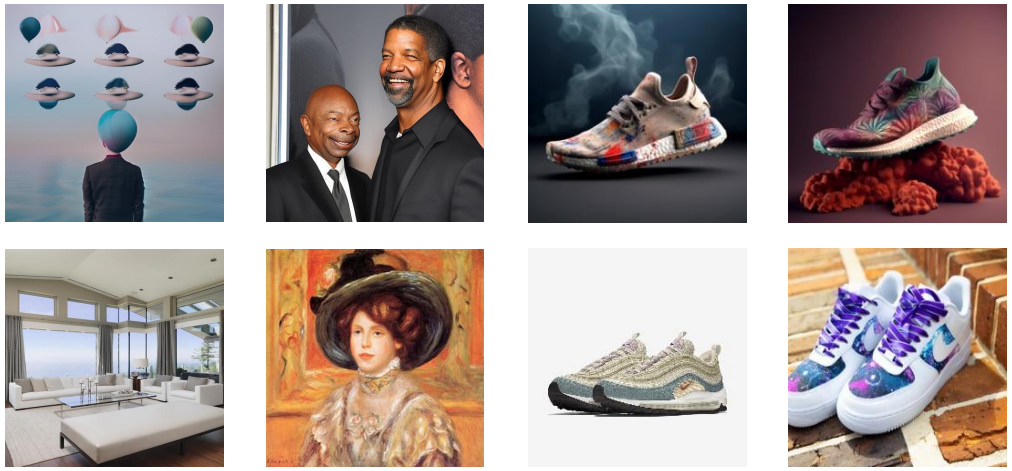
Faculty of Informatics, Mathematics, and Computer Science

Department of Information Systems and Technology

June, 2024

Synthetics Detection

Student: Alex Dadukin



Train split

Test split

Train

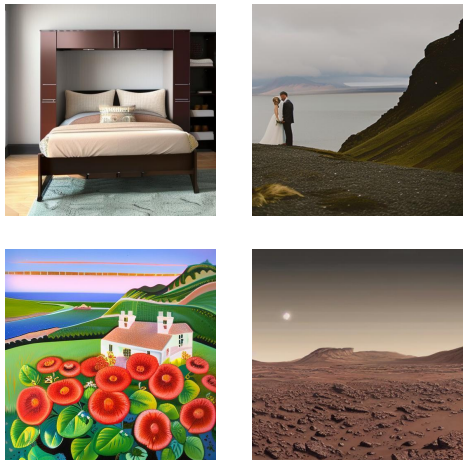
- about 18k images of size **512x512**
- 10k fake, 8k real (1.25:1)

Test

- about 5k images of size **240x240**
- 1.3k fake, 0.8 real (1.625:1)

Classes

- REAL
- FAKE



Fake vs. Real
Q: Can you spot real images?

Train

- There is no clear **visual** difference between fake and real images
- Real images contain drawings (including contemporary art)

Test

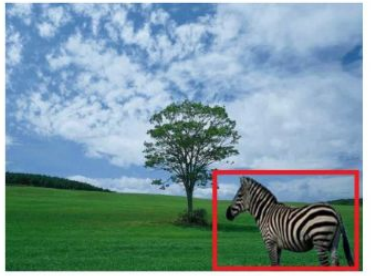
- Images are **4.5 times smaller** (2.3 times along single dimension) then train set
  - Downscaling **train** to **240x240**: potentially, losing important information
  - Upscaling **test** to **512x512**: potentially, adding noise/introducing
- Images are not as diverse as the **train split**
- Real class from both **train and test** have visual distortions which makes objects look less “real”
- Both test and train has slightly different fake to real ratio

Original aspect ratio is not preserved

J. Muindi et. al. Deep Fake Detector: Using ML techniques to Distinguish Real Images From Fakes // Stanford CS230, 2021

N. AlShariah et. al. Detecting Fake Images on Social Media using Machine Learning // IJACSA, 2019

(a) Authentic Image



(b) Fake Image

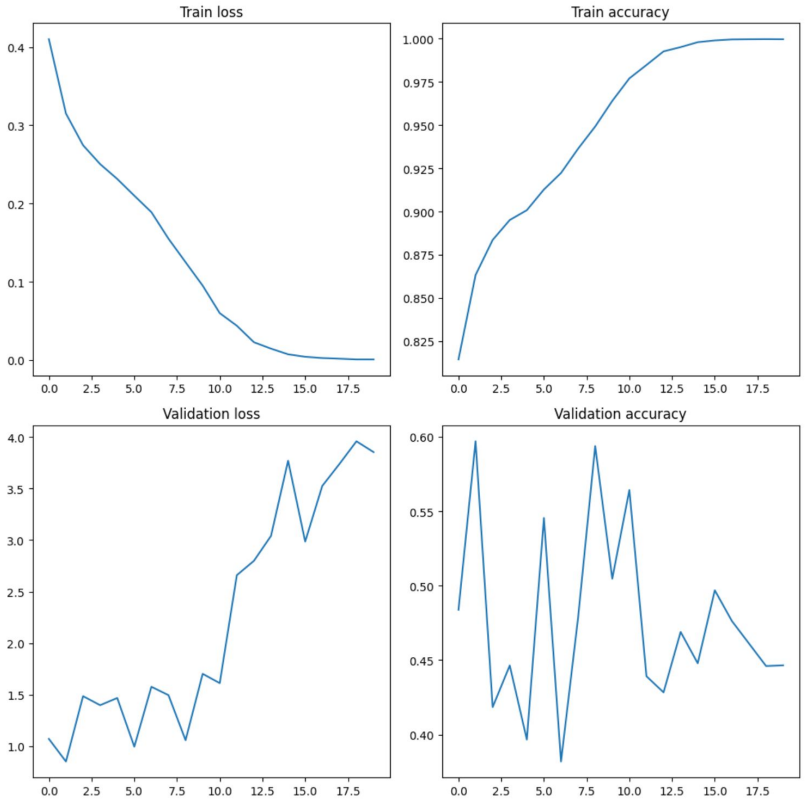
Fakes are defined more broadly: everything that is not a real photo

*"Of the numerous CNN architectures that we tested, we found that the **ResNet** architecture outperformed the others in terms of the dev set accuracy, with DenseNet coming in at a second..."*

Baseline accuracy: **75%**

Model architecture	Accuracy on raw
AlexNet	67%
VGG16	68%
Inception	64%
ResNet	75%
DenseNet	67%

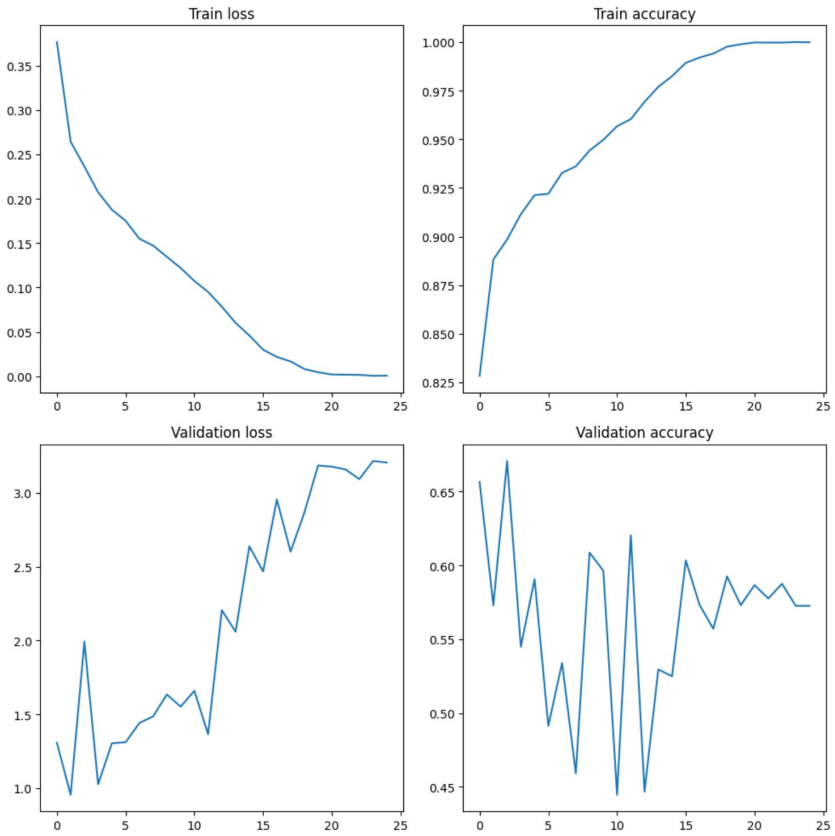
ResNet demonstrated superiority over other models



ResNet 18

- Batch Size: **8**
- Learning rate: **3e-4**
- Epochs: **20**

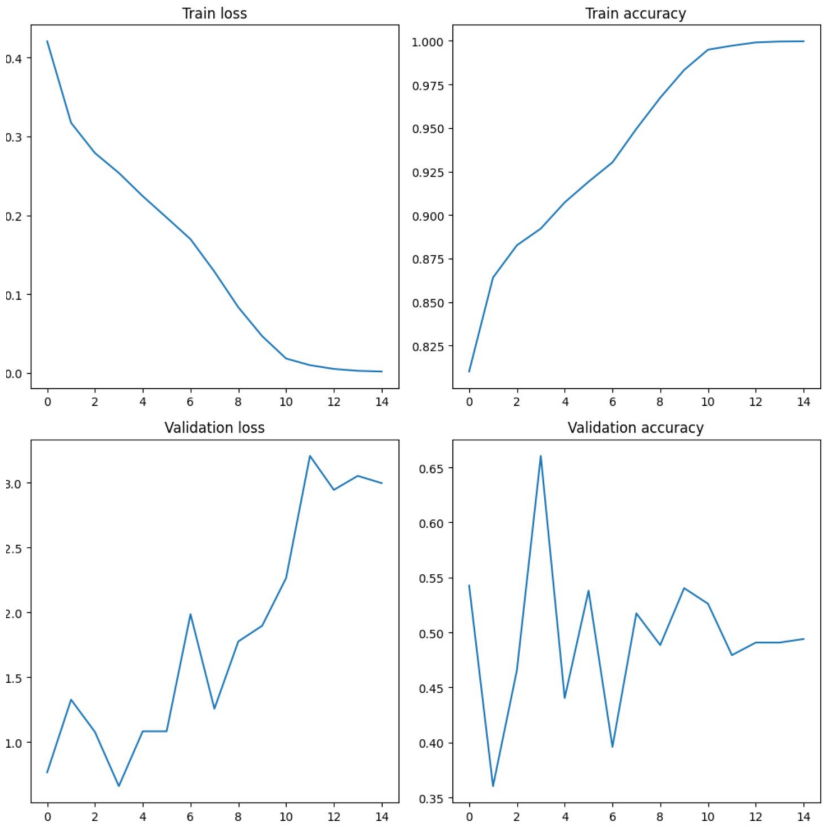
Best Accuracy: **0.60**



ResNet 18 (Grid Search)

- Batch Size: **32**
- Gradient Clip: **0.1**
- Weight Decay: **0.0003**
- Learning rate: **3e-4**
- Epochs: **25**

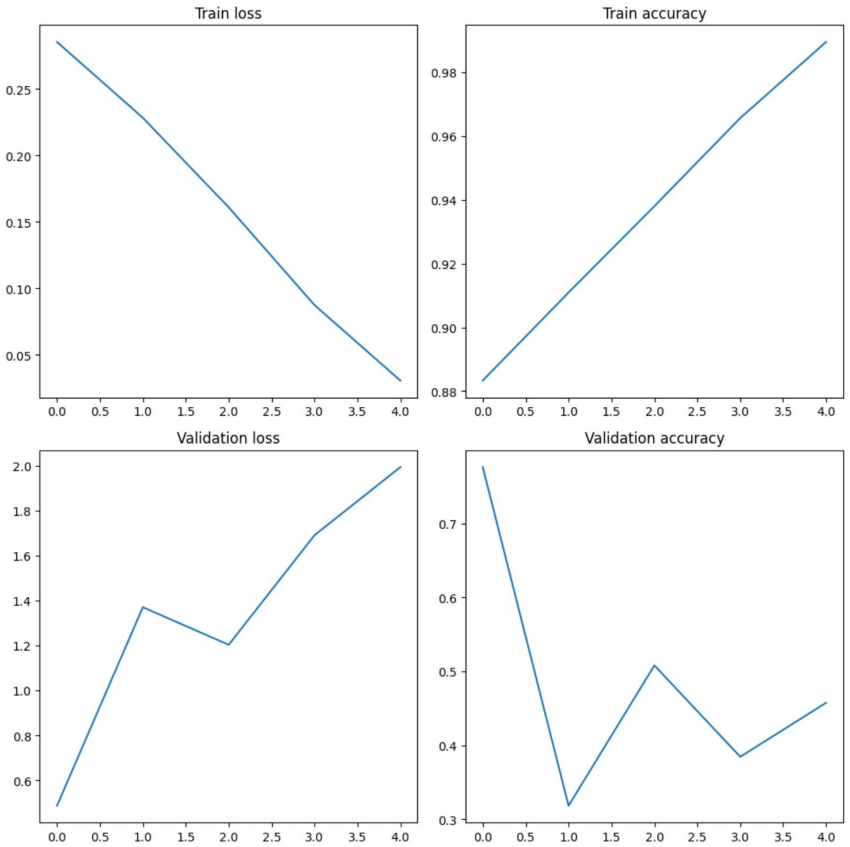
Best Accuracy: **0.67**



ResNet 18 (Balanced classes in train)

- Batch Size: **8**
- Learning rate: **3e-4**
- Epochs: **15**

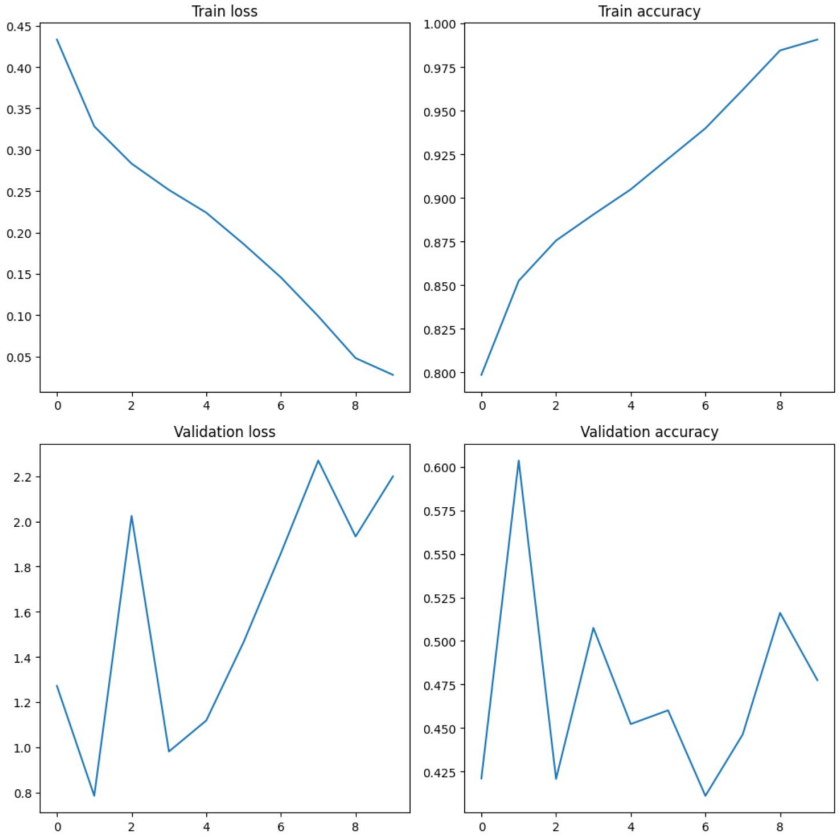
Best Accuracy: **0.66**



ResNet 18 (FT)

- ImageNet Weights (v1)
- Num Classes: **2**
- Batch Size: **8**
- Learning rate: **3e-4**
- Epochs: **5**

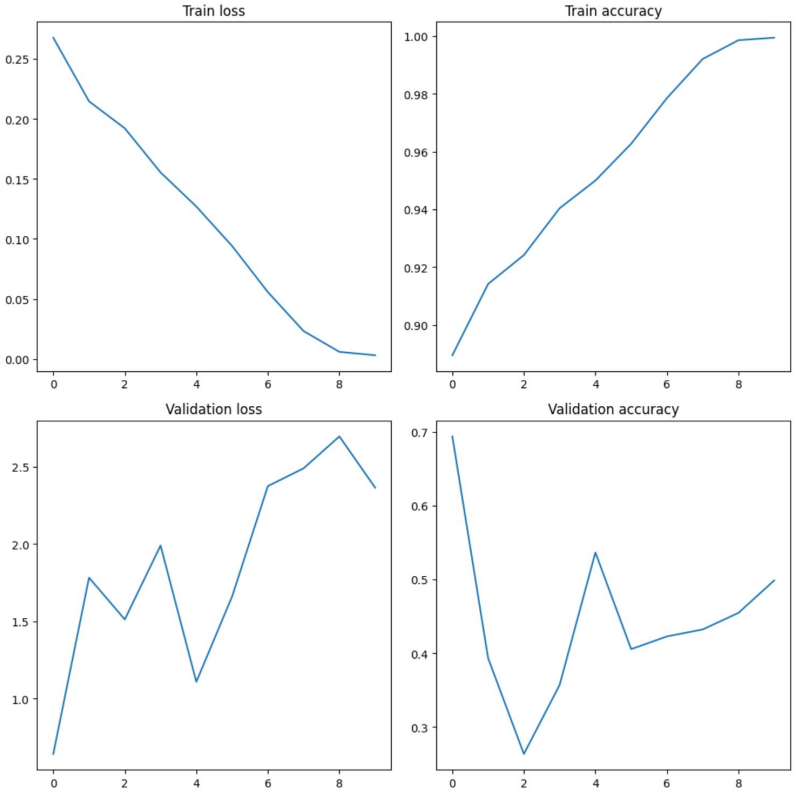
Best Accuracy: **0.78**



ResNet 34

- Batch Size: **8**
- Learning rate: **3e-4**
- Epochs: **10**

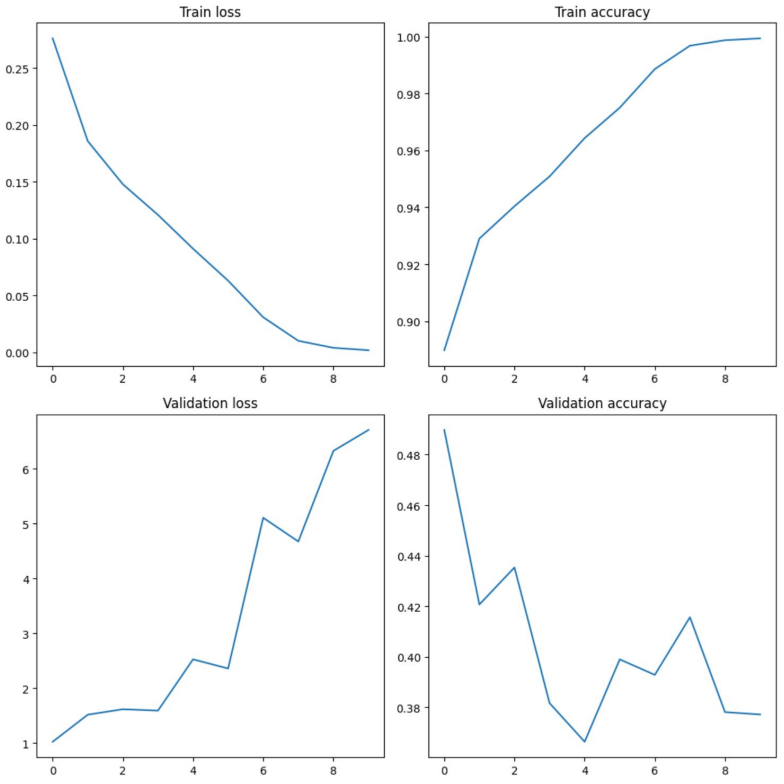
Best Accuracy: **0.60**



ResNet 34 (FT)

- ImageNet Weights
- Batch Size: **8**
- Learning rate: **3e-4**
- Epochs: **10**

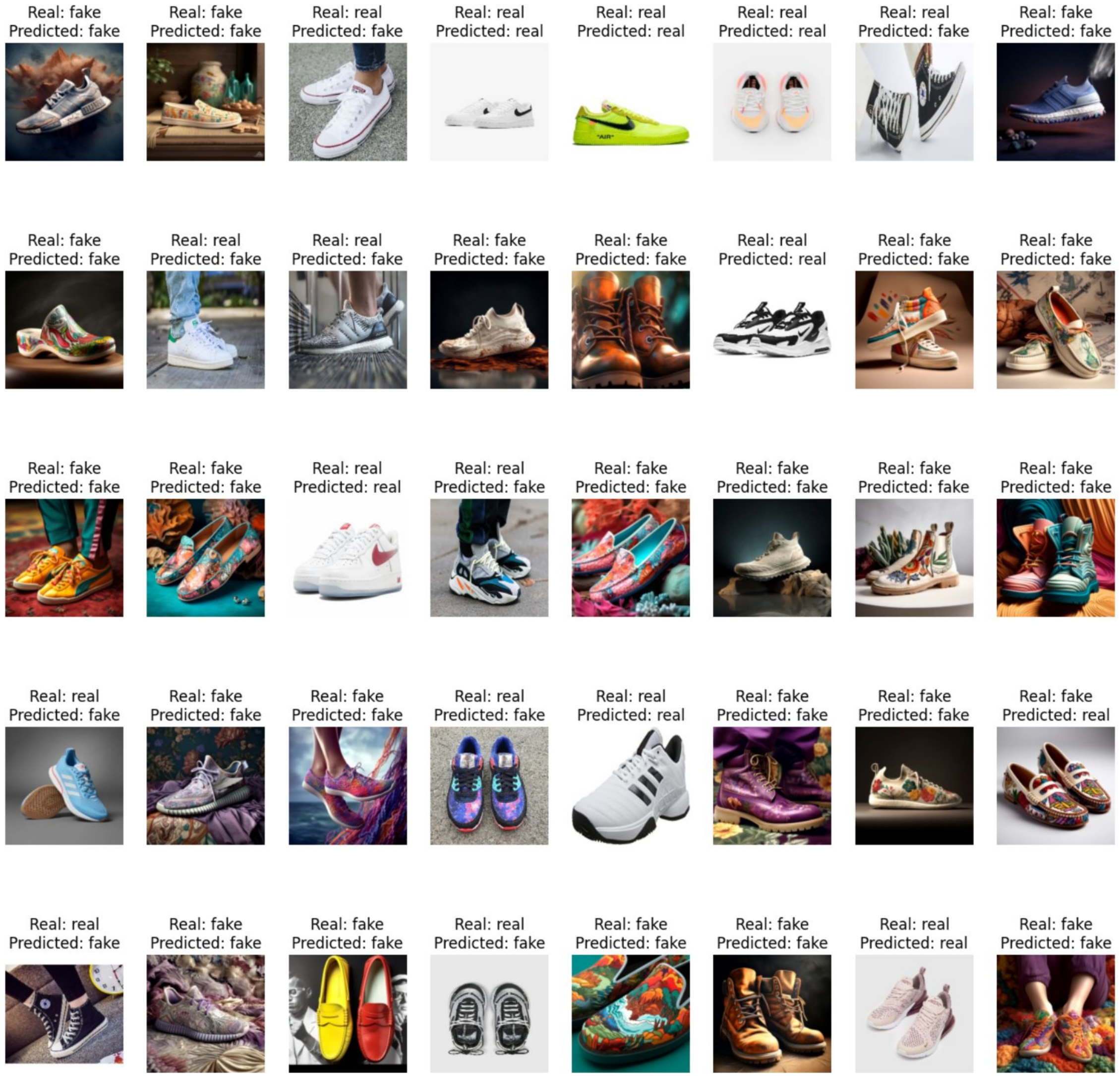
Best Accuracy: **0.69**

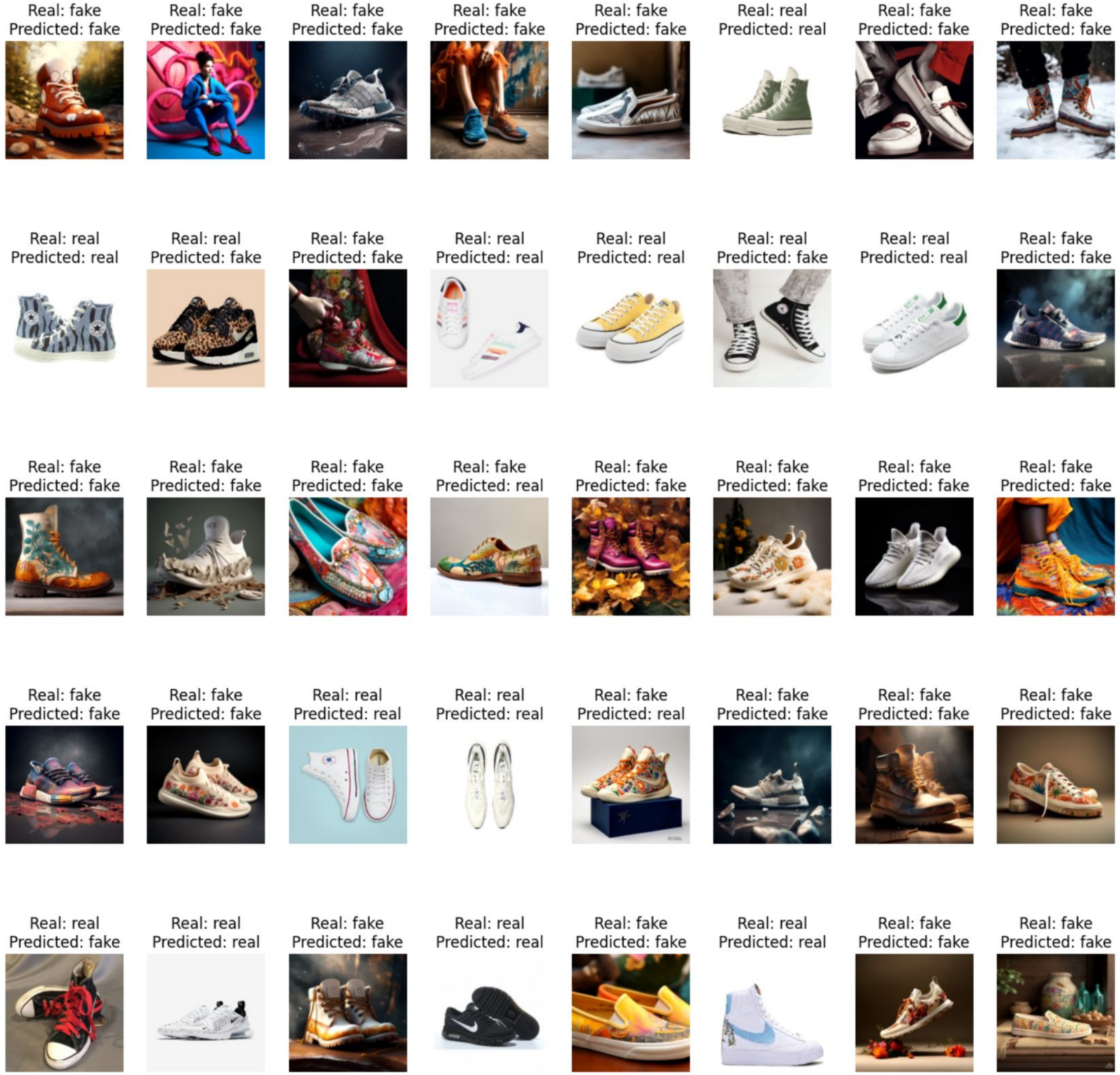


ResNet 50 (FT)

- ImageNet Weights
- Batch Size: **8**
- Learning rate: **3e-4**
- Epochs: **10**

Best Accuracy: **0.49**


Real: fake Predicted: fake
Real: fake Predicted: fake
Real: real Predicted: fake
Real: real Predicted: real
Real: real Predicted: real
Real: real Predicted: real
Real: real Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: real Predicted: fake
Real: real Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: real Predicted: real
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: real Predicted: real
Real: real Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: real Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: real Predicted: fake
Real: real Predicted: real
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: real
Real: real Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: real Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: real Predicted: real
Real: fake Predicted: fake


Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: real Predicted: real
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: real Predicted: real
Real: real Predicted: fake
Real: fake Predicted: fake
Real: real Predicted: real
Real: real Predicted: real
Real: real Predicted: fake
Real: real Predicted: real
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: real
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: real Predicted: real
Real: real Predicted: real
Real: fake Predicted: real
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: fake Predicted: fake
Real: real Predicted: fake
Real: real Predicted: real
Real: fake Predicted: fake
Real: real Predicted: real
Real: fake Predicted: fake
Real: real Predicted: real
Real: fake Predicted: fake
Real: fake Predicted: fake

Code

Releases 0.9.9 & 1.0.0: weights


HSE UNIVERSITY
EST 1992