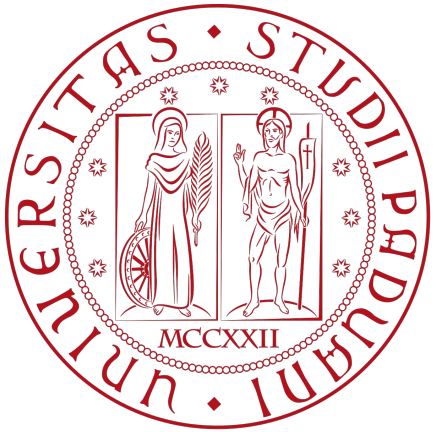
Deep 3D model optimisation for immersive interactive applications

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

Advisors: Prof. Simone Milani
Prof. Federica Battisti

Student: Francesco Pham



Introduction

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

- Increasing popularity of immersive technology
- Many applications:
  - Entertainment
  - Education
  - Health
- Three types of immersive technologies:
  - Virtual reality (VR)
  - Mixed reality (MR)
  - Augmented reality (AR)

Virtual Reality (VR)

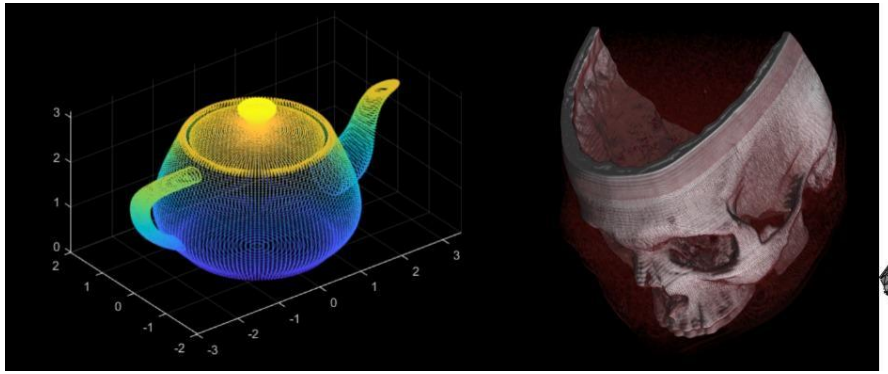
- Most common form of VR involves a HMD (*Head Mounted Display)* headset
- Stereoscopic vision
- 6 degrees-of-freedom movement

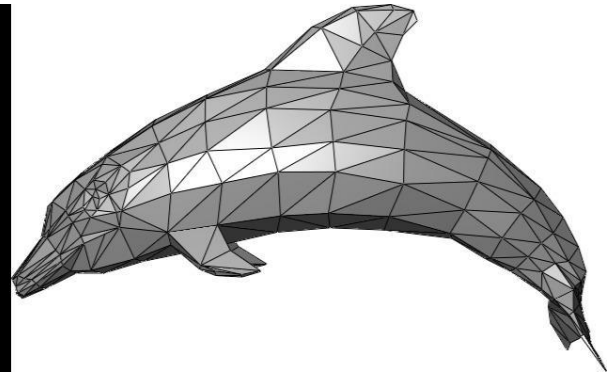
3D Computer Graphics

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

Different digital representation of 3D models:

- *Point cloud*
- *Voxel*: volumetric description
- ***Mesh*****:** object's surfaces description





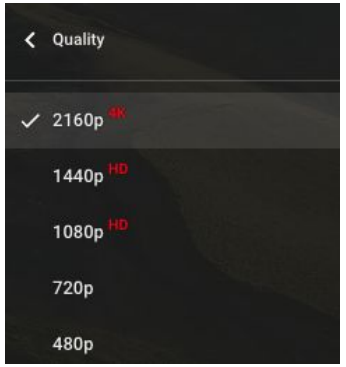
Quality of Experience

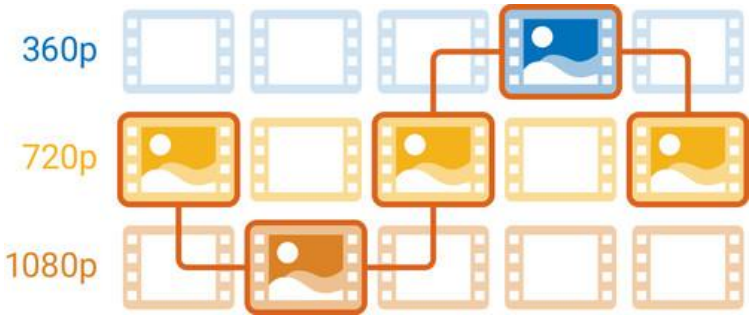
DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

Main factors impacting QoE:

- *Frame rate* (fps)
- Latency
- Visual quality

A video streaming analogy:



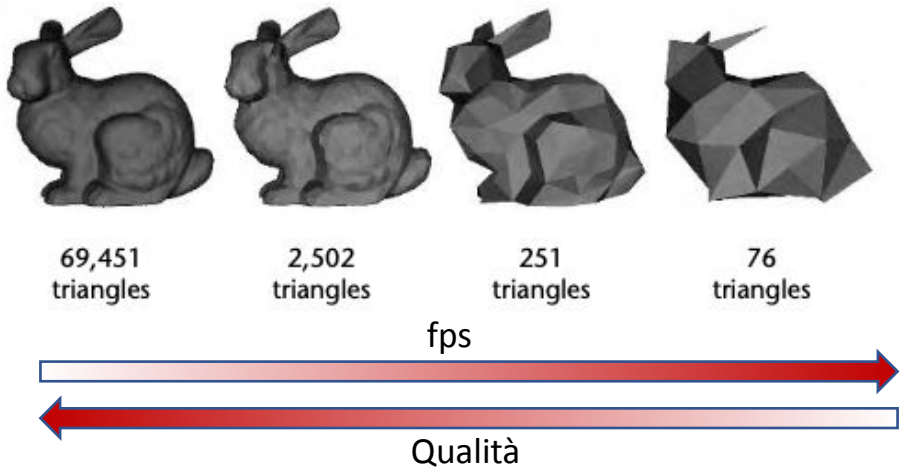


Quality metrics

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

How to select the optimal level-of-detail to obtain the best compromise between frame rate and visual quality ?

→ **Define a quality metric**



SSIM *(structural similarity index measure):* measures the structural similarity between two images

$$SSIM = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}$$

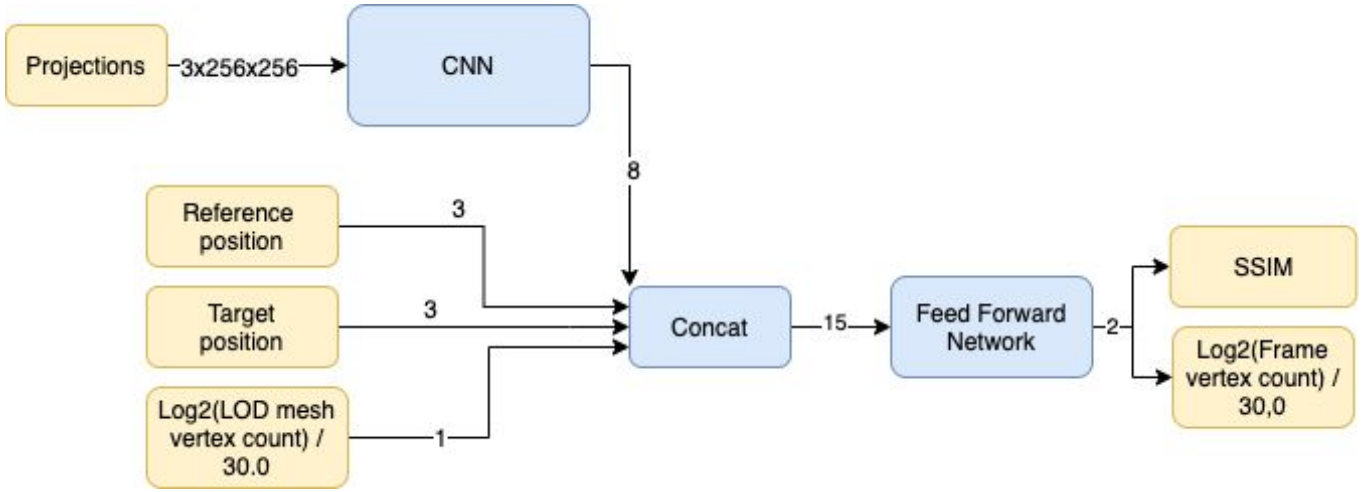
→ Evaluation of perceived quality by comparing the rendered images of original model and simplified model

Deep learning

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

Objective: real time prediction of the perceived quality of a mesh and the frame rate

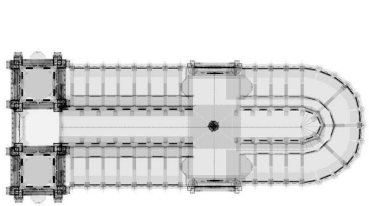
Deep Learning:
It's effectiveness in predicting data makes it suitable for predicting quality metrics



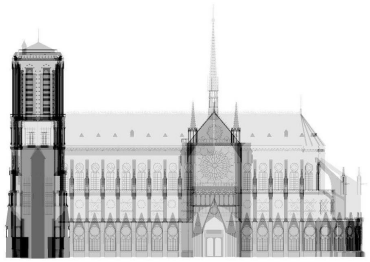
Dataset

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

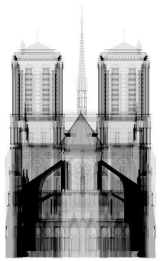
- Dataset generated using *Unity* rendering engine
- C# scripting for screenshot capturing from different point of view
- 9 3D models in 4 Level-of-Details
- 200 random viewpoints
- Orthographic surface count projections for each 3D model



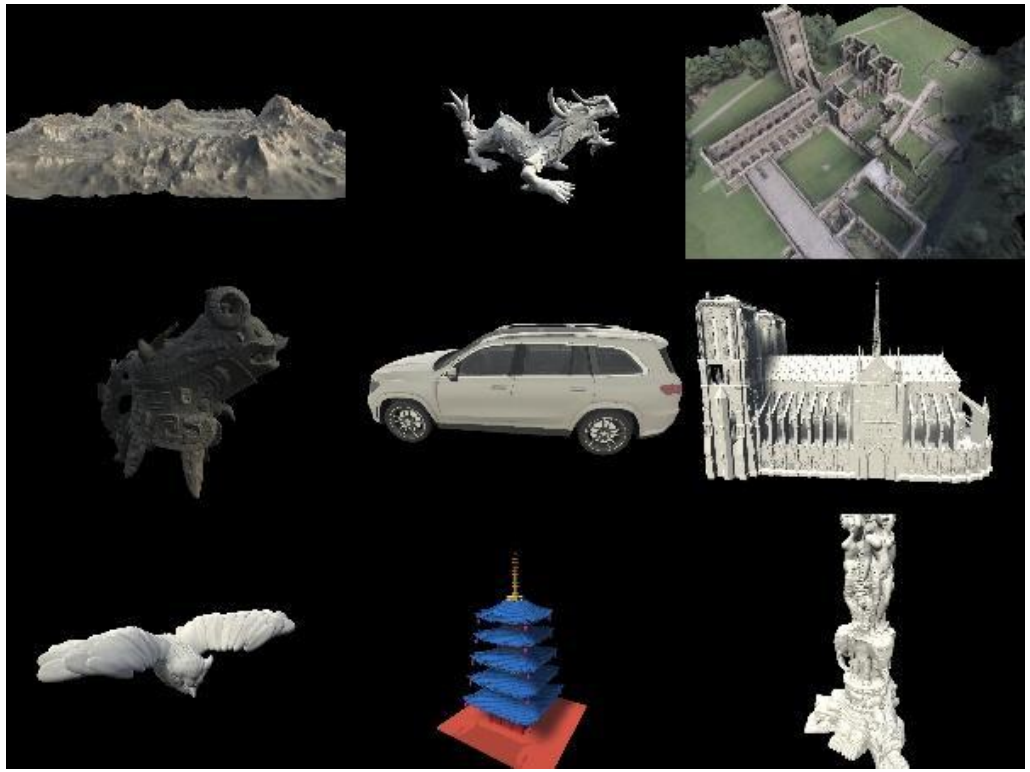
Piano Orizzontale



Piano Verticale



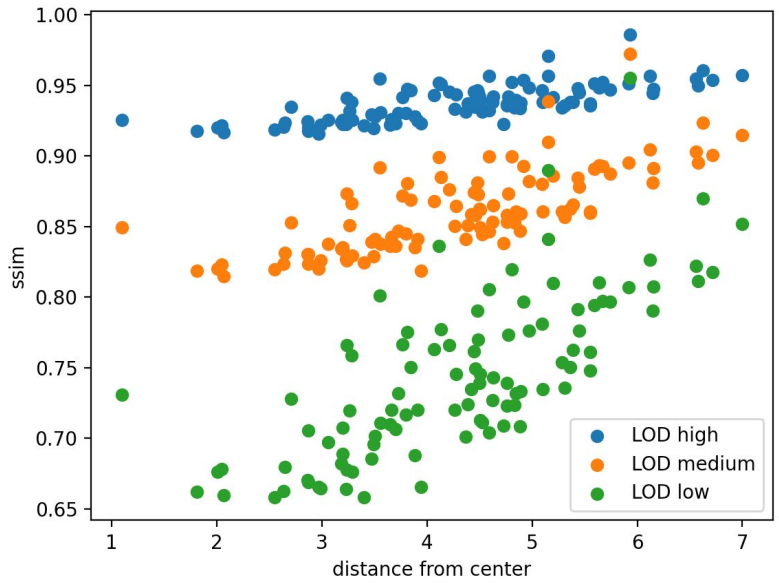
Piano Laterale



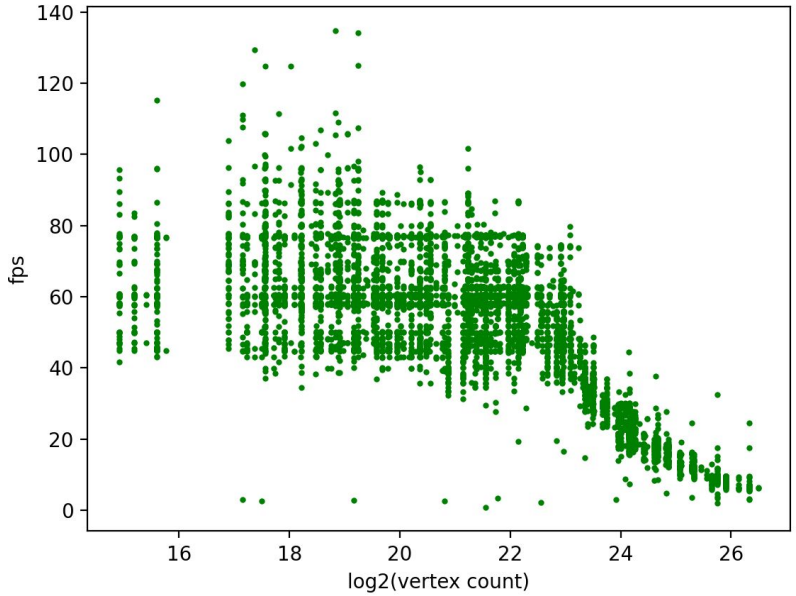
Data analysis

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

We observe an increase of the SSIM index with the distance from the object



Correlation between frame rate and the vertex count of the frame being rendered

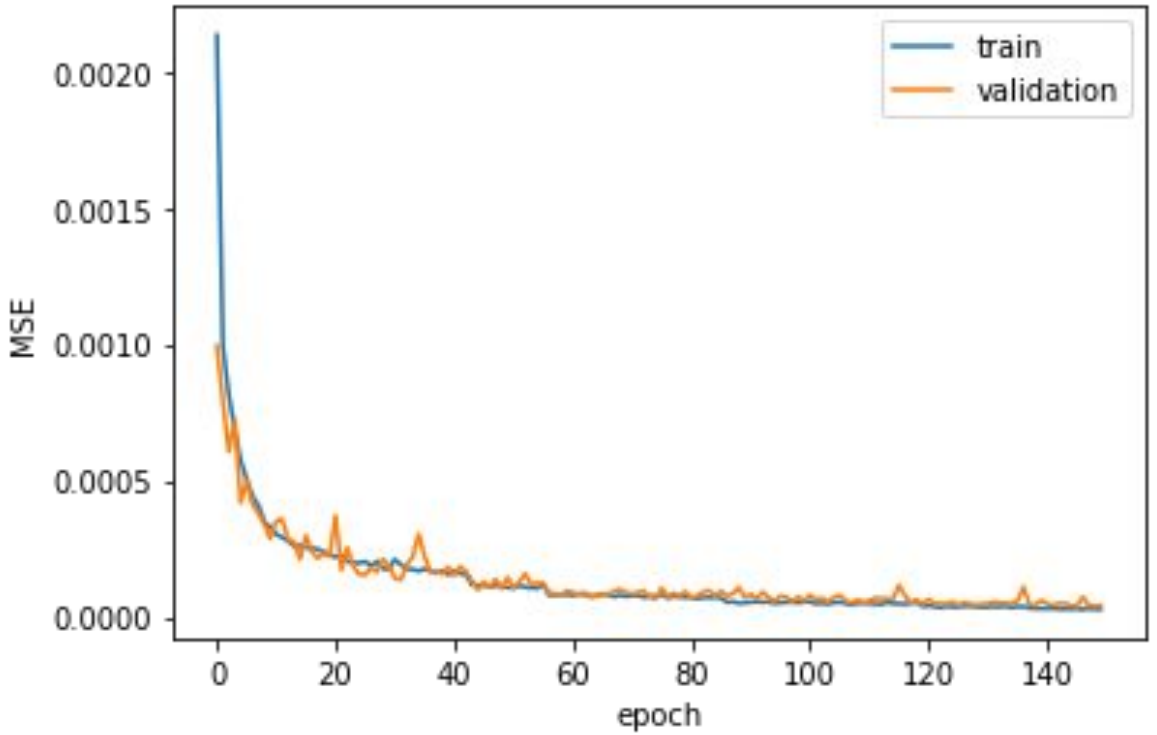


Experimental results

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

- The training curve shows an absence of overfitting/underfitting
- The test shows an accurate prediction of SSIM and FPS indices
- Generalisation to different 3D mesh

Loss at 150 epochs	MSE
Training	2.8e-5
Validation	4.2e-5
Test	5.69e-4



Conclusion

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

- We realised a deep learning model that accurately predicts objective quality metrics
- Our model provides a selection tool for the optimal level of detail to render
- The quality of experience can improve by addressing the main factors that impact VR immersion

Thank you

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE