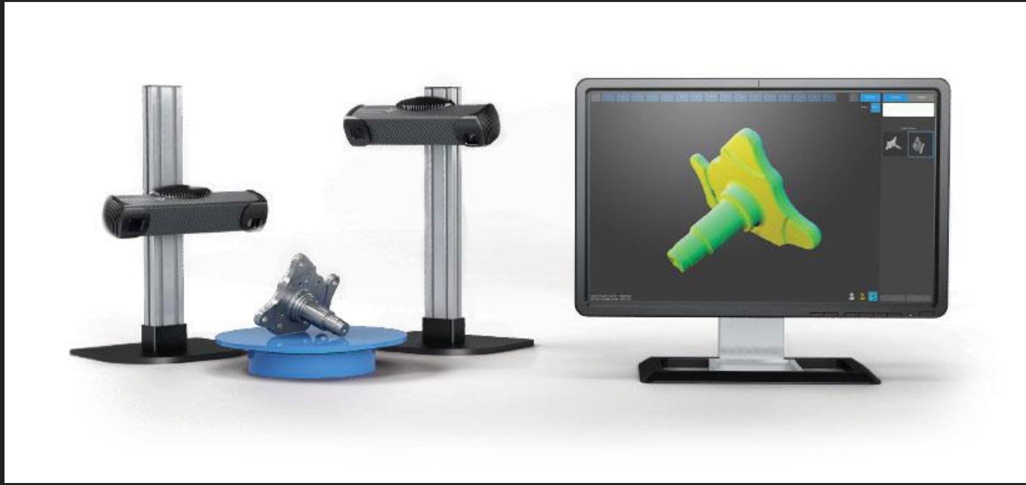
Super sampling of 3D scans using deep learning

Martin Melicherčík
Supervisor: doc. RNDr. Martin Madaras, PhD.

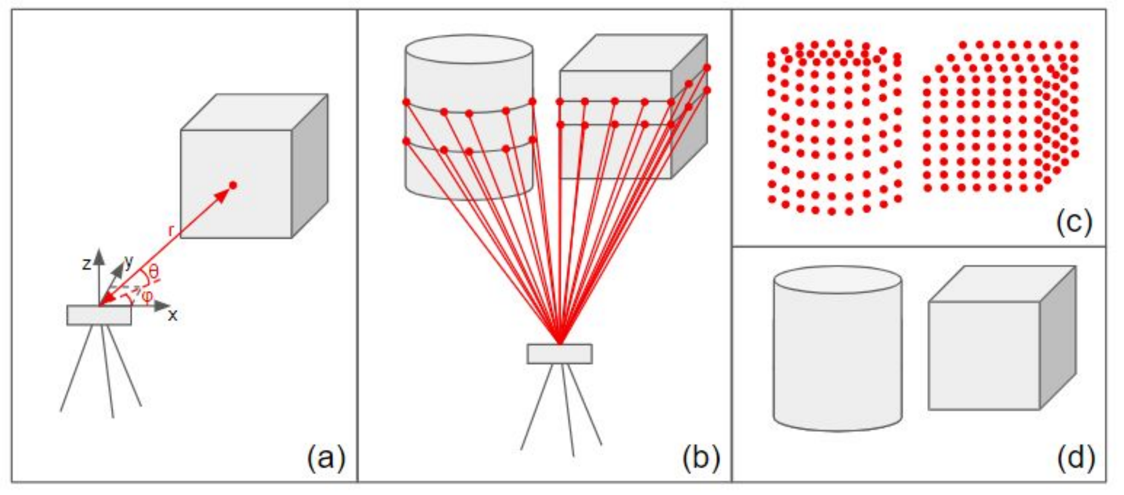
3D scanning

- Virtual representation of 3D objects
- Usage: product validation, analysis of archeological findings…



Representation of scanned scene

- Point cloud - set of points in 3D space
- Structured point cloud - 2D matrix where every element represents information about position in 3D space

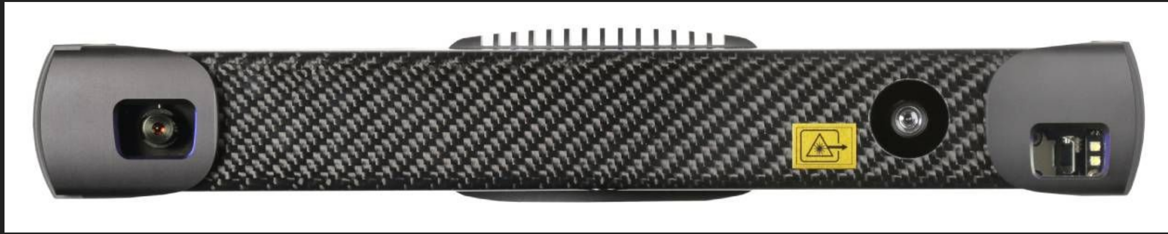


Data from MotionCam-3D

- Depth map (Structured point cloud)
- Intensity texture (Grayscale photo)

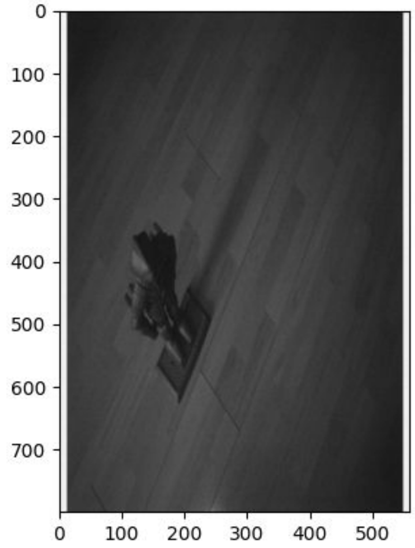
Camera

Laser

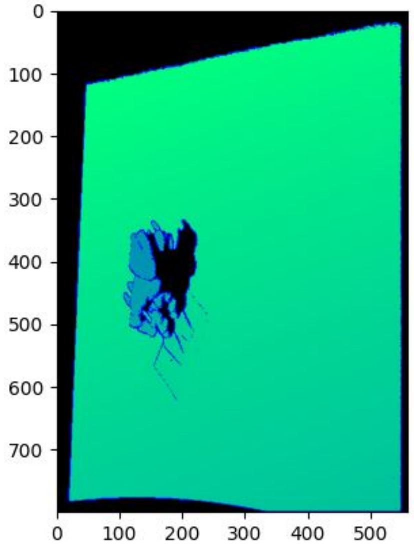


Photoneo MotionCam-3D

MotionCam-3D data example



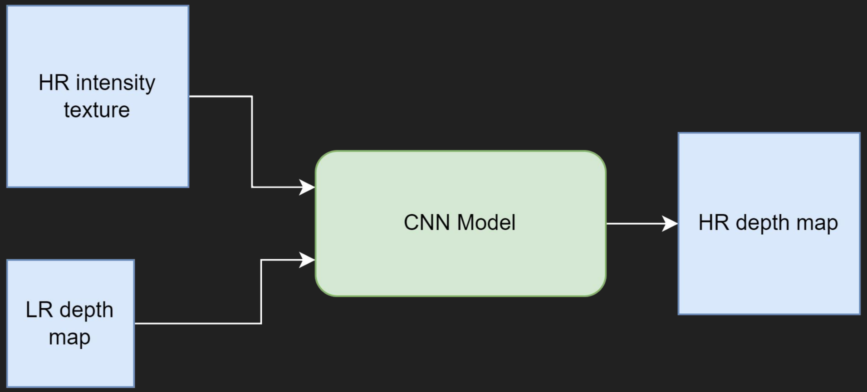
Grayscale



Linear scale **blue** to **green**

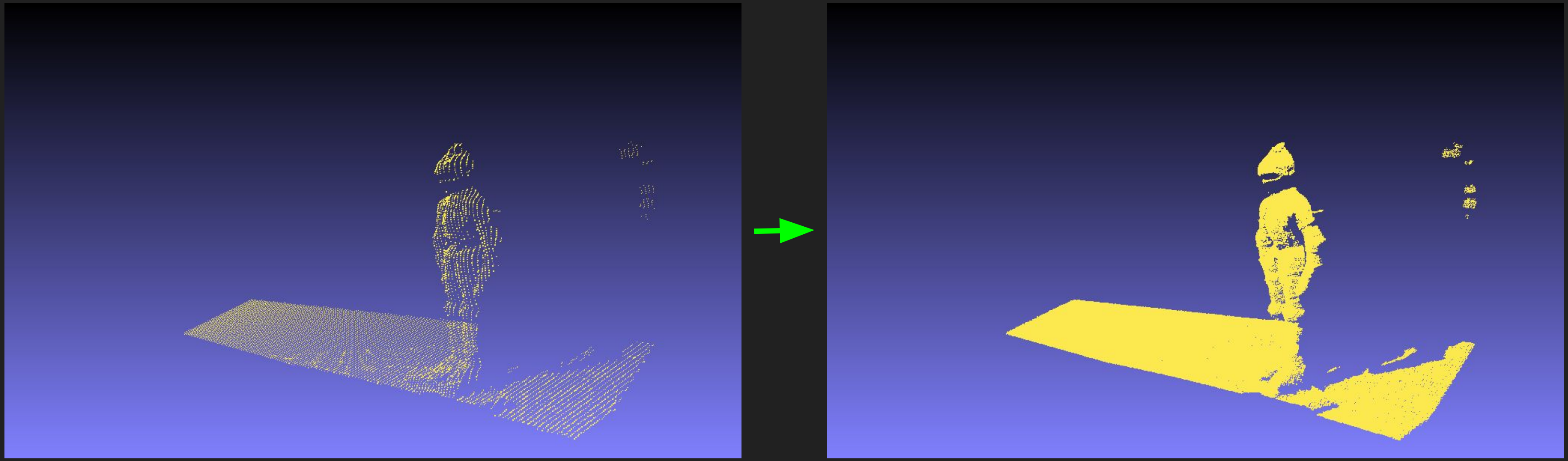
Problem definition

- Depth map super resolution
- INPUT: LR depth map, HR Intensity texture (grayscale image)
- OUTPUT: HR depth map
- Intensity texture is used as guidance image (additional input information)



Problem definition

- Point clouds generated from LR and HR depth maps



LR HR

Goals of thesis

- Create dataset using Photoneo MotionCam-3D
- Choose and implement CNN model for defined problem
- Propose metric for evaluation

Dataset

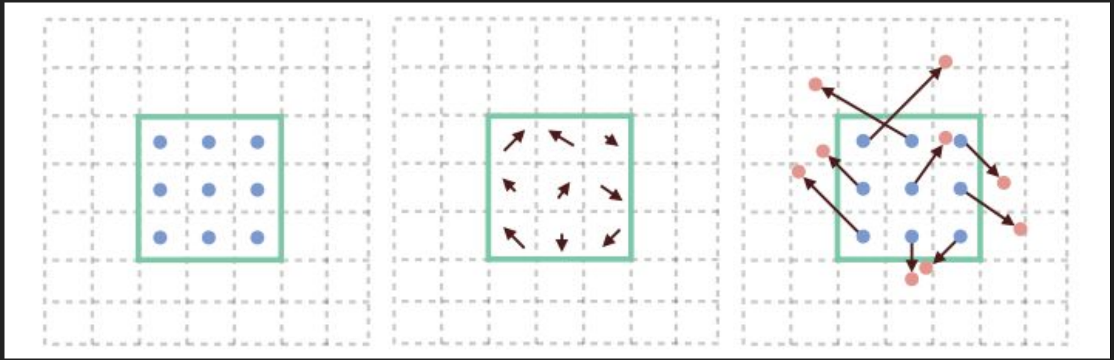
- Dataset sample is triplet: HR intensity texture, LR depth map, HR depth map
- Dataset size: 1200 samples
- Main focus on the 3D object fusion task
- Scene type: One object placed on the flat ground surface

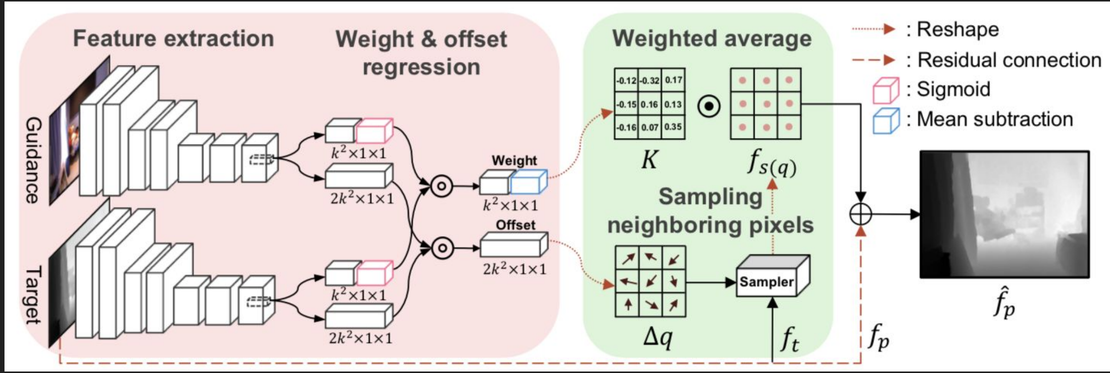
Implemented Depth map SR models

- FDSR : *Towards Fast and Accurate Real-World Depth Super-Resolution: Benchmark Dataset and Baseline, Lingzhi He et al. 2021*

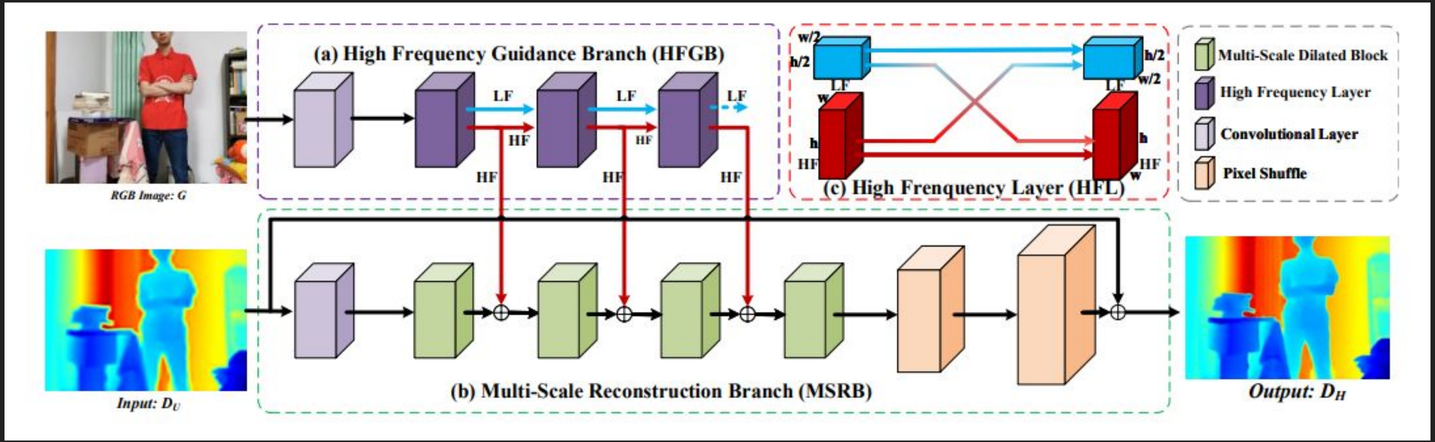
- DKN : *Deformable Kernel Networks for Joint Image Filtering: Beomjun Kim, Jean Ponce, Bumsub Ham 2019*

DKN model architecture



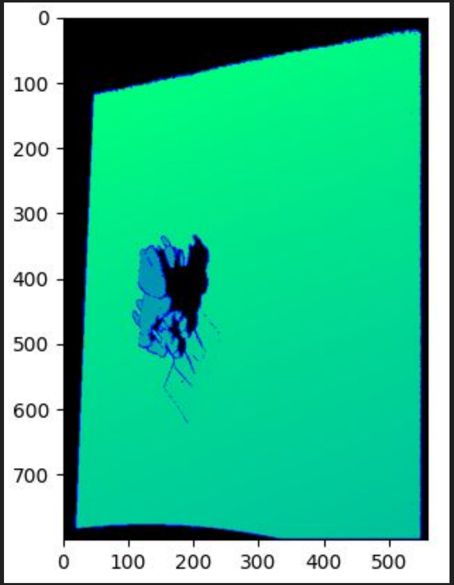


FDSR model architecture



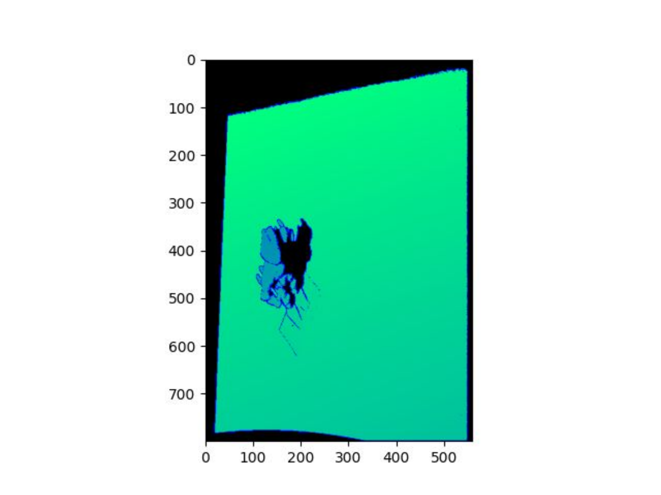
Undefined areas in depth maps

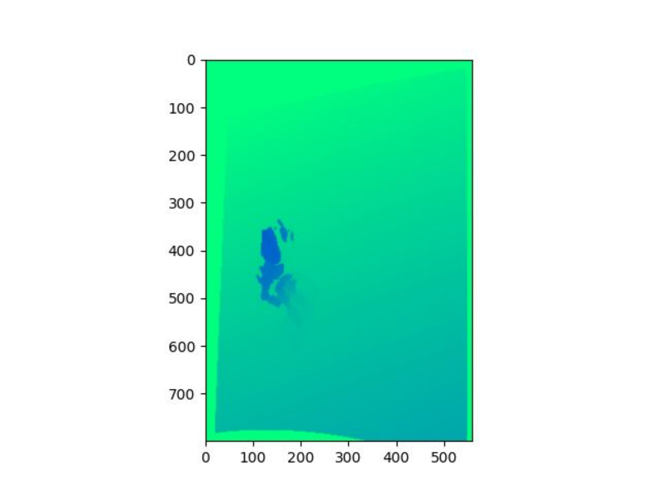
- Depth maps contain undefined pixels that form areas (Black pixels)
- Existing models work with filled depth maps
- *Hole-Filling of RealSense Depth Images Using a Color Edge Map, Ji-Min Cho, Soon-Yong Park, Sung-Il Chien 2020*

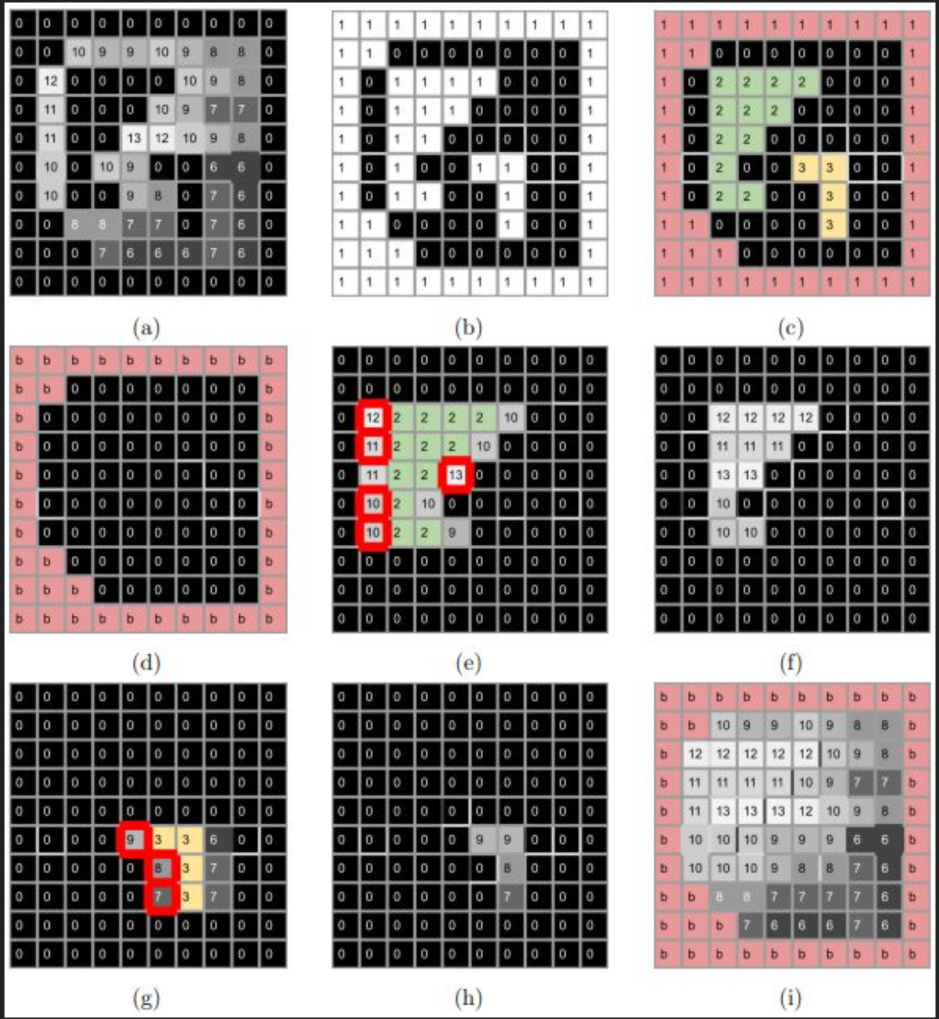


Filling depth maps

- We propose filling method that:
  - Separates object from background
  - Fills background
  - Fills small near object holes
- We have to care about undefined pixels while training otherwise they create large error

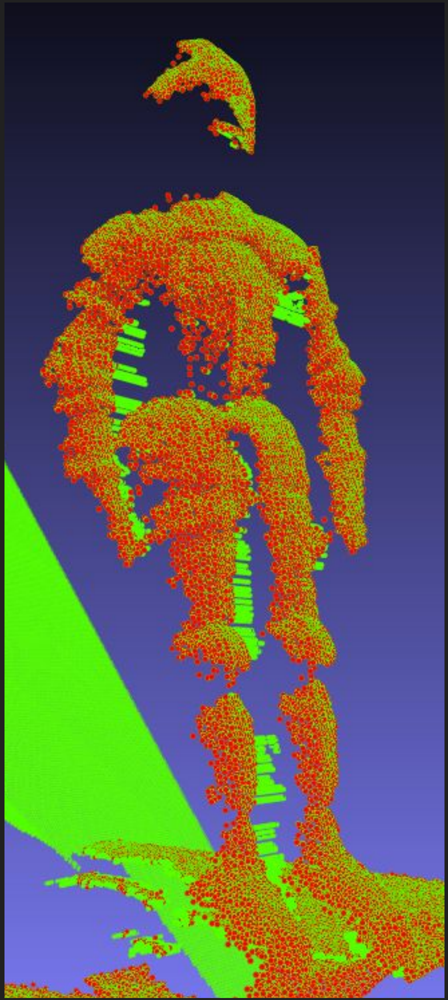




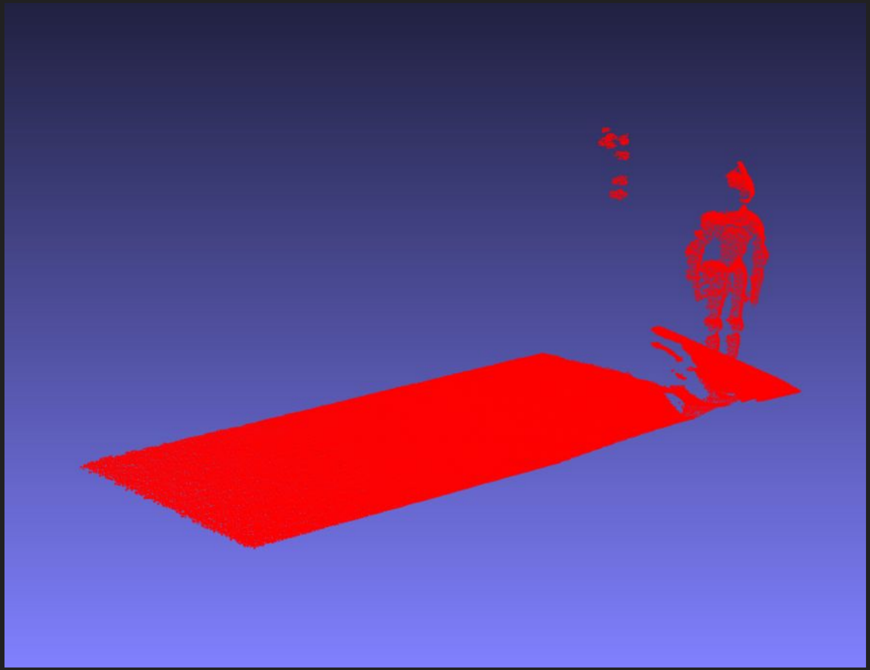

(a)
(b)
(c)
(d)
(e)
(f)
(g)
(h)
(i)

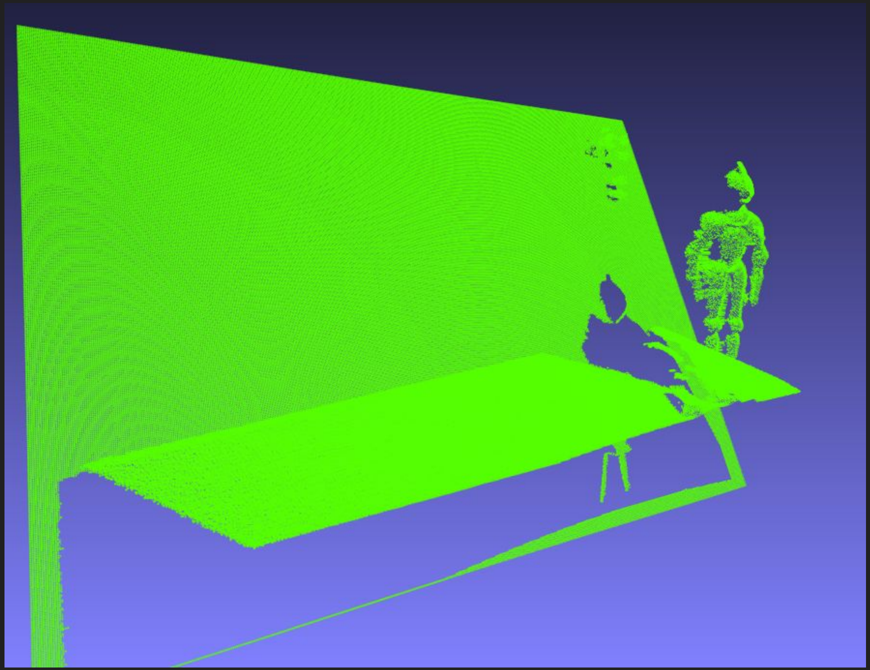
Depth map filling

- Small holes filled row-wise
- Hole row filled with maximum value of the row’s hole border pixels
- Approximate reasonable values for stable training



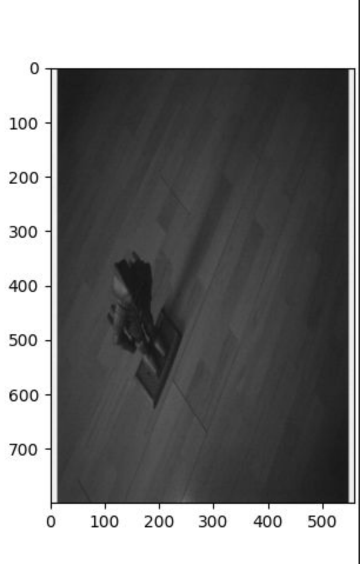
Depth map filling - background plane

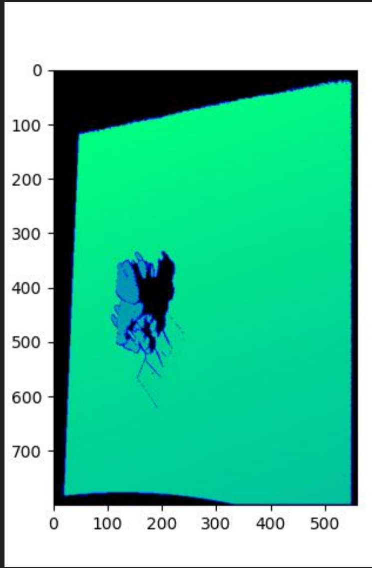




Texture augmentation

- Improve correlation between depth map and intensity texture
- Apply same filling procedure as for depth map undefined areas
- Removes surface structures from undefined areas

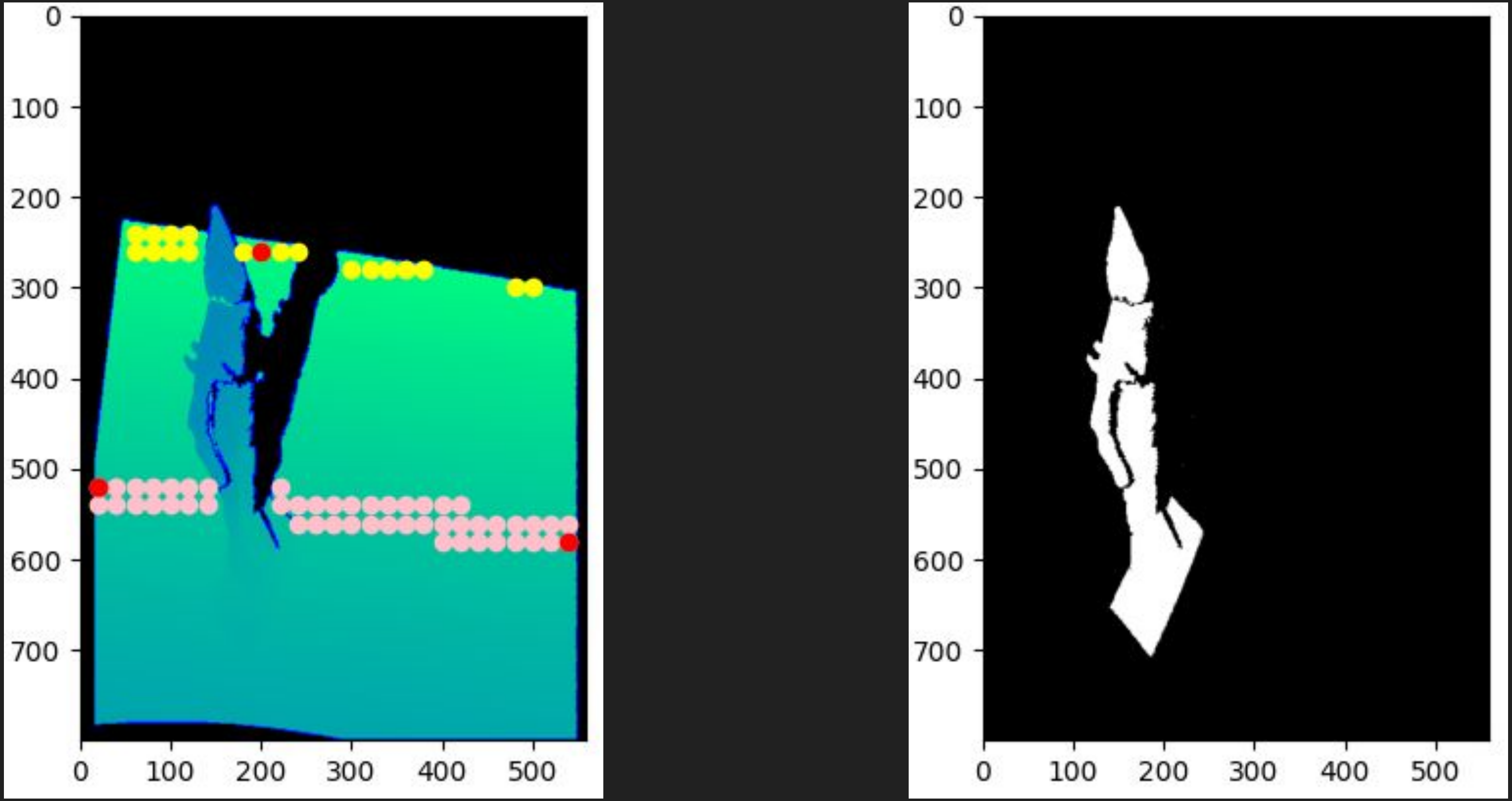




Modifying CNN models

- Halving DKN filters on each layer to fit our hardware limitations
- FDSR and DKN use L2 loss
- We propose loss function that gives attention to scanned object
- Object loss function:
    - Fitting plane to scene ground
    - Extracting set of pixels representing scanned object
    - Assign 100-times higher weight to the object’s pixels
    - Compute weighted mean squared error

Object loss function



Evaluation

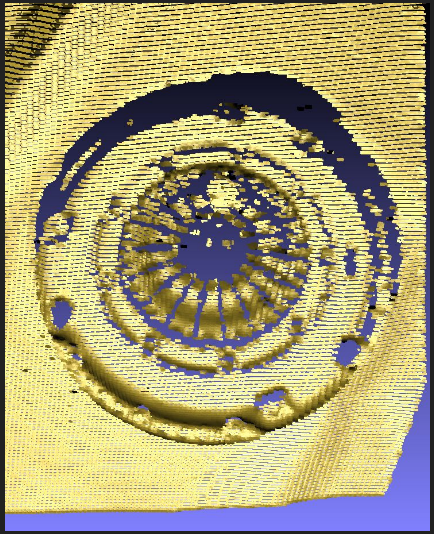
- Metrics
  - Depth map - RMSE (quantitative)
  - Point cloud
    - Analysis in Meshlab (qualitative)
    - Hausdorff’s distance of point clouds (quantitative)
  - Mesh - future work

Depth map metric -RMSE

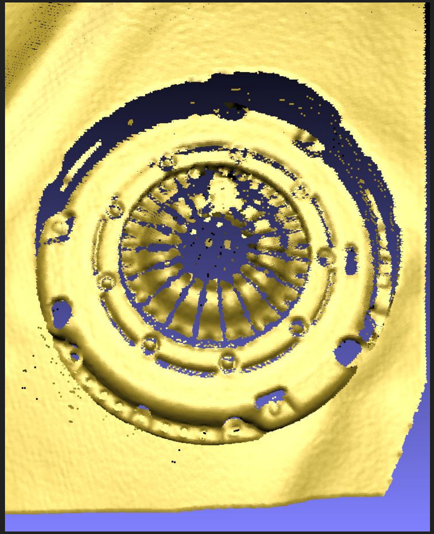
- RMSE - Root mean squared error
- Object RMSE - RMSE computed only from object's pixels
- Object loss - Our proposed loss function

Method	RMSE	Object RMSE	Object loss
FDSR	1.9537	4.4297	1.0427
DKN	2.2696	6.9191	2.3058
Nearest	3.0778	10.8156	7.0692

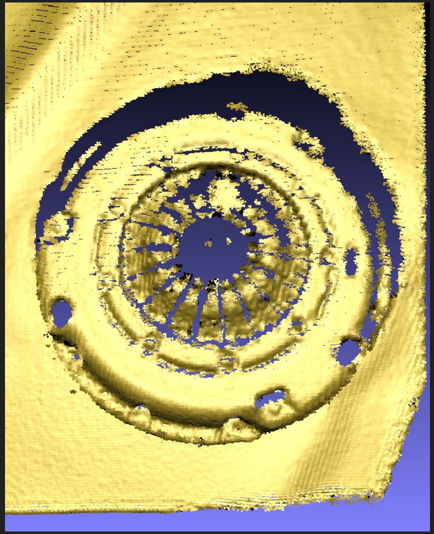
Point cloud metrics - Meshlab analysis



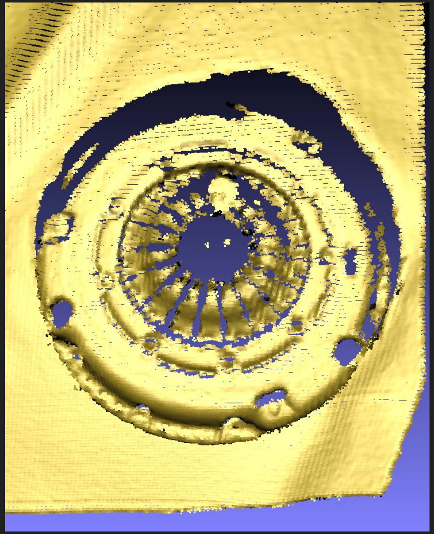
LR



HR



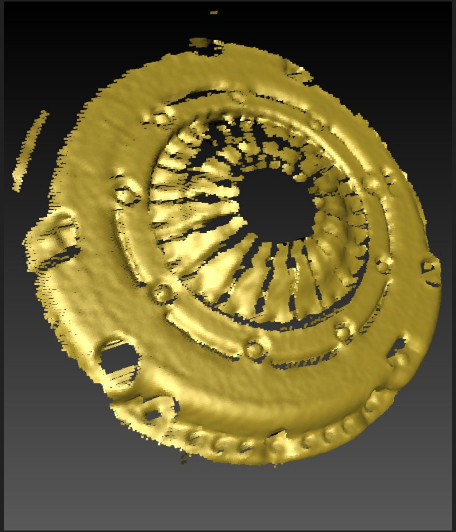
DKN



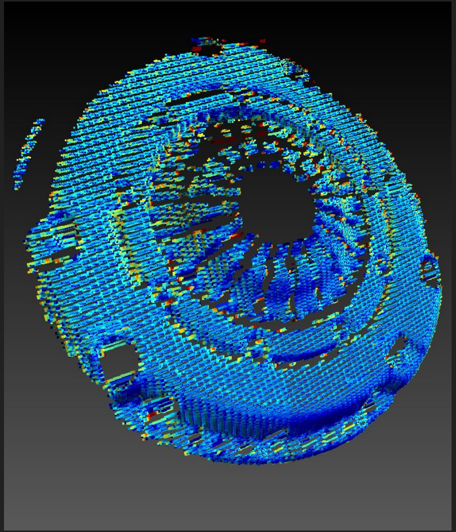
FDSR

Point cloud metrics - Hausdorff's distance

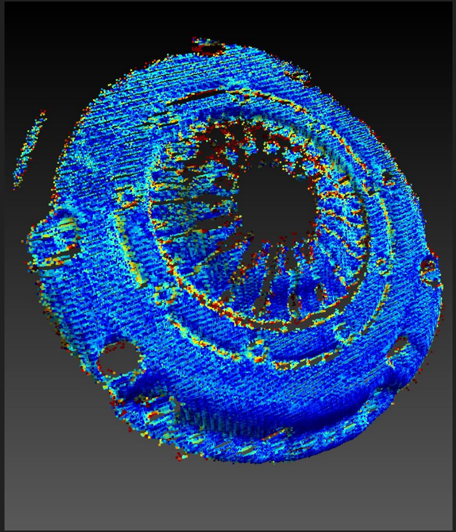
- Linear mapping of Hausdorff's distance to RGB spectre
- Mapping is from **blue** (small distance) through **green** to **red** (large distance) scale 0 - 2 mm



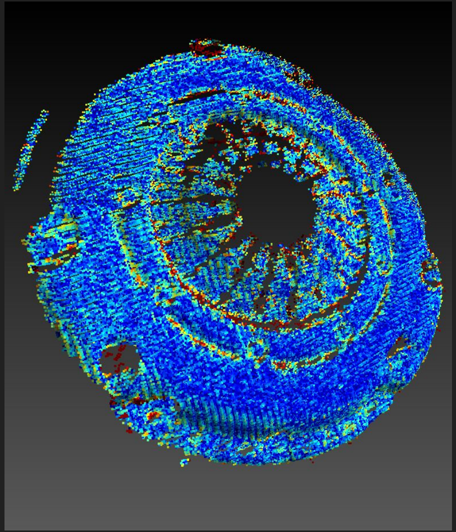
HR



LR



DKN



FDSR

Point cloud metrics - Hausdorff's distance

Method	min	max	mean	RMSE
FDSR	0.0000	9.8178	0.4650	0.7755
DKN	0.0000	10.3815	0.4449	0.7978
Nearest	0.0000	29.5586	0.4144	0.7455

Time measurements

- Pipeline computing time

Resolution [px]	140x200	560x800	1120x800	1680x1200
Time [s]	0.054	0.068	0.091	0.184

- Models computing time

Model	FDSR	DKN
Time [s]	0.007	0.634

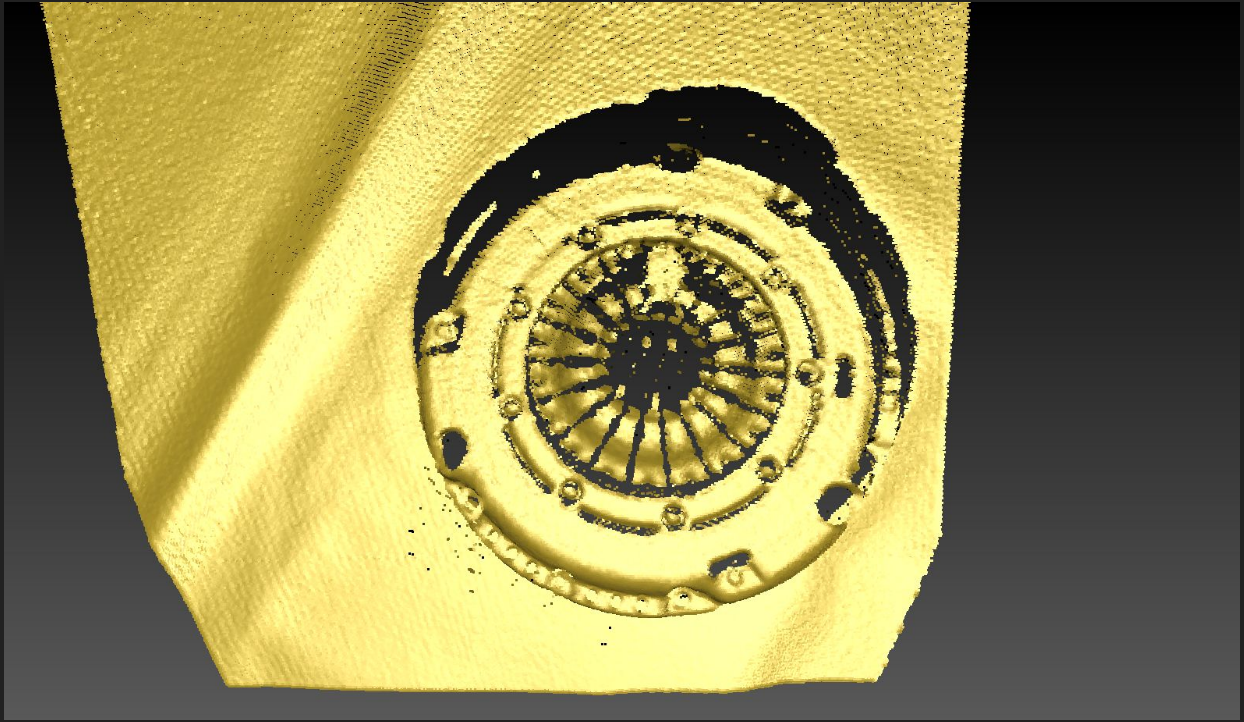
Thank you for your attention

Martin Melicherčík

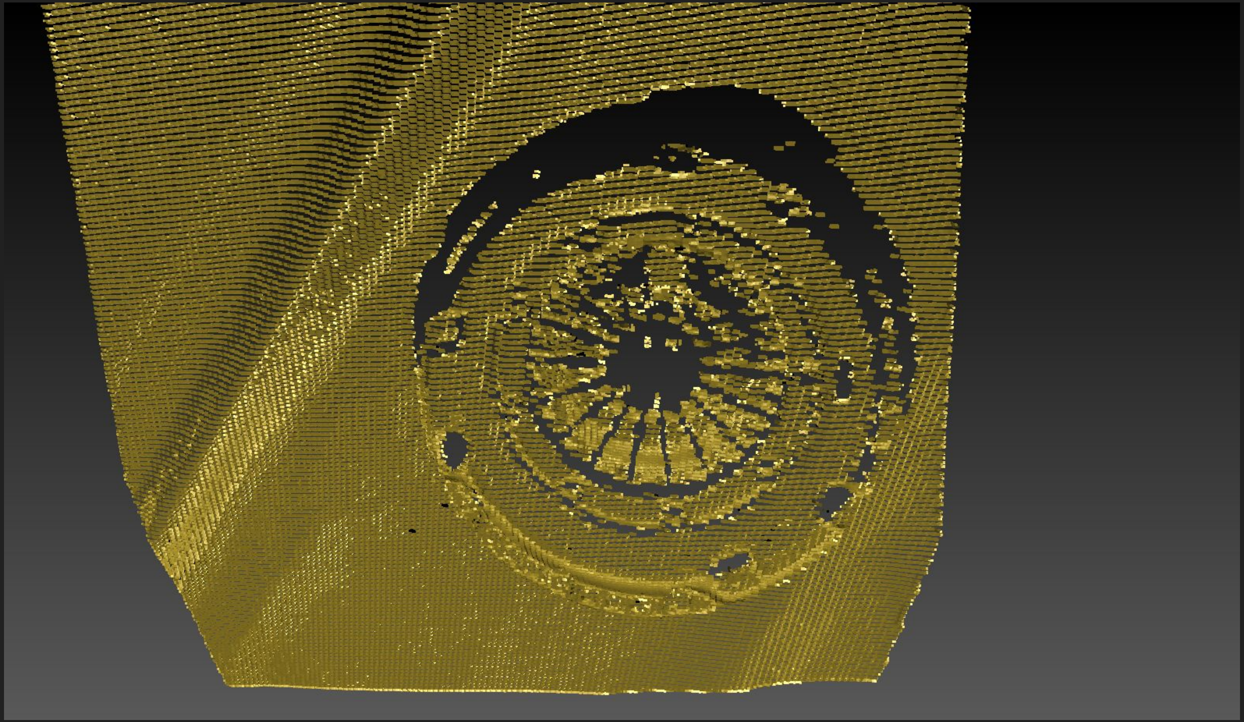
Downsampling vs. Upsampling time

Operation	FDSR Upsampling	DKN Upsampling	Downsampling
Time [s]	0.007	0.634	1.68

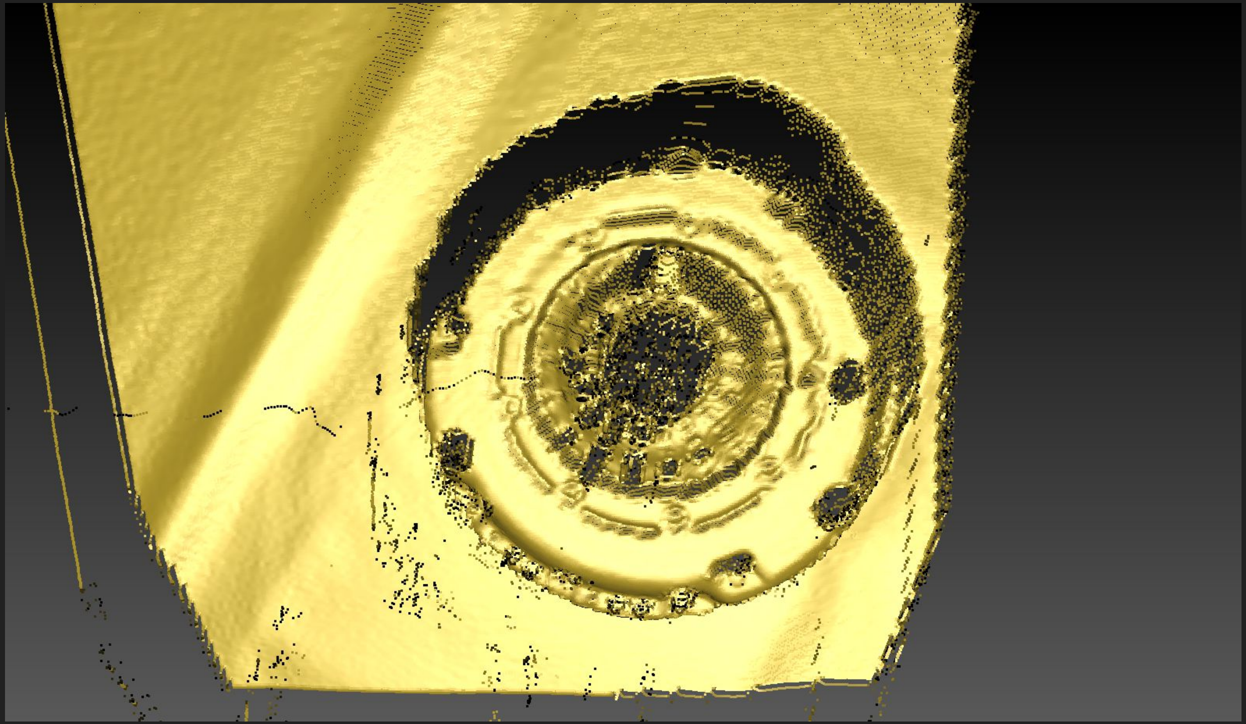
HR



LR - NEAREST NEIGHBOR

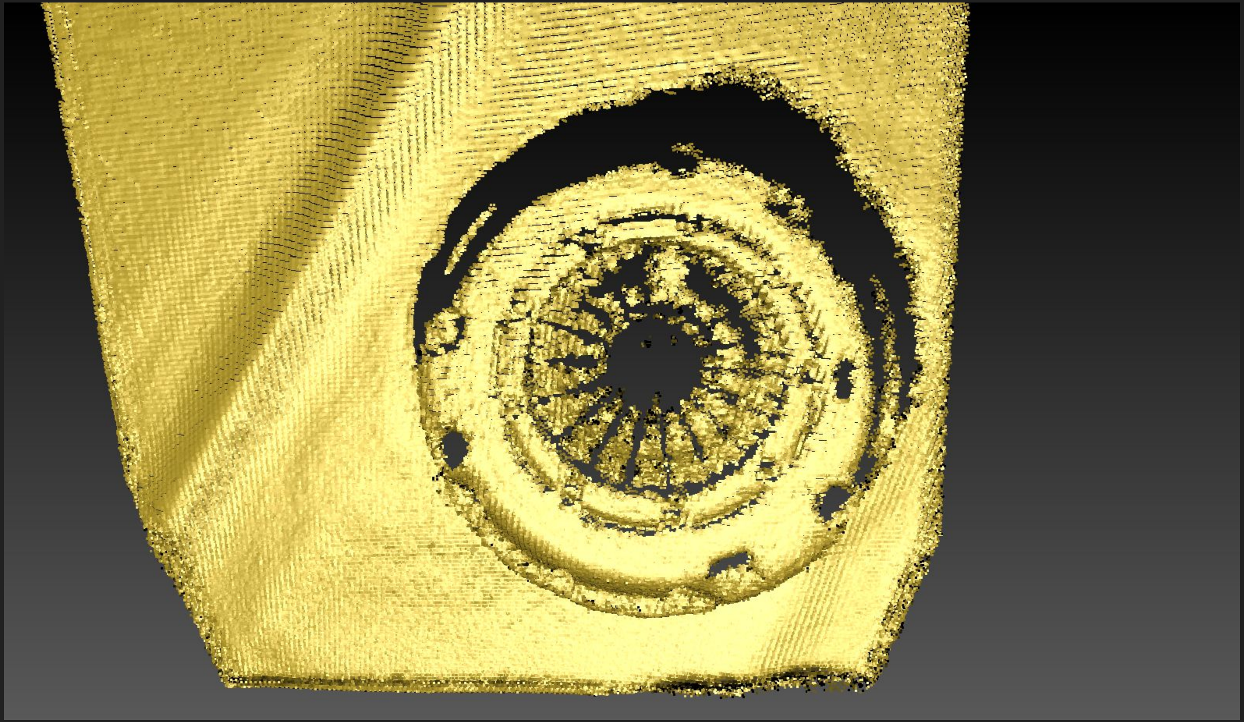


ImageMagic resize



DKN

FDSR



Sources

- https://www.photoneo.com/products/phoxi-scan-m/
- https://www.photoneo.com/3d-model-creation/
- https://www.bricsys.com/blog/point-clouds-whats-the-point
- https://github.com/jun0kim/DKN
- https://www.meshlab.net/
- https://arxiv.org/pdf/2104.06174v1.pdf