Generating images of thermal cats and dogs

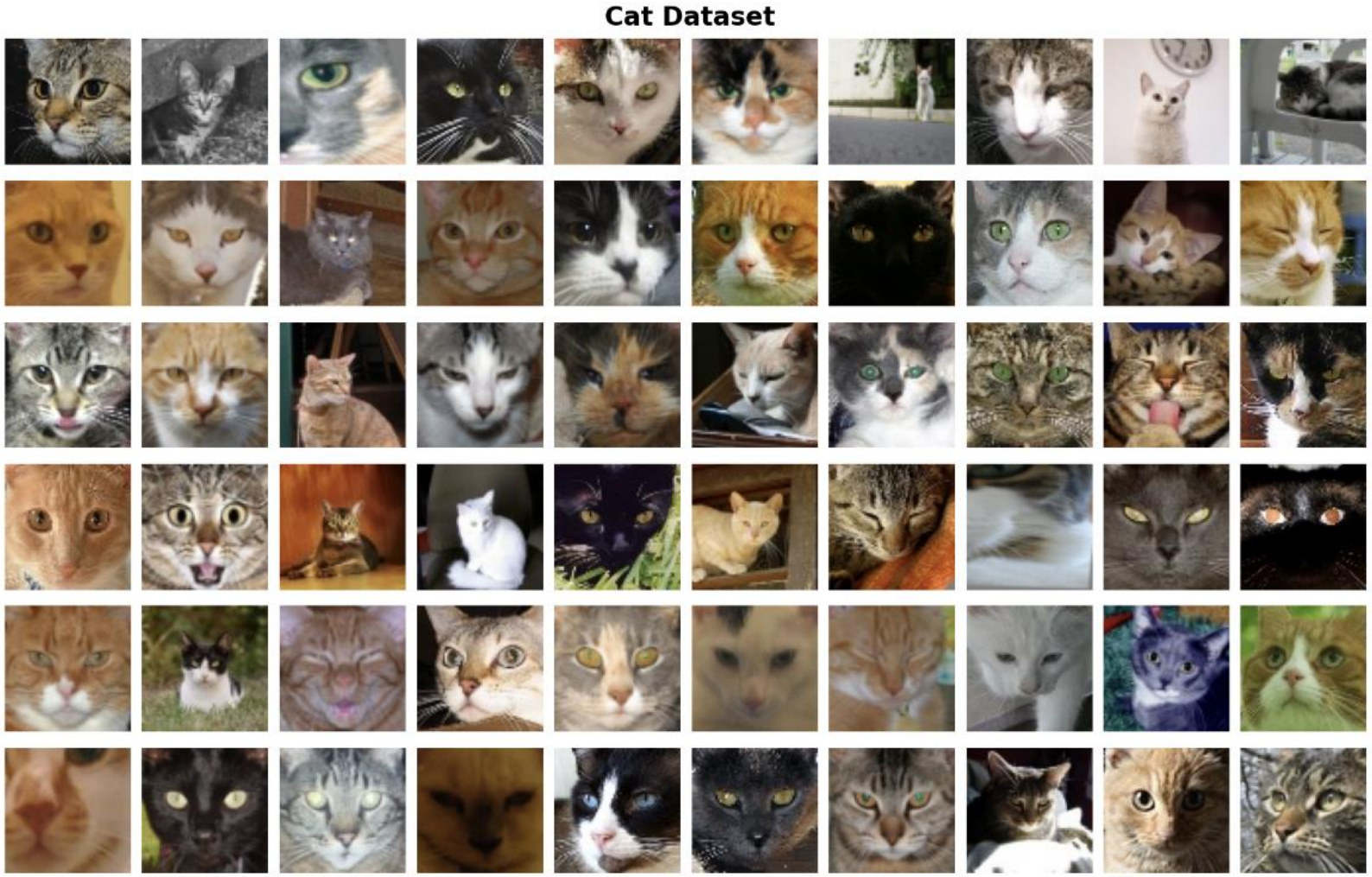
PAULINA KULCZYK, JAN POGŁÓD

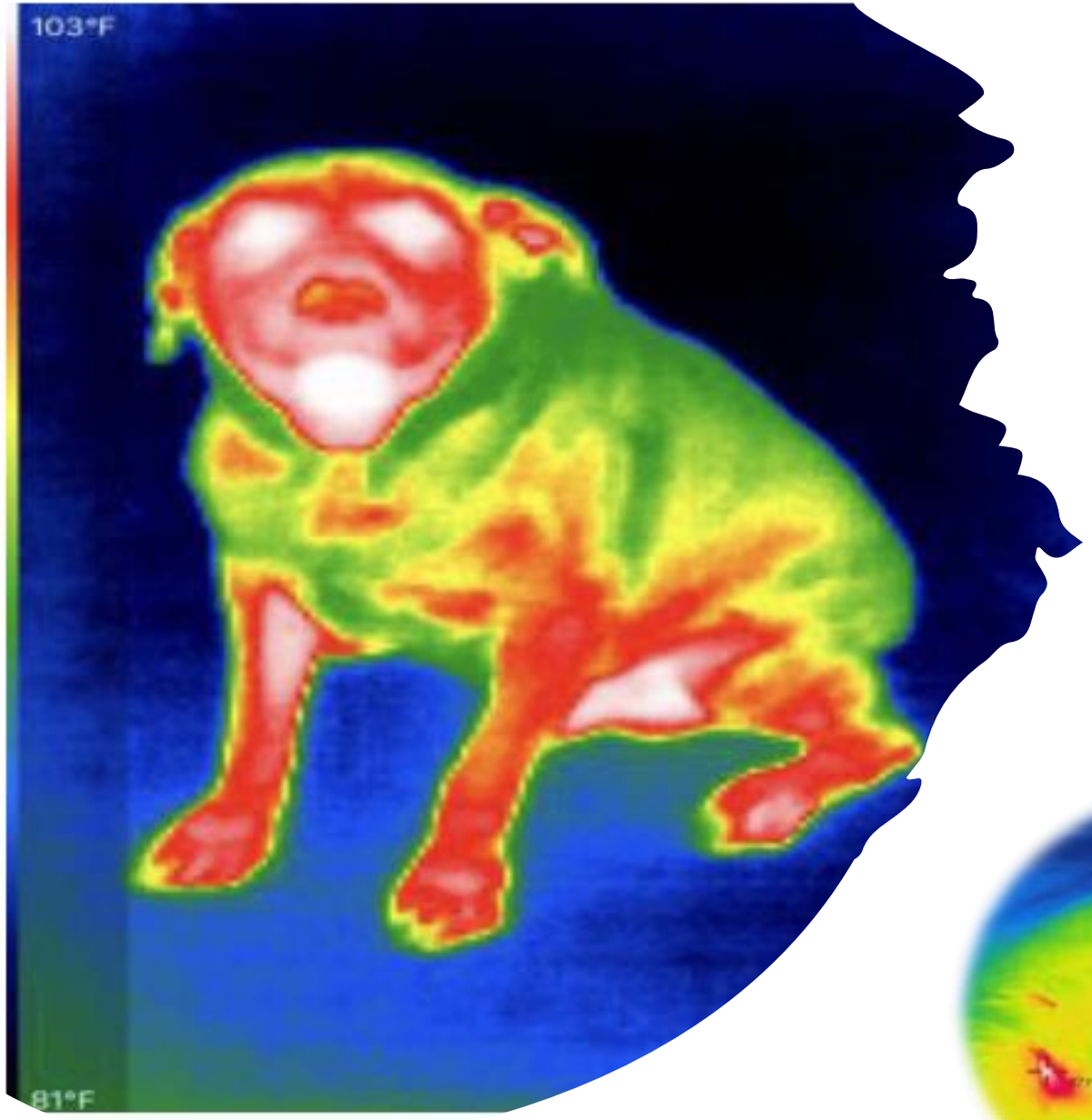
Agenda

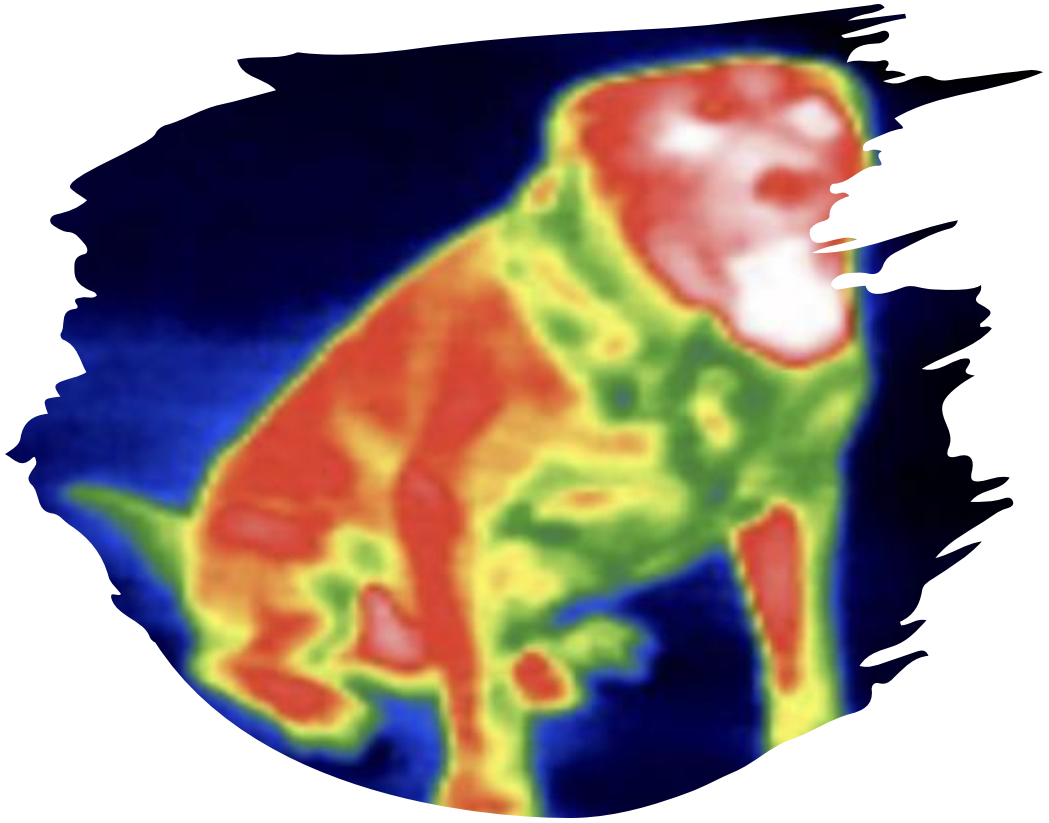
- Dataset
- Augmentation techniques
- Implementations
- Results

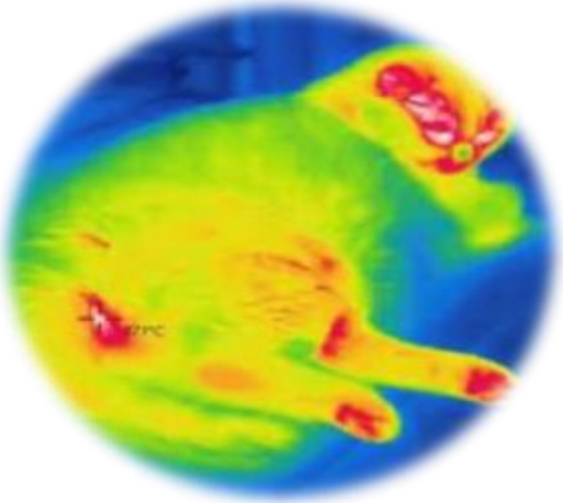
Dataset

- 29,843 RGB images
- 64x64 pixels





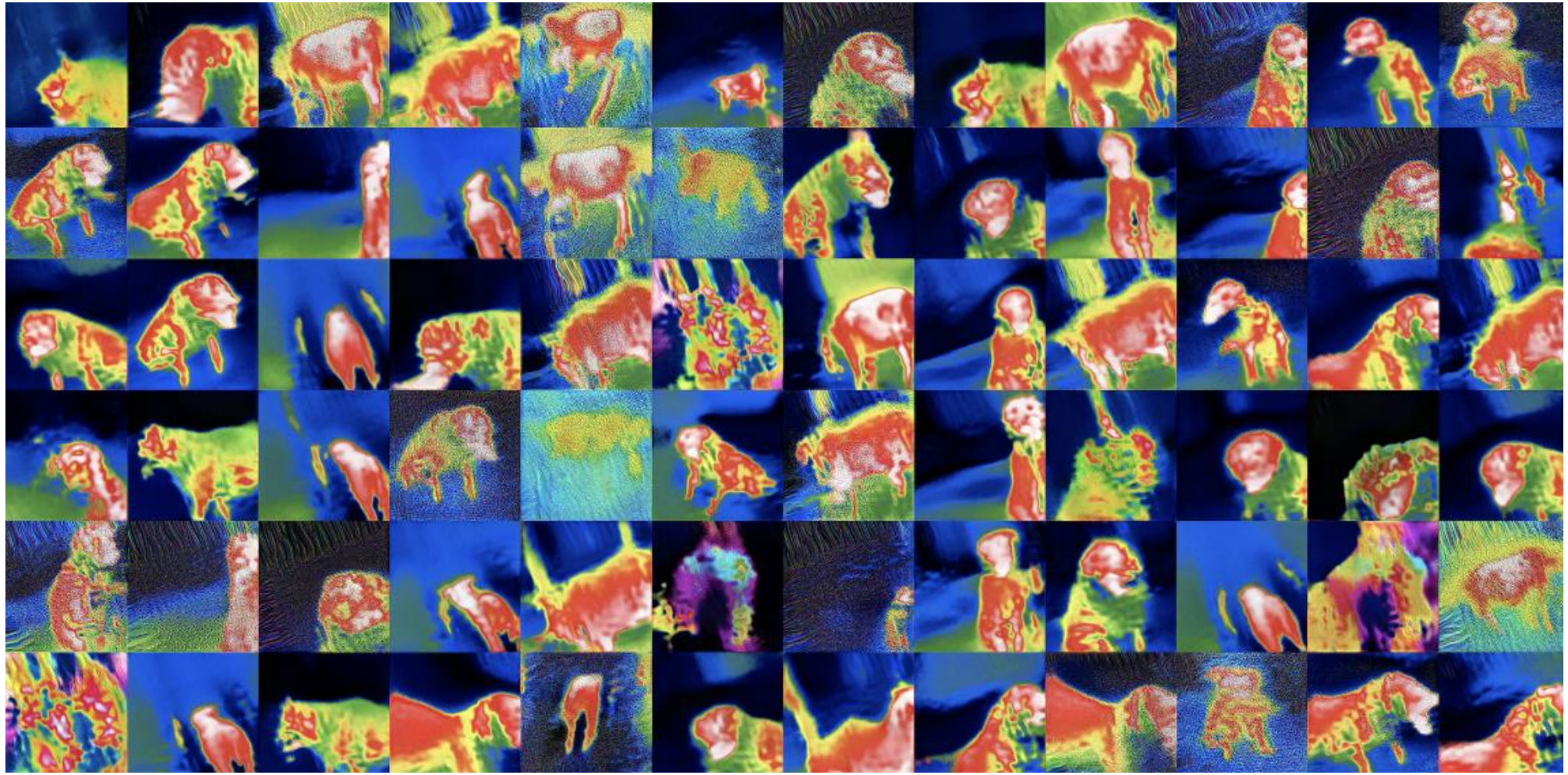




Dataset

- About 200 images
- Mostly 512x512 pixels

Augmentation

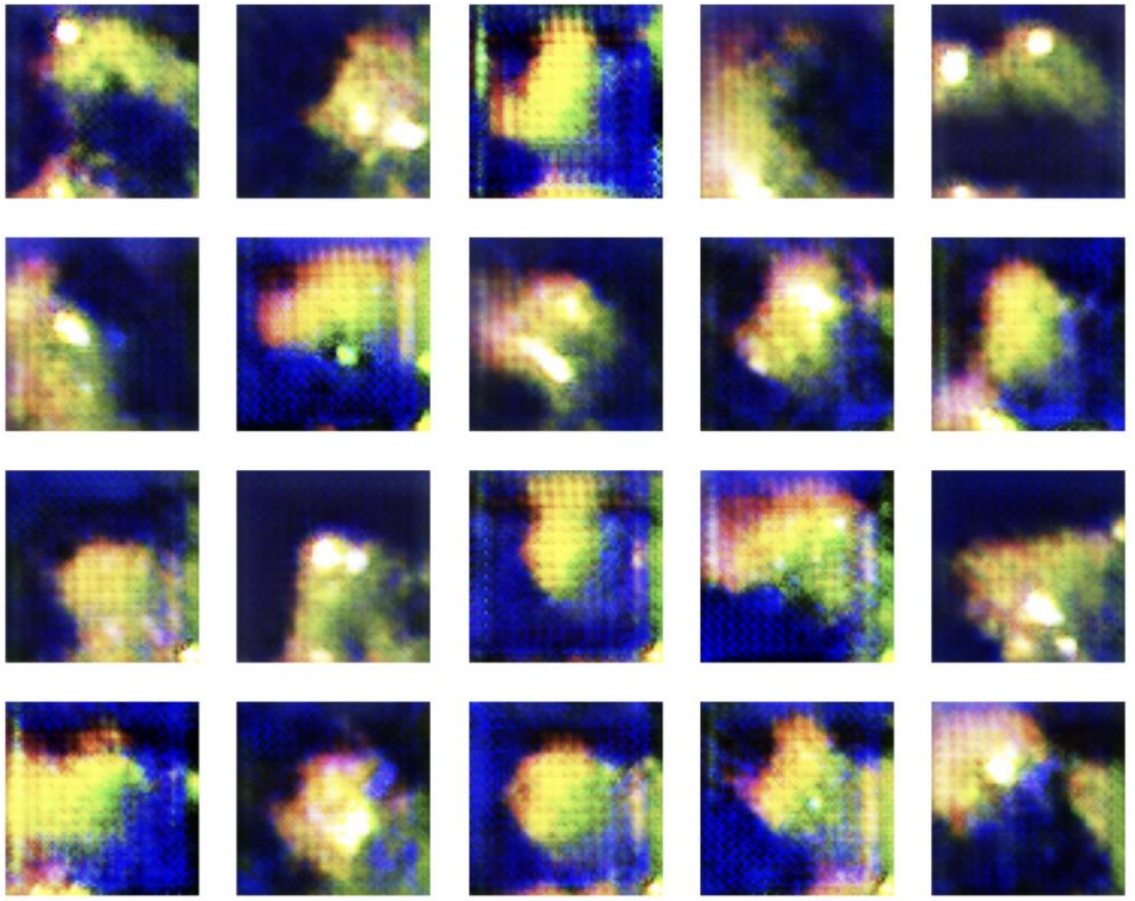


Implementations

- Cycle-Gan
- Difussion-Based model
- Style-Gan
- Style-Gan2-ADA
- For all models we changed photos to size to maintain the same size (128x128 or 256x256 depends on model)

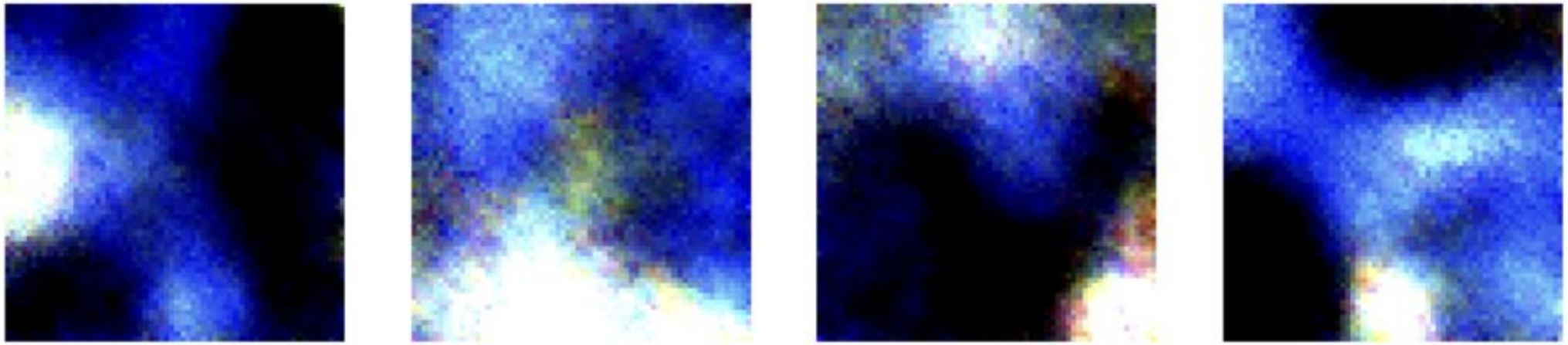
Implementations – Cycle-Gan

- Parameters:
  - Adam optimizers,
  - Learning rate of 2×10−4
  - Loss function : MSE



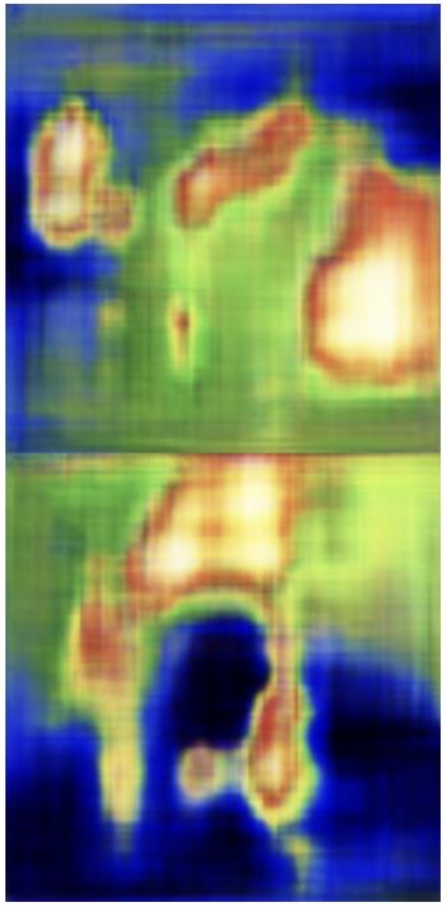
Implementations – Difussion-based

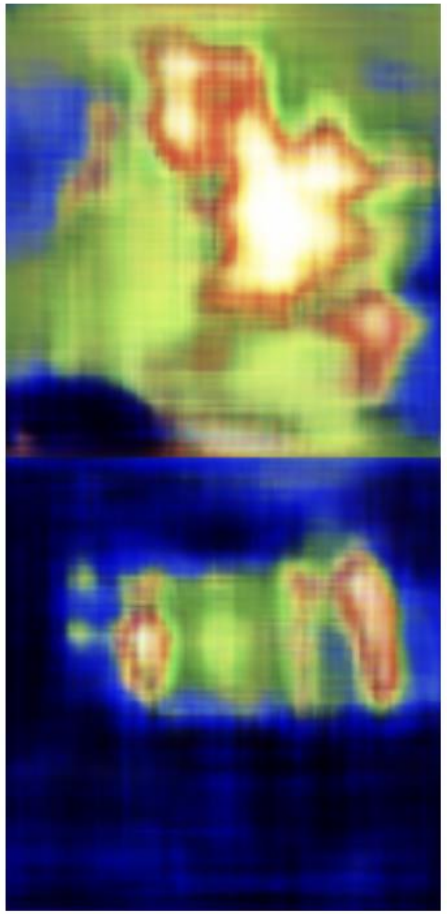
- Parameters:
  - 50 epochs using an Adam optimizer
  - learning rate of 1 × 10−4
  - Mean Squared Error (MSE) as the loss function.



Implementations – Style-Gan

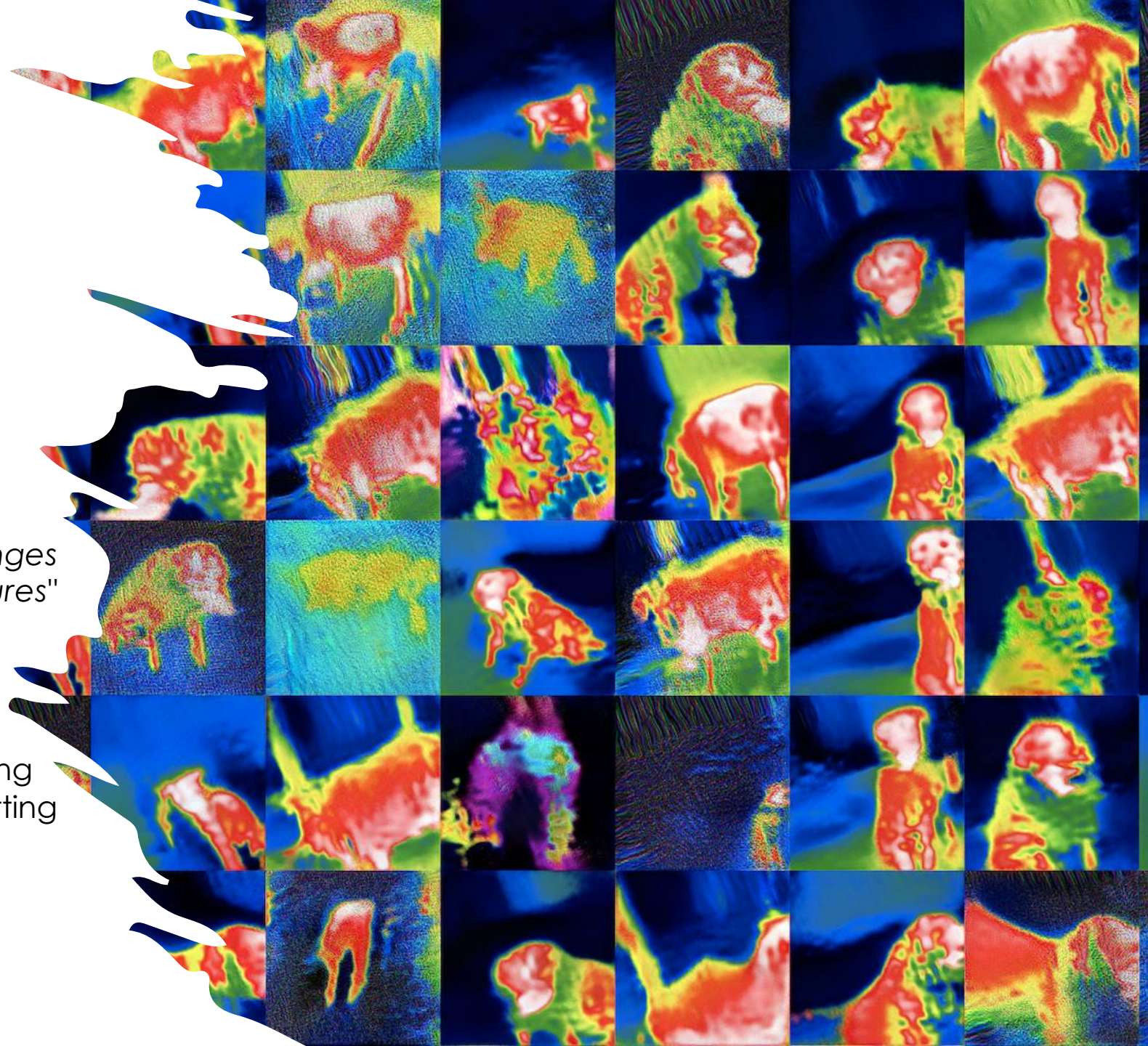
- Parameters:
  - 100 epochs,
  - batch size of 8,
  - Adam optimizers,
  - learning rate 2 × 10−4
  - Loss fun: binary cross-entropy loss





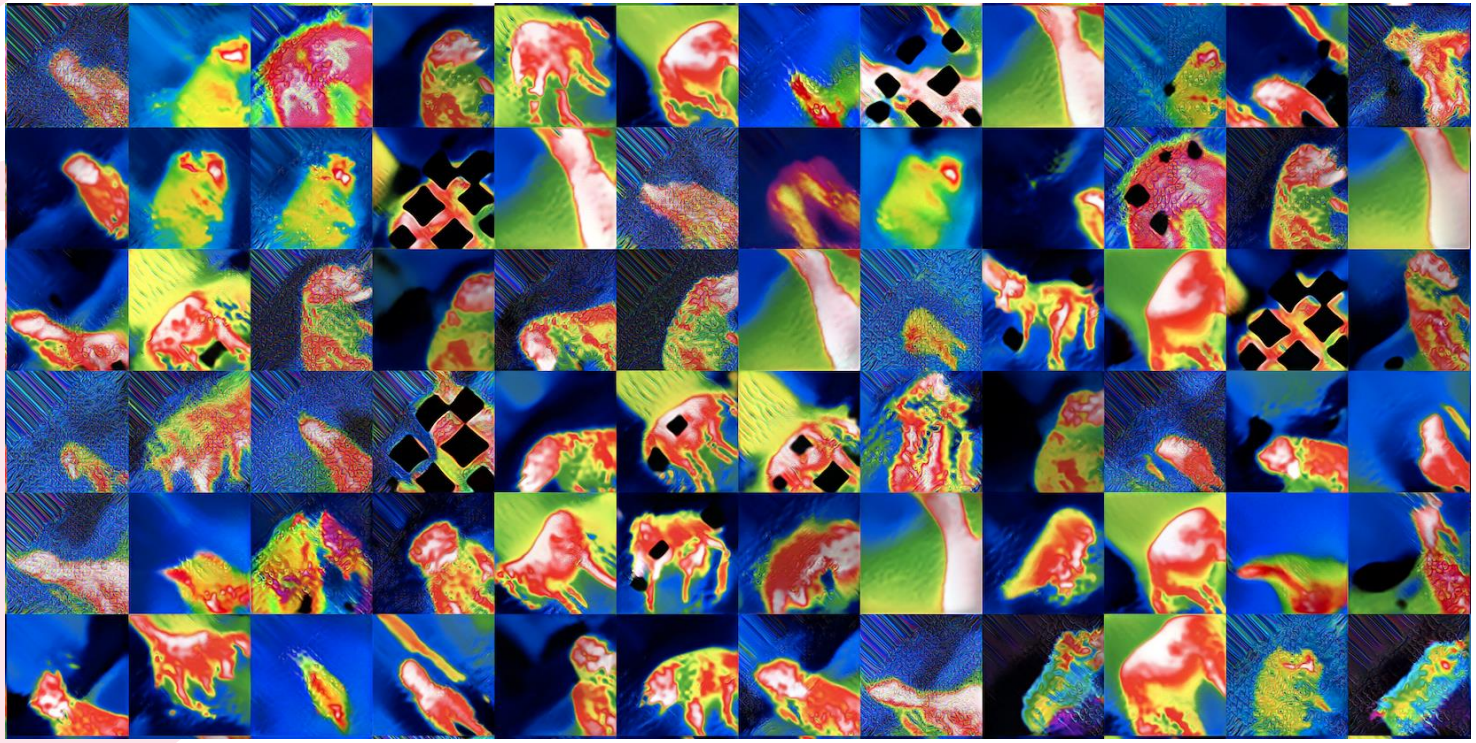
Implementations – StyleGan2-Ada

- Parameters:
  - 512-dimensional latent space,
  - 1 000 000 images
  - batch size of 16,
  - Adam optimizer
- "*The approach does not require changes to loss functions or network architectures*"
  - non-saturating logistic loss with R1 regularization
- dynamically adjust augmentation strength on the discriminator, improving training stability and mitigating overfitting on limited dataset.



StyleGan2-Ada on more aggressive augmentation

- larger and more frequent square occlusions,
- increased rotation angles,
- greater scaling variations

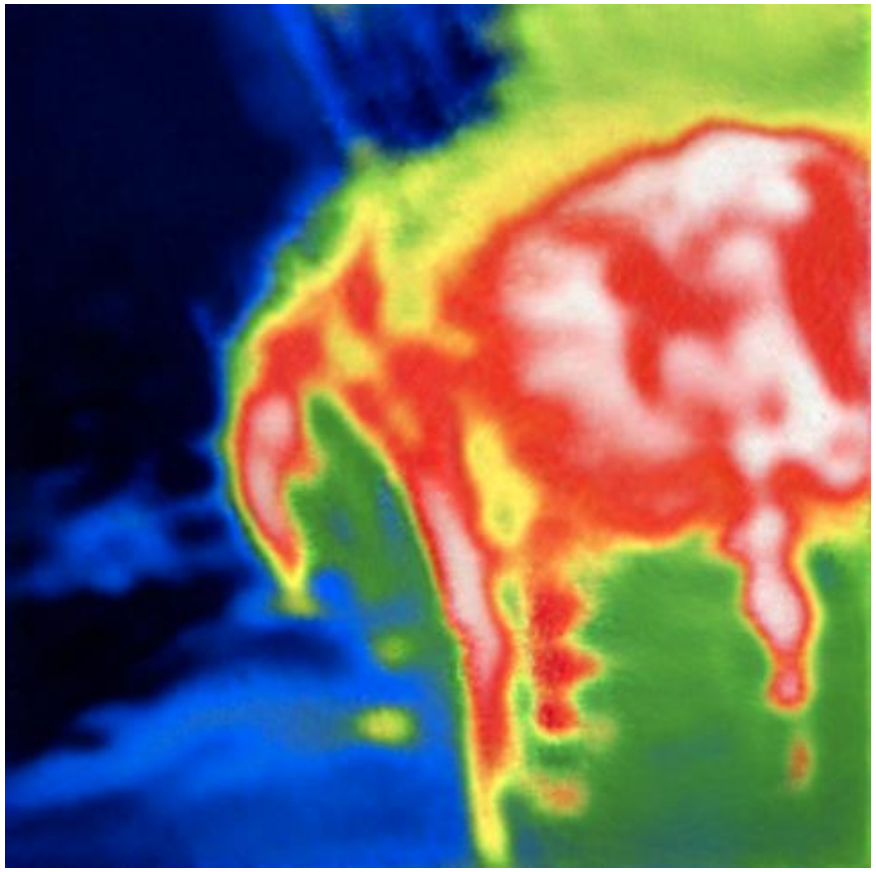


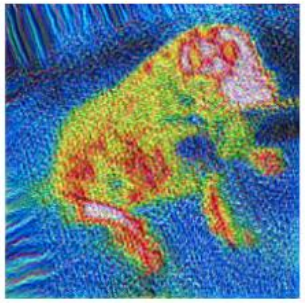
StyleGan2-Ada on RGB

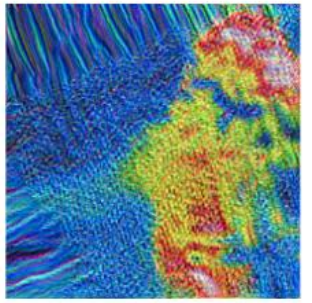
Style-mixing using StyleGan2-Ada models

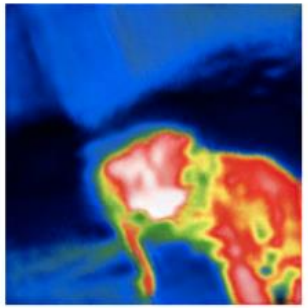
Variant	Description	Mixing ratio control	Synthesizer	Outputs
Single Seed	One shared latent vector (z)	Blending done by proportional layer mixing (e.g., 50% RGB, 50% Thermal)	Thermal Generator	1 Mixed Image
Independent Seeds	Two random latent seeds	Early layers from RGB, late from Thermal using hard cutoff (after layer 6).	Thermal Generator	RGB, Thermal, Mixed
Multi-Seed Shared	Same seed used in both models	Early structure layers from RGB, later ones from Thermal (after layer 4)	Thermal Generator	RGB, Thermal, Mixed

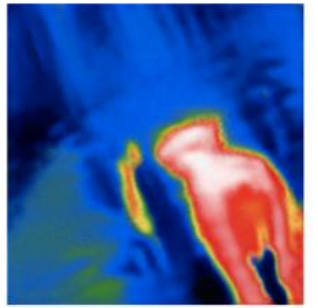
Results of style-mixing











Comparision of the best models – FID score

Model	FID Score
CycleGAN	350.37
StyleGAN	300
StyleGAN2-ADA	**101.37**
StyleGAN2-ADA (RGB photos only)	**8.12**


Thank you very much!
WE WILL BE WORKING FOR
BETTER THERMLA CATS...