Optics-free Image Classification with Deep Metric Learning

Prakhar Srivastava
University of Tartu

Supervised by:
Prof. Kallol Roy
University of Tartu, Estonia

Collaborations

Kallol Roy

Rajesh Menon

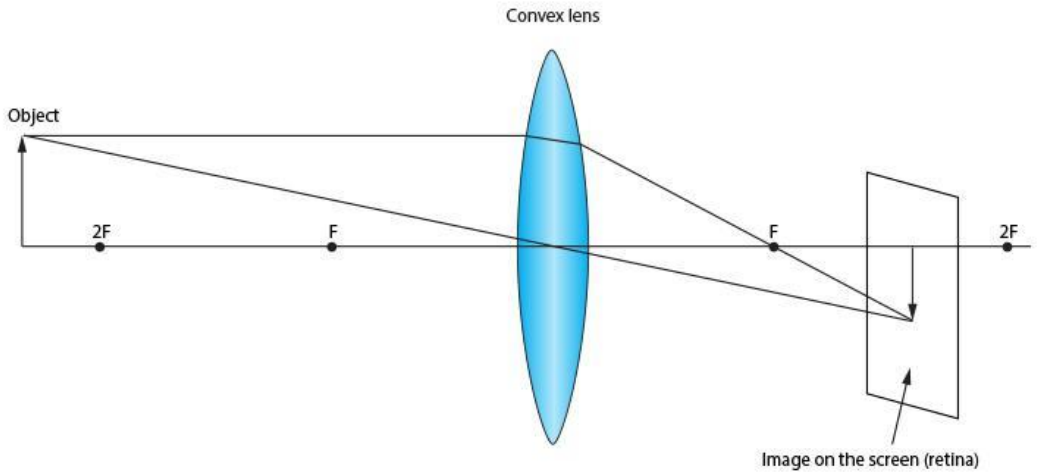
Soren Nelson

Content

- Introduction
- Camera fundamentals
- Optics-free imaging
- Image classification
- Methods
- Results
- Conclusion and Future work

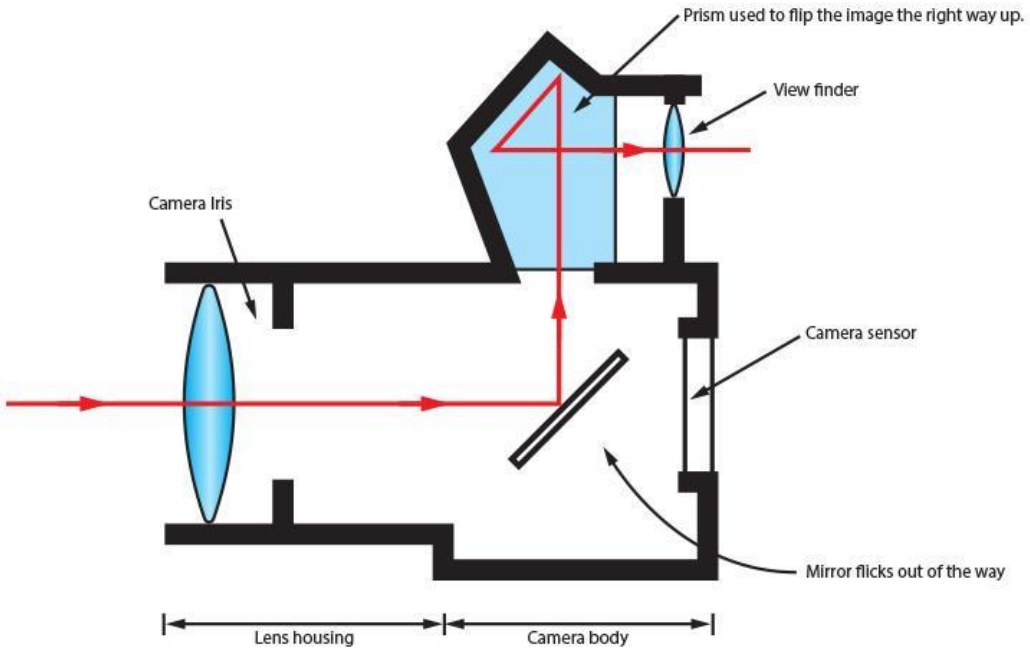
https://mammothmemory.net/physics/lenses/convex-lenses/convex-lens-uses-camera.html

Image formation basics



https://mammothmemory.net/physics/lenses/convex-lenses/convex-lens-uses-camera.html

Camera fundamentals



https://mammothmemory.net/physics/lenses/convex-lenses/convex-lens-uses-camera.html

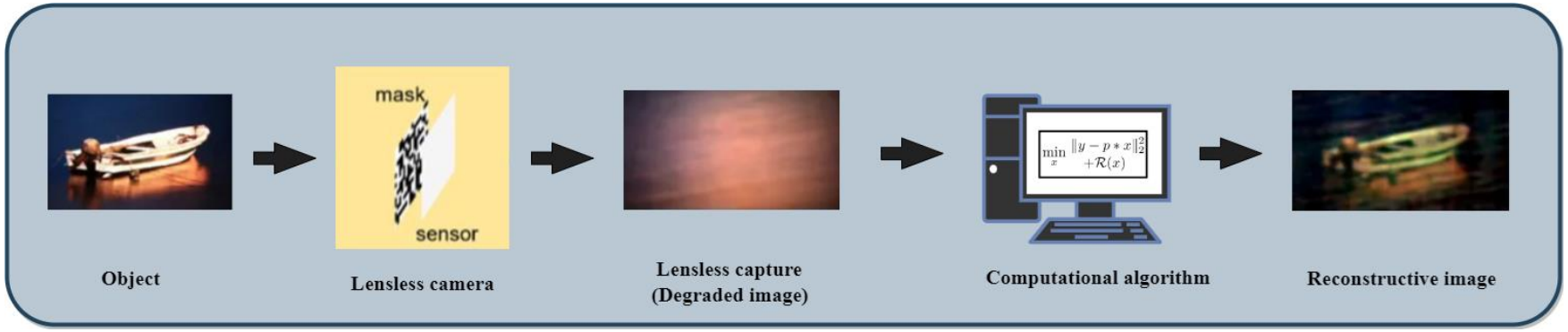
Need of miniaturization

https://www.outlookindia.com/website/story/india-news-iit-madras-and-rice-university-researchers-develop-algorithms-for-lensless-miniature-cameras/382684

Problem of miniaturization of lens

- Getting light inside miniature lenses is difficult

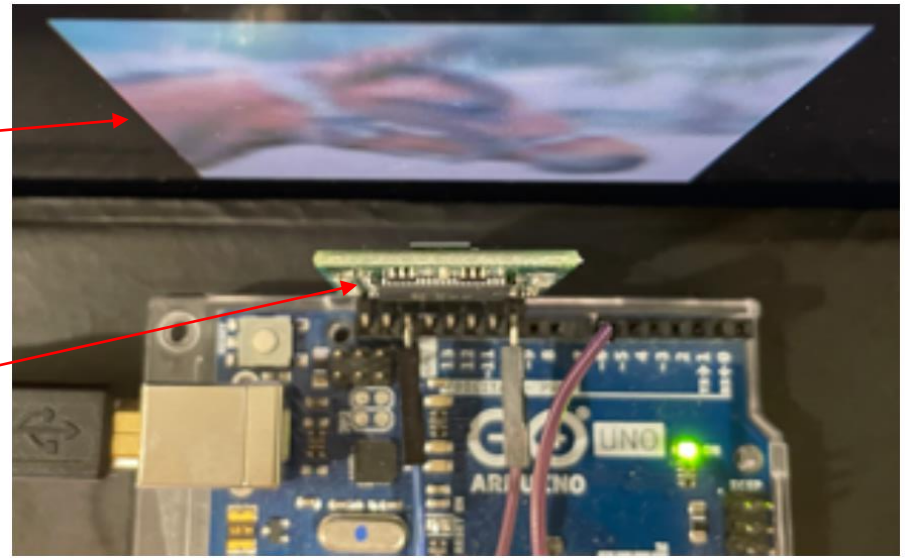
Lens replace with algorithm

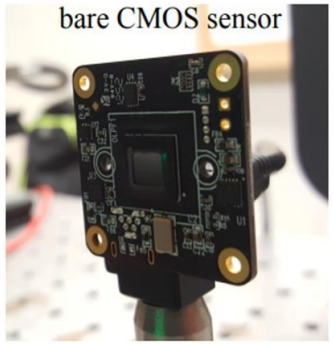


https://medium.com/@bezzam/lensless-imaging-with-the-raspberry-pi-and-python-diffusercam-473e47662857

LCD

CMOS Sensor





Arducam
without lens

Original Cifar10

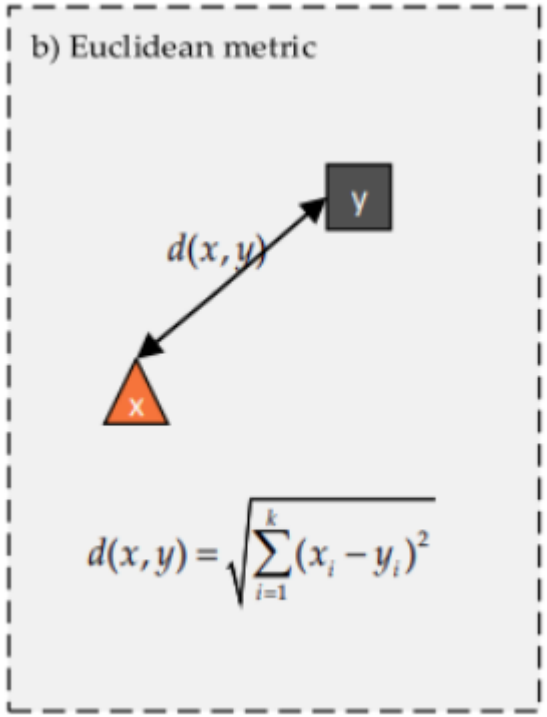
Lensless Cifar10 image
Obtained from CMOS
sensor

Menon, Rajesh; Nelson, Soren (2021): Supplementary document for Bijective-constrained cycle-consistent deep-learning for optics-free imaging and classification - 5399506.pdf.

“ Deep Metric Learning ”

What is Metric Learning?

Metric learning is a technique that focuses on mapping similar data points close to each other in the embedding space, while mapping dissimilar data points far away from each other.



KAYA, M.; BİLGE, H.Ş. Deep Metric Learning: A Survey. *Symmetry* 2019, *11*, 1066.

Feature is all you need

- Metric is the essential feature for the model to learn

Metric invariance

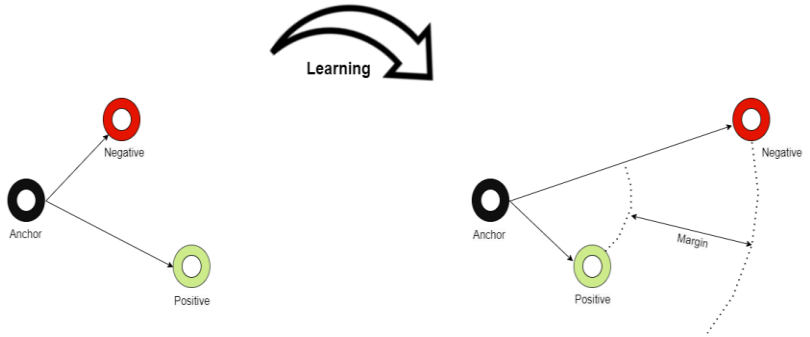
ideal condition

- Features stays invariant (constant)

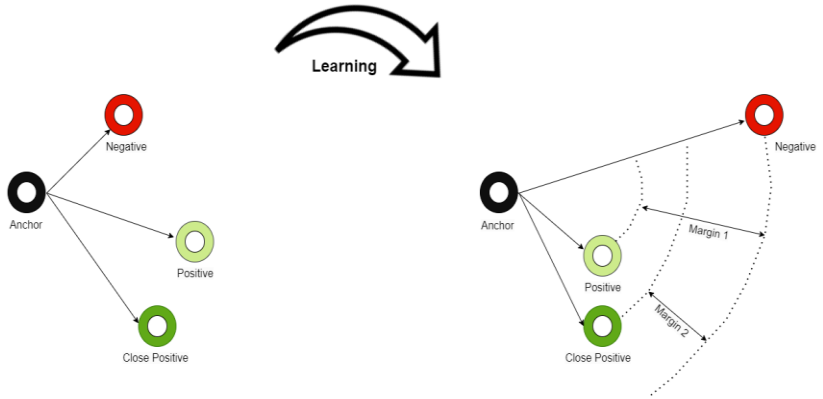
https://math.stackexchange.com/questions/2258389/how-can-a-mug-and-a-torus-be-equivalent-if-the-mug-is-chiral

Loss functions

Triplet loss



Quadruplet loss

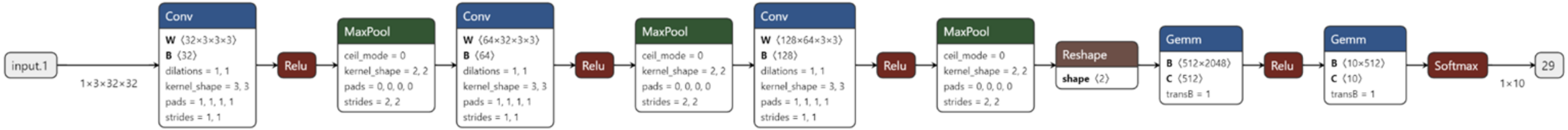


Building the dataset

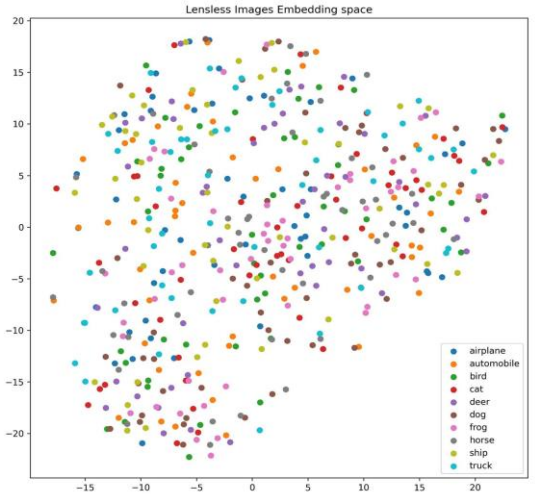
Data loader

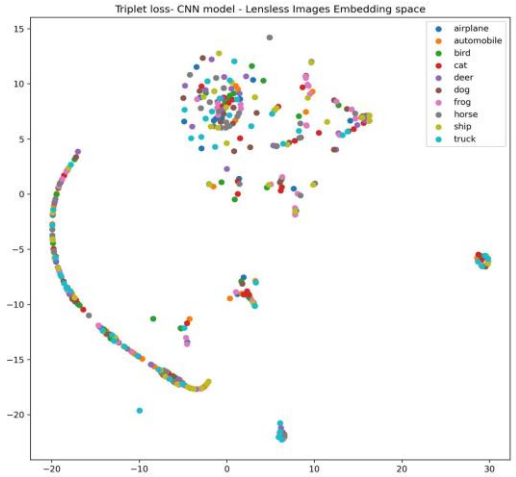
	anchor_img	positive_img	close_positive_img	negative_img	label	label_name
15394	cat_s_001459.jpeg	cat_s_001459.png	cat_s_001462.png	automobile_s_001307.png	3	cat
22494	fawn_s_000837.jpeg	fawn_s_000837.png	fawn_s_000844.png	pipit_s_001424.png	4	deer
10317	alauda_arvensis_s_001127.jpeg	alauda_arvensis_s_001127.png	alauda_arvensis_s_001145.png	red_deer_s_000983.png	2	bird
6895	convertible_s_000276.jpeg	convertible_s_000276.png	convertible_s_000281.png	gelding_s_001741.png	1	automobile
39343	tennessee_walker_s_001125.jpeg	tennessee_walker_s_001125.png	tennessee_walker_s_001132.png	moving_van_s_001754.png	7	horse
...	...	...	...	...	...	...
16514	felis_catus_s_000393.jpeg	felis_catus_s_000393.png	felis_catus_s_000394.png	coupe_s_000543.png	3	cat
14054	sparrow_s_000027.jpeg	sparrow_s_000027.png	sparrow_s_000028.png	wagon_s_001671.png	2	bird
24457	sika_s_000474.jpeg	sika_s_000474.png	sika_s_000475.png	passenger_ship_s_000477.png	4	deer
35950	broodmare_s_001448.jpeg	broodmare_s_001448.png	broodmare_s_001450.png	automobile_s_002579.png	7	horse
2784	jumbo_jet_s_000267.jpeg	jumbo_jet_s_000267.png	jumbo_jet_s_000268.png	bullfrog_s_000293.png	0	airplane

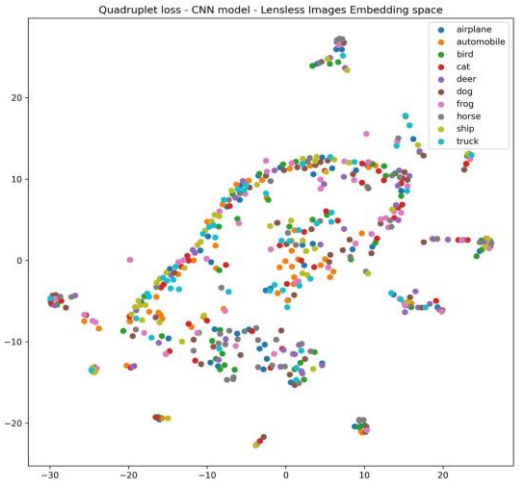
Model


input.1
1×3×32×32
Conv
W ⟨32×3×3×3⟩
B ⟨32⟩
dilations = 1, 1
kernel_shape = 3, 3
pads = 1, 1, 1, 1
strides = 1, 1
Relu
MaxPool
ceil_mode = 0
kernel_shape = 2, 2
pads = 0, 0, 0, 0
strides = 2, 2
Conv
W ⟨64×32×3×3⟩
B ⟨64⟩
dilations = 1, 1
kernel_shape = 3, 3
pads = 1, 1, 1, 1
strides = 1, 1
Relu
MaxPool
ceil_mode = 0
kernel_shape = 2, 2
pads = 0, 0, 0, 0
strides = 2, 2
Conv
W ⟨128×64×3×3⟩
B ⟨128⟩
dilations = 1, 1
kernel_shape = 3, 3
pads = 1, 1, 1, 1
strides = 1, 1
Relu
MaxPool
ceil_mode = 0
kernel_shape = 2, 2
pads = 0, 0, 0, 0
strides = 2, 2
Reshape
shape ⟨2⟩
Gemm
B ⟨512×2048⟩
C ⟨512⟩
transB = 1
Relu
Gemm
B ⟨10×512⟩
C ⟨10⟩
transB = 1
Softmax
1×10
29

Threshold for lensless images

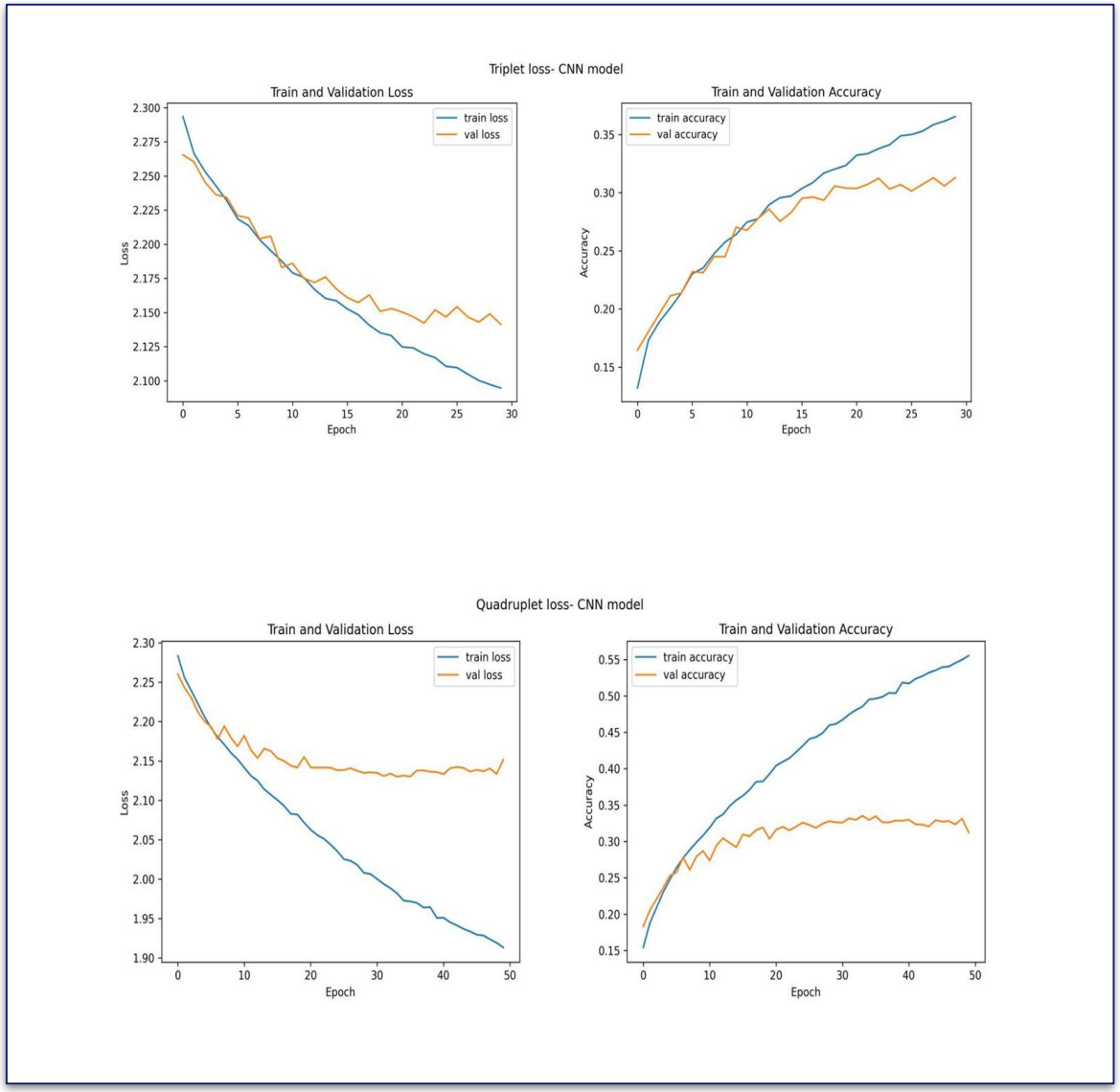






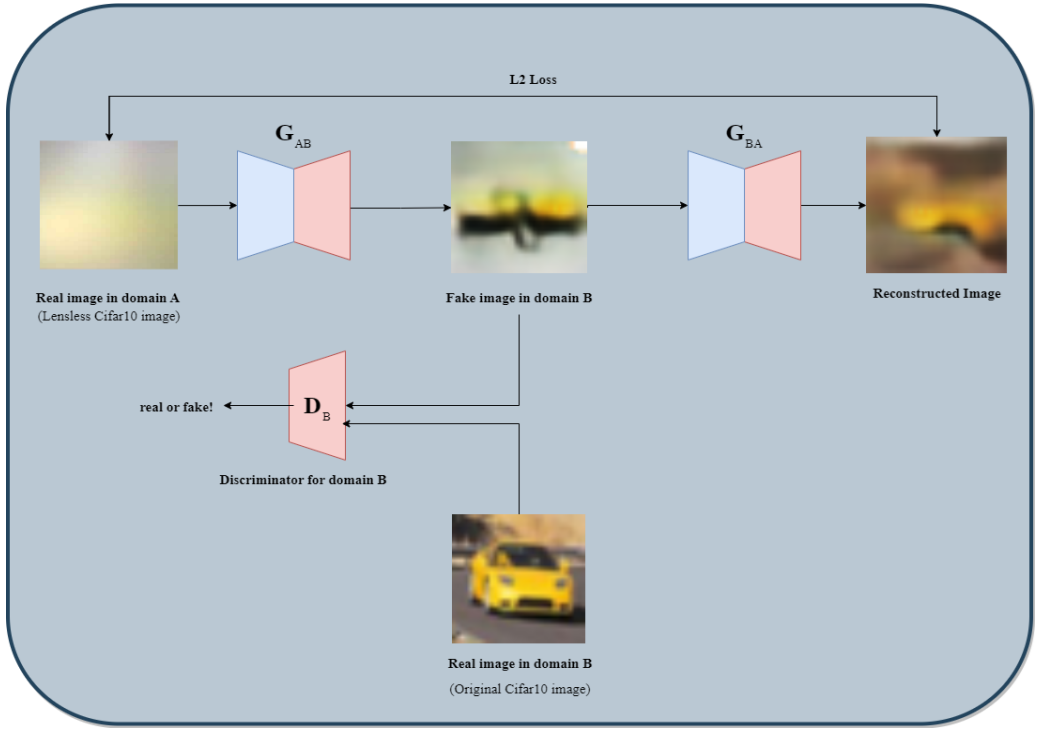
Results

Lensless images



How to improve?

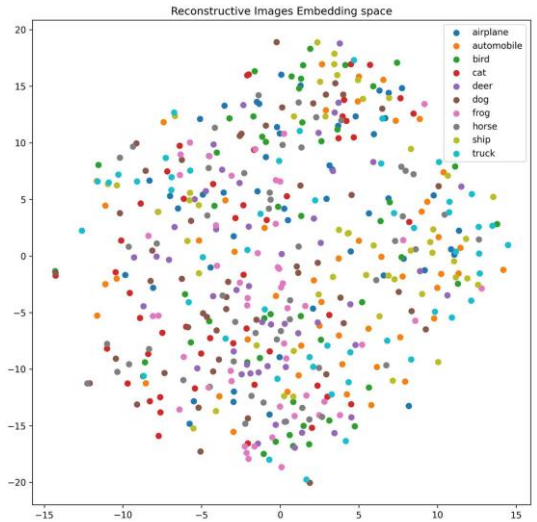
Reconstructive images using CycleGAN

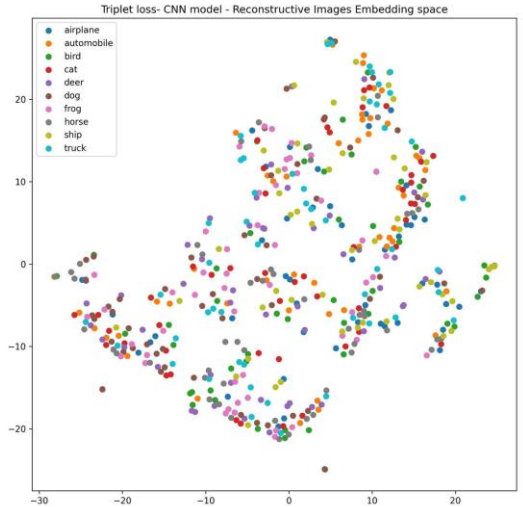


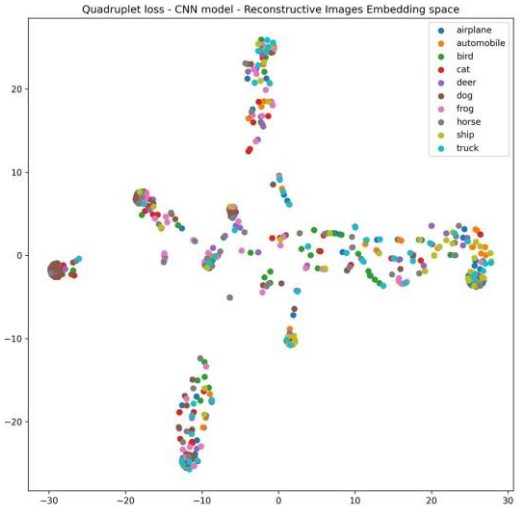
S. Nelson and R. Menon, "Bijective-constrained cycle-consistent deep learning for optics-free imaging and classification," Optica 9, 26-31 (2022).


Comparison of Image Quality Metrics: MSE, PSNR, and SSIM
Original Cifar10
Lensless Cifar10
MSE: 106.15
PSNR: 10.74
SSIM: 0.14
Reconstructive Cifar10
MSE: 82.13
PSNR: 18.68
SSIM: 0.59

Threshold for reconstructive images

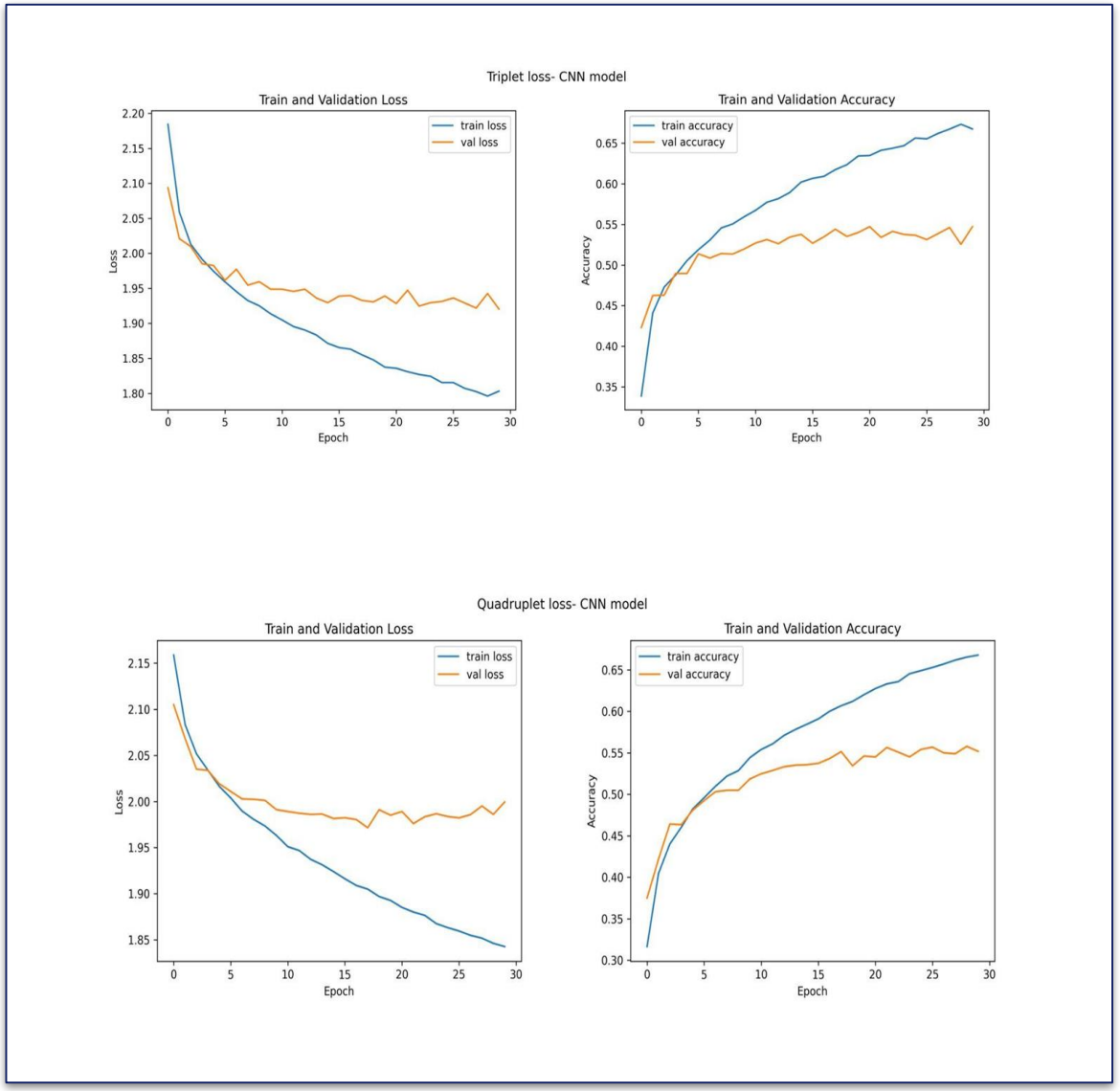






Results

Reconstructive images



Results

Table 1. Triplet Loss - Lensless images

Lensless Images	Accuracy	Precision	Recall	F1score	Epochs
test images	16.40	15.47	16.40	10.92	15
test images	30.73	30.37	30.73	30.22	30
test images	31.13	31.16	31.13	30.62	50

Table 2. Triplet Loss

Reconstructive Images	Accuracy	Precision	Recall	F1score	Epochs
test images	38.37	38.75	38.37	37.72	15
test images	39.58	39.89	39.58	39.41	30
test images	39.56	40.07	39.56	39.41	50

Table 3. Quadruplet Loss

Lensless Images	Accuracy	Precision	Recall	F1score	Epochs
test images	30.16	29.78	30.16	29.53	15
test images	31.84	31.57	31.84	31.24	30
test images	32.14	31.85	32.14	31.71	50

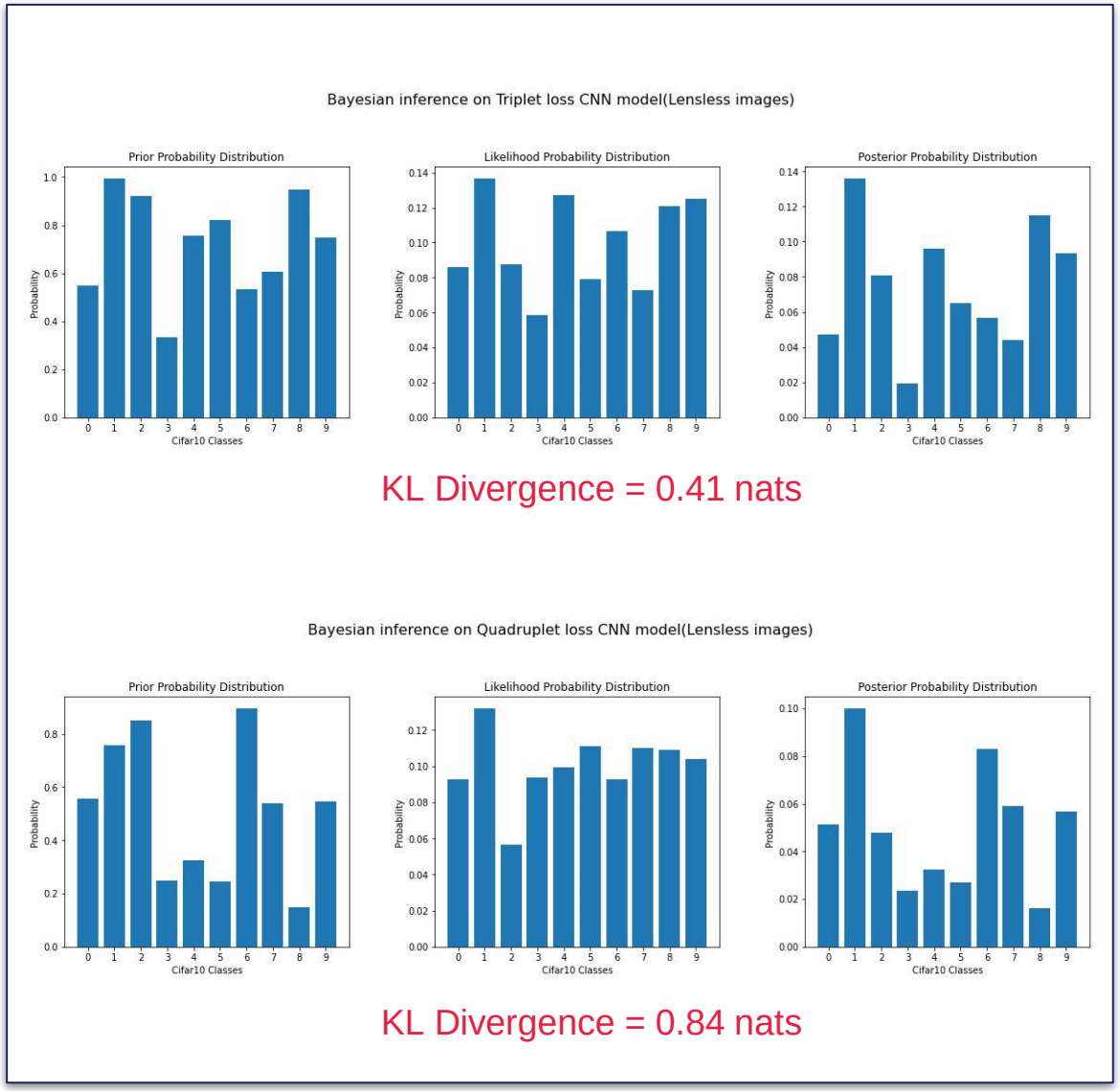
Table 4. Quadruplet Loss

Reconstructive Images	Accuracy	Precision	Recall	F1score	Epochs
test images	40.76	40.44	40.76	40.26	15
test images	40.78	40.51	40.78	40.44	30
test images	40.71	41.26	40.71	40.47	50

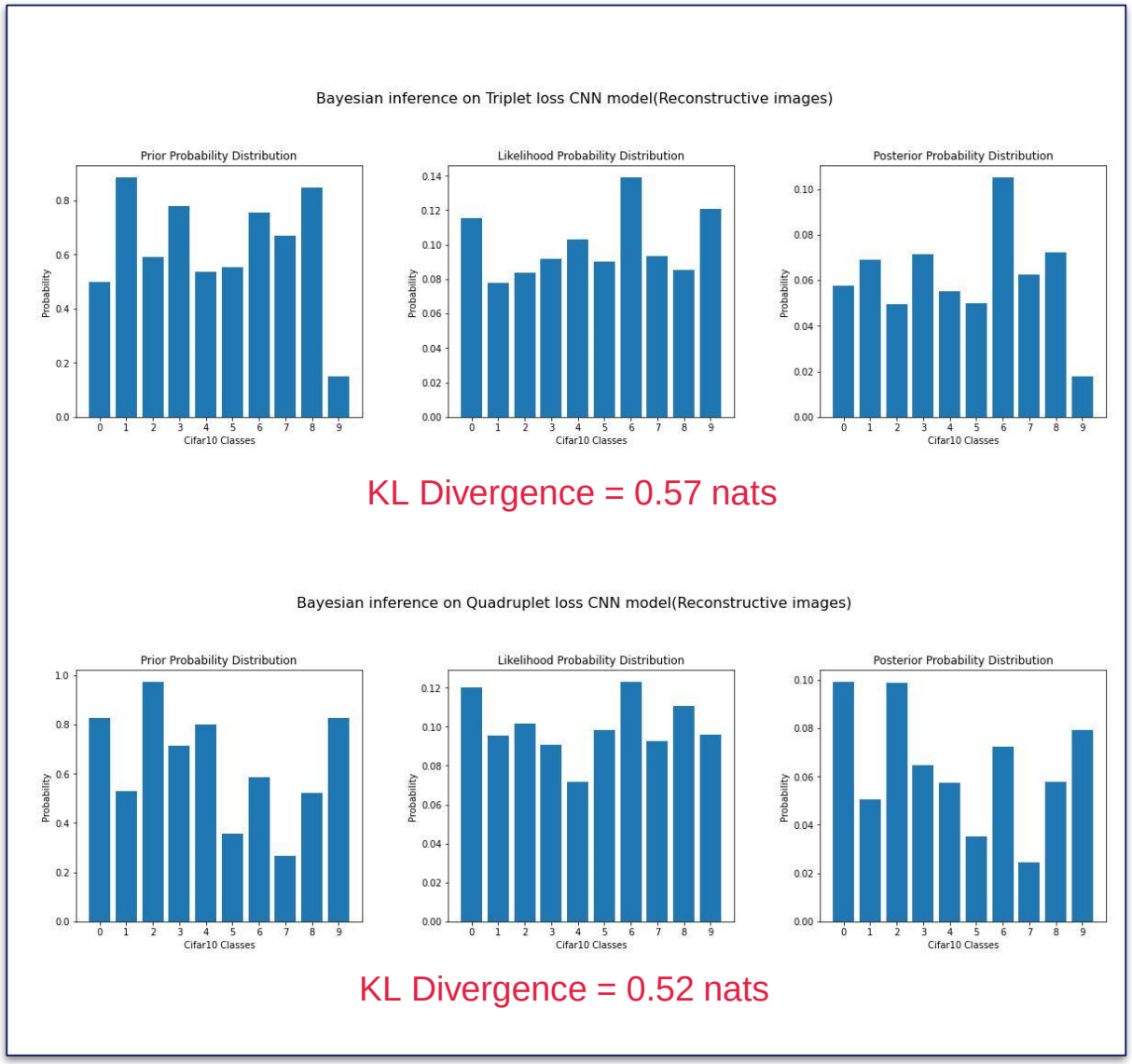
Bayesian Method

- Prior - initial belief or assumption
- Likelihood - prediction of the CNN
- Posterior – updated belief

Bayesian for Lensless images



Bayesian for Reconstructive images



Take away points

- Potential of making miniaturized cameras
- Images are noisy
- Some experiments didn’t converge

Future work

- Multimodality (text + images)
- Planned to use transformer architecture

UNIVERSITY OF TARTU

Thank you!

Reviewer's questions

1) How are hyperparameters chosen, does any experiments performed prior, and did any change in the performance of the accuracy in all the experiments?

-By trail and error.

2) What is the Computational complexity of the algorithms?

O(n)

3) How about software services and hardware configuration used in the experiments?

-Google colab, pro
-Hardware: GPU, High-RAM

4) What is the broader conclusion by comparing all the algorithms with a number of epochs considered?

-For lensless, when epochs are increased, accuracy also increases to a certain limit. But for reconstructive, when epochs are increased, accuracy decreases.

UNIVERSITY OF TARTU

Questions ?

unitartu

tartuuniversity