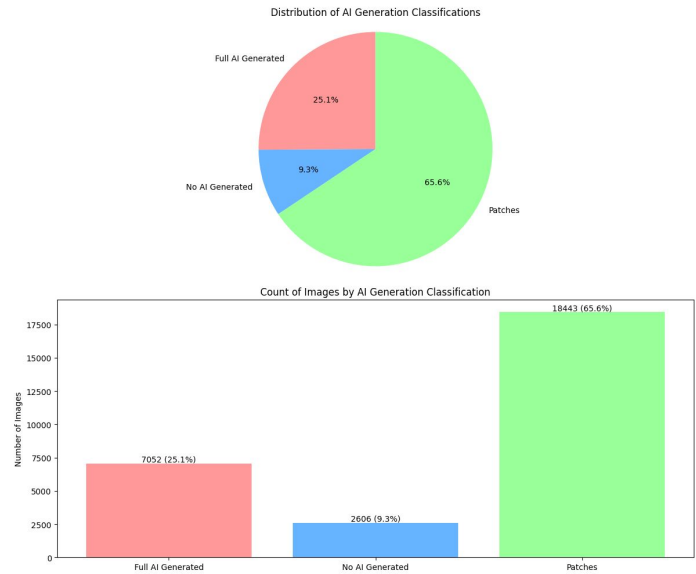
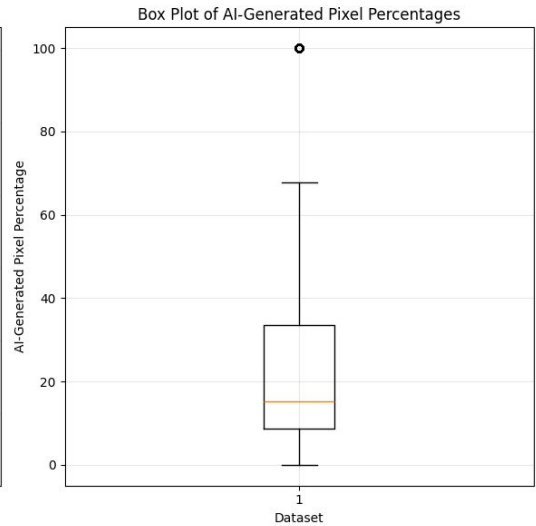
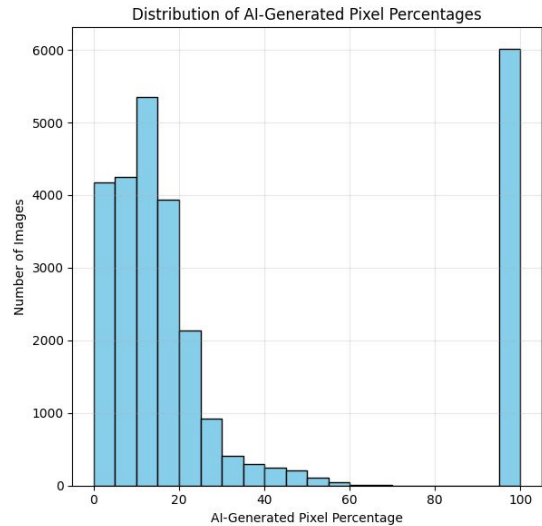


# Computer vision hackathon Aaltoes 2025

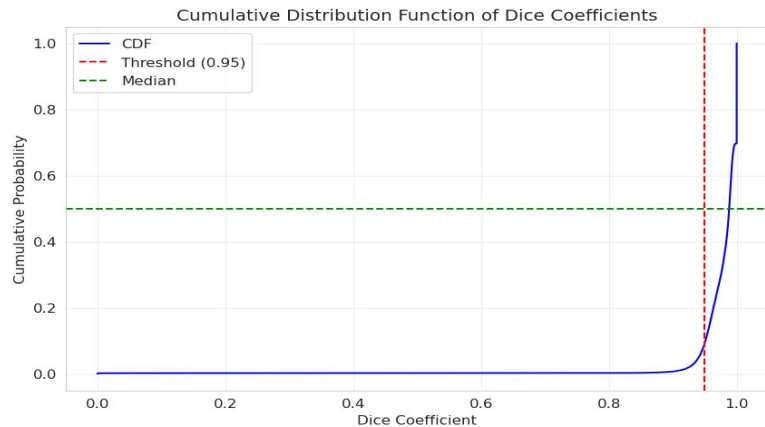
Edvard Ohlström & Jan Nyberg

# Exploratory data analysis



- Exploratory data analysis
  - understanding the problem
- Split dataset ->
  - full dataset, 10%, 5%, 1% and (full dataset + originals with mask)
- Research ->
  - requirements (code, env, weights, metrics) are available
- Two approaches
  - Baseline pytorch\_segmentation\_model and TruFor\* architecture
    - Based on Noiseprint++ and mit\_b5
- Made own training code
  - -> changed to larger pretrained encoder (mit\_b2 -> mit\_b5)
  - -> Base training: Dice + BCE Finetuning: Dice + Focal
  - -> from pretrained weight -> 0.89 to our training -> 0.96
- Fault analysis

\* <https://grip-unina.github.io/TruFor/>



## Fault analysis on our own validation split

Samples below threshold (0.95):

2597 (9.24%)

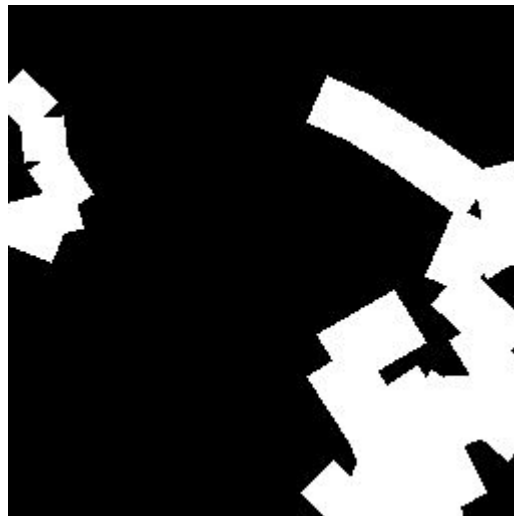
Samples above threshold (0.95):

25504 (90.76%)

Conclusion: Model fails to fit correct masks, AI/non-AI was not a problem



Ground truth ->  
<- Predicted



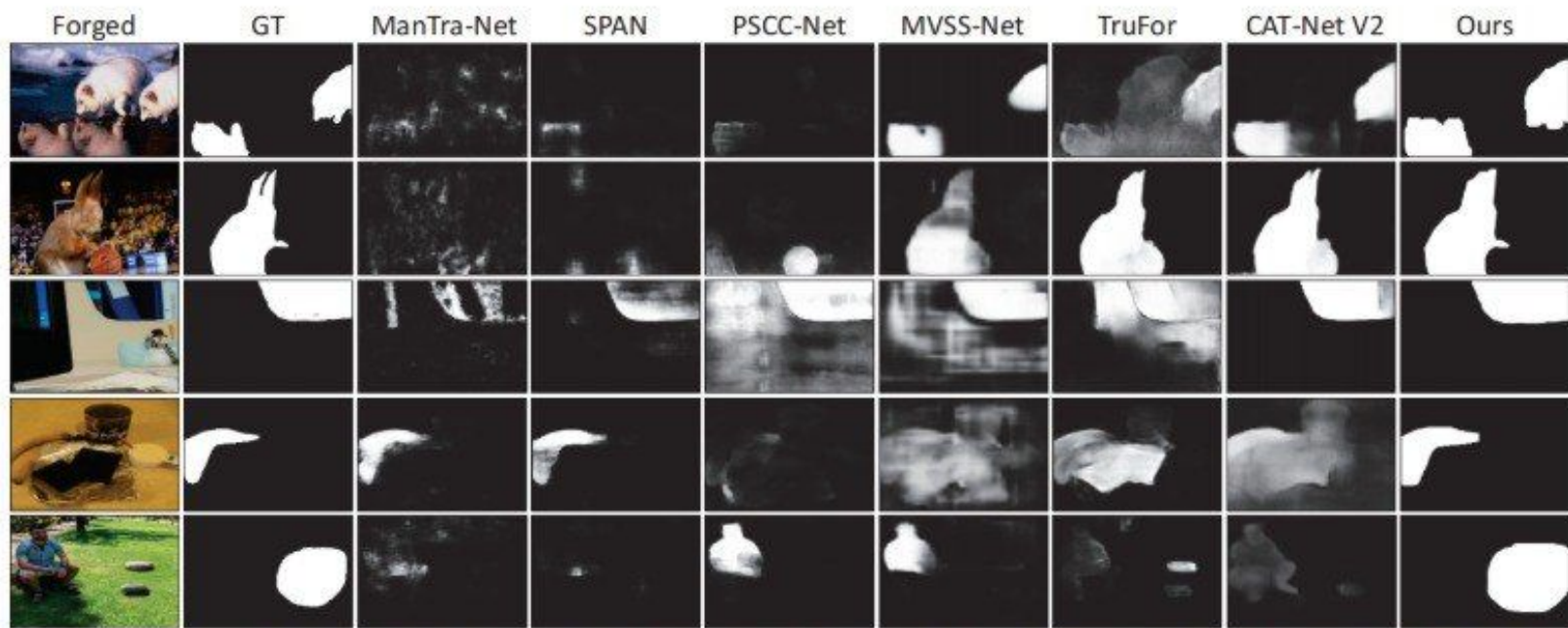


Figure 4. Qualitative comparison results. The first to fourth rows are respectively sourced from CASIA v1 [14], Columbia [26], Coverage [61], and IMD20 [46]. The last row is from the BDNIE dataset.

Runs (8)

&gt;&gt;

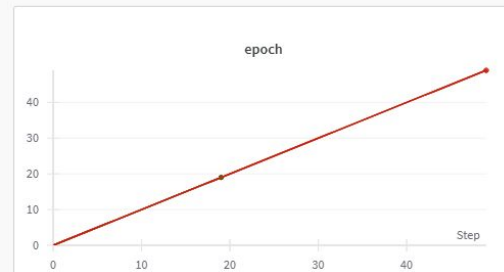
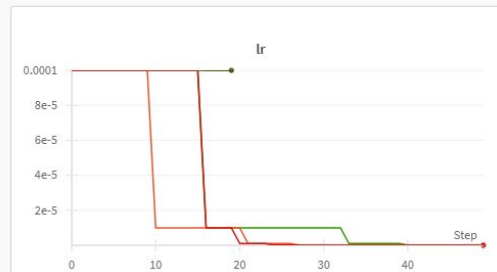
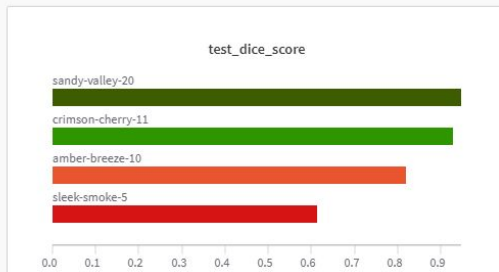
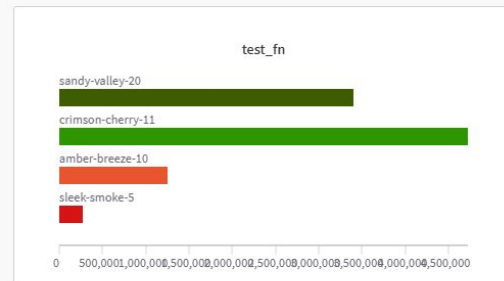
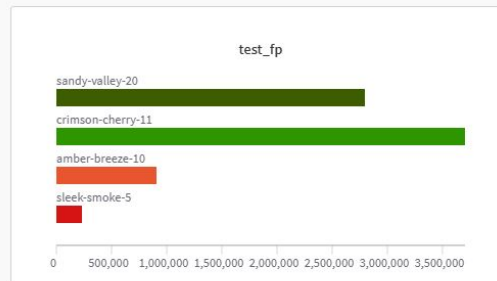
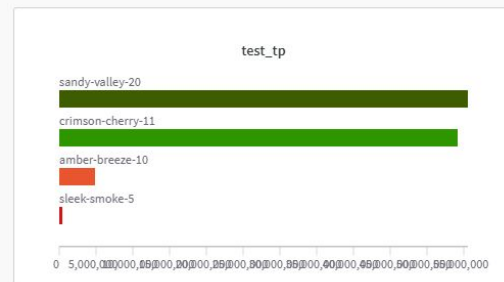
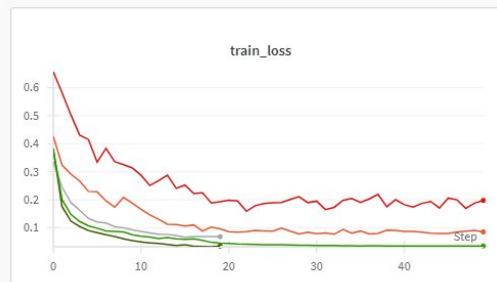
Search runs

.\*



Name (6 visualized)

- sandy-valley-20
- stilted-resonance-16
- ruby-water-13
- crimson-cherry-11
- amber-breeze-10
- smooth-planet-7
- sleek-smoke-5
- daily-blaze-3



Runs (22)



Search runs



Name (3 visualized)

- lunar-voice-35
- jumping-serenity-34
- clean-oath-33
- fluent-galaxy-30
- royal-terrain-29
- divine-cloud-26
- woven-terrain-25
- grateful-butterfly-24
- different-sun-19
- neat-snowflake-16
- gentle-yogurt-7
- smart-valley-6
- royal-grass-5
- fanciful-glade-4

Search panels with regex



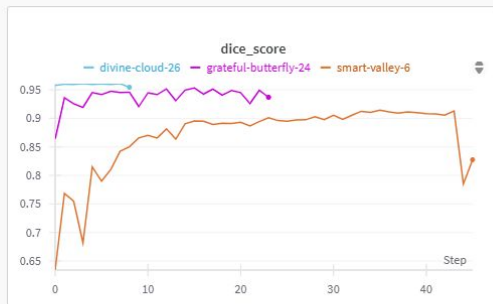
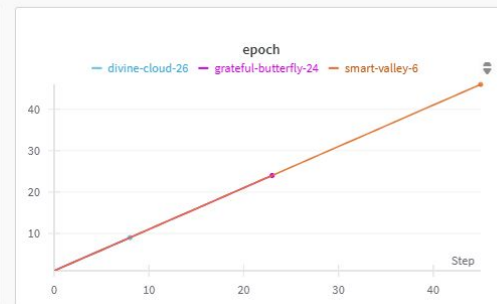
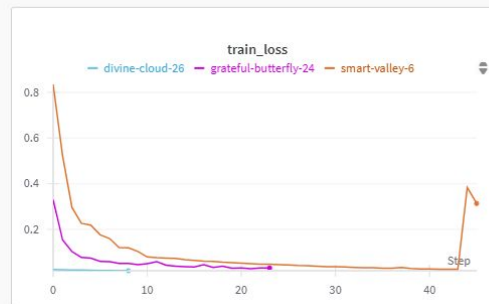
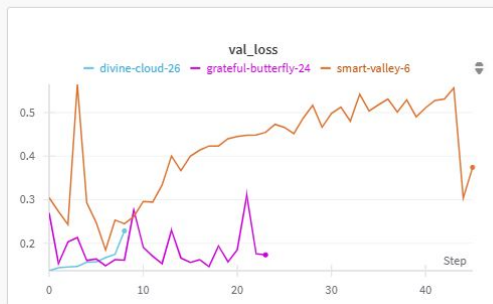
Settings

+ New report

+ Add panels

Charts 4

1-4 of 4



System 22

# Appendix

Transformer-based fusion architecture that combines the RGB image and a learned noise-sensitive fingerprint

Train/val = 80/20

Batch size = 32/64

Epochs = 50

Learning rate  $1e-4$

Did not work or improve performance

- Dice alone
- Weight decay
- Tried cosine scheduler and 1 cycle scheduler
- Threshold 0.5 , changing this made results worse
- Ensemble with different size model or different training did not improve model health the same



# Architecture

