Generative Adversarial Networks

PETR CEZNER

Obsah

Generativní modely

GAN

Odvozené modely

Evaluační metriky

Aplikace

Metody učení

- S učitelem
- Bez učitele

Učení s učitelem

- Data (x, y)
  - x – data
  - y – label
- Cílem: Naučit funkci, která mapuje $x \rightarrow y$
- Například:
  - Klasifikace
  - Regrese
  - Detekce objektů
  - …

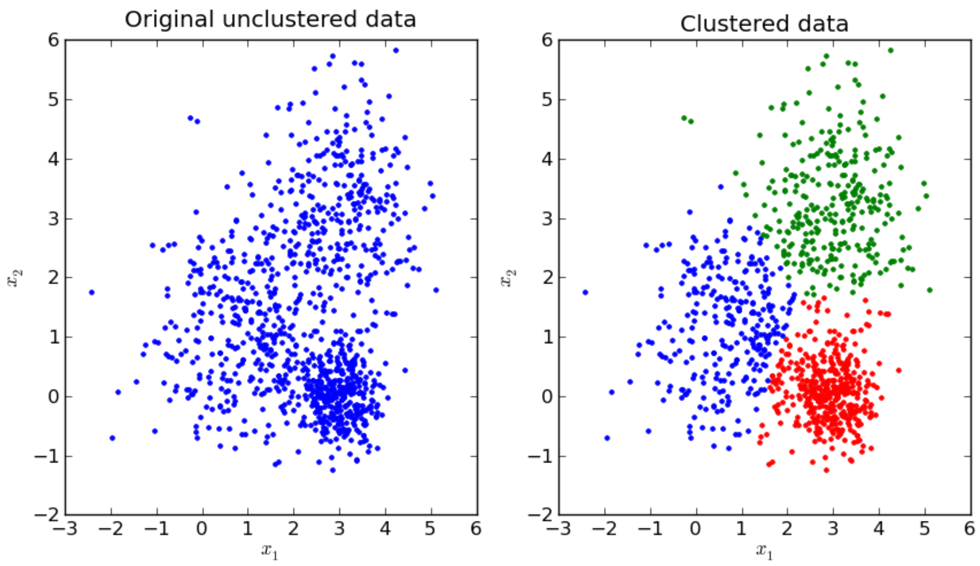
→ Kočka

Klasifikace

Učení bez učitele

- Data: x
  - Máme pouze data, bez labelů

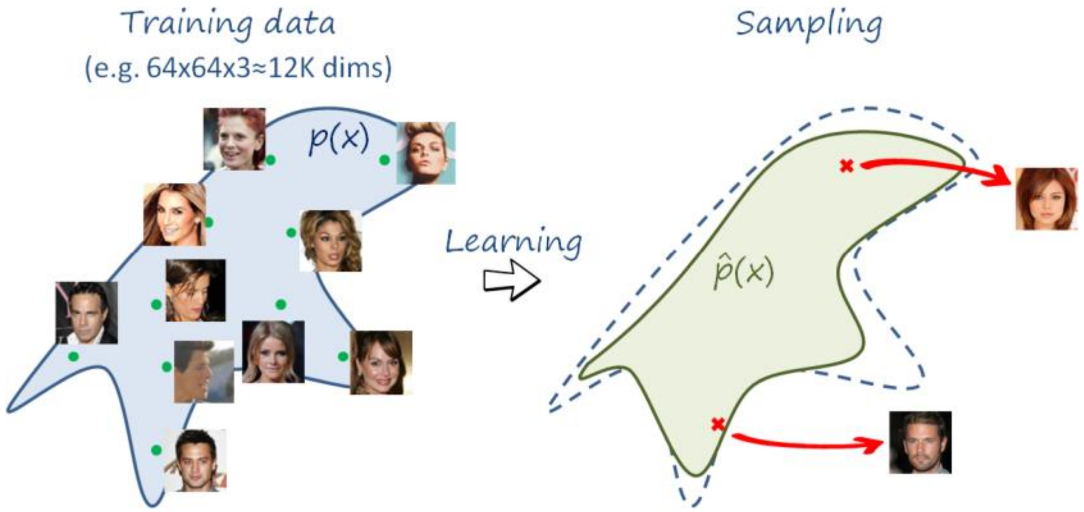
- Cíl: Naučit se nějaké základní, skryté informace z dat
- Například:
  - Clusterizace
  - Redukce dimenze
  - Naučení featur
  - Odhad hustoty pravděpodobnosti
  - …

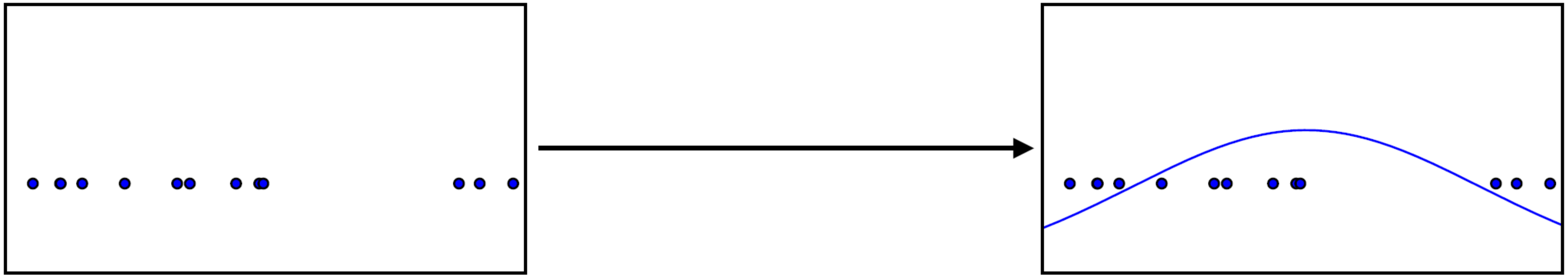


Clusterizace

Generativní modely

- Trénovací vzorek $\boldsymbol{x}$ pochází z neznámé distribuce $p_{data}(\boldsymbol{x})$
- Cílem generativních modelů je naučit se aproximaci $p_{data}(\boldsymbol{x})$ pomocí $p_{model}(\boldsymbol{x})$





Generativní modely: Maximum Likelihood Estimate

- Odhad pravděpodobnostního rozdělení pomocí parametru $\boldsymbol{\theta}$
- Aproximaci $p_{data}(\boldsymbol{x})$ pomocí $p_{model}(\boldsymbol{x};\boldsymbol{\theta})$

$$\begin{aligned}\boldsymbol{\theta}^* &= \arg\max_{\boldsymbol{\theta}} \prod_{i=1}^{m} p_{\text{model}}\left(\boldsymbol{x}^{(i)};\boldsymbol{\theta}\right)\\ &= \arg\max_{\boldsymbol{\theta}} \log \prod_{i=1}^{m} p_{\text{model}}\left(\boldsymbol{x}^{(i)};\boldsymbol{\theta}\right)\\ &= \arg\max_{\boldsymbol{\theta}} \sum_{i=1}^{m} \log p_{\text{model}}\left(\boldsymbol{x}^{(i)};\boldsymbol{\theta}\right)\end{aligned}$$

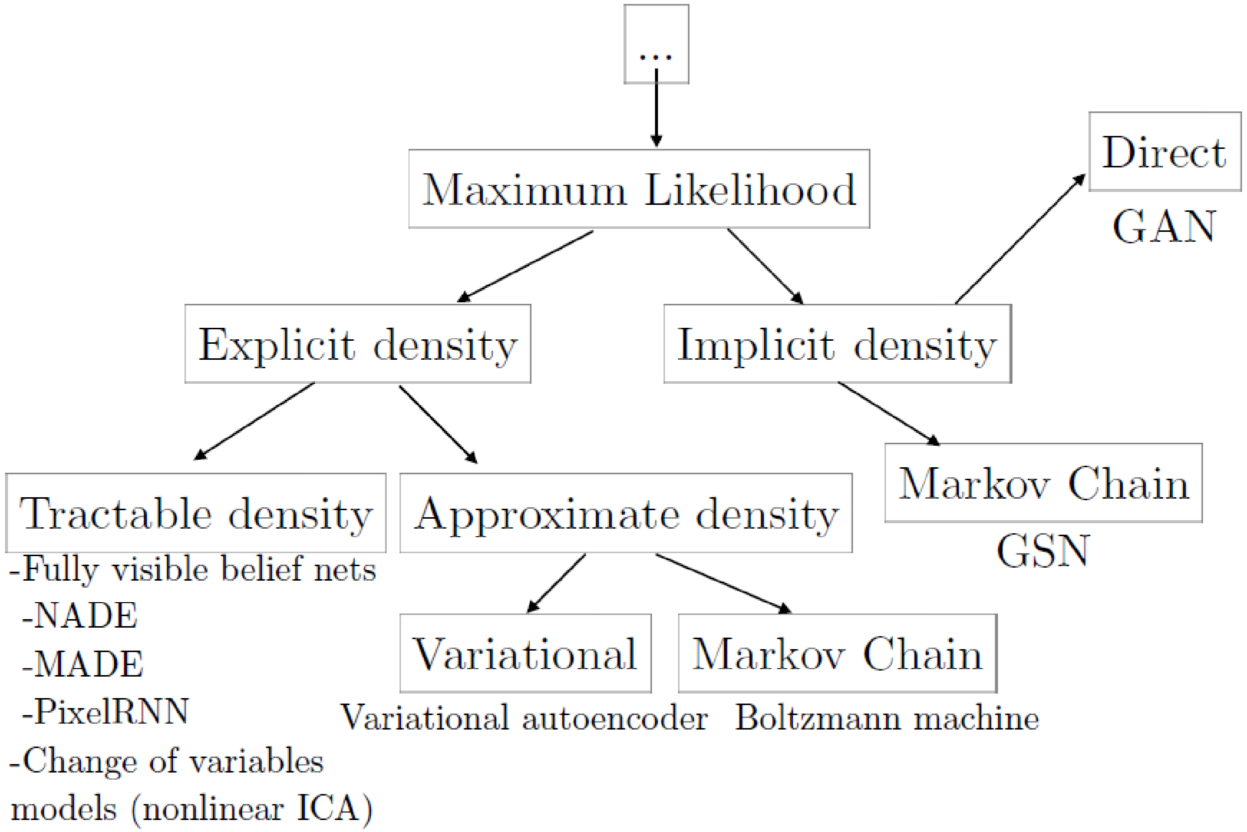
- ML odhad se dá taky definovat jako minimalizace KL (Kullback - Leibler) divergence mezi daty generující distribuci a modelem

$$\boldsymbol{\theta}^* = \arg\min_{\boldsymbol{\theta}} D_{\text{KL}}\left(p_{\text{data}}(\boldsymbol{x}) \| p_{\text{model}}(\boldsymbol{x};\boldsymbol{\theta})\right)$$

Generativní modely: Maximum Likelihood Estimate

- Obě dvě varianty počítají to samé
- Výhodou je to že můžeme MLE použít jako proxy pro odhad skutečného rozložení
- To by pomocí přímé metody nebylo možné, protože skutečné rozložení je neznámé

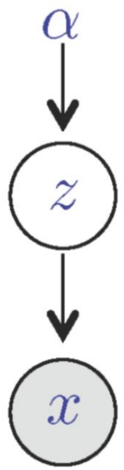
Generativní modely: Dělení Generativních modelů



Goodfellow et Al.

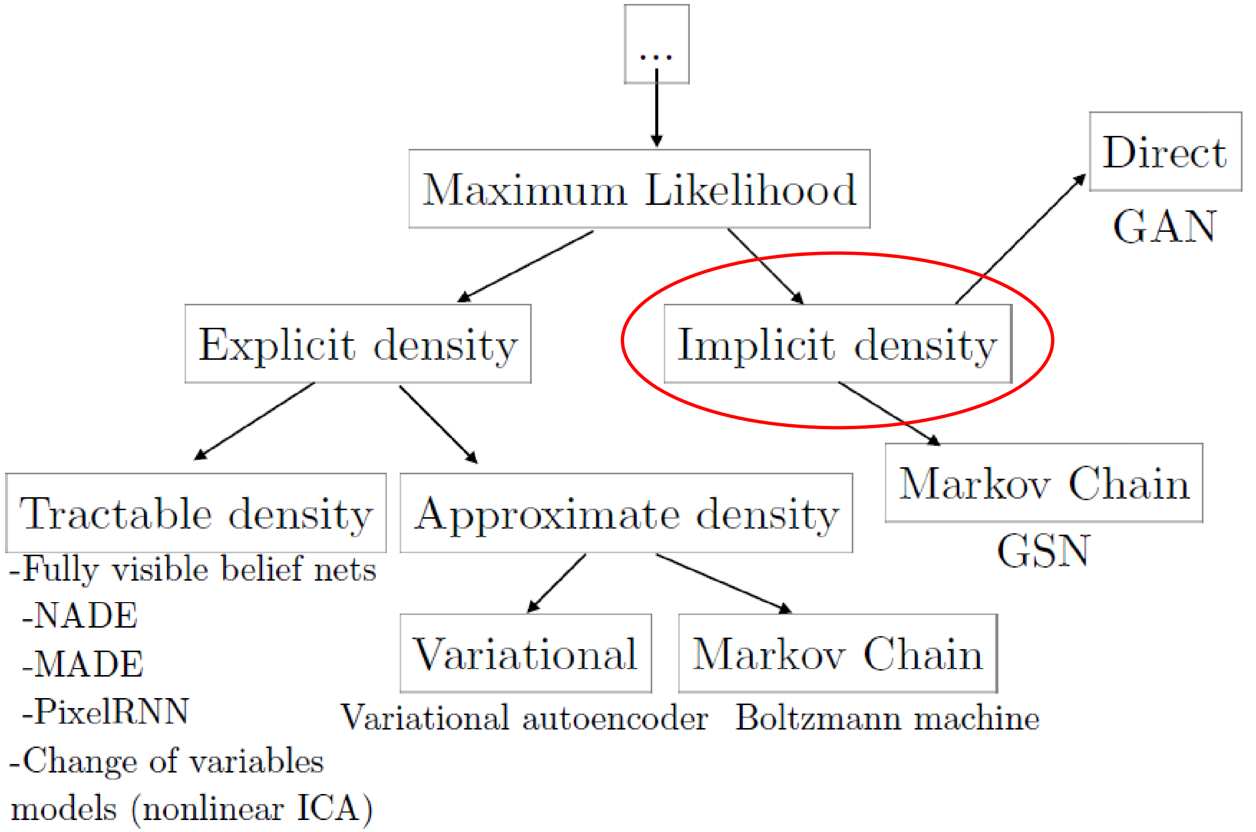
Generativní modely: Explicitní modely

- Poskytují explicitní parametrickou specifikaci distribuce dat
- Jsou schopné přiřadit pravděpodobnost vzorku
- Hlavním problémem je navržení modelu, který je schopný zachytit veškerou komplexnost data a zachovat si výpočetní složitost



$$p(x, z|\alpha) = p(x|z)p(z|\alpha)$$

Generativní modely:
Dělení Generativních modelů

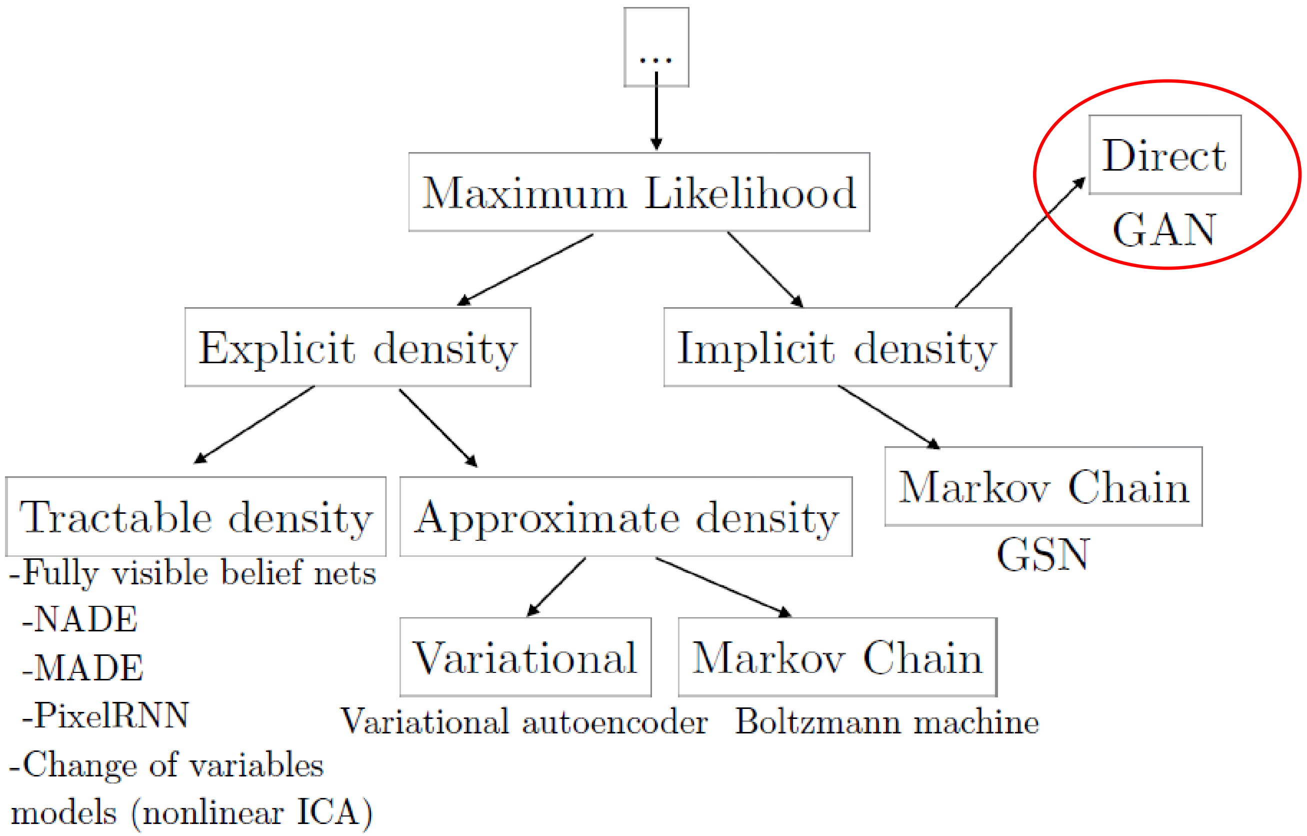


Goodfellow et Al.

Generativní modely: Implicitní modely

- Modely, které nepotřebují explicitně definovat funkci hustoty pravděpodobnosti (densiti function)
- Specifikují stochastický proces, který po naučení náhodně vybírá vzorky z distribuce dat
- Simulují data
- Protože nepotřebují specificky znát distribuci, implicitní modely nepotřebují

Generativní modely:
Dělení Generativních modelů

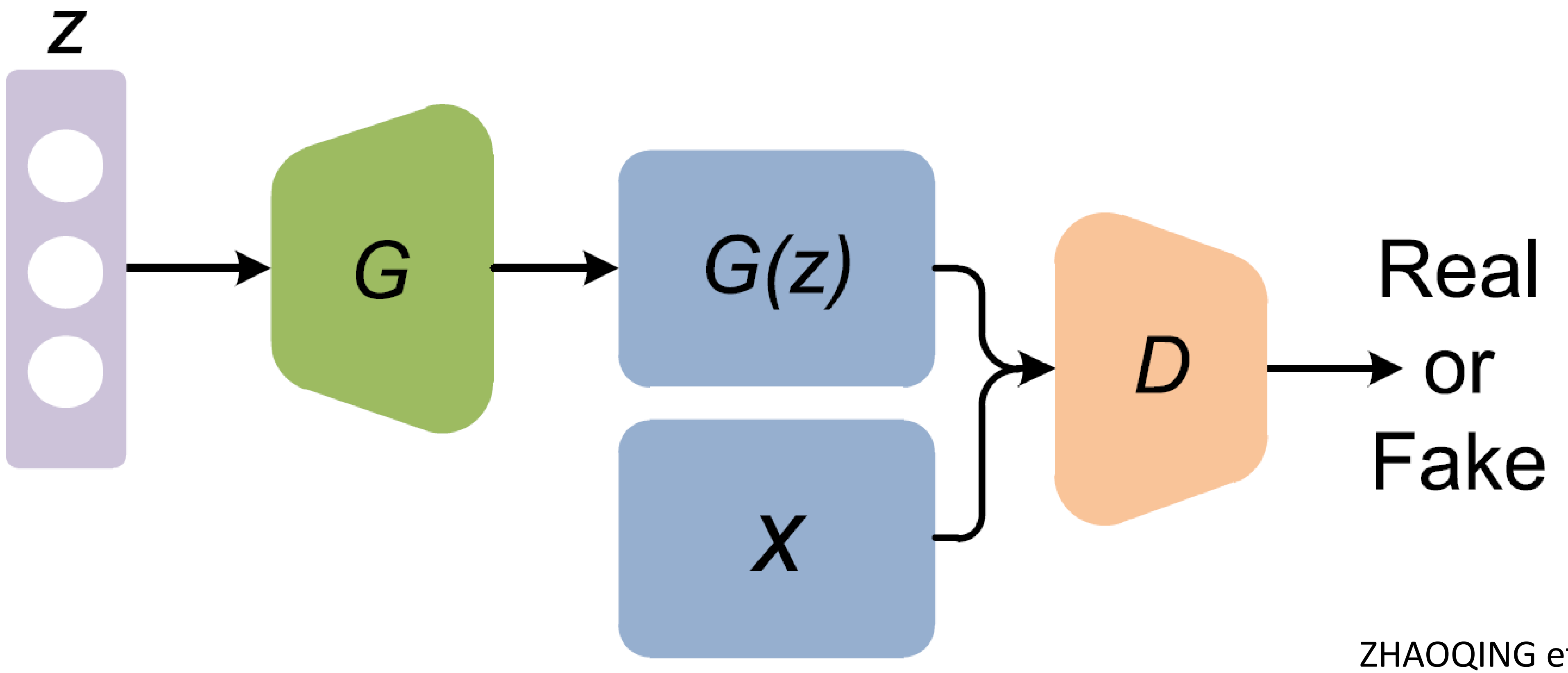


Goodfellow et Al.

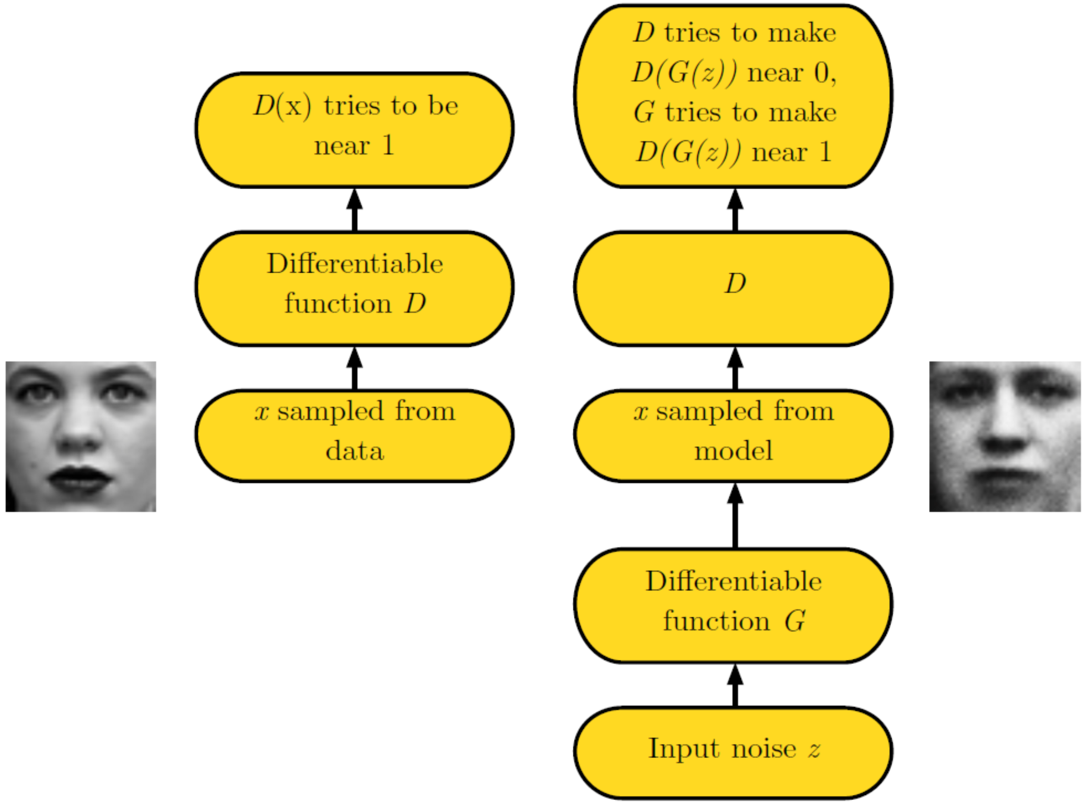
GAN: Historie

- Poprvé publikováno v roce 2014
- *Goodfellow, Ian; Pouget-Abadie, Jean; Mirza, Mehdi; Xu, Bing; Warde-Farley, David; Ozair, Sherjil; Courville, Aaron; Bengio, Yoshua (2014). Generative Adversarial Networks (PDF). Proceedings of the International Conference on Neural Information Processing Systems (NIPS 2014). pp. 2672–2680*
- Ian Goodfellow
  - Pracoval v Googlu, OpenAI
  - Pracuje v Applu
  - Napsal: *Deep Learning*

GAN



ZHAOQING et Al.



Goodfellow et Al.

GAN: Matematická formulace

- Každý hráč optimalizuje svojí loss funkci: $J^{(D)}, J^{(G)}$

$$J^{(D)}(\boldsymbol{\theta}^{(D)}, \boldsymbol{\theta}^{(G)}) = -\frac{1}{2}\mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}} \log D(\boldsymbol{x}) - \frac{1}{2}\mathbb{E}_{\boldsymbol{z}} \log\left(1 - D\left(G(z)\right)\right)$$

$$J^{(G)} = -J^{(D)}$$

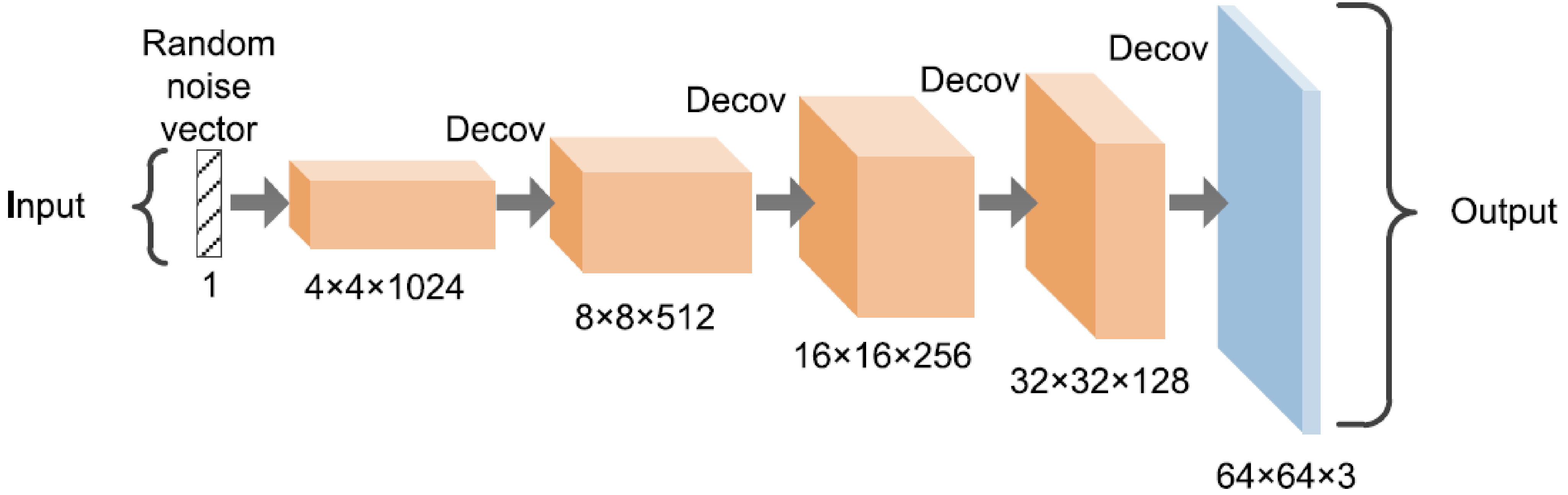
$$\min_G \max_D V(D, G) = \mathbb{E}_{\boldsymbol{x}\sim p_{\text{data}}(\boldsymbol{x})}[\log D(\boldsymbol{x})] + \mathbb{E}_{\boldsymbol{z}\sim p_{\boldsymbol{z}}(\boldsymbol{z})}[\log(1 - D(G(\boldsymbol{z})))]$$

Odvozené modely

Architecture Optimization Based GANs	Convolution based GANs	DCGAN [12]
	Condition based GANs	CGANs [13]; InfoGAN [14]; ACGAN [15]
	Autoencoder based GANs	AAE [16]; BiGAN [17]; ALI [18]; AGE [19]; VAE-GAN [20]
Objective Function Optimization Based GANs	unrolled GAN [21]; f-GAN [22]; Mode-Regularized GAN [23]; Least-Square GAN [24]; Loss-Senstive GAN [25]; EBGAN [26]; WGAN [27]; WGAN-GP [28]; WGAN-LP [29]	

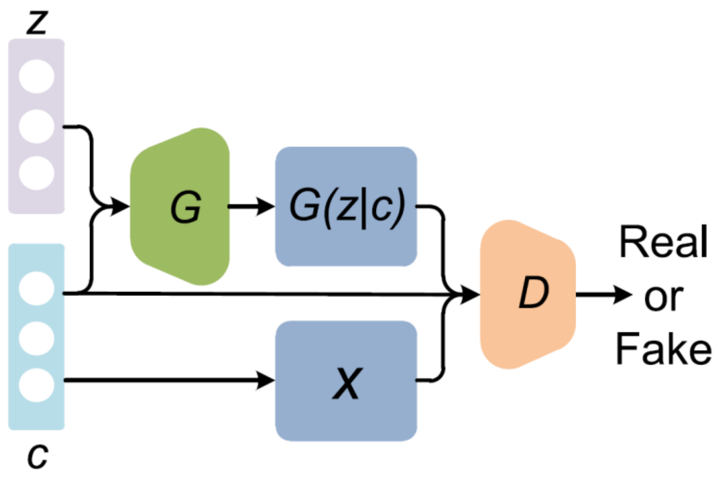
ZHAOQING et Al.

Odvozené modely: Convolution based GANs



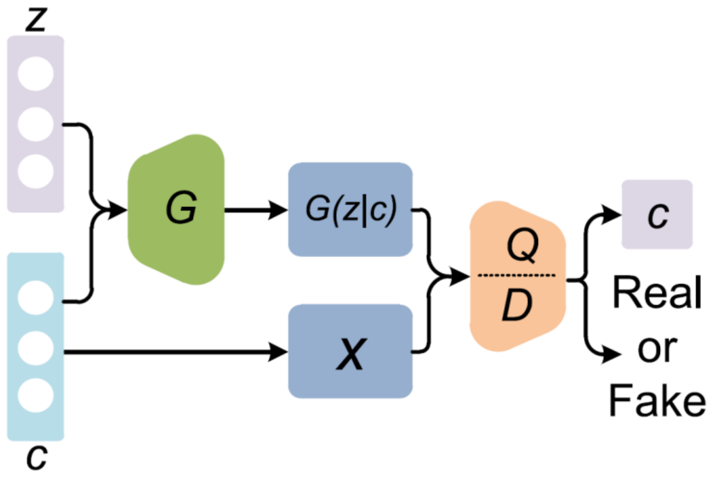
ZHAOQING et Al.

Odvozené modely:
Condition based GANs



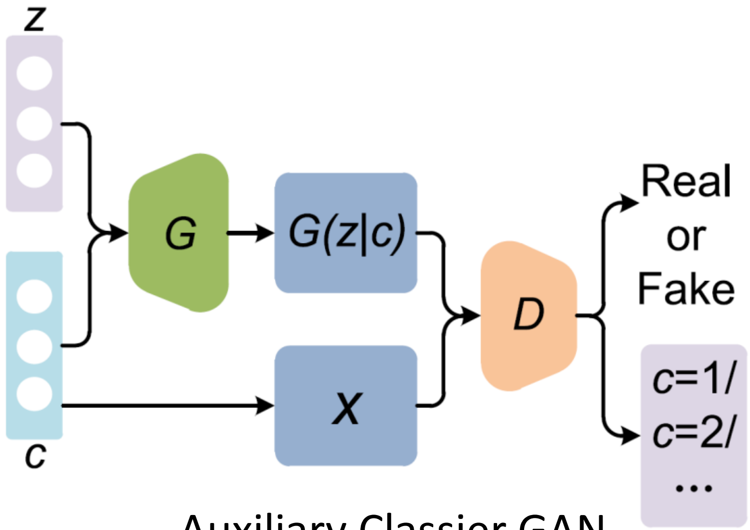
Conditional Generative Adversarial Networks
(CGAN)

$$\min_G \max_D V(D,G) = \mathbb{E}_{x\sim pdata(x)}\left[\log D(x|c)\right] + \mathbb{E}_{z\sim p(z)}\left[\log\left(1 - D\left(G\left(z|c\right)\right)\right)\right]$$



InfoGAN

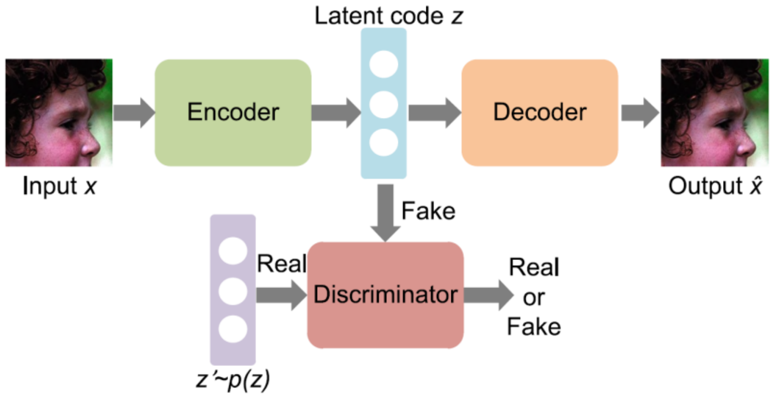
$$\min_G \max_D V(D,G) - \lambda I\left(c, G\left(z,c\right)\right)$$



Auxiliary Classier GAN
(ACGAN)

ZHAOQING et Al.

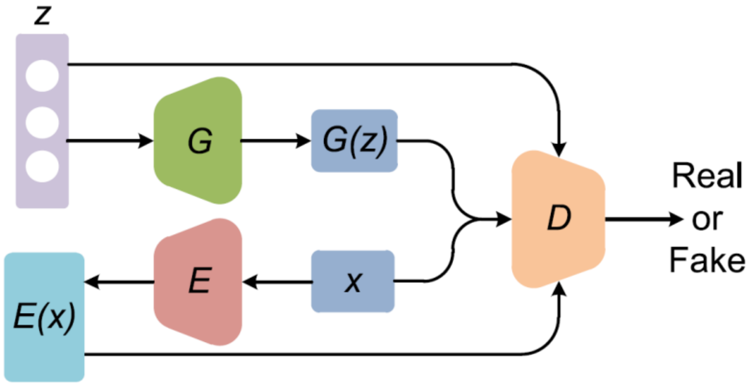
Odvozené modely: AutoEncoder Based GANs



Adversarial Autoencoder
(AAG)

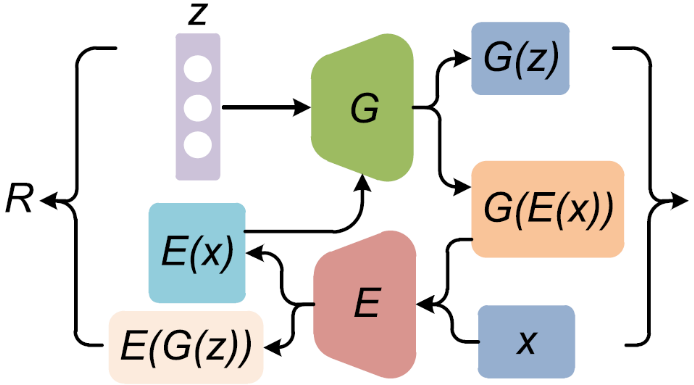
$$\mathcal{L} = \mathbb{E}_x\left[\underbrace{\mathbb{E}_{q(z|x)}[-\log p(x|z)]}_{\text{Reconstruction Error}}\right] + \mathbb{E}_x\left[\underbrace{\mathrm{KL}(q(z|x)||p(z))}_{\text{KL Regularizer}}\right]$$

Replaced by adversarial loss in AAE



Bidirectional Generative Adversarial
Networks
(BiGAN)

$$V(D,E,G) := \mathbb{E}_{\mathbf{x}\sim p_{\mathbf{X}}}\Big[\underbrace{\mathbb{E}_{\mathbf{z}\sim p_E(\cdot|\mathbf{x})}\left[\log D(\mathbf{x},\mathbf{z})\right]}_{\log D(\mathbf{x},E(\mathbf{x}))}\Big] + \mathbb{E}_{\mathbf{z}\sim p_{\mathbf{Z}}}\Big[\underbrace{\mathbb{E}_{\mathbf{x}\sim p_G(\cdot|\mathbf{z})}\left[\log\left(1-D(\mathbf{x},\mathbf{z})\right)\right]}_{\log(1-D(G(\mathbf{z}),\mathbf{z}))}\Big]$$



Adversarial Generator-Encoder Network
(AGE)

ZHAOQING et Al.

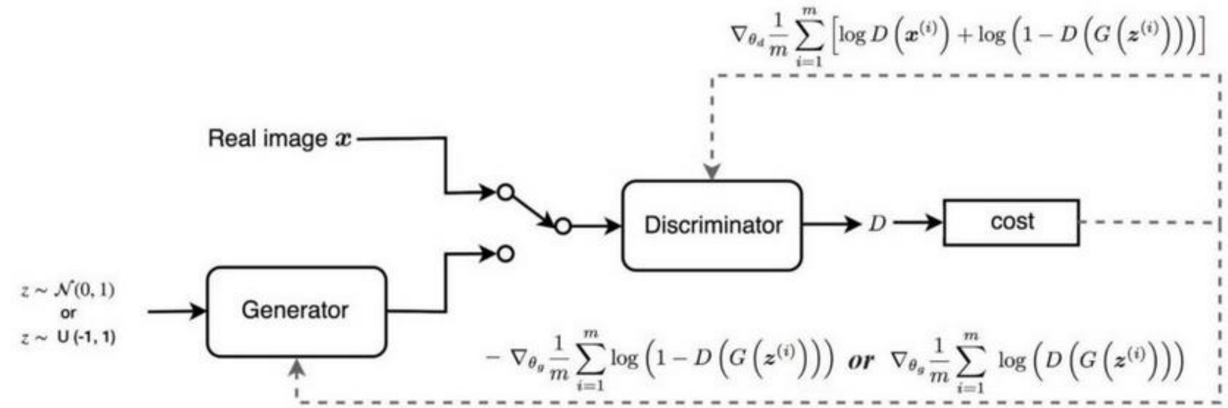
Odvozené modely

Architecture Optimization Based GANs	Convolution based GANs	DCGAN [12]
	Condition based GANs	CGANs [13]; InfoGAN [14]; ACGAN [15]
	Autoencoder based GANs	AAE [16]; BiGAN [17]; ALI [18]; AGE [19]; VAE-GAN [20]
Objective Function Optimization Based GANs	unrolled GAN [21]; f-GAN [22]; Mode-Regularized GAN [23]; Least-Square GAN [24]; Loss-Senstive GAN [25]; EBGAN [26]; WGAN [27]; WGAN-GP [28]; WGAN-LP [29]	

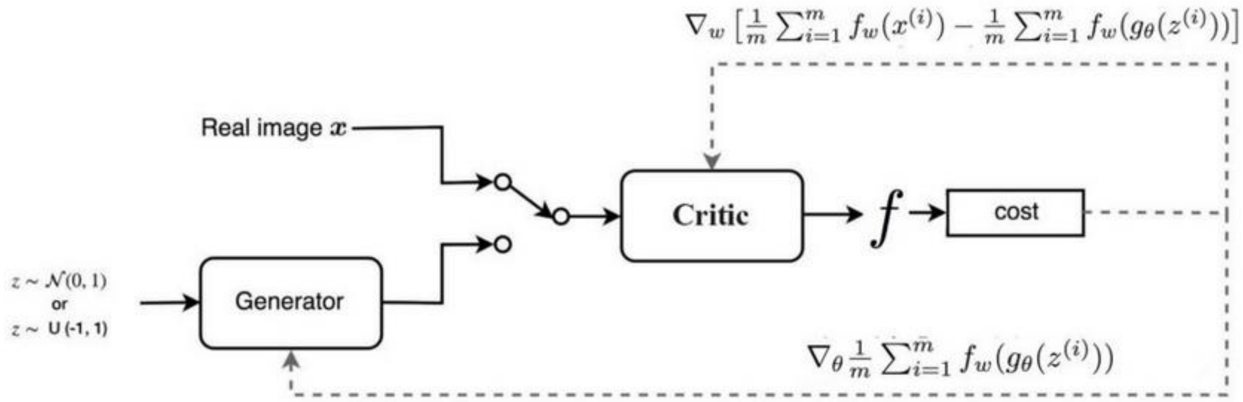
ZHAOQING et Al.

Odvozené modely:
Objective Function Optimization Based GANs

GAN



WGAN

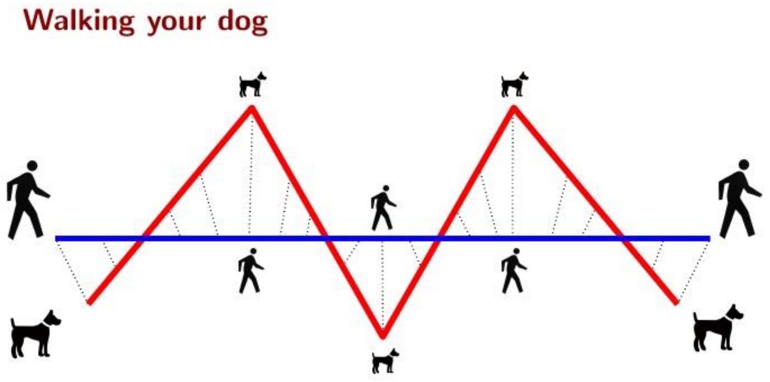


Evaluační metriky

- Inception Score (IS)
  - Na vstupu má seznam obrázků a vrátí jedno číslo, skóre
  - K tomu využívá Inception model
  - Hodně používaná metrika, měří jak moc je výstup z Generátoru podobný s lablem:
    - Pomocí Inception modelu spočítá $p(x|y)$
    - Celkové skóre je pak dáno: $\exp(\mathbb{E}_x KL(p(y|x)||p(y)))$
  - Vysoké IS skóre indikuje to, že generovaný model je schopný kvalitní vzorky.
  - Ale pokud model spadne do Mode Collapse, tak metrika stále vrací vysoké skóre.
- Mode Score (MS)
  - Založeno na IS, ale v potaz prior labelů z dat
  - Schopné rozlišovat různost a vizuální kvalitu generovaných vzorků
  - $\exp(\mathbb{E}_x KL(p(y|x)||p\left(y^{train}\right)) - \mathbb{E}_x KL(p(y)||p\left(y^{train}\right)))$

Evaluační metriky

- FRéCHET Inception Distance (FID)
  - Založeno na IS
  - Počítá jak jsou si dvě skupiny obrázku podobné na základě featur raw obrázků spočítané pomocí Inception V3 modelu
  - Specificky používá poslední poolovací vrstvu před výstupem klasifikátoru
  - Nižší skóre znamená, že jsou si skupiny více podobné



The **Fréchet distance** between the curves is the minimum leash length that permits such a walk

	Measure	Description
Quantitative	1. Average Log-likelihood [18, 22]	• Log likelihood of explaining realworld held out/test data using a density estimated from the generated data (*e.g.* using KDE or Parzen window estimation). $L = \frac{1}{N}\sum_i \log P_{model}(\mathbf{x}_i)$
	2. Coverage Metric [33]	• The probability mass of the true data "covered" by the model distribution $C := P_{data}(dP_{model} > t)$ with t such that $P_{model}(dP_{model} > t) = 0.95$
	3. Inception Score (IS) [3]	• KLD between conditional and marginal label distributions over generated data. $\exp\left(\mathbb{E}_{\mathbf{x}}\left[\mathbb{KL}\left(p\left(y \mid \mathbf{x}\right) \parallel p\left(y\right)\right)\right]\right)$
	4. Modified Inception Score (m-IS) [34]	• Encourages diversity within images sampled from a particular category. $\exp(\mathbb{E}_{\mathbf{x}_i}[\mathbb{E}_{\mathbf{x}_j}[(\mathbb{KL}(P(y\|\mathbf{x}_i)\|\|P(y\|\mathbf{x}_j))]])$
	5. Mode Score (MS) [35]	• Similar to IS but also takes into account the prior distribution of the labels over real data. $\exp\left(\mathbb{E}_{\mathbf{x}}\left[\mathbb{KL}\left(p\left(y \mid \mathbf{x}\right) \parallel p\left(y^{train}\right)\right)\right] - \mathbb{KL}\left(p\left(y\right) \parallel p\left(y^{train}\right)\right)\right)$
	6. AM Score [36]	• Takes into account the KLD between distributions of training labels vs. predicted labels, as well as the entropy of predictions. $\mathbb{KL}(p(y^{\mathrm{train}}) \parallel p(y)) + \mathbb{E}_{\mathbf{x}}\left[H(y\|\mathbf{x})\right]$
	7. Fréchet Inception Distance (FID) [37]	• Wasserstein-2 distance between multi-variate Gaussians fitted to data embedded into a feature space $FID(r,g) = \|\|\mu_r - \mu_g\|\|_2^2 + Tr(\Sigma_r + \Sigma_g - 2(\Sigma_r\Sigma_g)^{\frac{1}{2}})$
	8. Maximum Mean Discrepancy (MMD) [38]	• Measures the dissimilarity between two probability distributions P_r and P_g using samples drawn independently from each distribution. $M_k(P_r,P_g) = \mathbb{E}_{\mathbf{x},\mathbf{x}'\sim P_r}[k(\mathbf{x},\mathbf{x}')] - 2\mathbb{E}_{\mathbf{x}\sim P_r,\mathbf{y}\sim P_g}[k(\mathbf{x},\mathbf{y})] + \mathbb{E}_{\mathbf{y},\mathbf{y}'\sim P_g}[k(\mathbf{y},\mathbf{y}')]$
	9. The Wasserstein Critic [39]	• The critic (*e.g.* an NN) is trained to produce high values at real samples and low values at generated samples $\hat{W}(\mathbf{x}_{test},\mathbf{x}_g) = \frac{1}{N}\sum_{i=1}^N \hat{f}(\mathbf{x}_{test}[i]) - \frac{1}{N}\sum_{i=1}^N \hat{f}(\mathbf{x}_g[i])$
	10. Birthday Paradox Test [27]	• Measures the support size of a discrete (continuous) distribution by counting the duplicates (near duplicates)
	11. Classifier Two Sample Test (C2ST) [40]	• Answers whether two samples are drawn from the same distribution (*e.g.* by training a binary classifier)
	12. Classification Performance [1, 15]	• An indirect technique for evaluating the quality of unsupervised representations (*e.g.* feature extraction; FCN score). See also the GAN Quality Index (GQI) [41].
	13. Boundary Distortion [42]	• Measures diversity of generated samples and covariate shift using classification methods.
	14. Number of Statistically-Different Bins (NDB) [43]	• Given two sets of samples from the same distribution, the number of samples that fall into a given bin should be the same up to sampling noise
	15. Image Retrieval Performance [44]	• Measures the distributions of distances to the nearest neighbors of some query images (*i.e.* diversity)
	16. Generative Adversarial Metric (GAM) [31]	• Compares two GANs by having them engaged in a battle against each other by swapping discriminators or generators. $p(\mathbf{x}\|y=1;M_1')/p(\mathbf{x}\|y=1;M_2') = (p(y=1\|\mathbf{x};D_1)p(\mathbf{x};G_2))/(p(y=1\|\mathbf{x};D_2)p(\mathbf{x};G_1))$
	17. Tournament Win Rate and Skill Rating [45]	• Implements a tournament in which a player is either a discriminator that attempts to distinguish between real and fake data or a generator that attempts to fool the discriminators into accepting fake data as real.
	18. Normalized Relative Discriminative Score (NRDS) [32]	• Compares n GANs based on the idea that if the generated samples are closer to real ones, more epochs would be needed to distinguish them from real samples.
	19. Adversarial Accuracy and Divergence [46]	• Adversarial Accuracy. Computes the classification accuracies achieved by the two classifiers, one trained on real data and another on generated data, on a labeled validation set to approximate $P_g(y\|\mathbf{x})$ and $P_r(y\|\mathbf{x})$. Adversarial Divergence: Computes $\mathbb{KL}(P_g(y\|\mathbf{x}), P_r(y\|\mathbf{x}))$
	20. Geometry Score [47]	• Compares geometrical properties of the underlying data manifold between real and generated data.
	21. Reconstruction Error [48]	• Measures the reconstruction error (*e.g.* L_2 norm) between a test image and its closest generated image by optimizing for z (*i.e.* $min_{\mathbf{z}}\|\|G(\mathbf{z}) - \mathbf{x}^{(test)}\|\|^2$)
	22. Image Quality Measures [49, 50, 51]	• Evaluates the quality of generated images using measures such as SSIM, PSNR, and sharpness difference
	23. Low-level Image Statistics [52, 53]	• Evaluates how similar low-level statistics of generated images are to those of natural scenes in terms of mean power spectrum, distribution of random filter responses, contrast distribution, etc.
	24. Precision, Recall and F_1 score [23]	• These measures are used to quantify the degree of overfitting in GANs, often over toy datasets.
Qualitative	1. Nearest Neighbors	• To detect overfitting, generated samples are shown next to their nearest neighbors in the training set
	2. Rapid Scene Categorization [18]	• In these experiments, participants are asked to distinguish generated samples from real images in a short presentation time (*e.g.* 100 ms); *i.e.* real v.s fake
	3. Preference Judgment [54, 55, 56, 57]	• Participants are asked to rank models in terms of the fidelity of their generated images (*e.g.* pairs, triples)
	4. Mode Drop and Collapse [58, 59]	• Over datasets with known modes (*e.g.* a GMM or a labeled dataset), modes are computed as by measuring the distances of generated data to mode centers
	5. Network Internals [1, 60, 61, 62, 63, 64]	• Regards exploring and illustrating the internal representation and dynamics of models (*e.g.* space continuity) as well as visualizing learned features

Borji

Aplikace

- Počítačové vidění
  - Image Super-Resolution
  - Image Translation
  - Texture Synthesis
  - Face Synthesis
- NLP
  - Pro tvorbu hudby, poezie a textů
- Ostatní

Image Super-Resolution

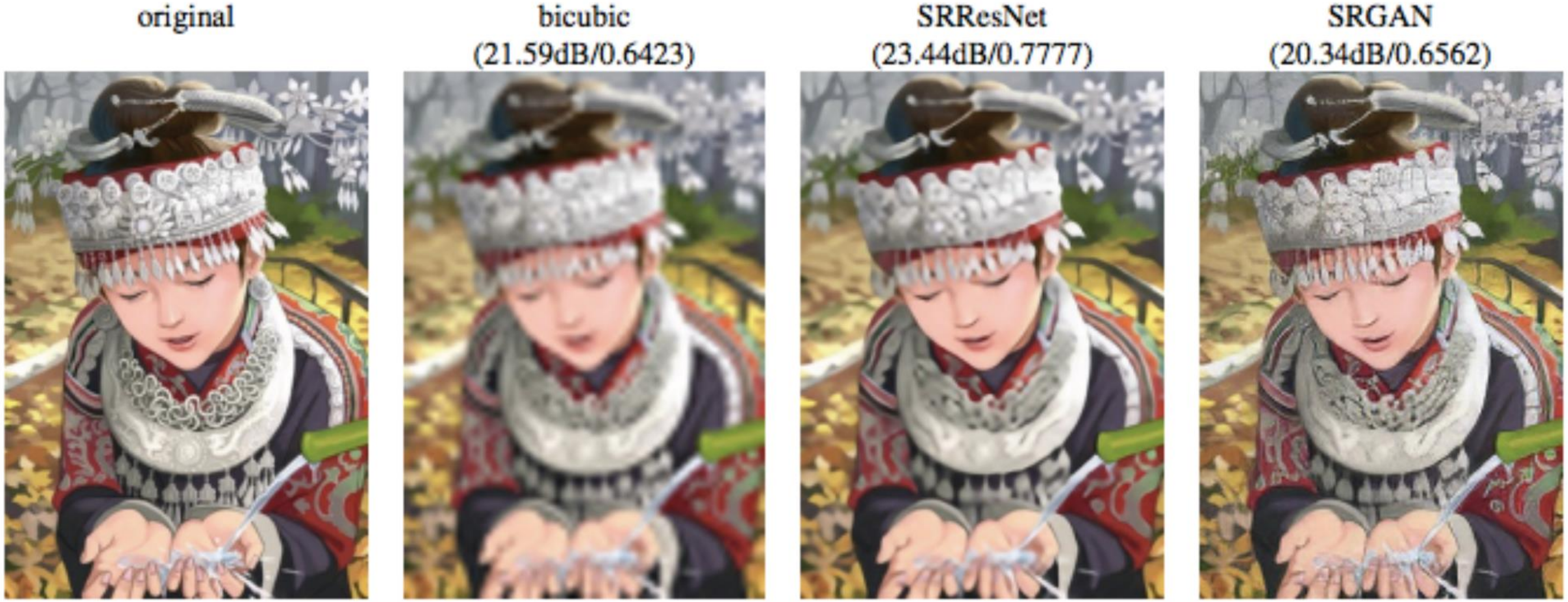


Image Super-Resolution

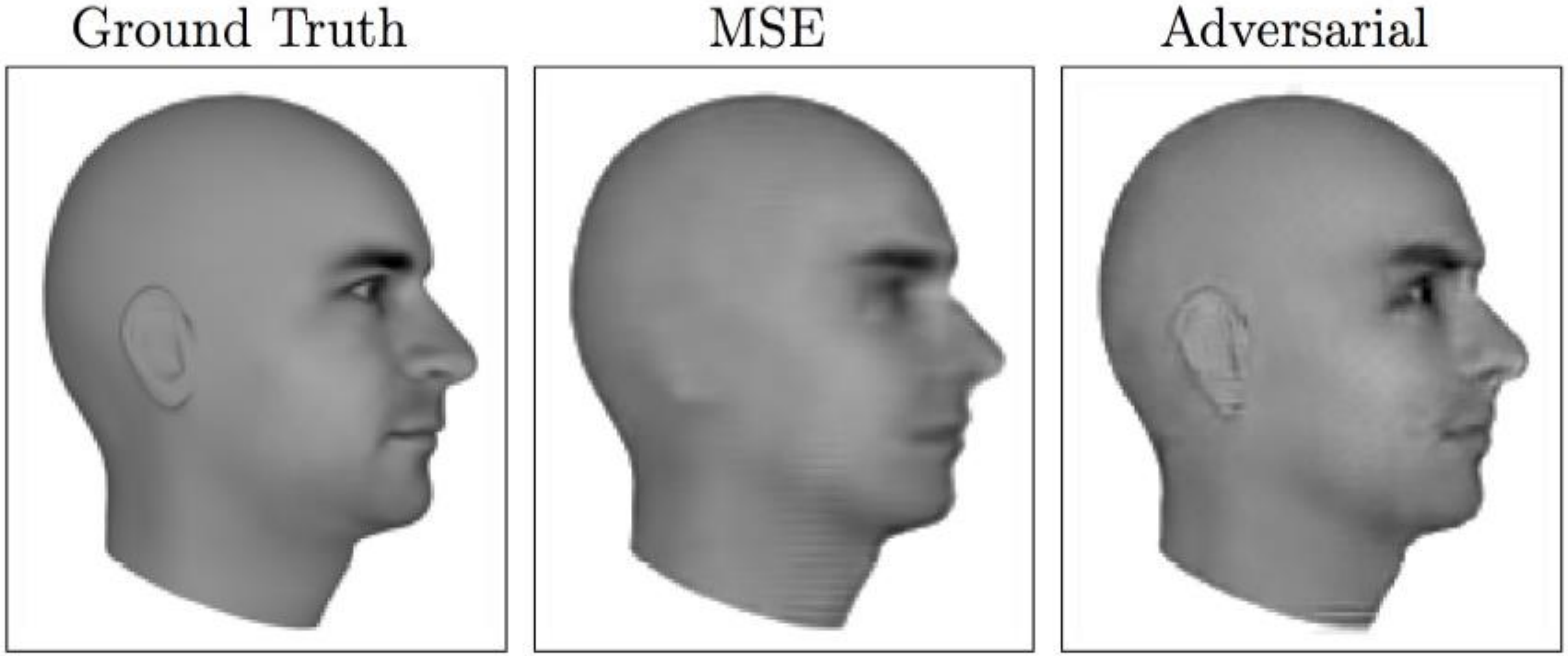
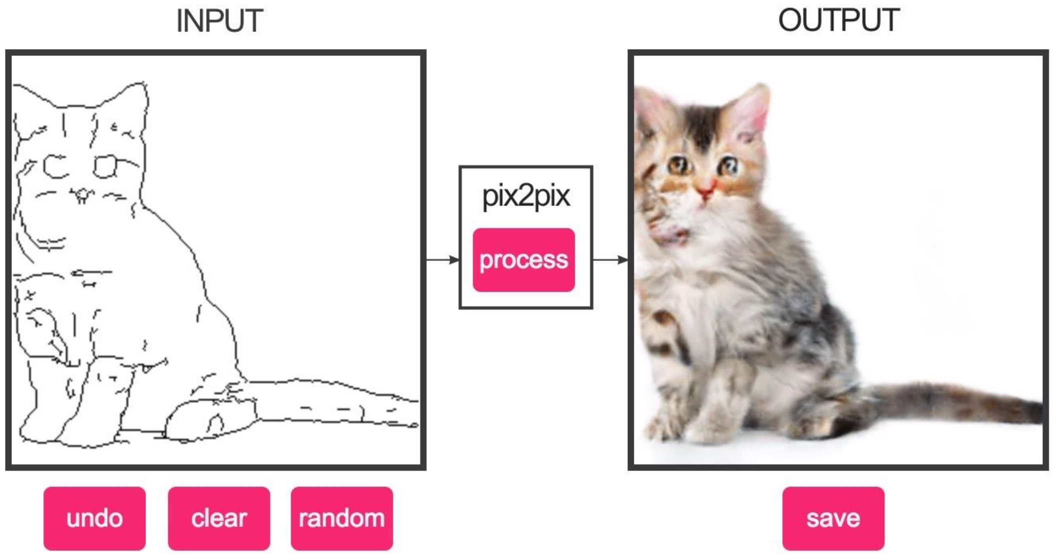
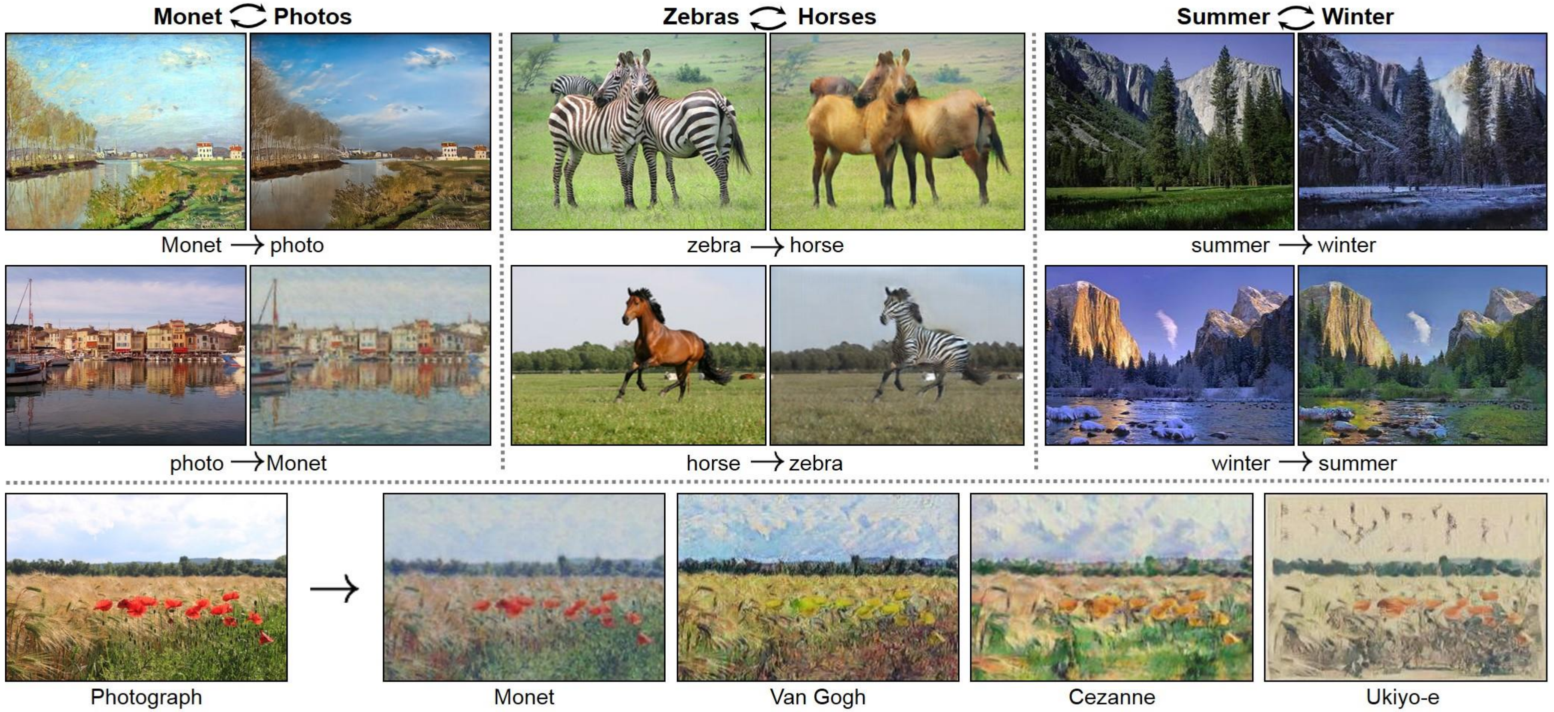


Image Translation

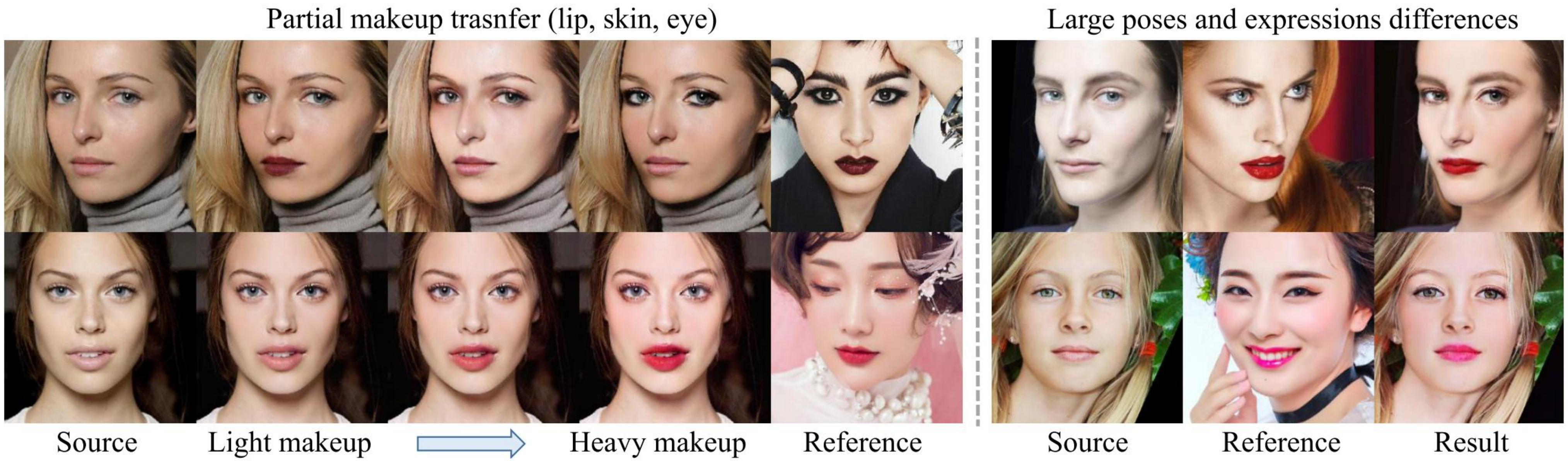


pix2pix (DeepNude)



CycleGAN

Texture Synthesis

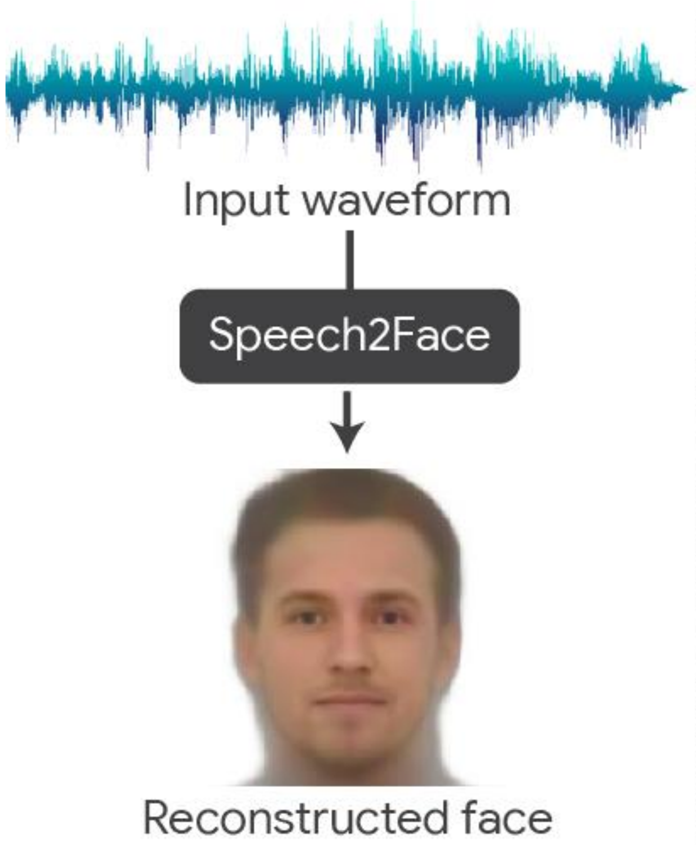


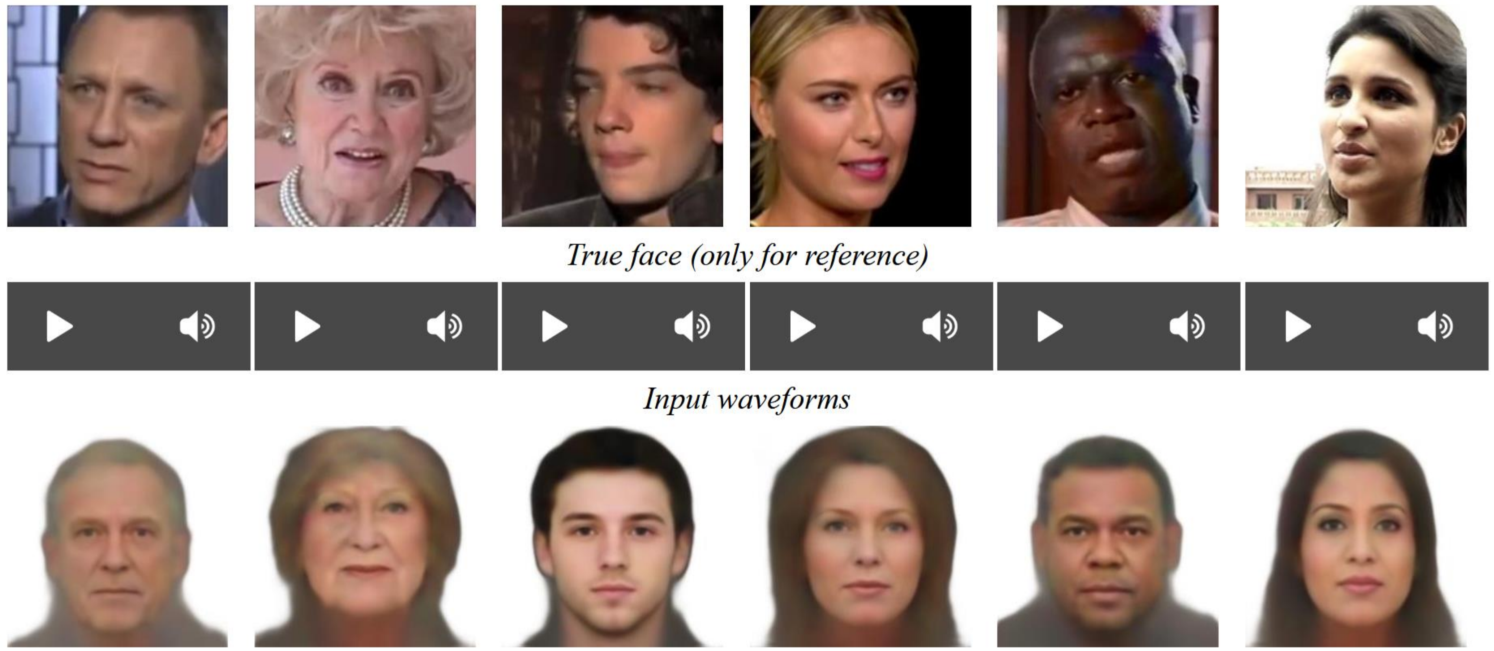
Periodic Spatial GAN (PSGAN)

Face Synthesis

Two-Pathway Generative Adversarial Network (TP-GAN)

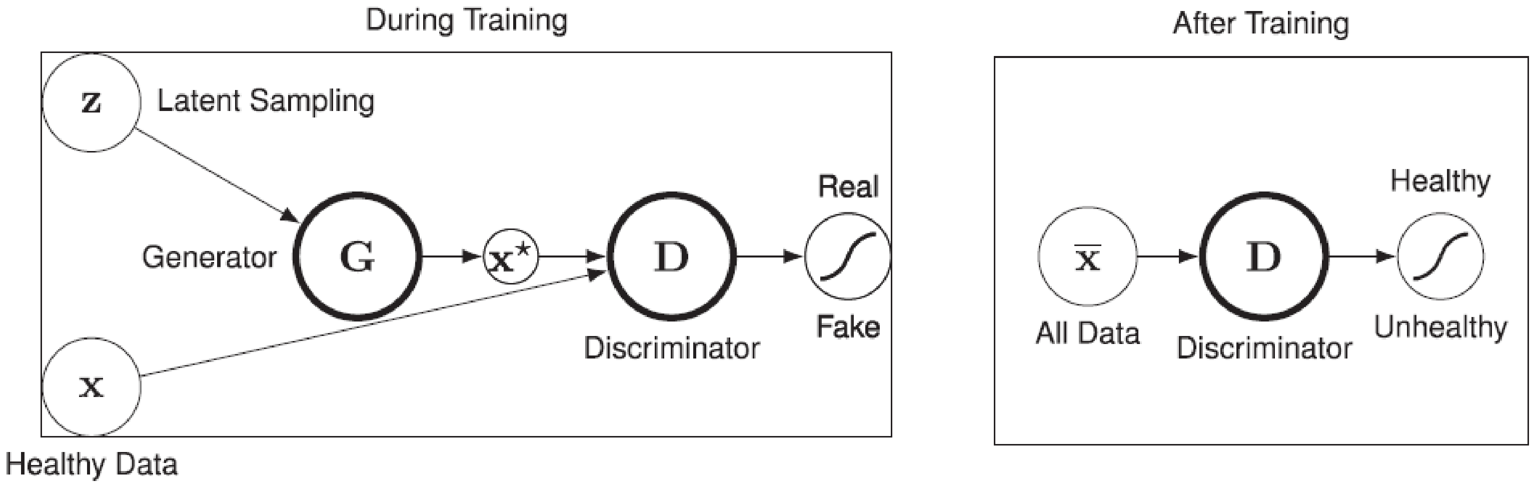
NLP


Input waveform
Speech2Face
Reconstructed face


True face (only for reference)
Input waveforms

Ostatní

- Lze použít k detekci poruchy
- Booyse, W., Wilke, D., & Heyns, S. (2020, February 01). Deep digital twins for detection, diagnostics and prognostics. Retrieved November 12, 2020, from https://www.sciencedirect.com/science/article/abs/pii/S0888327019308337



Shrnutí

- GAN je hra dvou hráčů, tzv. protivníků:
  - Diskriminátor a Generátor
- Diskriminátor se snaží dát vysoké skóre reálním obrázkům a nízké skoré falešným (generovaným) obrázkům
- Generátor se snaží generovat takové obrázky, aby vypadali reálně.
  - Snaží se modifikovat svůj výstup tak, aby získal vysoké skóre od Diskriminátoru
- Cílem celé hry (min-max) pro Generátor je, aby distribuce generovaných dat byla shodná s distribucí reálných dat: $p_G(x) \approx p_{data}(x)$

Zdroje

- Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., . . . Bengio, Y. (2014, December 01). Generative adversarial nets. Retrieved November 11, 2020, from https://dl.acm.org/doi/10.5555/2969033.2969125
- Goodfellow, I. (2017, April 03). NIPS 2016 Tutorial: Generative Adversarial Networks. Retrieved November 11, 2020, from https://arxiv.org/abs/1701.00160
- Zhaoqing, P., Weijie, Y., Xiaokai, Y., Asifullah, K., Feng, Y., & Yuhui, Z. (2019, March 14). Recent Progress on Generative Adversarial Networks (GANs): A Survey. Retrieved November 11, 2020, from https://ieeexplore.ieee.org/document/8667290
- Borji, A. (2018, October 24). Pros and Cons of GAN Evaluation Measures. Retrieved November 11, 2020, from https://arxiv.org/abs/1802.03446
- Salimans, T., Goodfellow, I., Zaremba, W., Cheung, V., Radford, A., & Chen, X. (2016, June 10). Improved Techniques for Training GANs. Retrieved November 11, 2020, from https://arxiv.org/abs/1606.03498
- Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2018, February 26). Progressive Growing of GANs for Improved Quality, Stability, and Variation. Retrieved November 11, 2020, from https://arxiv.org/abs/1710.10196
- Barratt, S., & Sharma, R. (2018, June 21). A Note on the Inception Score. Retrieved November 11, 2020, from https://arxiv.org/abs/1801.01973

Zdroje

- Booyse, W., Wilke, D., & Heyns, S. (2020, February 01). Deep digital twins for detection, diagnostics and prognostics. Retrieved November 12, 2020, from https://www.sciencedirect.com/science/article/abs/pii/S0888327019308337
- Oh, T., Dekel, T., Kim, C., Mosseri, I., Freeman, W., Rubinstein, M., & Matusik, W. (2019, May 23). Speech2Face: Learning the Face Behind a Voice. Retrieved November 12, 2020, from https://arxiv.org/abs/1905.09773
- Brownlee, J. (2019, October 10). How to Implement the Frechet Inception Distance (FID) for Evaluating GANs. Retrieved November 11, 2020, from https://machinelearningmastery.com/how-to-implement-the-frechet-inception-distance-fid-from-scratch/
- Brownlee, J. (2019, July 12). How to Evaluate Generative Adversarial Networks. Retrieved November 11, 2020, from https://machinelearningmastery.com/how-to-evaluate-generative-adversarial-networks/
- Brownlee, J. (2020, September 01). How to Develop a Wasserstein Generative Adversarial Network (WGAN) From Scratch. Retrieved November 11, 2020, from https://machinelearningmastery.com/how-to-code-a-wasserstein-generative-adversarial-network-wgan-from-scratch/
- Papers with Code - BiGAN Explained. (n.d.). Retrieved November 11, 2020, from https://paperswithcode.com/method/bigan

Zajimavé odkazy

- Tutoriály: https://machinelearningmastery.com/category/generative-adversarial-networks/
- GAN Zoo: https://github.com/hindupuravinash/the-gan-zoo