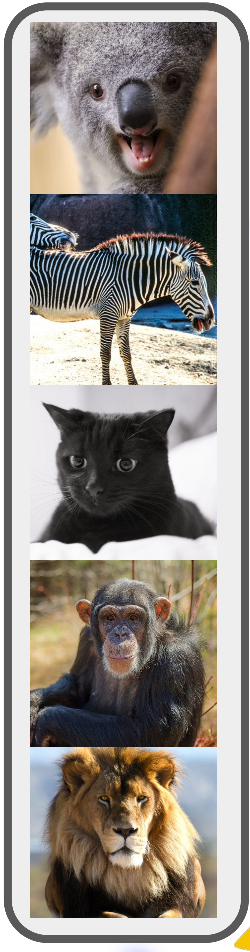
TOOPLOOX AI

My story at Tooplox

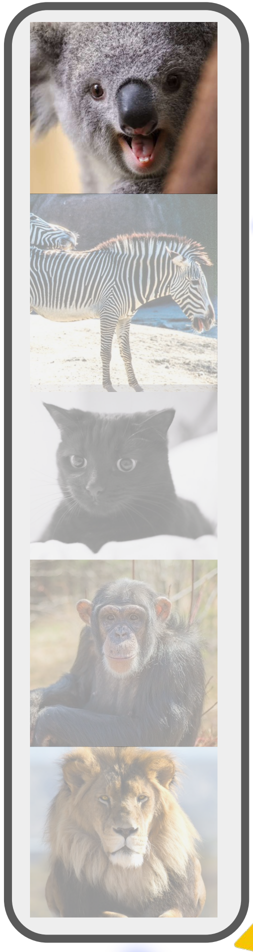
Maciej Zięba

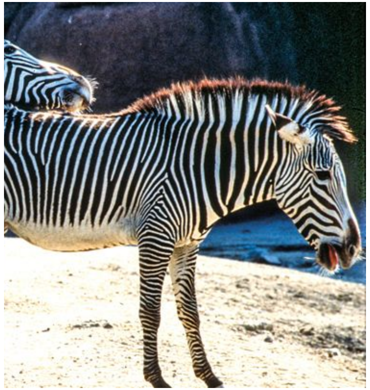
My first research project:
Learning binary representations in unsupervised manner

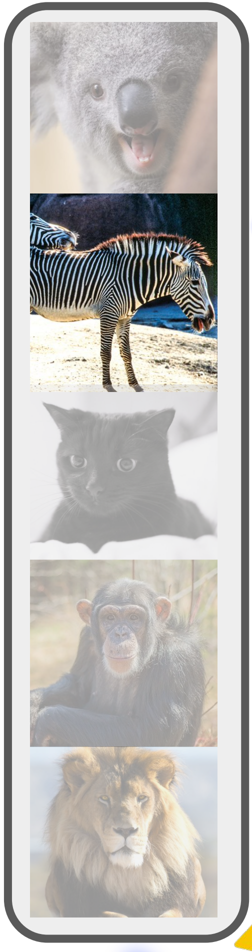
Google Images



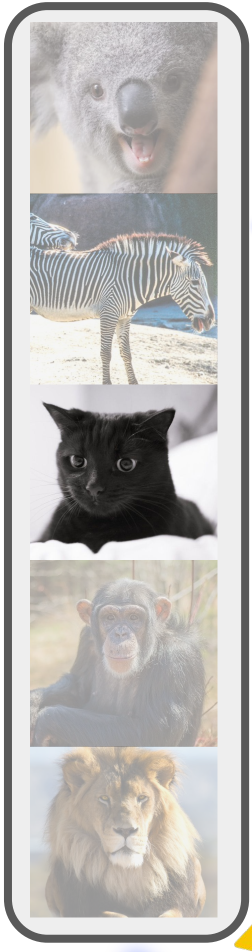
$($  $-$  $)^2$



$\Big($  $-$  $\Big)^2$

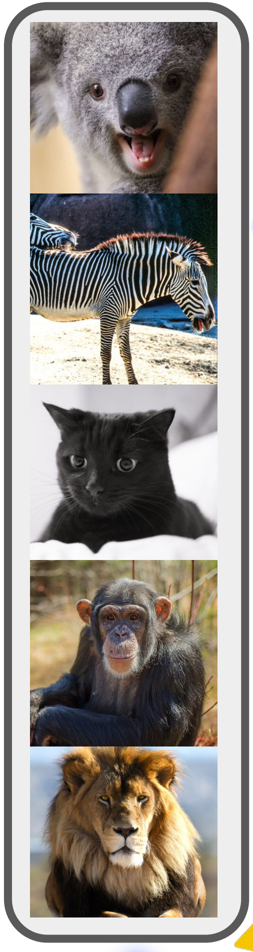


( − )^{2}

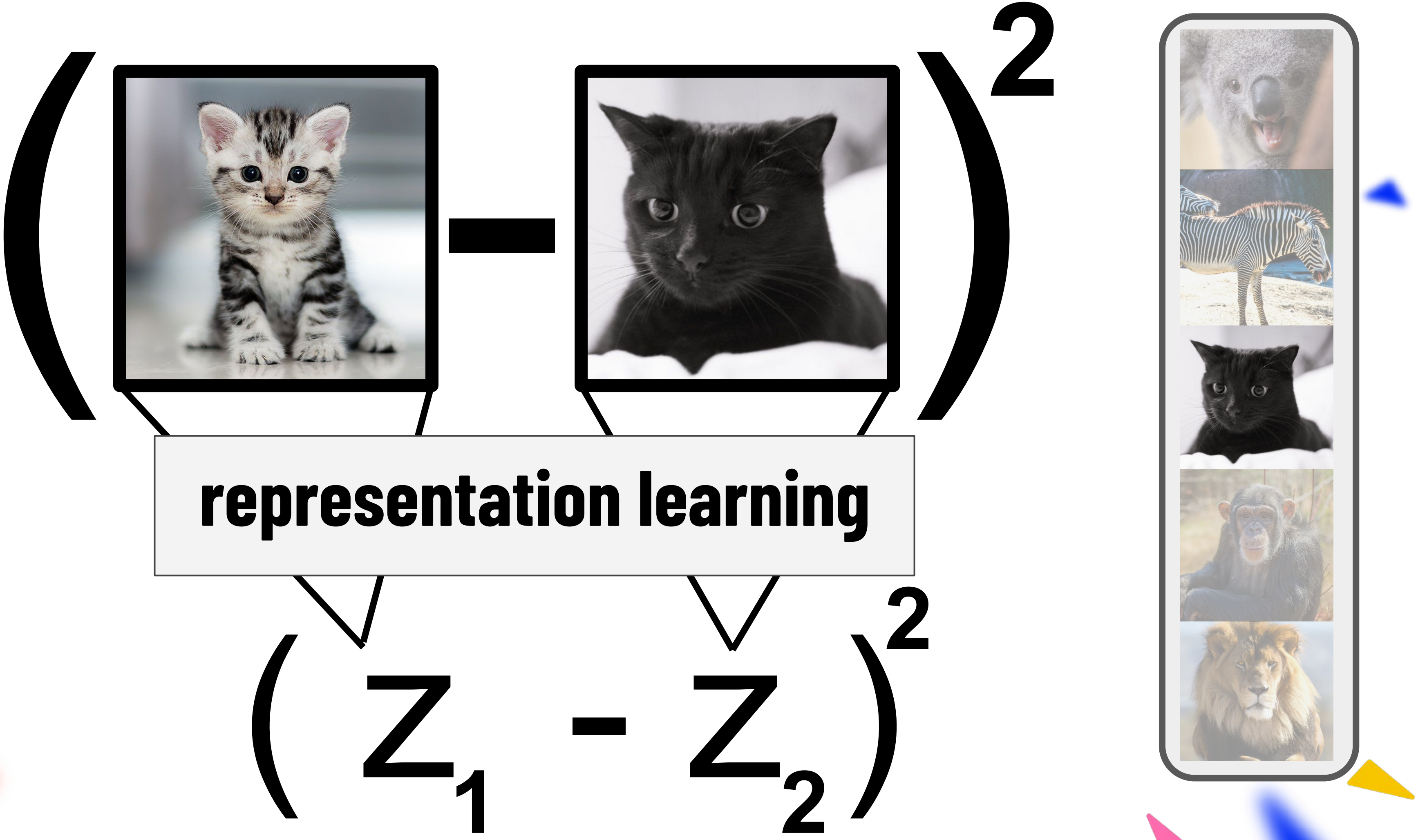


Google Images

Visually similar images

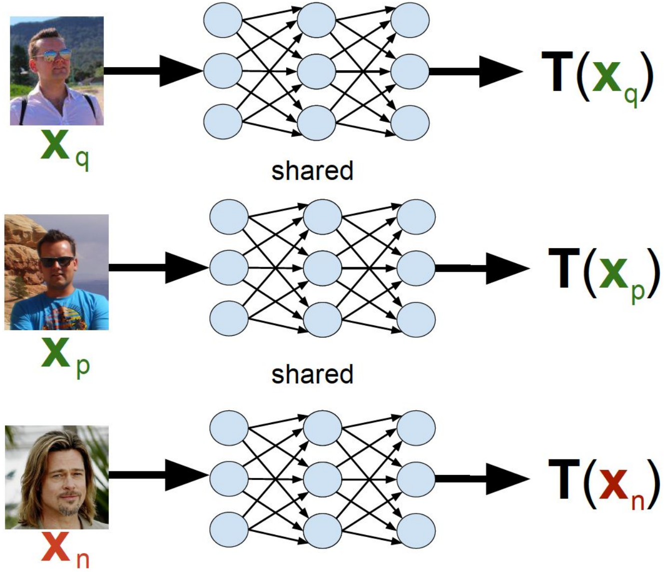


#epicfailure


representation learning
$(z_1 - z_2)^2$

If we have supervised data

What we need is triplet network...



Expected outcome:

$$d_+ < d_-,$$

where:

$$d_+ = d(T(\mathbf{x}_q), T(\mathbf{x}_+))$$
$$d_- = d(T(\mathbf{x}_q), T(\mathbf{x}_-))$$

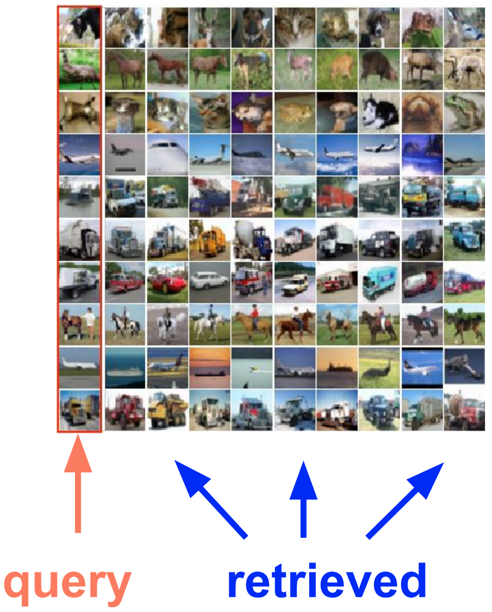
and $d(\cdot, \cdot)$ is some distance function.

Binary descriptors

What we need is compact and binary representation for data examples...

Why binary representation is important ?

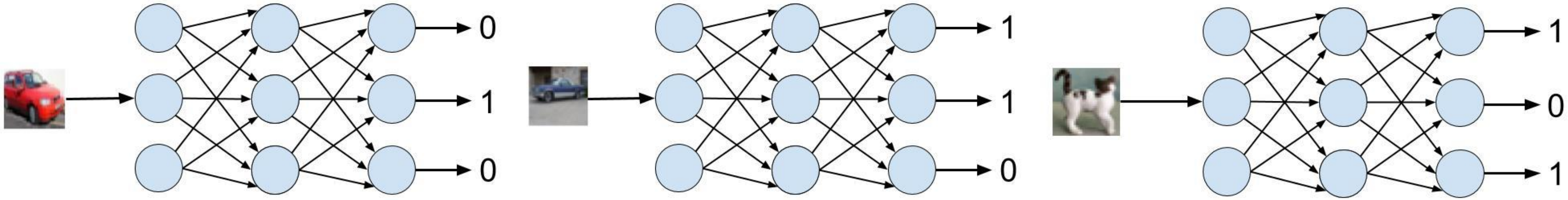
- data compressing,
- effective image retrieval,
- hashing images.



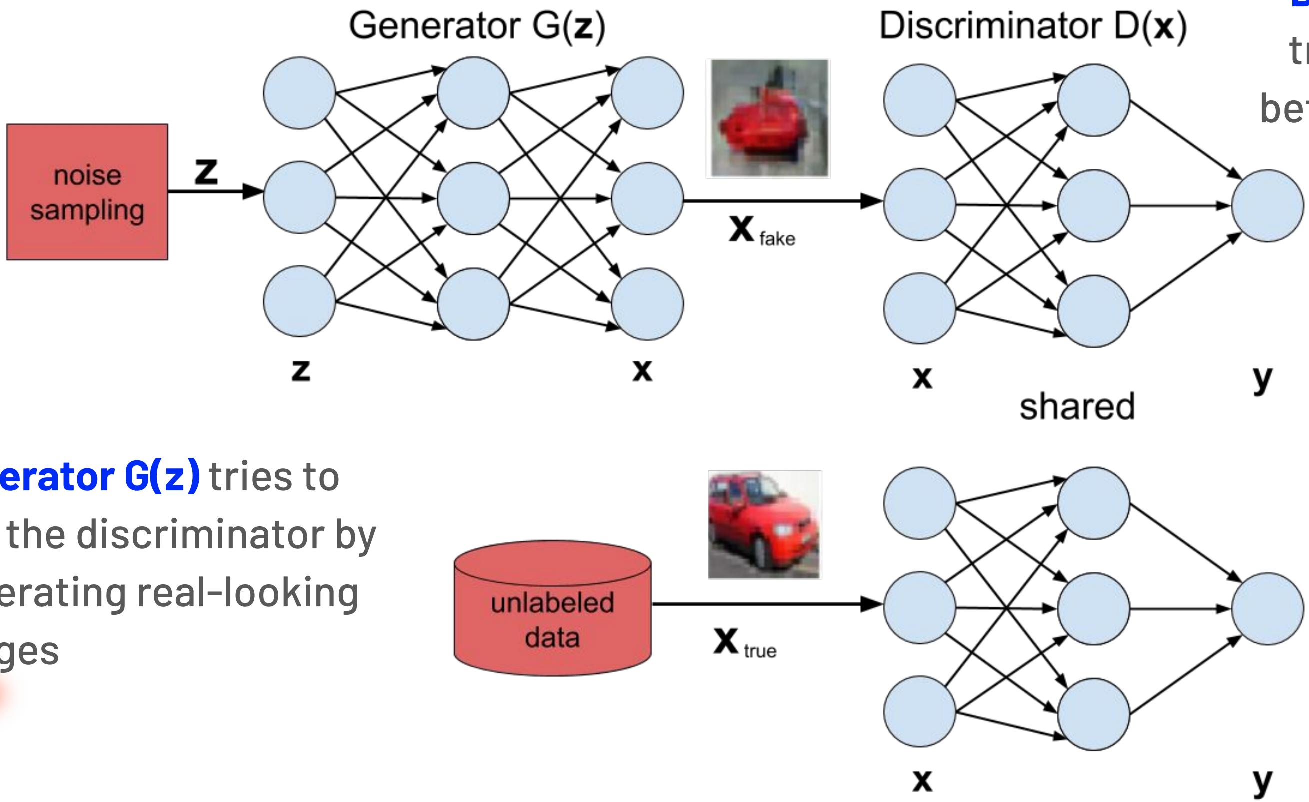
Binary descriptors - unsupervised

What we need is compact and binary representation for data examples...

How can we represent data using binary codes computed with deep nets without the need for costly and imperfect data labelling?



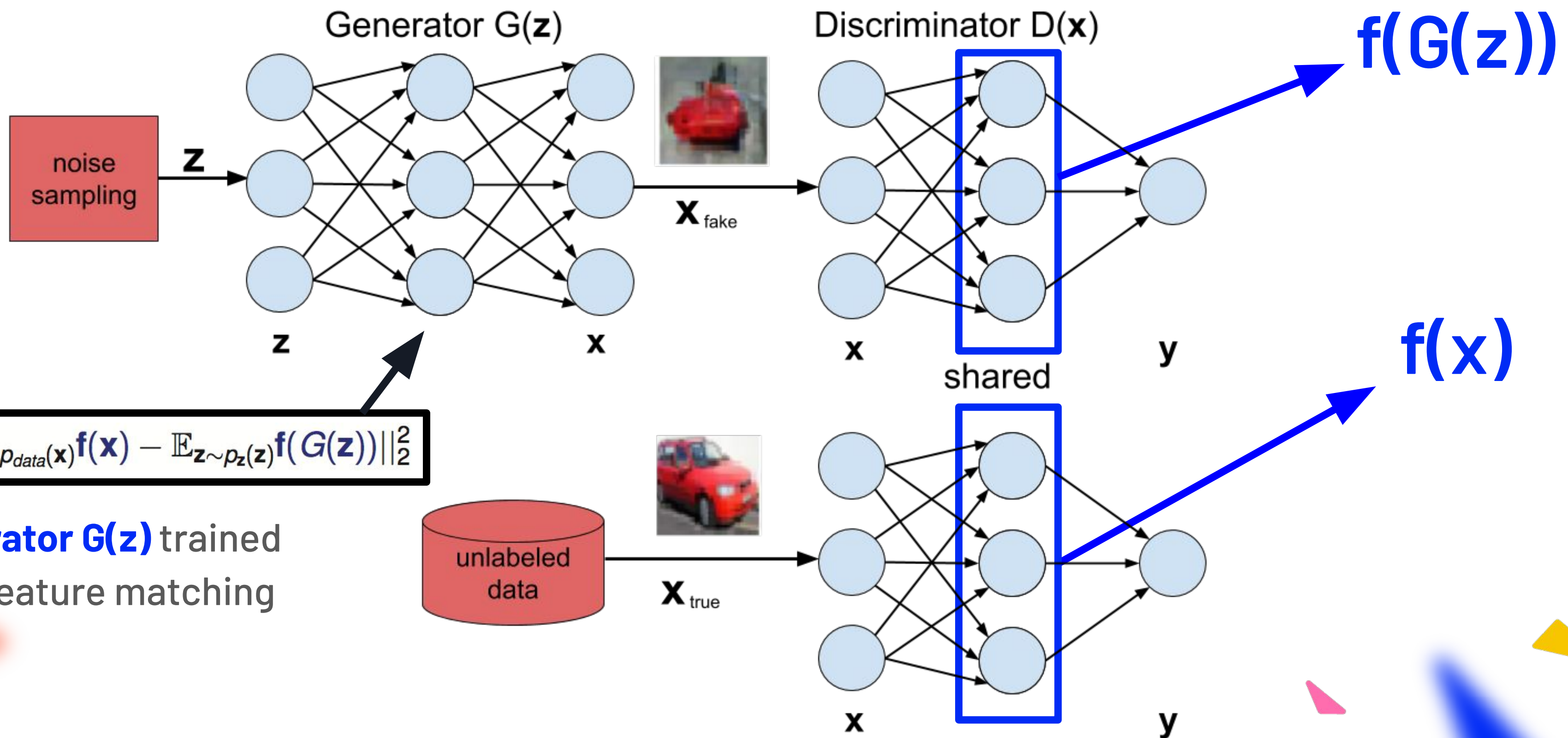
GANs for unsupervised representation learning



Discriminator D(x) tries to distinguish between real and fake images

Generator G(z) tries to fool the discriminator by generating real-looking images

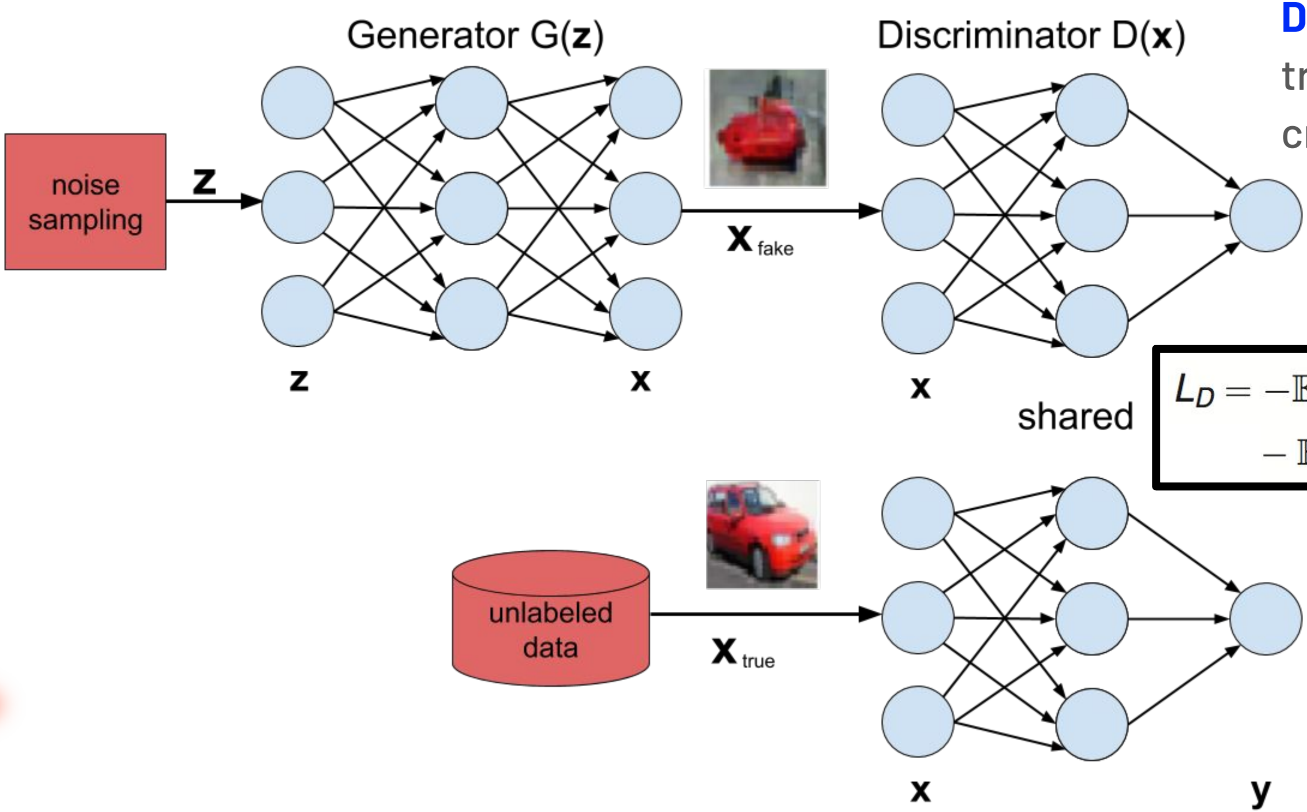
How to train the GAN ?



$$||\mathbb{E}_{\mathbf{x}\sim p_{data}(\mathbf{x})}\mathbf{f}(\mathbf{x}) - \mathbb{E}_{\mathbf{z}\sim p_{\mathbf{z}}(\mathbf{z})}\mathbf{f}(G(\mathbf{z}))||_2^2$$

Generator G(z) trained with feature matching loss

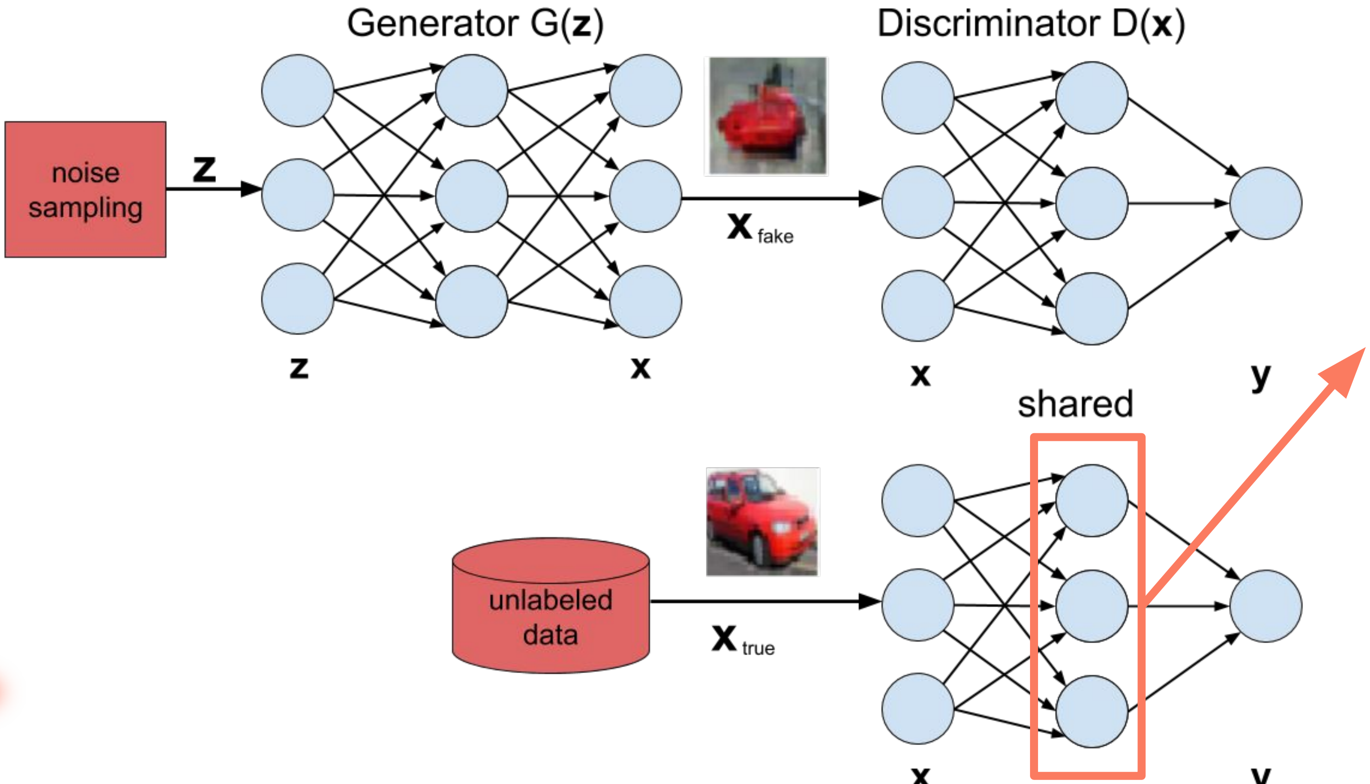
How to train the GAN ?



Discriminator D(x)
trained with standard cross entropy

$$L_D = -\mathbb{E}_{\mathbf{x}\sim p_{data}(\mathbf{x})}[\log(D(\mathbf{x}))] - \mathbb{E}_{z\sim p_{\mathbf{z}}(\mathbf{z})}[\log(1 - D(G(\mathbf{z})))]$$

Discriminator to represent the data



Nice data representation from adversarial training

How to get binary codes from discriminator ?

Discriminator D(x)

loss

$$\boxed{\begin{aligned} L_D = & -\mathbb{E}_{\mathbf{x}\sim p_{data}(\mathbf{x})}[\log(D(\mathbf{x}))] \\ & -\mathbb{E}_{z\sim p_{\mathbf{z}}(\mathbf{z})}[\log(1 - D(G(\mathbf{z})))] \end{aligned}}$$

BinGAN loss

$$L = L_D + \lambda_{BRE} \cdot L_{BRE} + \lambda_{DMR} \cdot L_{DMR}$$

How to get binary codes from discriminator ?

Discriminator D(x)
loss

$$\boxed{\begin{aligned} L_D = &-\mathbb{E}_{\mathbf{x}\sim p_{data}(\mathbf{x})}[\log(D(\mathbf{x}))] \\ &-\mathbb{E}_{z\sim p_{\mathbf{z}}(\mathbf{z})}[\log(1 - D(G(\mathbf{z})))] \end{aligned}}$$

BinGAN loss

$$L = L_D + \lambda_{BRE} \cdot L_{BRE} + \lambda_{DMR} \cdot L_{DMR}$$

Binary Representation Entropy (BRE)

How to get binary codes from discriminator ?

Discriminator D(x)
loss

$$L_D = -\mathbb{E}_{\mathbf{x}\sim p_{data}(\mathbf{x})}[\log(D(\mathbf{x}))] - \mathbb{E}_{z\sim p_{\mathbf{z}}(\mathbf{z})}[\log(1 - D(G(\mathbf{z})))]$$

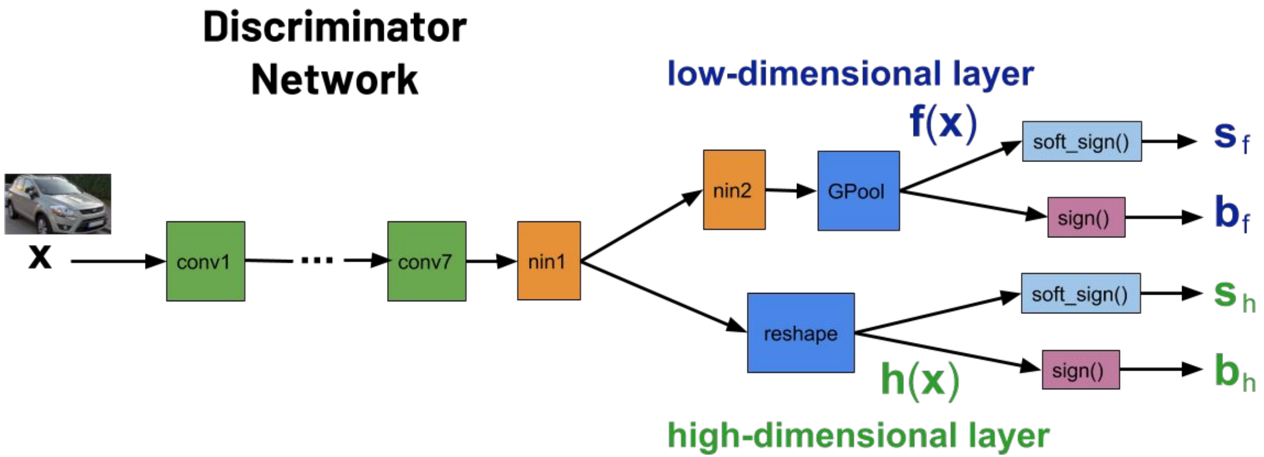
BinGAN loss

$$L = L_D + \lambda_{BRE} \cdot L_{BRE} + \lambda_{DMR} \cdot L_{DMR}$$

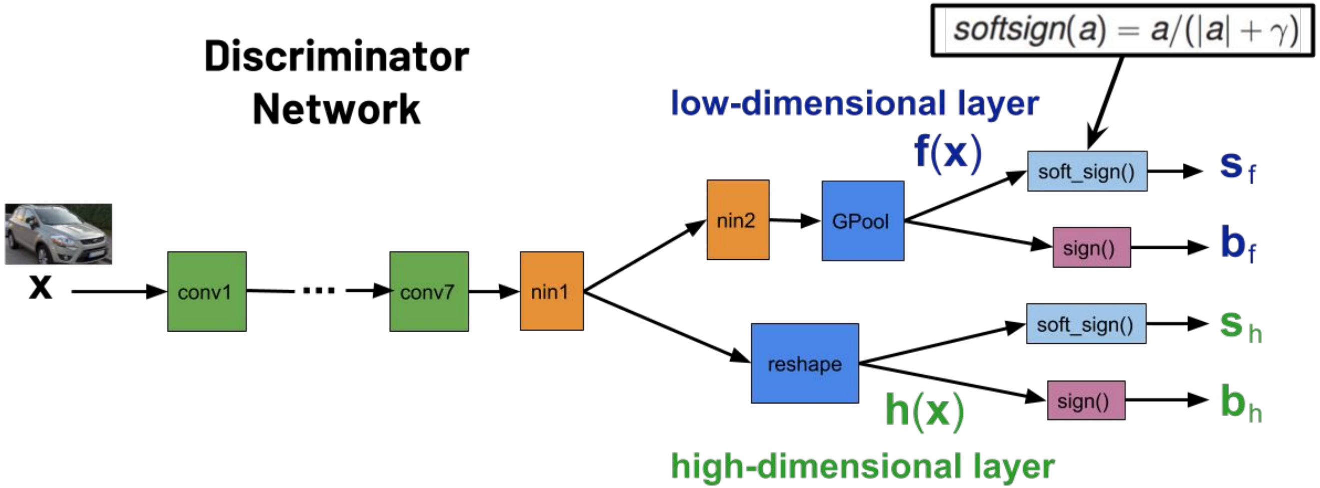
Binary Representation Entropy (BRE)

Distance Matching Regularizer (DMR)

BRE Regularizer



BRE Regularizer



BRE Regularizer

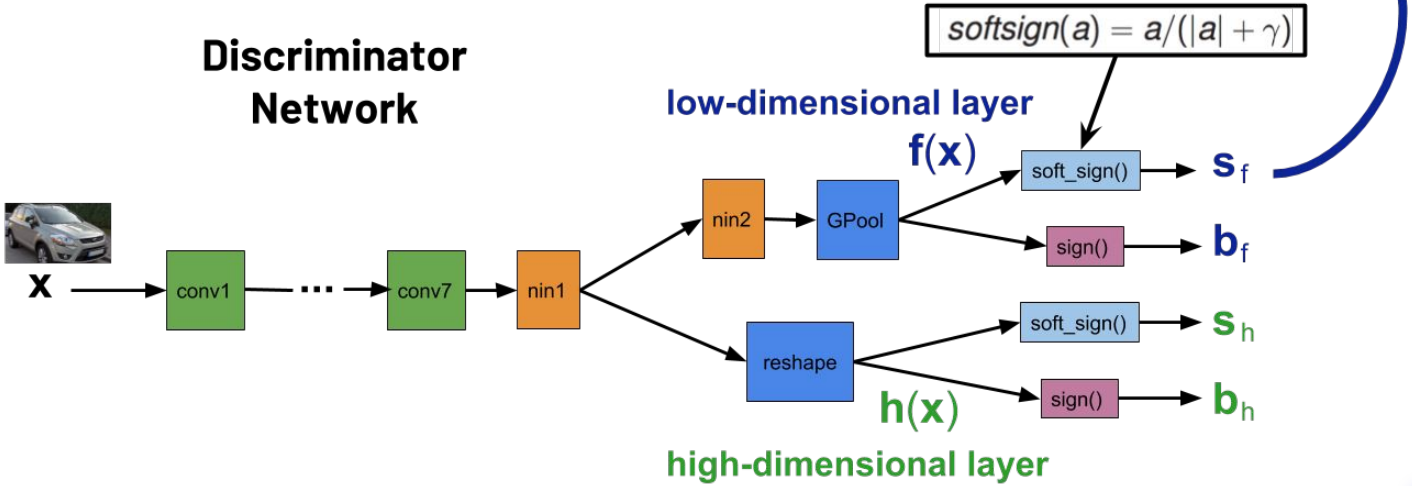
Increase diversity for descriptors with M/2 distance

$$\exp\left\{\frac{-|\mathbf{b}_{h,k}^{T}\mathbf{b}_{h,j}|}{\beta\cdot M}\right\}$$

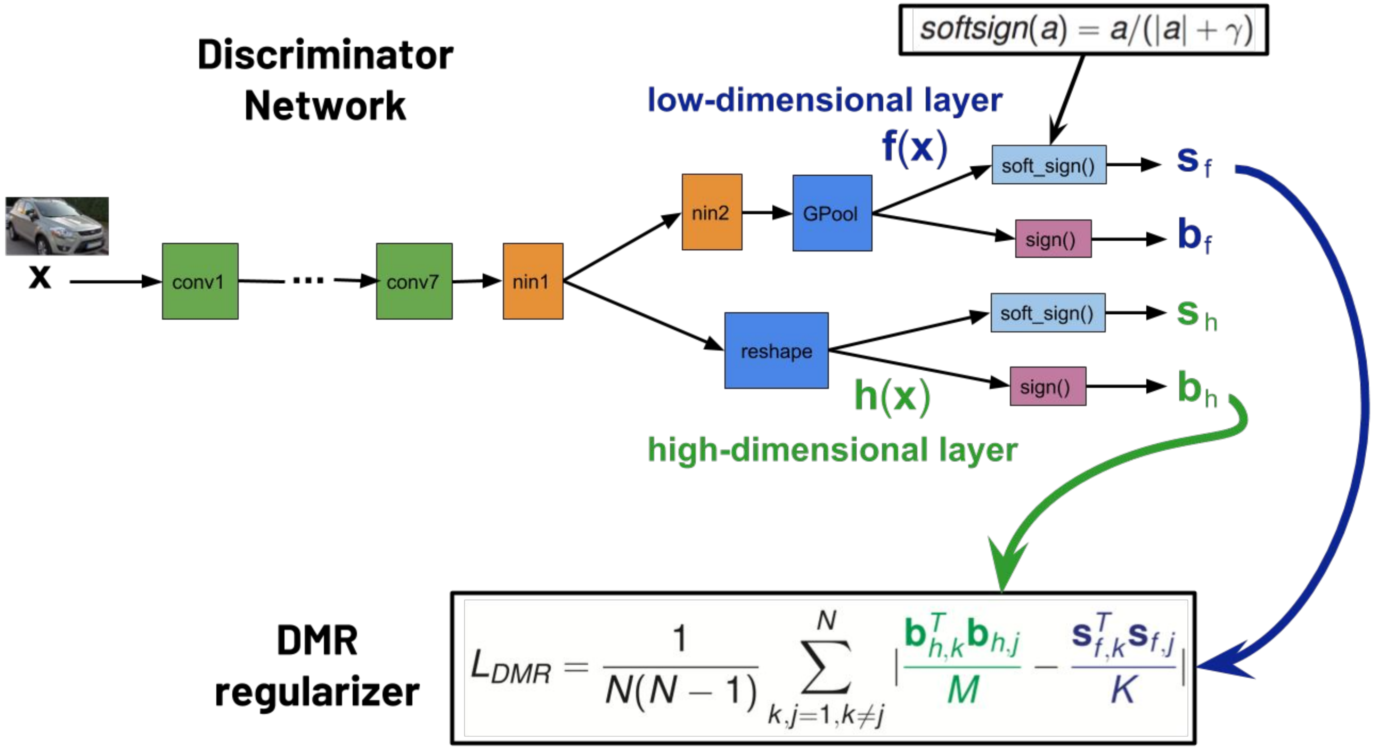
BRE regularizer

$$L_{BRE} = L_{ME} + L_{MAC} = \frac{1}{K}\sum_{k=1}^{K}(\bar{s}_{f,k})^2 + \sum_{k,j=1,k\neq j}^{N} \boxed{w_{k,j}} \frac{|\mathbf{s}_{f,k}^{T}\cdot \mathbf{s}_{f,j}|}{K}$$

Discriminator Network



DMR Regularizer



Ablation study

Train Test	Yosemite Notre Dame	 Liberty	Notre Dame Yosemite	 Liberty	Liberty Notre Dame	 Yosemite	Average FPR@95%
$\lambda_{DMR} = \lambda_{BRE} = 0$	32.72	39.44	**39.44**	27.92	27.24	50.48	36.21
$\lambda_{DMR} = 0 \quad \lambda_{BRE} = \mathbf{0.01}$	30.12	36.28	44.2	**24.28**	**26.44**	51.88	35.53
$\lambda_{DMR} = \mathbf{0.05} \quad \lambda_{BRE} = 0$	24.68	26.96	40.16	27.00	27.28	**45.28**	31.90
$\lambda_{DMR} = \mathbf{0.05} \quad \lambda_{BRE} = \mathbf{0.01}$	**16.88**	**26.08**	40.80	25.76	27.84	47.64	**30.76**

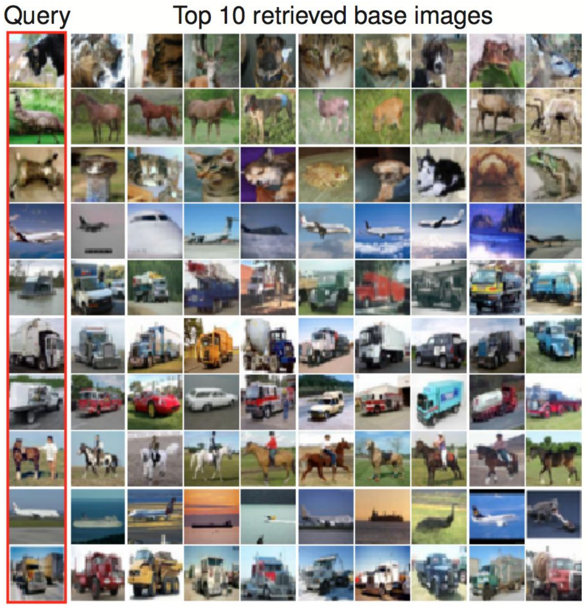
Results - image matching

Train	Yosemite		Notre Dame		Liberty		Average
Test	Notre Dame	Liberty	Yosemite	Liberty	Notre Dame	Yosemite	FPR@95%
Supervised							
LDAHash (16 bytes)	51.58	49.66	52.95	49.66	51.58	52.95	51.40
D-BRIEF (4 bytes)	43.96	53.39	46.22	51.30	43.10	47.29	47.54
BinBoost (8 bytes)	14.54	21.67	18.96	20.49	16.90	22.88	19.24
RFD (50-70 bytes)	11.68	19.40	14.50	19.35	13.23	16.99	15.86
Binary L2-Net (32 bytes)	2.51	6.65	4.04	4.01	1.9	5.61	4.12
Unsupervised							
SIFT (128 bytes)	28.09	36.27	29.15	36.27	28.09	29.15	31.17
BRISK (64 bytes)	74.88	79.36	73.21	79.36	74.88	73.21	75.81
BRIEF (32 bytes)	54.57	59.15	54.96	59.15	54.57	54.96	56.23
DeepBit (32 bytes)	29.60	34.41	63.68	32.06	26.66	57.61	40.67
DBD-MQ (32 bytes)	27.20	33.11	57.24	31.10	**25.78**	57.15	38.59
BinGAN (32 bytes)	**16.88**	**26.08**	**40.80**	**25.76**	27.84	**47.64**	**30.76**

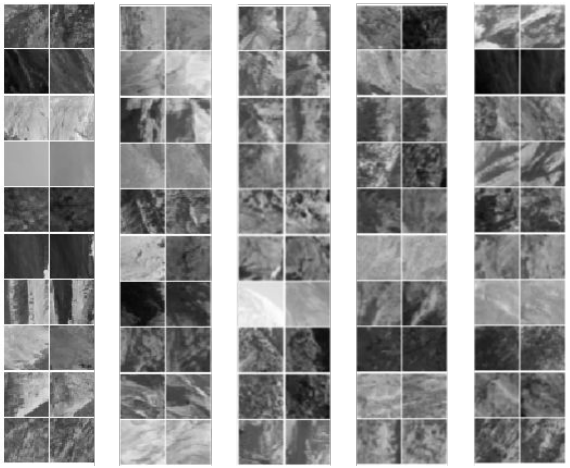
Results - image retrieval

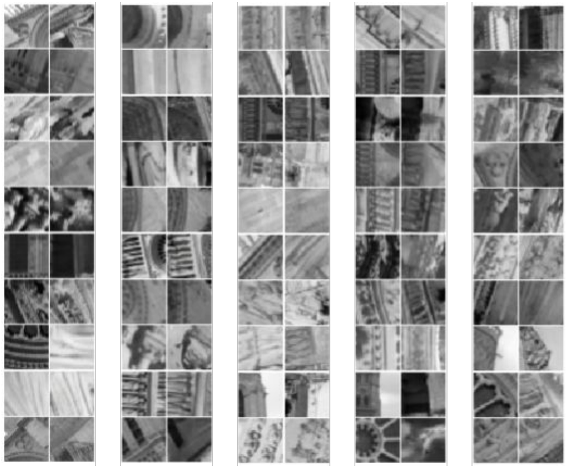
Mean Average Precision (**mAP**) - **top 1000**.

Method	16 bit	32 bit	64 bit
KHM	13.59	13.93	14.46
SphH	13.98	14.58	15.38
SpeH	12.55	12.42	12.56
SH	12.95	14.09	13.89
PCAH	12.91	12.60	12.10
LSH	12.55	13.76	15.07
PCA-ITQ	15.67	16.20	16.64
DH	16.17	16.62	16.96
DeepBit	19.43	24.86	27.73
DBD-MQ	21.53	26.50	31.85
BinGAN	**30.05**	**34.65**	**36.77**



BinGAN for data augmentation?

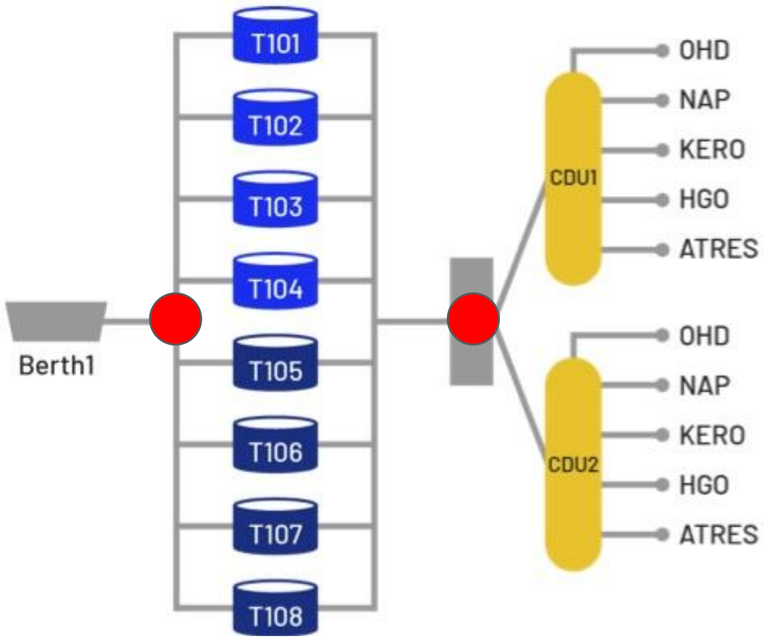




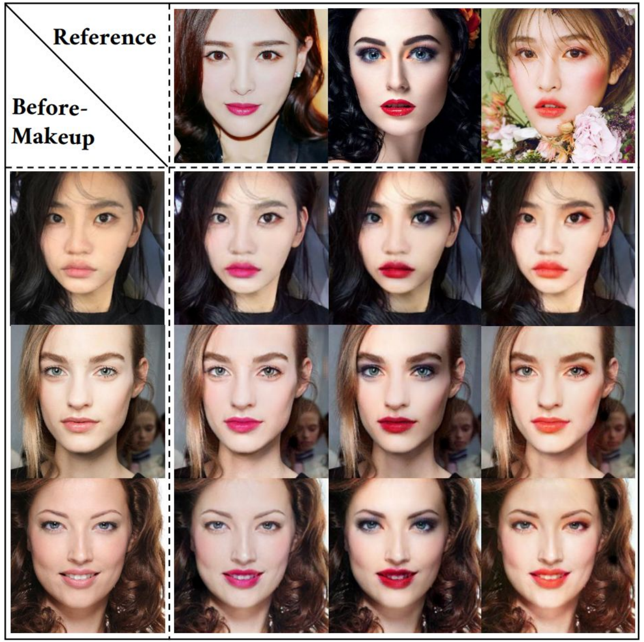
My involvement in commercial AI projects at Tooplоox

Reinforcement Learning

Dynamic System Modeling



Makeup transfer



Machine Learning - Classification

Predicting authorisation after phone contact

Wypełnij formularz

* Pola oznaczone gwiazdką są obowiązkowe

Kwota *

Imię *

Nazwisko *

Pesel *

Telefon *

Email

Akceptuję wszystkie poniższe zgody:

Bank od Kredytów | Santander Consumer Bank

&

TOOPLOOX AI

Call or not ?

TOOPLOOX AI

Thank you