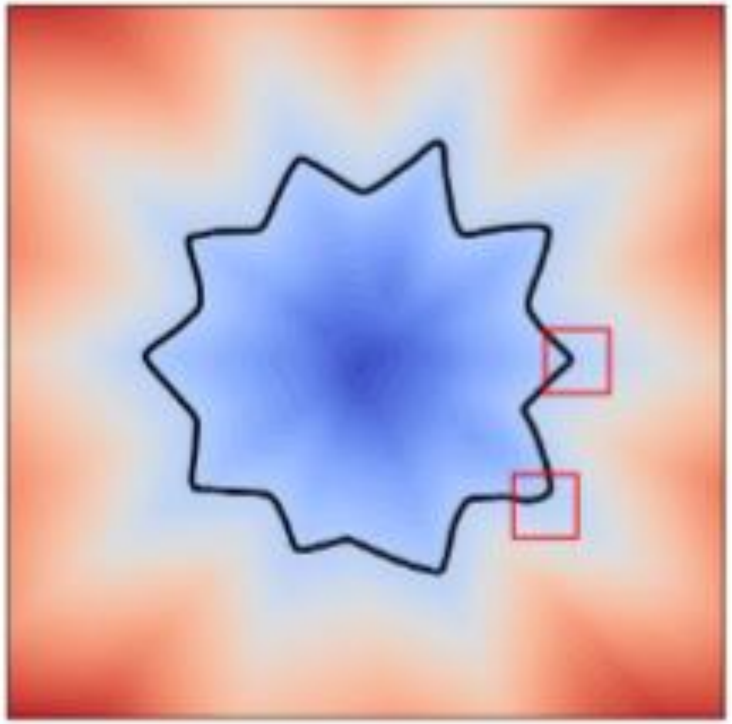


Groundtruth

Are MLP-based Representations the Right Ones?



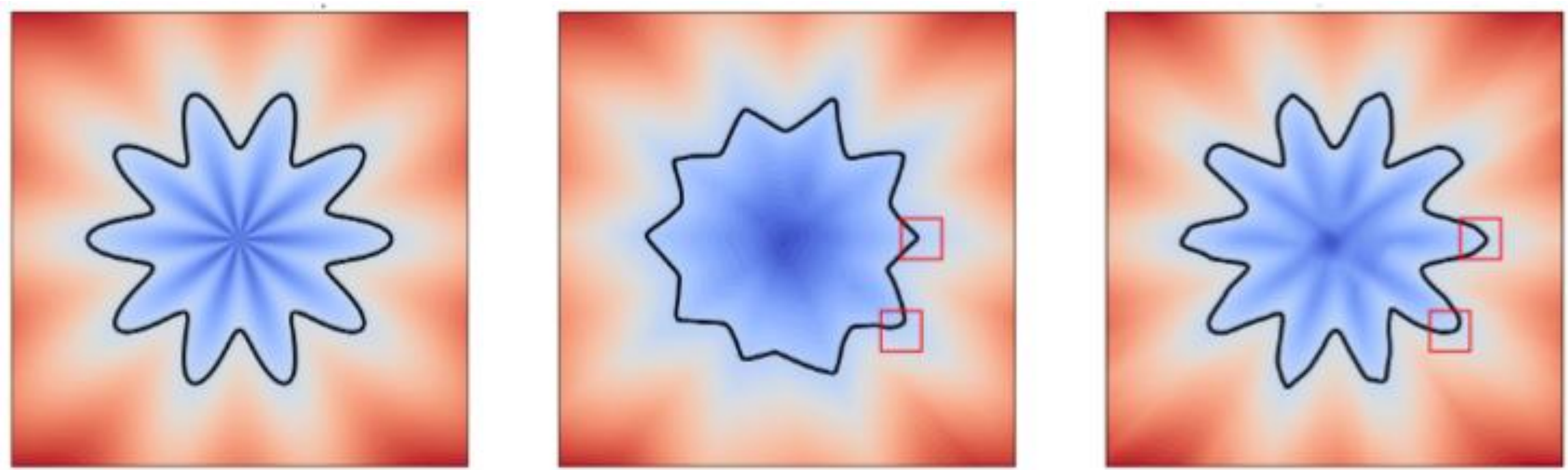
Groundtruth



Traditional ReLU-based MLP

Are MLP-based Representations the Right Ones?

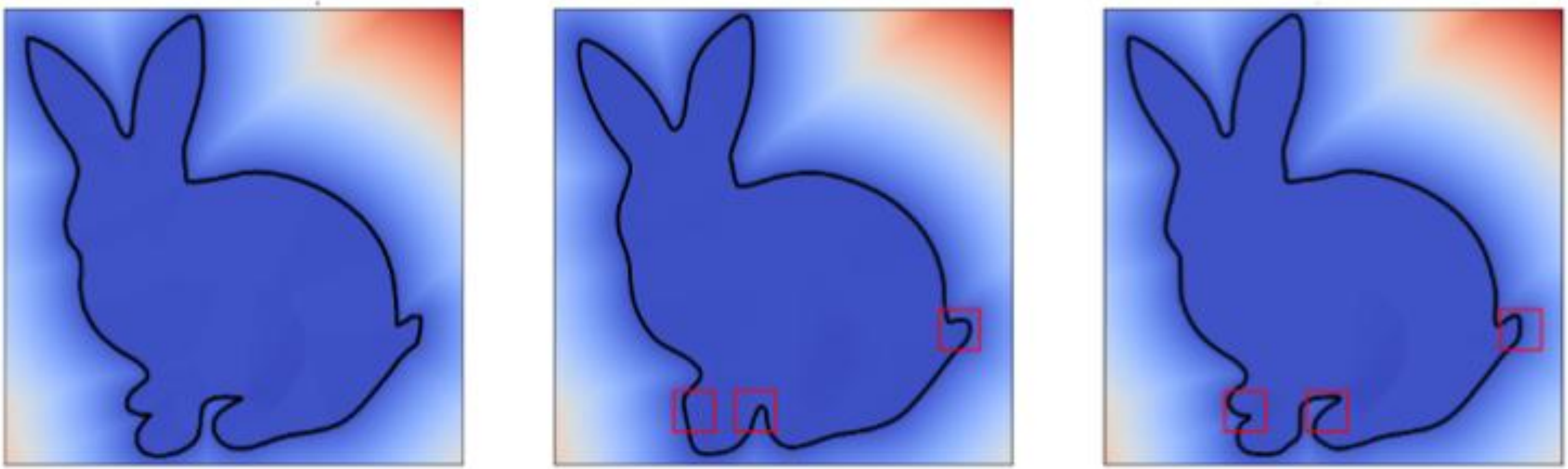
- Significantly improves the reconstruction quality



Groundtruth | Traditional ReLU-based MLP | Quadratic MLP

Are MLP-based Representations the Right Ones?

- Significantly improves the reconstruction quality



Groundtruth | Traditional ReLU-based MLP | Quadratic MLP

MLPs with Quadratic Layers

- Significantly improves the reconstruction quality

Input Image

Reconstruction using
traditional ReLU-based MLP

Reconstruction using
an MLP where the last layer is quadratic

MLPs with Quadratic Layers

- For more information (paper, results and code)

https://neqis.github.io/

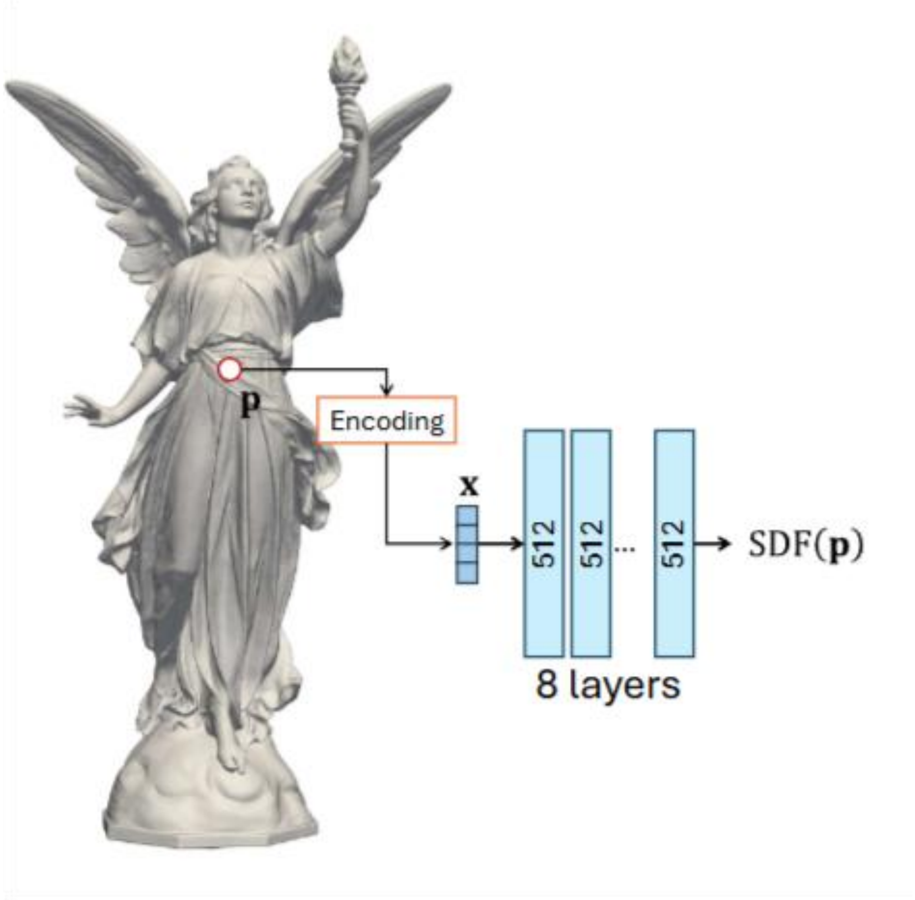
Aarya Patel, Hamid Laga and Ojaswa Sharma.

NeQIS: Neural Quadratic Implicit Surfaces (To Appear)

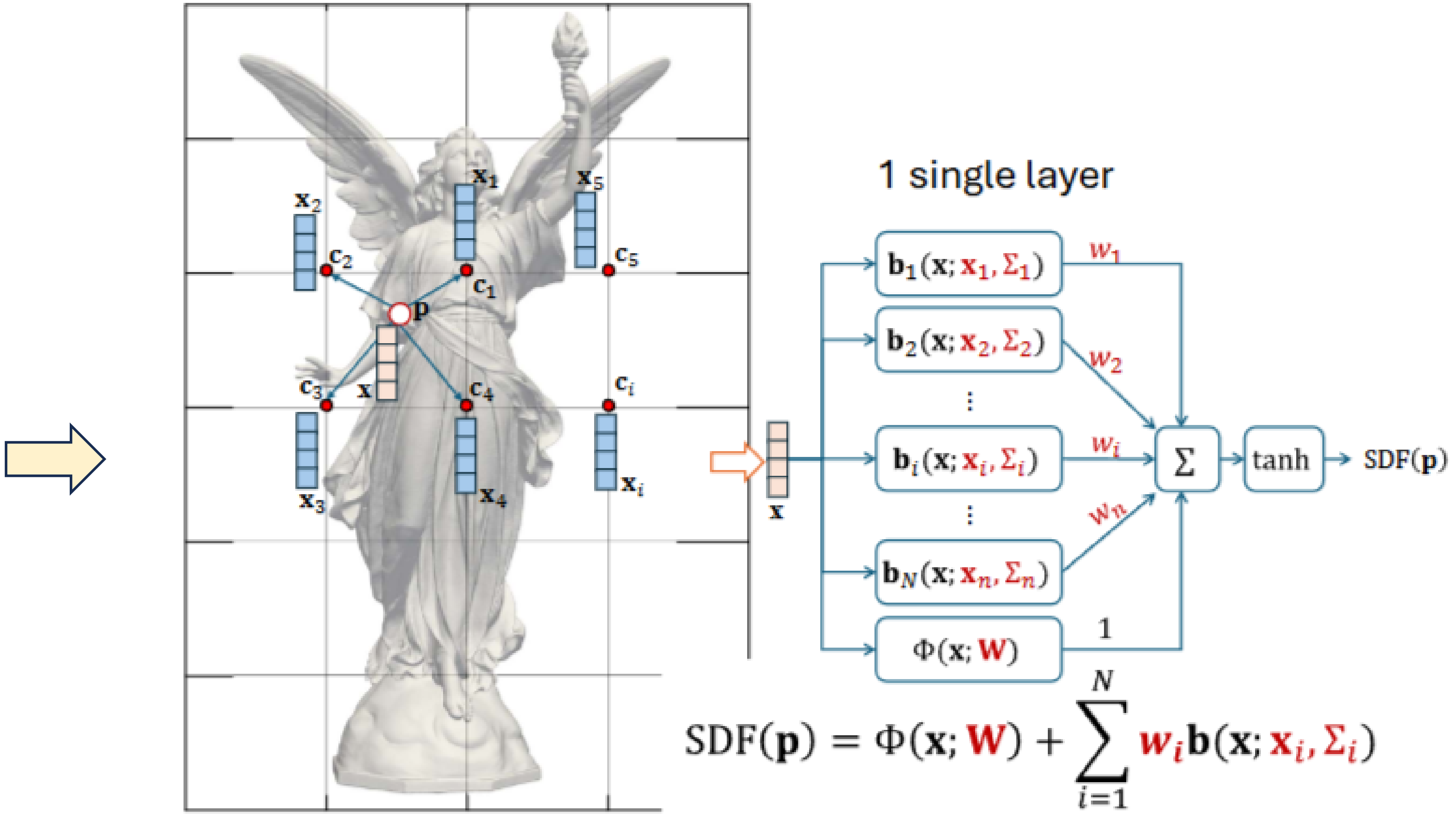
Can We do Better

- Quadratic layers significantly improve accuracy
- Still expensive (memory and computation)

The Power of the Gaussian Functions



Traditional Deep SDF



$$\mathrm{SDF}(\mathbf{p}) = \Phi(\mathbf{x}; \mathbf{W}) + \sum_{i=1}^{N} w_i \mathbf{b}(\mathbf{x}; \mathbf{x}_i, \Sigma_i)$$

Proposed Gaussian RBF networks

The Power of the Gaussian Functions

- Benefits
  - Very fast at training and inference
  - Highly accurate
  - It can represent and reconstruct signals of arbitrary dimensions

Abdelaziz Bouzidi, Hamid Laga, Hazem Wannous, Ferdous Sohel. GNF: Gaussian Neural Fields for Multidimensional Signal Representation and Reconstruction. In Computer Graphics Forum (2025).

Best Student Paper Award

https://grbfnet.github.io/

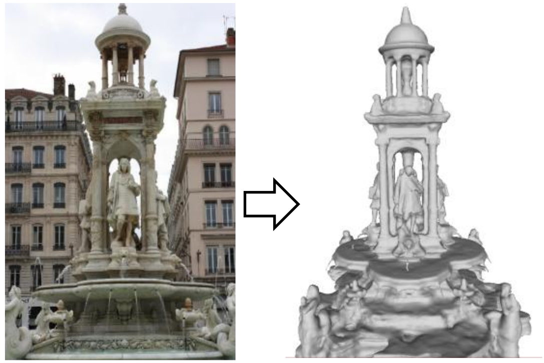
In Summary

- Neural Networks allow us to represent high dimensional signals as continuous functions
  - Images, 3D/4D geometry, and Radiance Fields (3D geometry + appearance)
- Beyond traditional MLPs
  - Gaussian and Quadratic layers can significantly improve accuracy, training/inference speed, and memory usage

https://sn-nir.github.io/ https://neqis.github.io/ https://grbfnet.github.io/

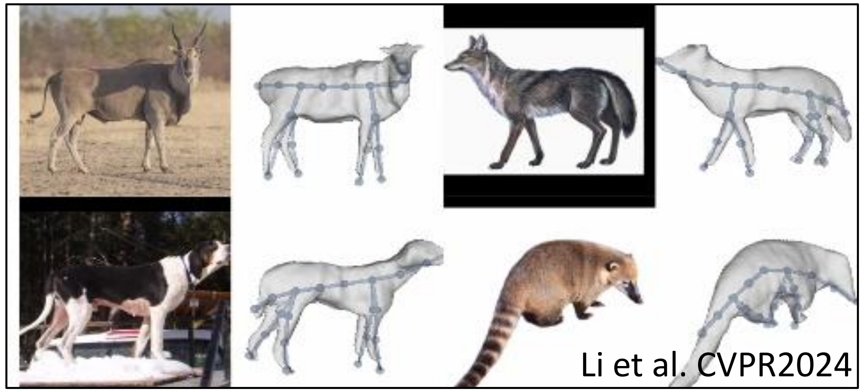
My Current Research Focus

Mathematical models that reproduce the physical world and the underlying processes that control Geometry, Appearance, Behaviors and Interactions, and Growth

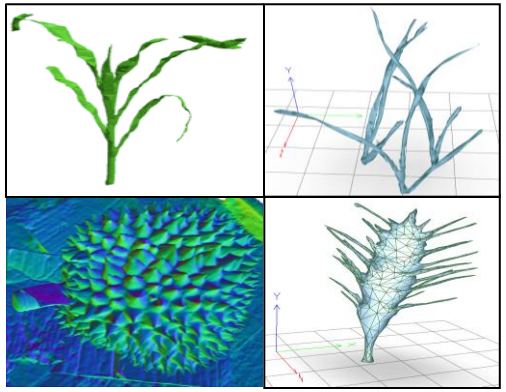


Digital twins of manmade objects

Virtual Humans



Virtual Animals (and insects)



Virtual Plants

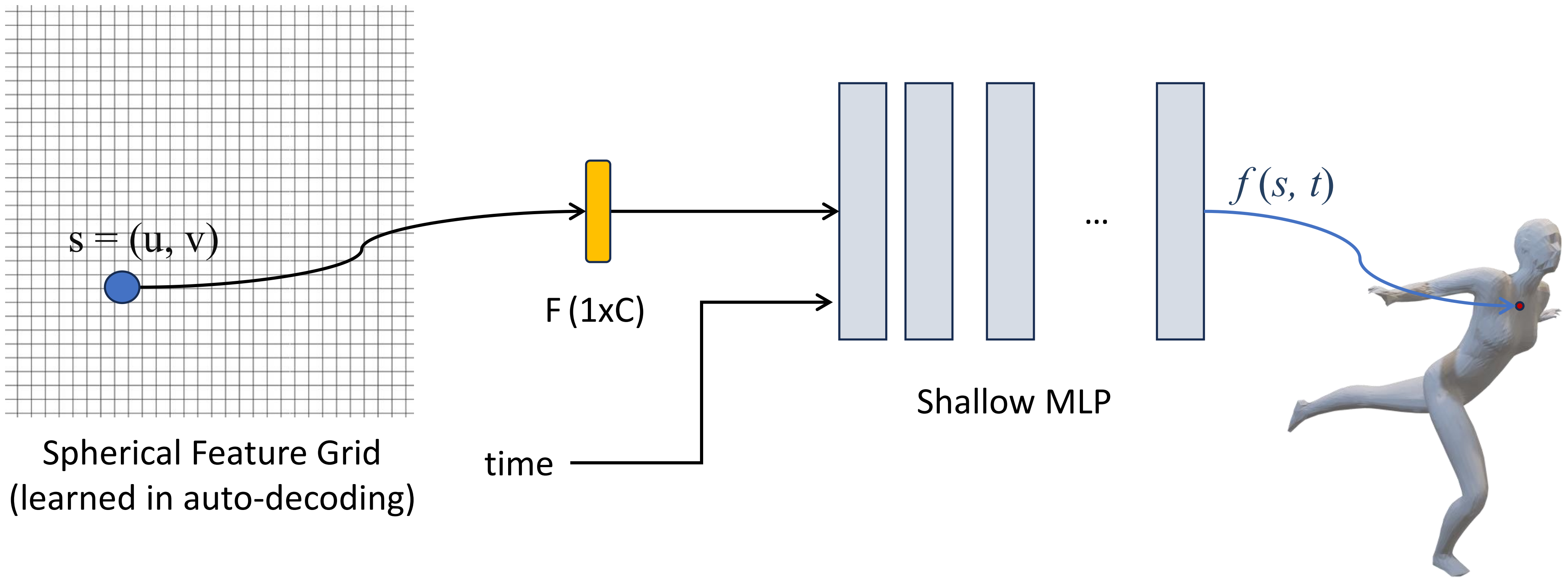
The Virtual Human Project

- Can we create/simulate virtual avatars that
  - Look like real humans (in terms of Geometry, Appearance, and Motion)?
  - Integrate the real (physical) world, move around and behave as if they are real?
  - Perceive and understand the real (physical) world as we do?
  - Think and reason the same way humans do?

- Many applications
  - Health care, Education
  - Virtual assistants and companions, …

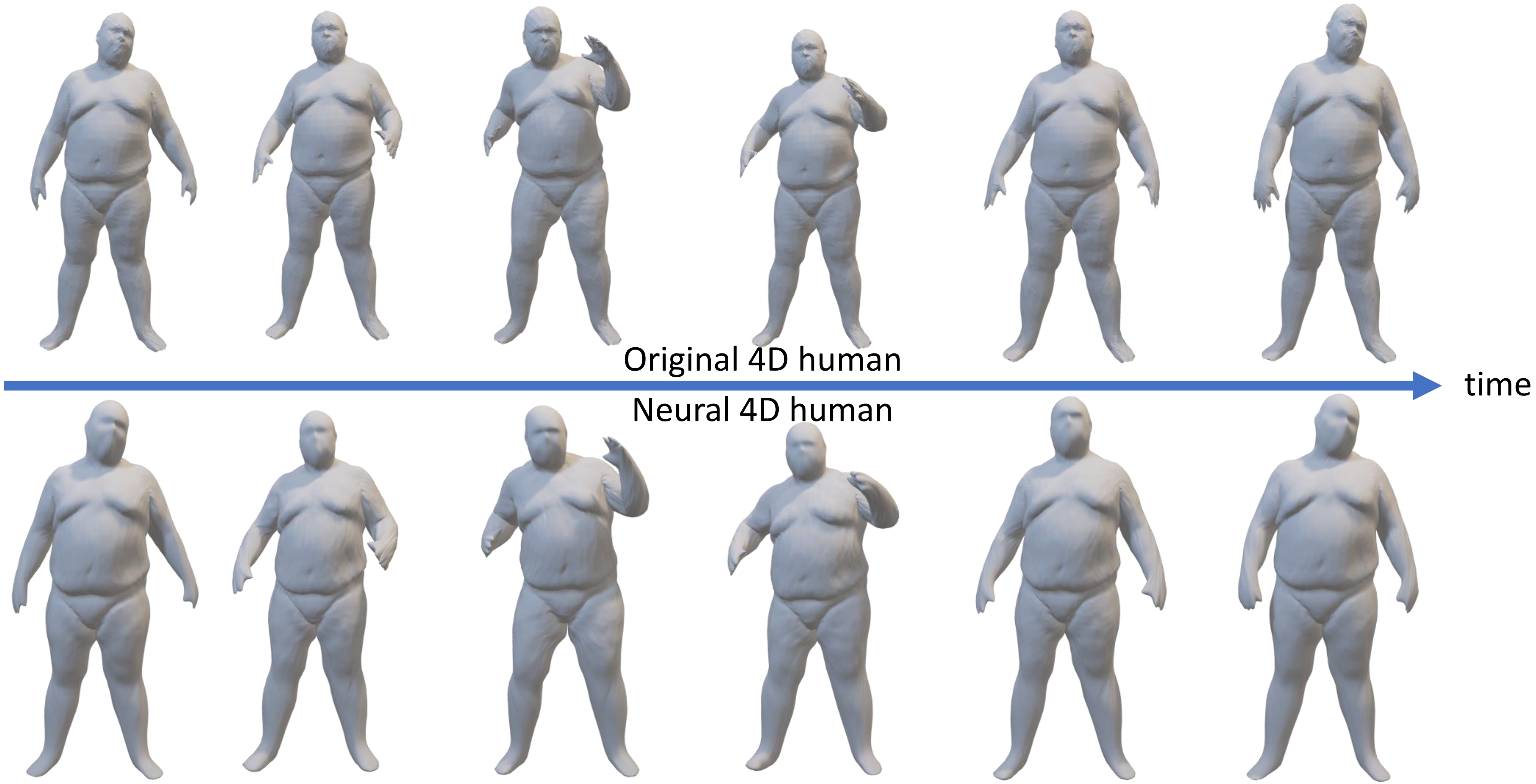
The Virtual Human Project (Funded by ARC Discovery)

- Neural 4D Humans
  - Model the 3D geometry and motion/deformation of humans using a Neural Network



The Virtual Human Project (Funded by ARC Discovery)

- Neural 4D Humans



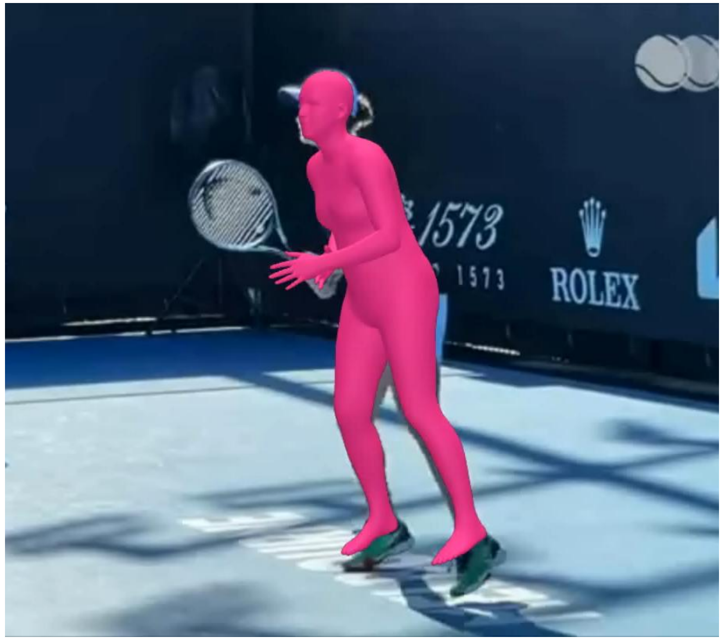
The Virtual Human Project (Funded by ARC Discovery)

- Neural 4D Humans

A Nizamani, Hamid Laga et al. (IEEE CVPR 2025). Dynamic Neural Surfaces for Elastic 4D Shape Representation and Analysis

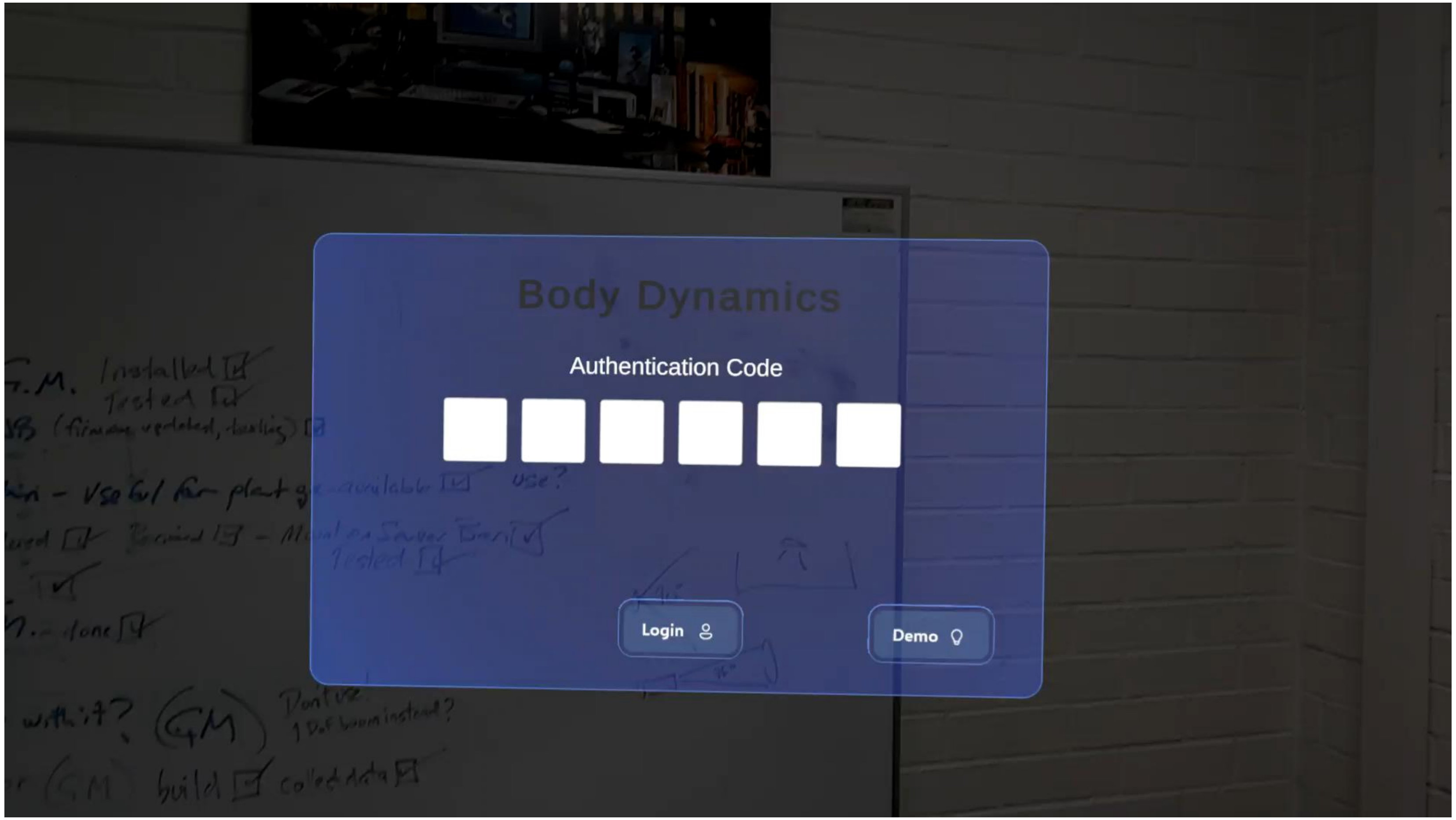
https://4d-dsns.github.io/DSNS/

The Virtual Human Project – Body Dynamics



4D reconstruction from a single RGB video

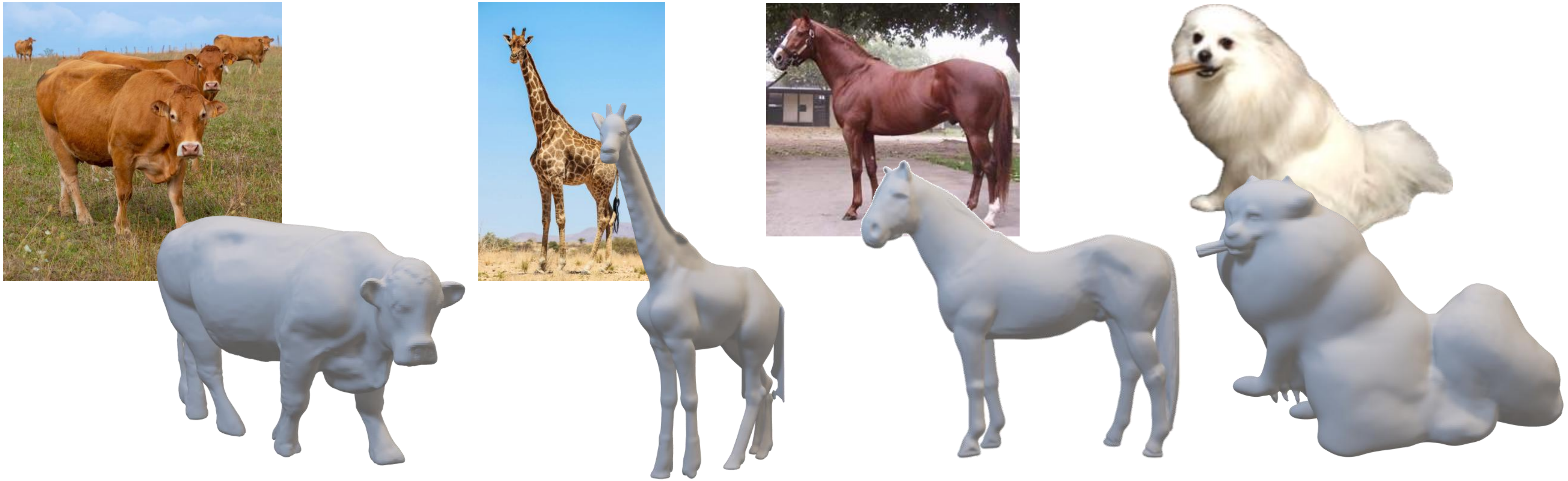
The Virtual Human Project – BodyDynamics



[OFFICIAL]

The Virtual Animal Project

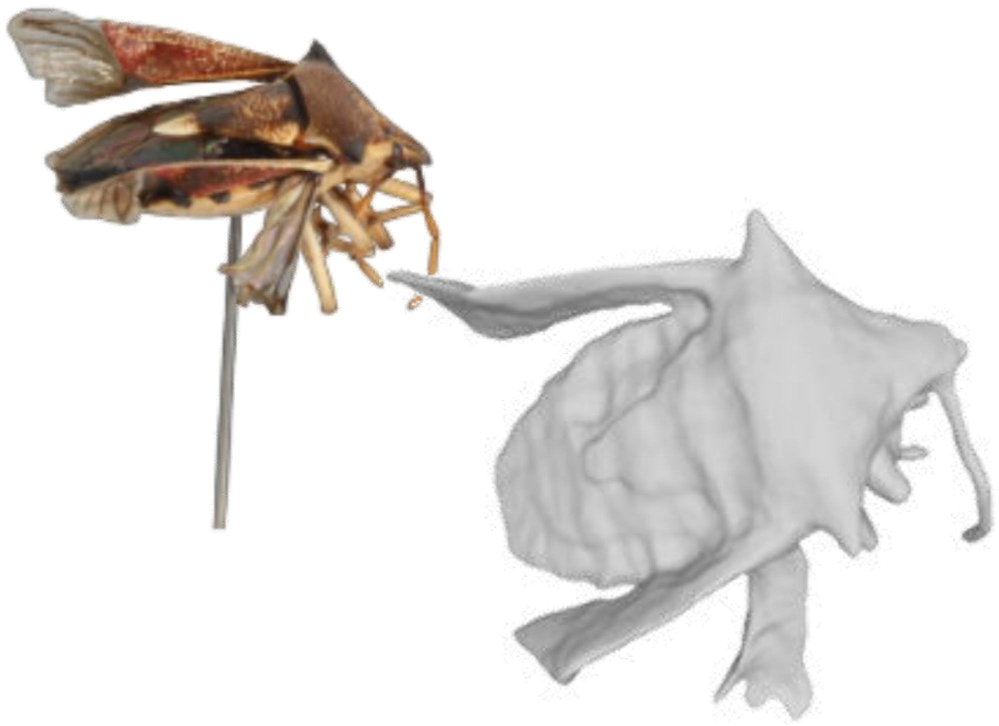
- How to reconstruct the 3D and 4D shape of animals from images and videos
  - Animals are not cooperative



Ziqi Li, Abderraouf Amrani, Shri Rai and Hamid Laga

The Virtual Animal Project

Recovering the 3D geometry of Insects from RGB images



Abderraouf Amrani, Melissa Thomas, Volker Hameau and Hamid Laga.
Dynamic Adaptive Sampling for Accurate Image-based 3D Insect Reconstruction Using Neural Implicit Surfaces. In DICTA2025

Invasive Ant Identifier (IdAntifier)

- Automatic detection and identification of 7 invasive ant species

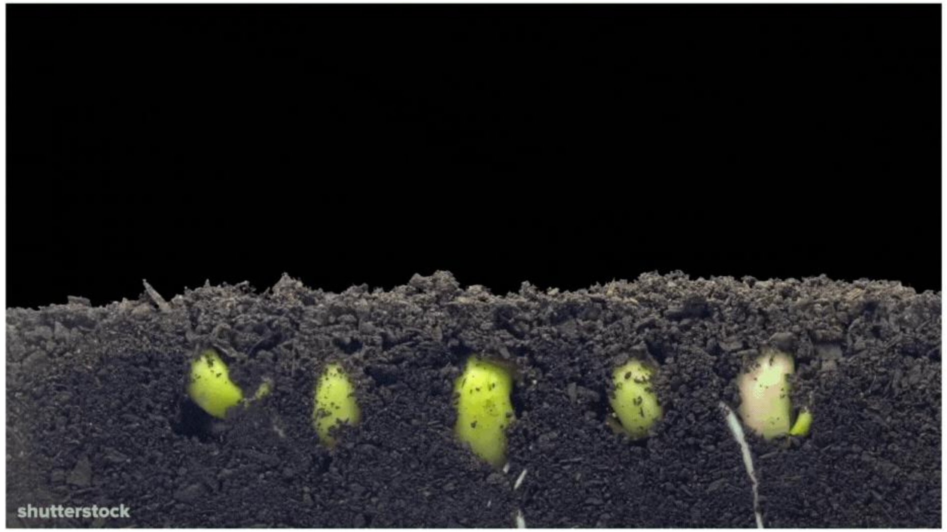
Pheidole sp. – invasive versus native

Pheidole megacephala - invasive

Pheidole metallescans - native

Pheidole megacephala (Big Headed ant) - caste variation

The Virtual Plant Project



The Virtual Plant Project

https://4dtreeshapegeometry.github.io/

In Summary

- AI and Machine Learning are revolutionizing all aspects of STEM
- In this presentation
  - How deep learning has changed the field of 3D and 4D computer vision
  - Digital life (digital humans, animals and plants) will soon be a reality
  - A wide range of applications
    - Health, agriculture, biosecurity, mining, etc.
- A lot of potential but also risks and challenges (at an unprecedented scale)
  - Cyber security
  - Ethical, responsible and explainable AI

Open for Collaborations

- We are always looking for PhD candidates
  - Send me an email with CV and list of publications (H.Laga@murdoch.edu.au)
- Open for Collaborations

Acknowledgement

- Collaborators
  - Ferdous Sohel, Melissa Thomas, Shri Rai, Guanjin Wang (Murdoch)
  - Anuj Srivastava and Eric Klassen (Florida State Uni)
  - Sebastian Kurtek (Ohio State Uni)
  - Ian H. Jermyn (Durham Uni)
  - Hedi Tabia (Paris Saclay)
  - Hazem Wannous (IMT Lille, France)
  - Guan Wang and Xie Ning (UESTC, China)
  - Mohammed Bennamoun, Farid Boussaid (UWA)
- Postdoc
  - Abderraouf Amrani (Murdoch)
  - Lian Xu (UWA)
- Current PhD Students
  - Tahmina Khanam
  - Awais Nizamani
  - Ziqi Li
  - Uchitha Rajapaksha
  - Abdelaziz Bouzidi (MU – IMT Lille France)
  - Aarya Patel
- Funding for these specific projects
  - ARC Discovery DP21 and DP22