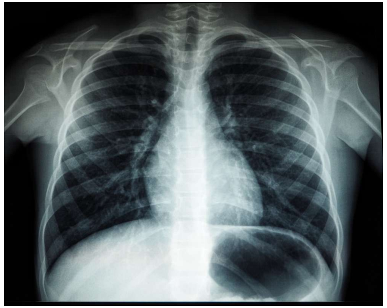
# Synthetic Image Generation With a Fine-Tuned Latent Diffusion Model For Organ on Chip Cell Image Classification

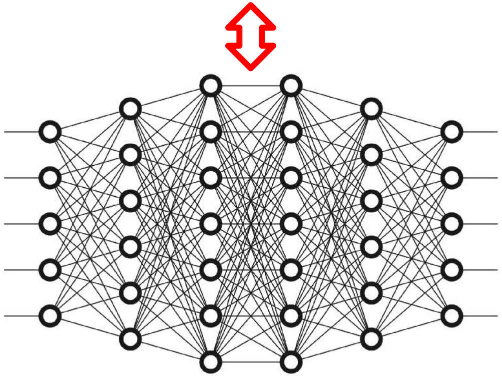
Maksims Ivanovs[1], Laura Leja[1], Kārlis Zviedris[1], Roberts Rimsa[2], Karina Narbute[3], Valerija Movcana[3], Felikss Rumnieks[3], Arnis Strods[3], Kevin Gillois[3], Gatis Mozolevskis[2], Arturs Abols[2,3], Roberts Kadiķis[1]

[1] Institute of Electronics and Computer Science (EDI), Riga, Latvia
[2] Cellbox Labs, Riga, Latvia
[3] Latvian Biomedical Research and Study Centre, Riga, Latvia

# Introduction: DNN for Biomedical Image Analysis

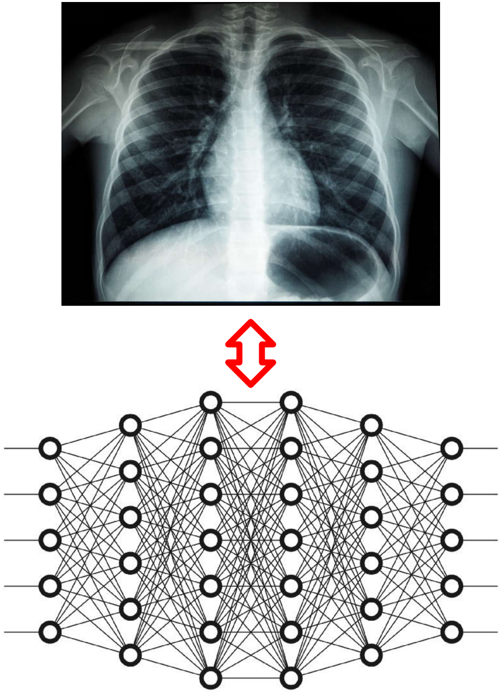
- AI is increasingly used in biomedical research
- State-of-the-art AI methods for image analysis: DNNs
- Sample applications:
  - segmentation of endoscopy images
  - segmentation of nuclei microscopy images
  - classification of X-ray images
- Advantages of DNN:
  - high accuracy
  - learn from the data -> no need for feature engineering
- Disadvantages of DNN:
  - require a lot of computing power
  - ‘black boxes’
  - need a lot of data for training





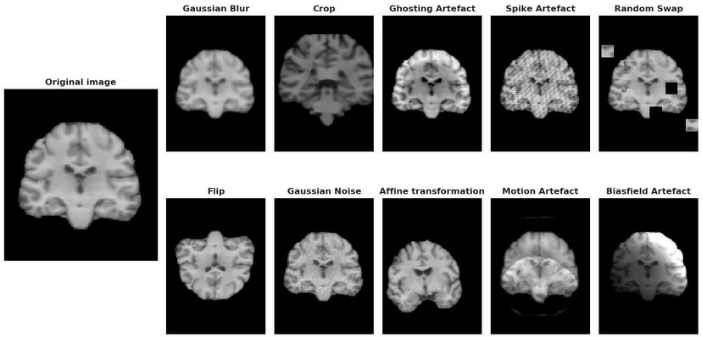
# Introduction: DNN for Biomedical Image Analysis

- AI is increasingly used in biomedical research
- State-of-the-art AI methods for image analysis: DNNs
- Sample applications:
  - segmentation of endoscopy images
  - segmentation of nuclei microscopy images
  - classification of X-ray images
- Advantages of DNN:
  - high accuracy
  - learn from the data -> no need for feature engineering
- Disadvantages of DNN:
  - require a lot of computing power
  - ‘black boxes’
  - **need a lot of data for training**
    **NB! biomedical datasets tend to be small**

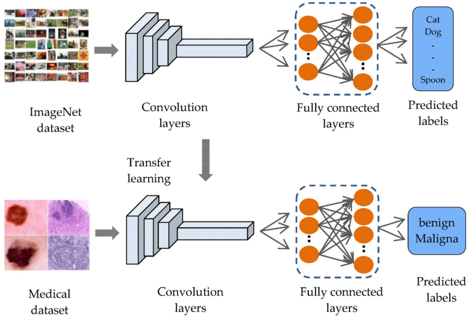


# Data (un)availability problem: possible solutions

- How to deal with the lack of available data for training DNN models?

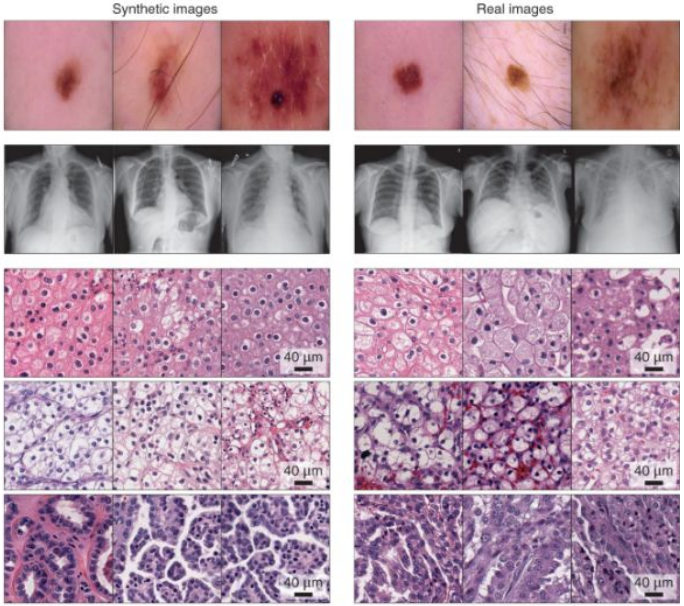



Solution #1: data augmentation
*figure: Dufumier et al, 2021*



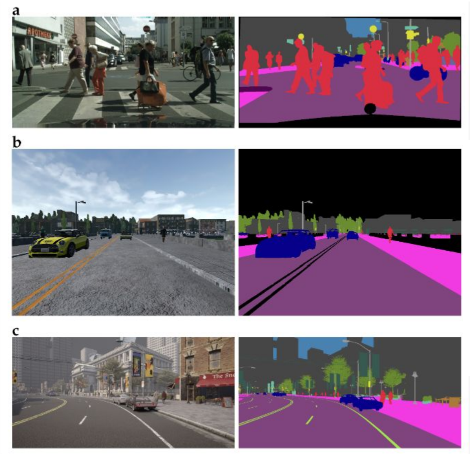

Solution #2: transfer learning
*figure: Mukhlif et al, 2023*

Solution #3: synthetic data
*figure: Chen et al, 2021*



# Synthetic data for training DNN

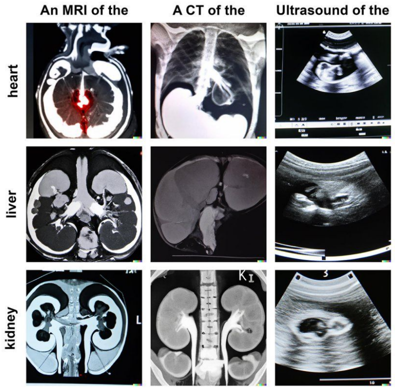
- Widely used approach in different domains:
  - easy to modify output
  - as soon as generative pipeline is created, it's easy to generate a lot of data
- Main challenges:
  - creating the generative pipeline
  - synthetic ↔real gap
- State-of-the-art large generative models:
  - Midjourney
  - DALL-E
  - Stable Diffusion
- Can we use such models for generating biomedical images?
  - Biomedical data are too specific for general-purpose models
  - Retraining from ground up is costly and time-consuming
  - Possible solution: **fine-tuning**



*Ivanovs et al, 2022*


(a) Real image (b) Synthetic image

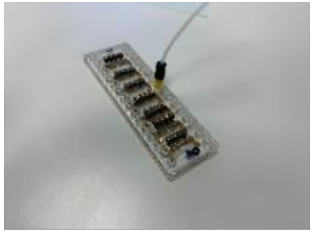
*Duplevska et al, 2022*

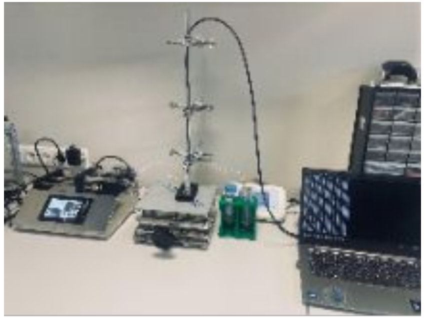


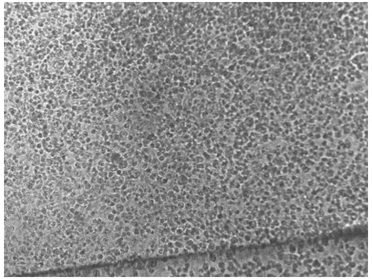
*Adams et al, 2023*

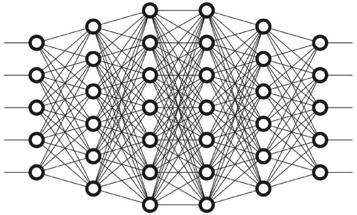
# Goals and hypotheses

- Goal: to design a classifier for Organ on Chip (OOC) microscopy images







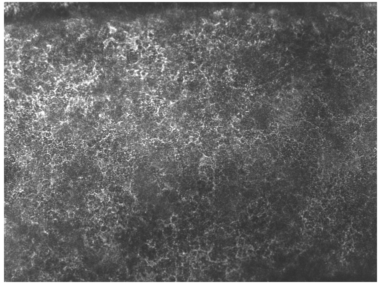


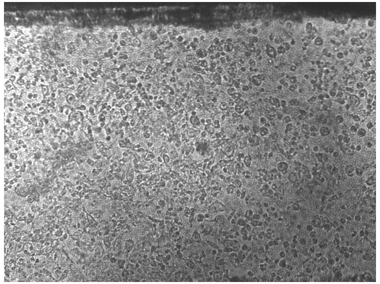
good
acceptable
bad

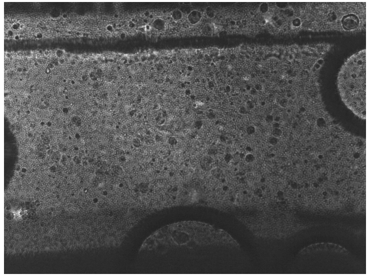
- Hypotheses:
  - *H1:* by by using DNN, better-than-naive accuracy can be achieved
  - *H2:* augmenting dataset with synthetic data will result in improved accuracy

# Methodology: dataset of microscopy images

- 822 JPG images:
  - color:
    - 32 RGB
    - 790 grayscale
  - resolution:
    - 810 images: 2048x1536 pixels
    - 12 images: 640x480 pixels
  - 5 cell lines: HUVEC, HSAEC, A549, CACO, HPMEC
  - classes:
    - ‘good’: 500 images
    - ‘acceptable’: 212 images
    - ‘bad’: 110 images


‘good’

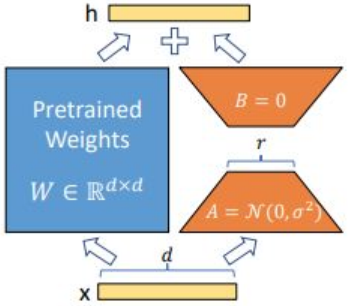

‘acceptable’


‘bad’

# Methodology: synthetic data generation

- Pipeline: Stable Diffusion (Rombach et al, 2022) fine-tuned with LoRA (Hu et al, 2021)
- LoRA parameters:
  - 2 repeats per image
  - 10 epochs
  - batch size = 2
  - U-Net learning rate = 5E-4
  - text encoder learning rate = 1E-4
- Stable Diffusion parameters:
  - Euler A sampler
  - 20 sampling steps
  - CGF score = 7
  - random seed
- Generated 2 datasets:
  - with LoRA weight = 1.0
  - with LoRA weight = 0.8

Higher weight -> higher similarity to original images




*Fine-tuning pretrained model with LoRA*
figure: Hu et al, 2021

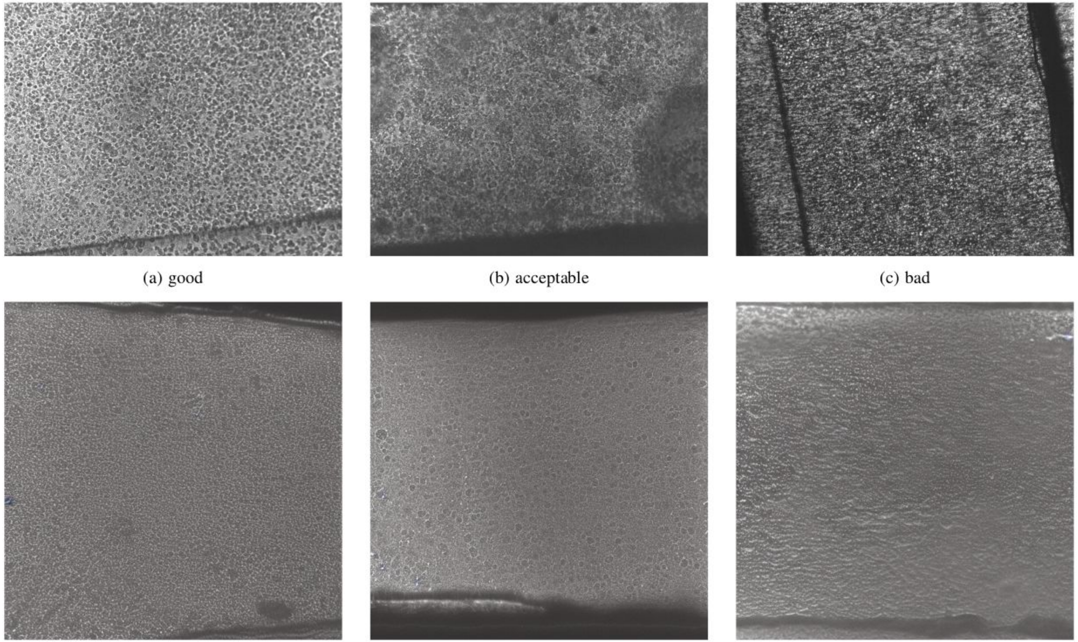
# Methodology: training and validating DNN

- DNN model:
  - EfficientNet B7 pretrained on imagenet
  - Architecture:
    - input layer: 600x600 pixels
    - data augmentation layers
    - basic EfficientNet B7 with frozen weights
    - GlobalAvgPooling2D layer
    - BatchNormalization layer (dropout=0.2)
    - Dense layer (3 neurons, softmax)
- Training:
  - 30 epochs
  - Adam optimizer (lr=0.001)
  - categorical crossentropy loss
- Dataset split: 5 folds with 1 different fold being a holdout fold for cross-validation each time

# Results: synthetic images of the cells

original

synthetic



(a) good (b) acceptable (c) bad

(d) good (e) acceptable (f) bad

# Results: performance of DNN classifier

TABLE I: Classification Results on Synthetic Data with LoRA weight=1.0. The best result for each metric is in bold.

| Dataset | Accuracy | Precision | Recall |
|---|---|---|---|
| Baseline: Real-world | **0.729** | **0.731** | **0.715** |
| Real-world & 10% synthetic | 0.721 | 0.729 | 0.701 |
| Real-world & 25% synthetic | 0.707 | 0.719 | 0.699 |
| Real-world & 50% synthetic | 0.710 | 0.720 | 0.697 |
| Real-world & 75% synthetic | 0.696 | 0.702 | 0.680 |
| Real-world & 100% synthetic | 0.699 | 0.707 | 0.687 |
| Synthetic only (100%) | 0.614 | 0.617 | 0.608 |

TABLE II: Classification Results on Synthetic Data with LoRA weight=0.8. The best result for each metric is in bold.

| Dataset | Accuracy | Precision | Recall |
|---|---|---|---|
| Baseline: Real-world | **0.729** | **0.731** | **0.715** |
| Real-world & 10% synthetic | 0.718 | 0.729 | 0.702 |
| Real-world & 25% synthetic | 0.704 | 0.709 | 0.69 |
| Real-world & 50% synthetic | 0.693 | 0.698 | 0.679 |
| Real-world & 75% synthetic | 0.707 | 0.718 | 0.696 |
| Real-world & 100% synthetic | 0.701 | 0.709 | 0.685 |
| Synthetic only (100%) | 0.621 | 0.63 | 0.595 |

# Summary, conclusions and future work

- Main findings:
  - *H1* was confirmed: best accuracy = 72.9% (naive classifier acc. = 60.8%)
  - Synthetic images look somewhat similar to the authentic ones
  - Yet *H2* was not confirmed: augmentation with synthetic data resulted in worse accuracy
- Further improvements of synthetic data generation methods are needed, e.g.:
  - search for optimal parameters for LoRA
  - generation of synthetic images corresponding to specific cell lines rather than generic images
  - img2img translation with Stable Diffusion

# Thank you for your attention!

This work was supported by the project ’AI-improved organ on chip cultivation for personalised medicine (AimOOC)’ (contract with Central Finance and Contracting Agency of Republic of Latvia no. 1.1.1.1/21/A/079; the project is cofinanced by REACT-EU funding for mitigating the consequences of the pandemic crisis).

NATIONAL DEVELOPMENT PLAN 2020

EUROPEAN UNION
European Regional Development Fund