

# Koneoppimiseen perustuva kuvapohjainen kolmiulotteinen rekonstruktio monesta näkymästä (aihe-esittely)

*Juho-Matti Kilpeläinen*

*11.03.2025*

*Ohjaaja: Juho Kannala*

*Valvoja: Ahti Salo*

Työn saa tallentaa ja julkistaa Aalto-yliopiston avoimilla verkkosivuilla. Muilta osin kaikki oikeudet pidätetään.

# Tausta

- 3D-rekonstruktion ideana on muodostaa sensori- tai kuvadatan pohjalta 3D-malli jostain esineestä tai alueesta
- Ennen koneoppimista parhaat tekniikat hyödynsivät pitkälti pistepilviä ja perinteisiä meshejä tai voxeleita
- NeRF-tekniikka yksi ensimmäisistä fotorealistisista koneoppimismalleista, mutta koulutus kesti kymmeniä tunteja ja malli voitiin renderöidä noin 1 FPS (Mildenhall et al., 2020)
- Gaussian Splatting –niminen tekniikka pudotti koulutusajan murto-osaan ja nosti renderöinnin 100 FPS tasolle (Kerbl et al., 2023).

# Tavoitteet

- Työn tavoitteena on kattava kirjallisuuskatsaus koneoppimiseen perustuvasta 3D-rekonstruktioista
- Työn kokeellinen osuus keskittyy NoPoSplat –malliin ja sen laajentamiseen uusille dataseteille

# Rajaukset

- Työssä keskitytään oppimispohjaisiin menetelmiin vrt. perinteisemmät algoritmit tai SfM pipeline osat

# Konkretia ja aikataulu

- Työn kokeellinen osuus rakentuu NoPoSplat-mallin päälle (*Ye et al., 2024*)
- Alkuperäinen malli on benchmarkattu vain kahdella datasetillä → hankala verrata muihin malleihin
- Olen jo ohjelmoinut tuen ETH3D-datasetille
- Huhtikuun ja toukokuun aikana tuki myös muutamalle muulle datasetille
- Uusilla dataseiteillä sekä syvyys että rekonstruktio – benchmarkkeja
- Noposplat ei hyödynnä syvyysdataa koulutuksessa -> syvyys benchmarkit mielenkiintoisia

# Työkalut

- Noposplat-malli, jolla benchmarkkaukset suoritetaan non koulutettu Aallon Triton-klusterilla RE10k-datasetillä
- Evaluointiin käytetään RTX 4090 –näytönohjainta
- Keskeisiä ohjelmistoja ja kirjastoja muun muassa Pytorch Lightning, Hydra, OpenCV, Wandb sekä lukuisat muut kirjastot pienemmässä osassa

# Tietolähteet/Aineistot

- Mildenhall, Ben, Pratul P. Srinivasan, Matthew Tancik, Jonathan T. Barron, Ravi Ramamoorthi, and Ren Ng. 'NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis'. arXiv, 3 August 2020. <https://doi.org/10.48550/arXiv.2003.08934>.
- Kerbl, Bernhard, Georgios Kopanas, Thomas Leimkuehler, and George Drettakis. '3D Gaussian Splatting for Real-Time Radiance Field Rendering'. ACM Transactions on Graphics 42, no. 4 (August 2023): 1–14. <https://doi.org/10.1145/3592433>.
- Ye, Botao, Sifei Liu, Haofei Xu, Xueting Li, Marc Pollefeys, Ming-Hsuan Yang, and Songyou Peng. 'No Pose, No Problem: Surprisingly Simple 3D Gaussian Splats from Sparse Unposed Images'. arXiv, 31 October 2024. <https://doi.org/10.48550/arXiv.2410.24207>.