



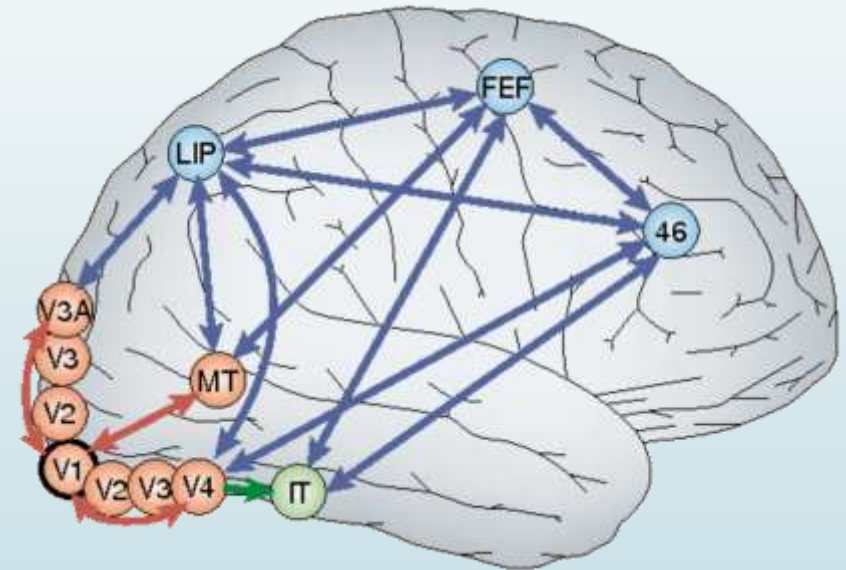
Hlboké učenie a obraz

Mgr. Matej Pecháč (mpechac@tachyum.com)

6.12.2022 FIIT

Vizuálny cortex

- ▶ Ventrálna cesta – „Čo“
 - ▶ Rozpoznanie formy a reprezentácia objektu
- ▶ Dorsálna cesta – „Kde“
 - ▶ Priestorové rozloženie a vzťahy objektov



Klasifikácia – „Čo je na obrázku?“



n07745940 (949)



n07749582 (951)



n01749939 (64)



n01677366 (39)



n02788148 (421)



n02423022 (353)



n02692877 (405)

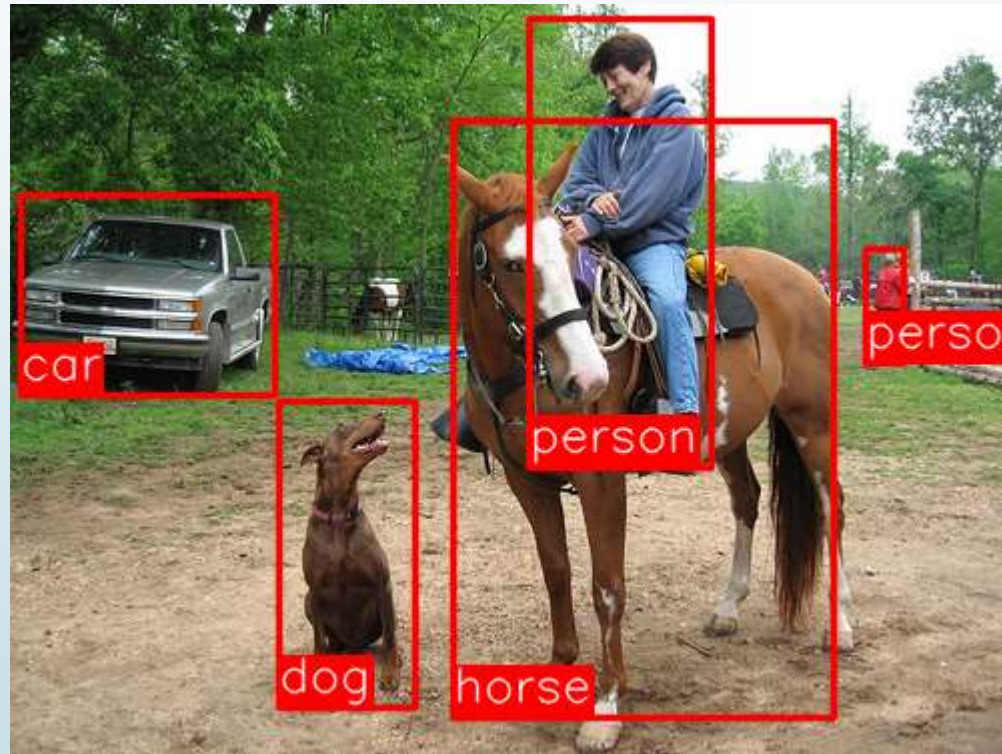


n02490219 (377)



n03443371 (572)

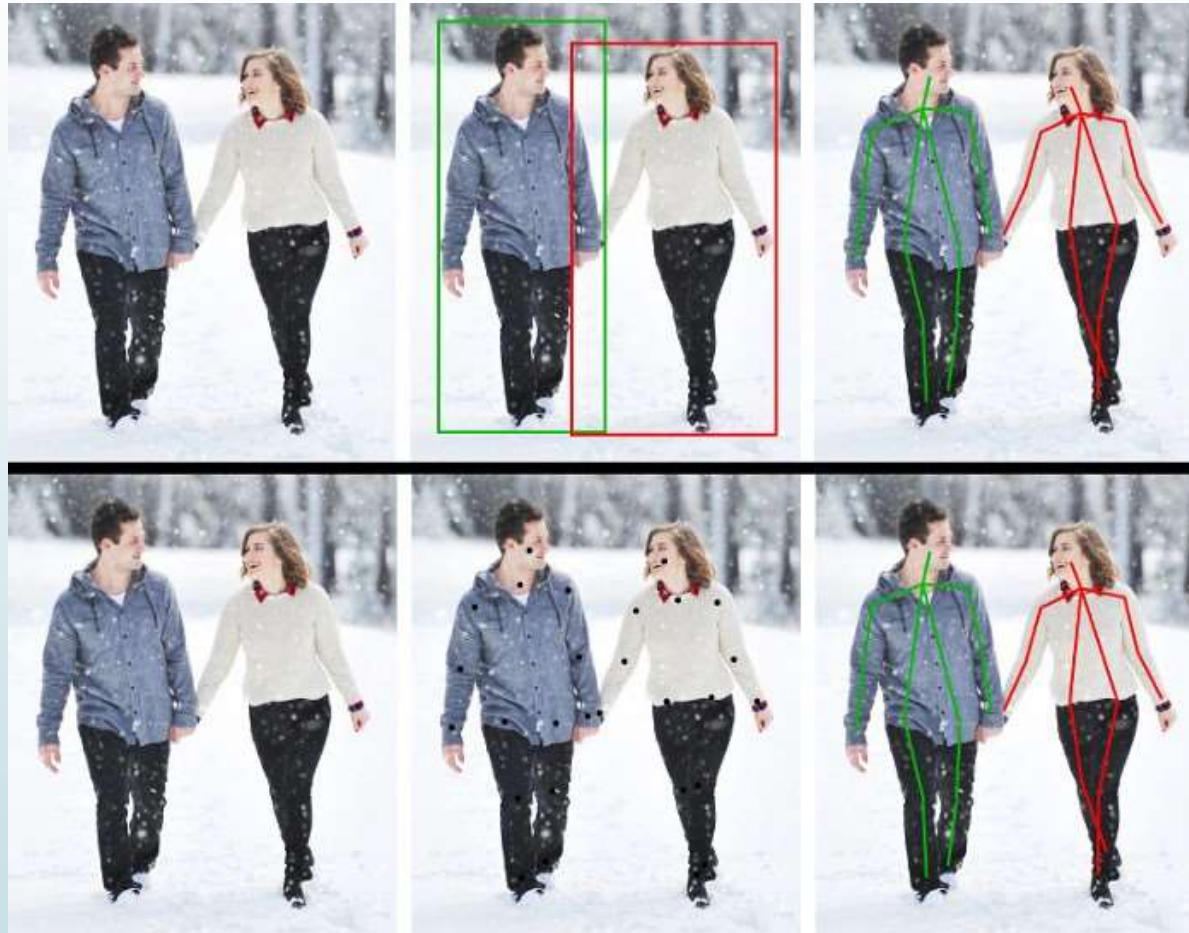
Detekcia – „Čo a kde je na obrázku?“



Detekcia a segmentácia



Odhadovanie pózy



Revolúcia v spracovaní obrazu



ImageNet Large Scale Visual Recognition Challenge (**ILSVRC**)

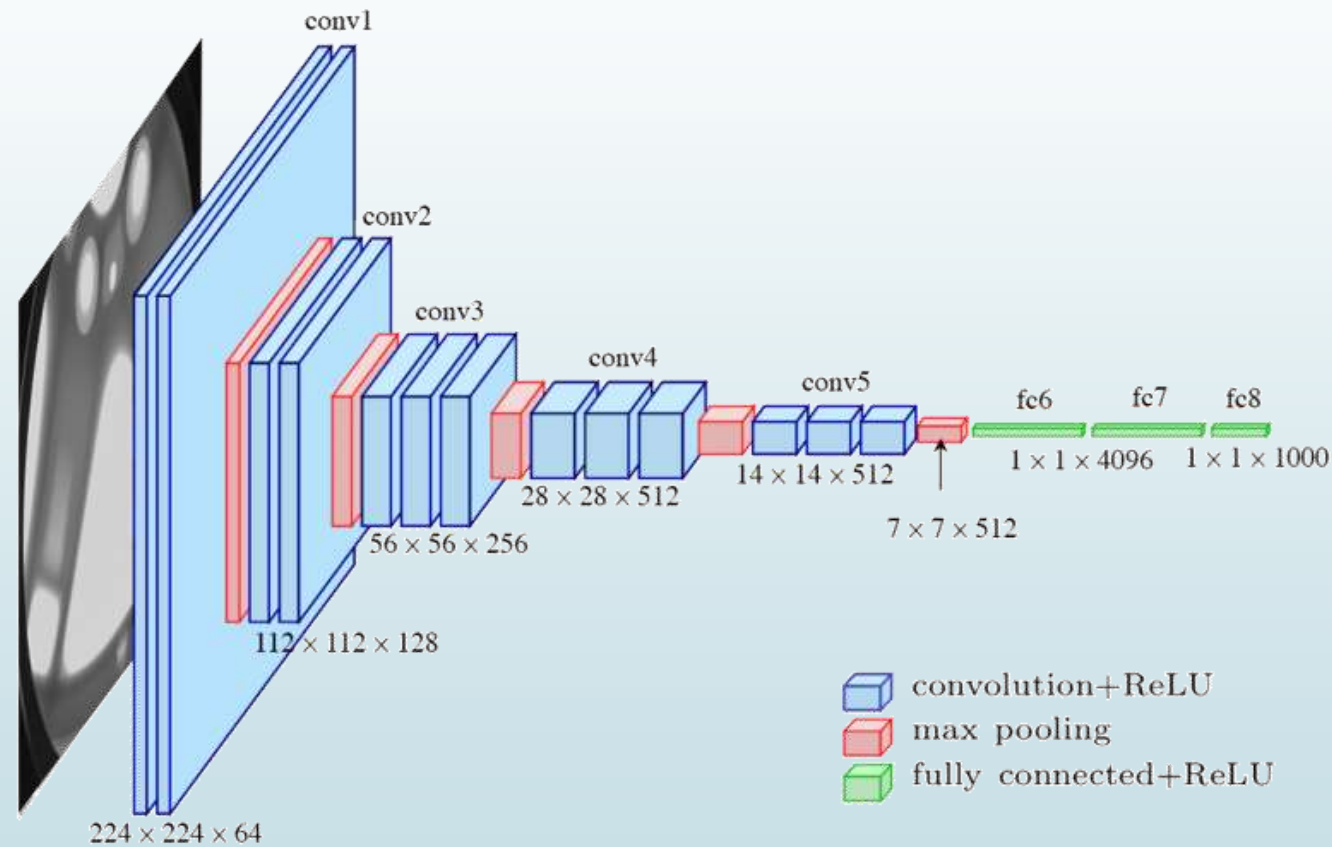
Classification Results (CLS)



Imagenet classification with deep convolutional neural networks

Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton, 2012

Konvolučná sieť



Konvolučná vrstva

- Fukushima 1979 - Neocognitron
- Nový typ neurónov – **konvolučnú vrstvu** (operátor)
- Vie zachytiť priestorové vzťahy vstupných dát a fungovať ako filter

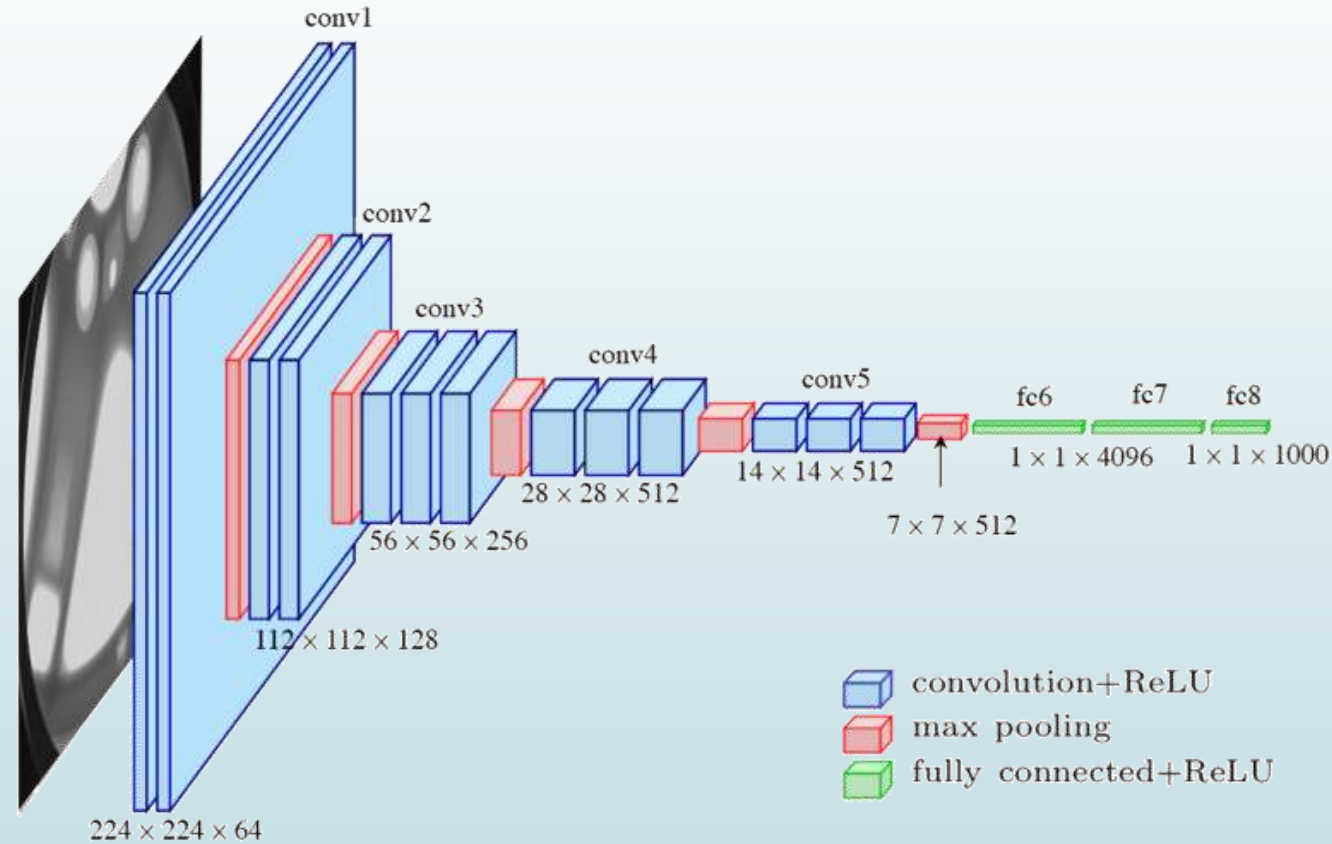
1 _{x1}	1 _{x0}	1 _{x1}	0	0
0 _{x0}	1 _{x1}	1 _{x0}	1	0
0 _{x1}	0 _{x0}	1 _{x1}	1	1
0	0	1	1	0
0	1	1	0	0

Image

4		

Convolved
Feature

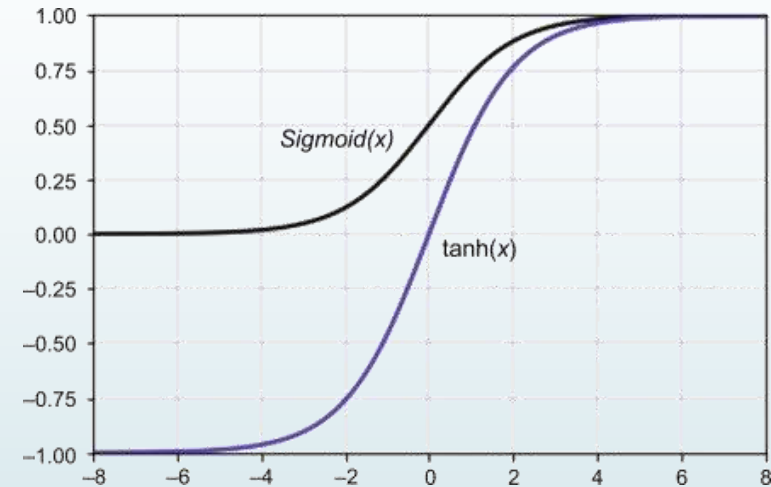
Konvolučná sieť = hlboká sieť



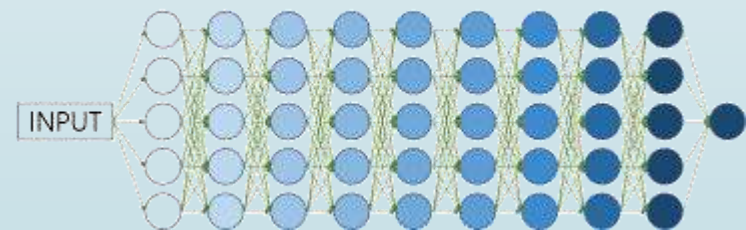
Problém s hĺbkou sieťí

► Pri vysokých hodnotách dochádza k saturácii a derivácia takejto funkcie sa blíži k 0

► Pri učení dochádza k zmiznutiu alebo výbuchu gradientu siete

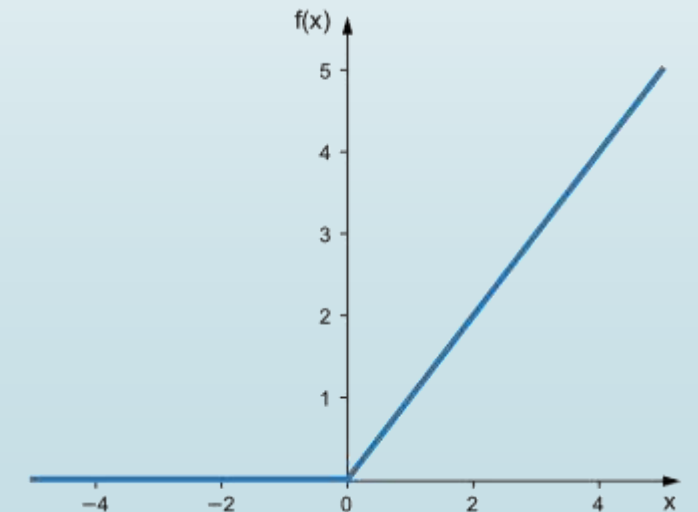
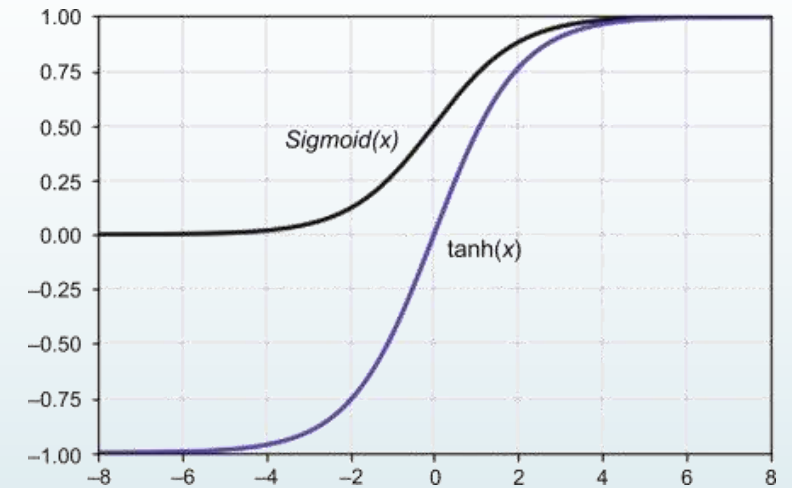


VANISHING GRADIENT PROBLEM

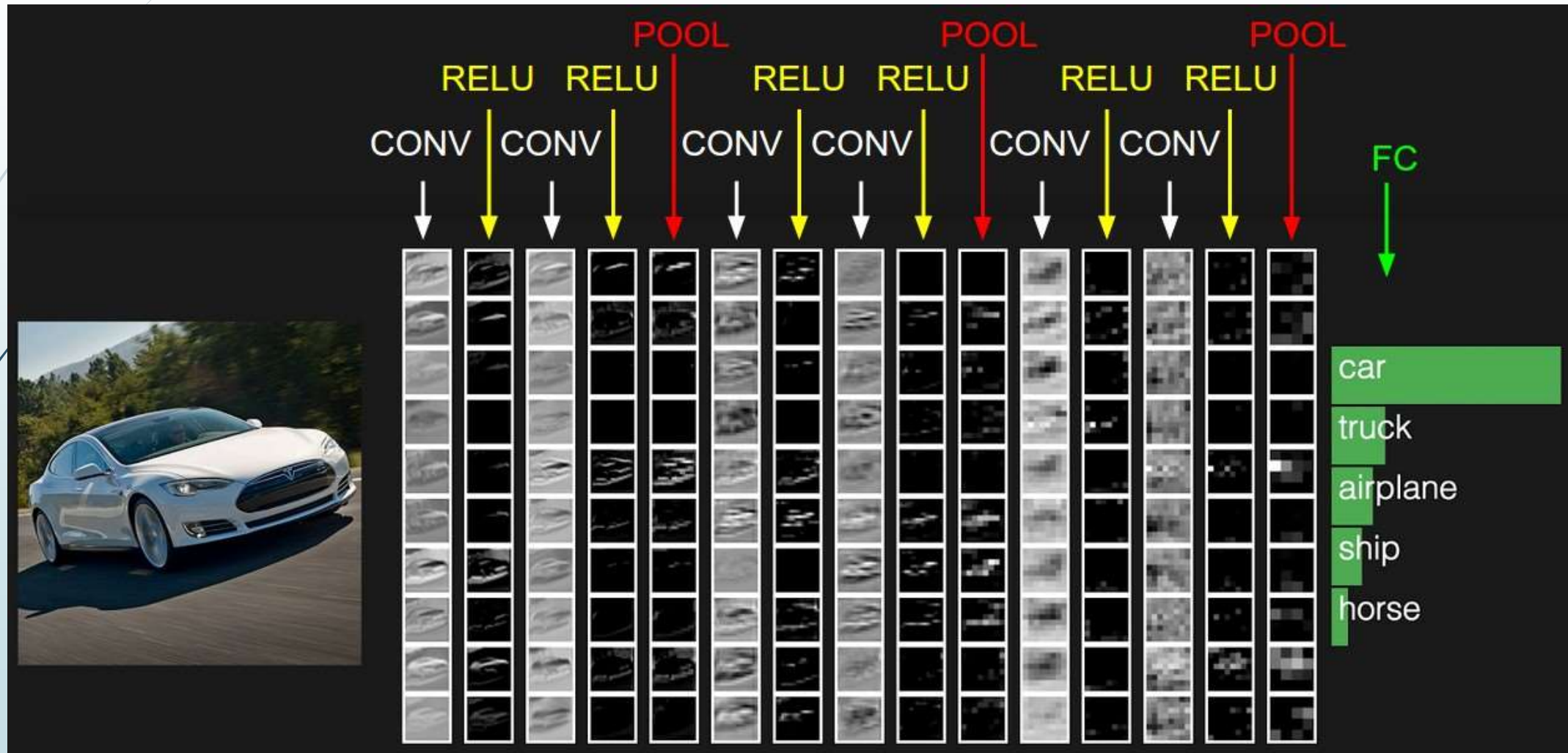


Riešenia pre hlboké siete

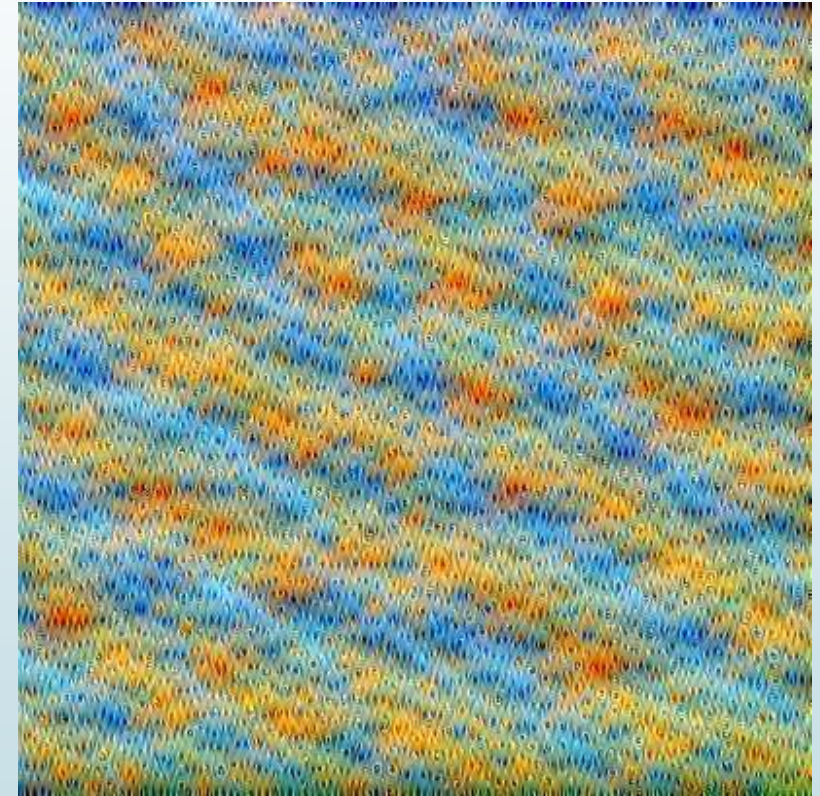
- Nepoužiť gradientovú metódu optimalizácie
- Použiť optimalizáciu pomocou derivácií 2. rádu (1989)
- Použiť pred-trénovanie siete (2007)
- Použiť rýchlejšiu optimalizáciu pomocou derivácií 2. rádu (2010)
- **Použiť inú aktivačnú funkciu (2010)**



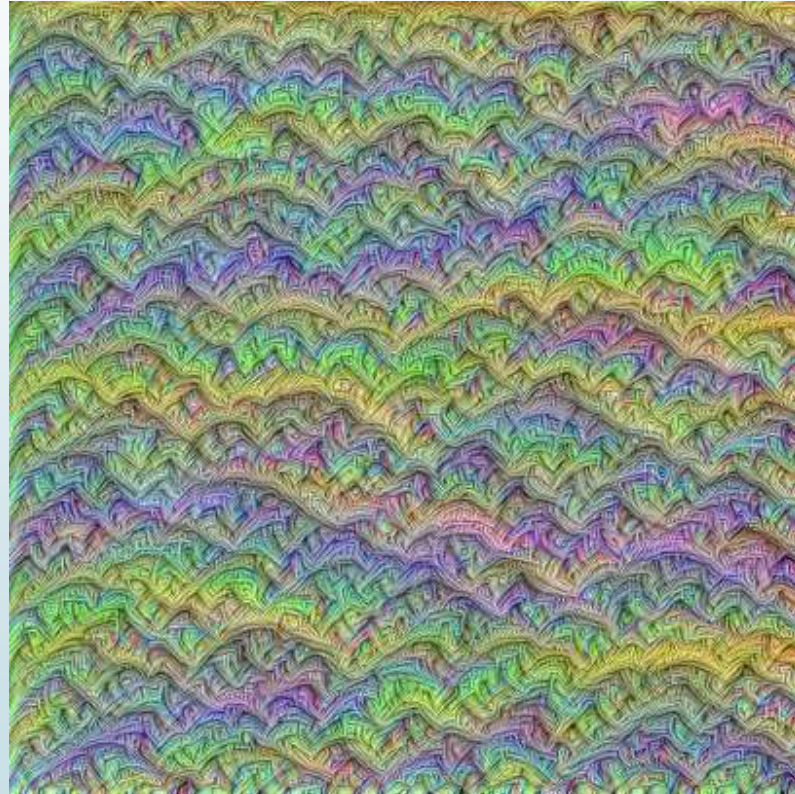
Pohľad dovnútra konvolučnej siete



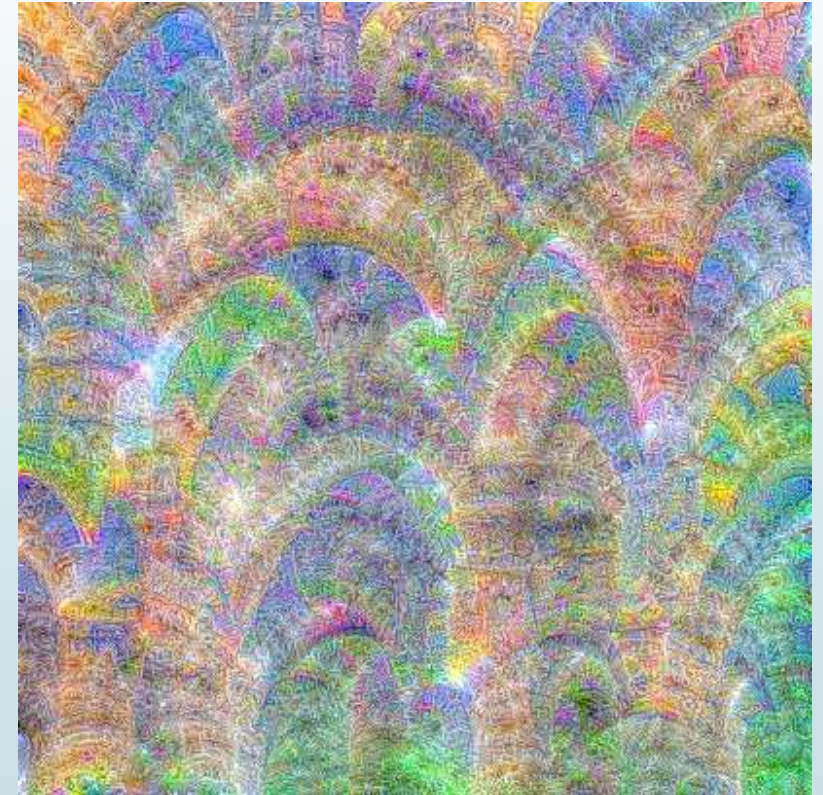
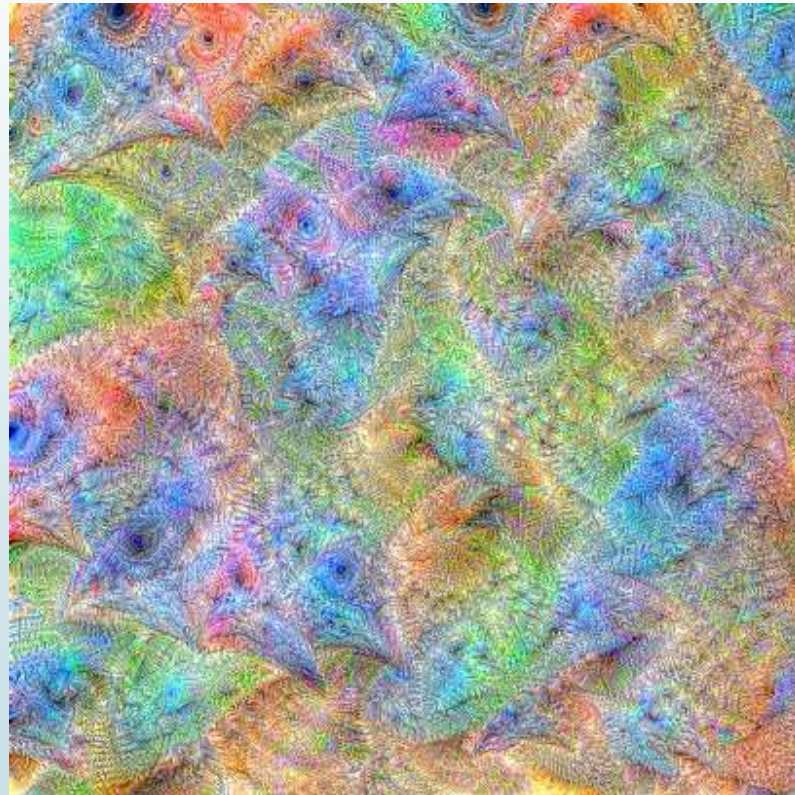
Pohľad dovnútra konvolučnej siete: nízko abstraktné reprezentácie



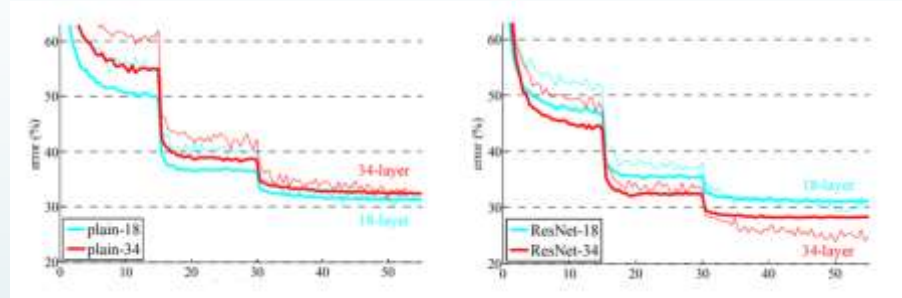
Pohľad dovnútra konvolučnej siete: stredne abstraktné reprezentácie



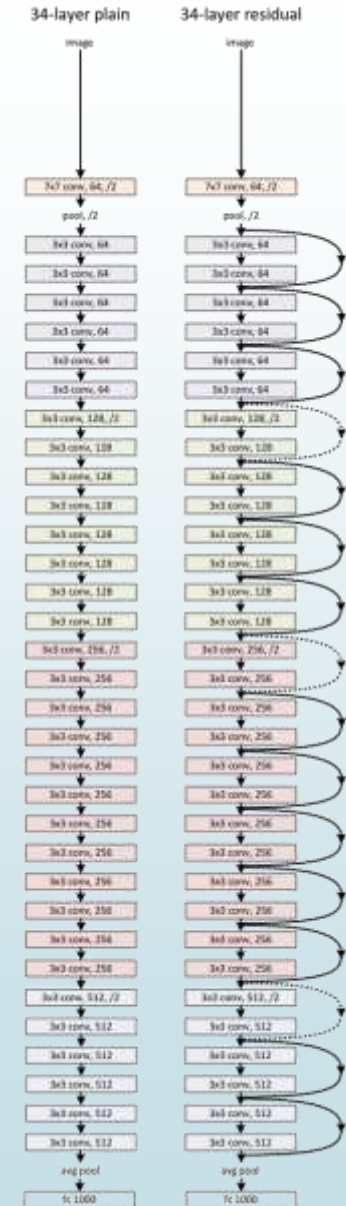
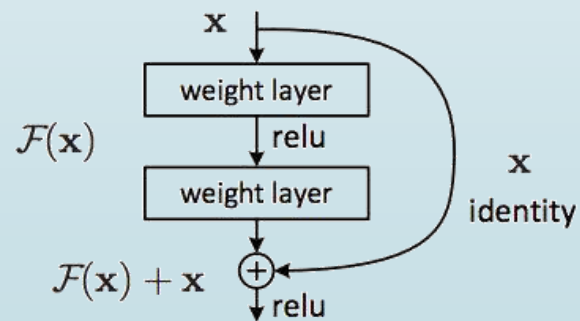
Pohľad dovnútra konvolučnej siete: vysoko abstraktné reprezentácie



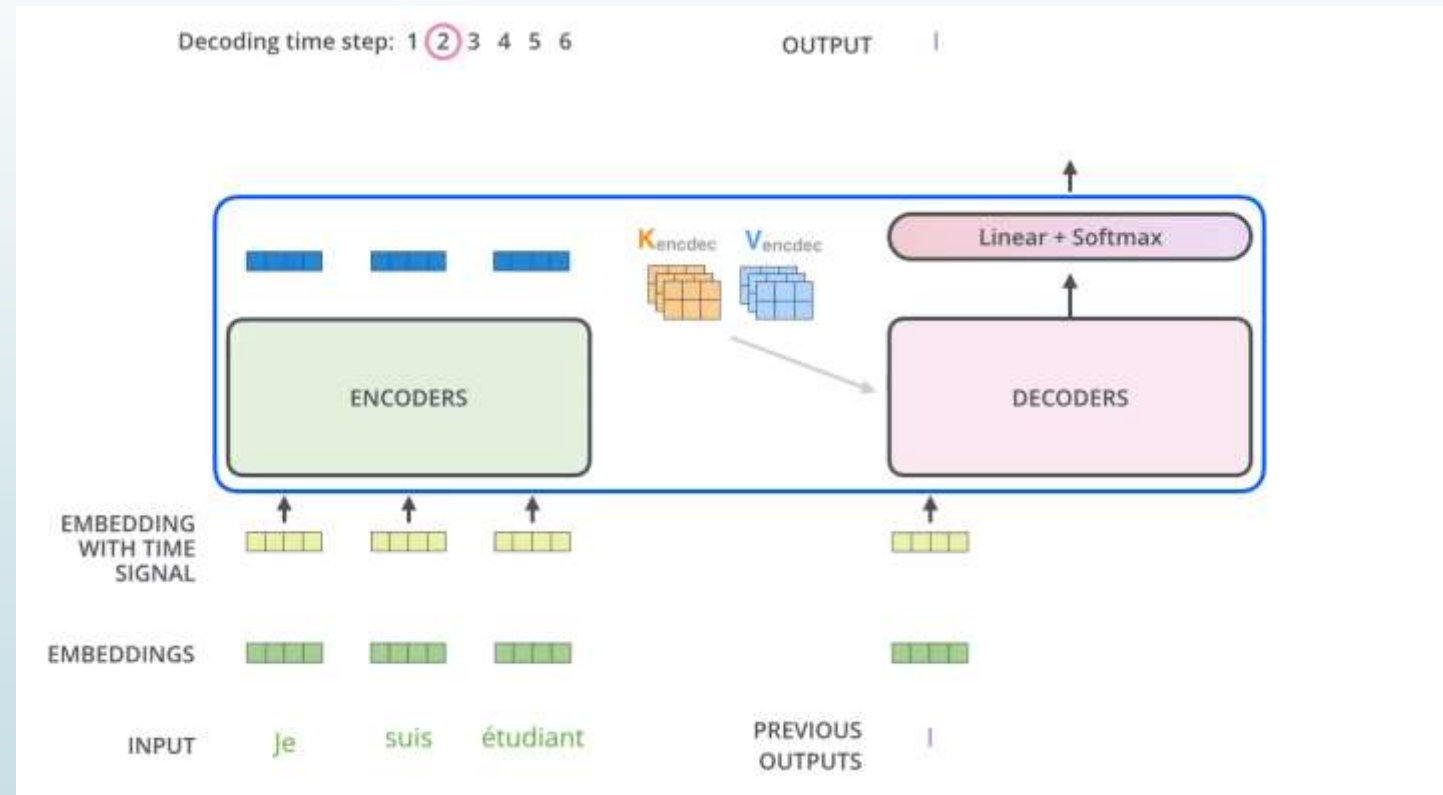
Reziduálne konvolučné siete



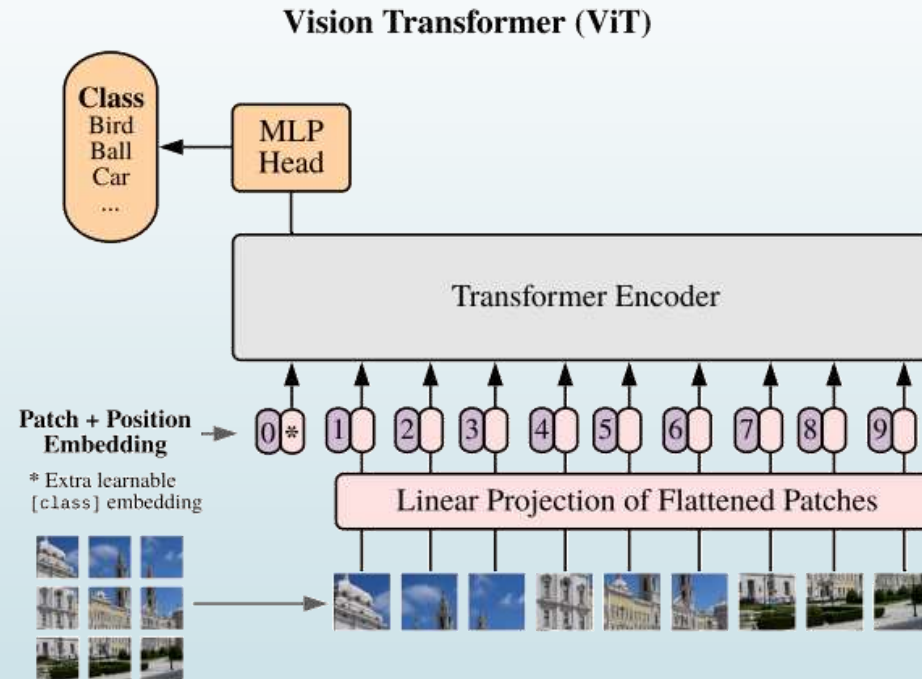
- Prvé konvolučné siete narazili na limitovanú hĺbku
- Riešením bolo pridať reziduálne spojenie



Vizuálne transformery (ViT)

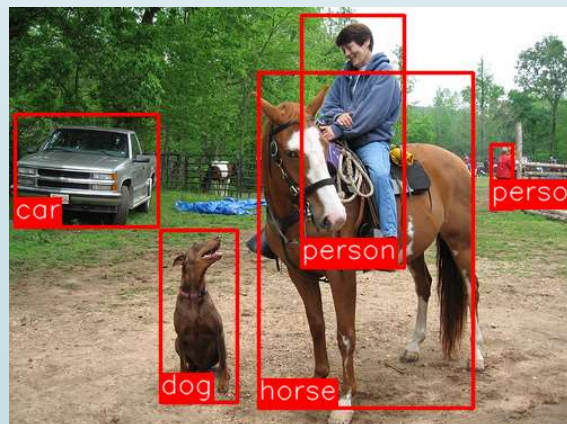


Vizuálne transformery (ViT)

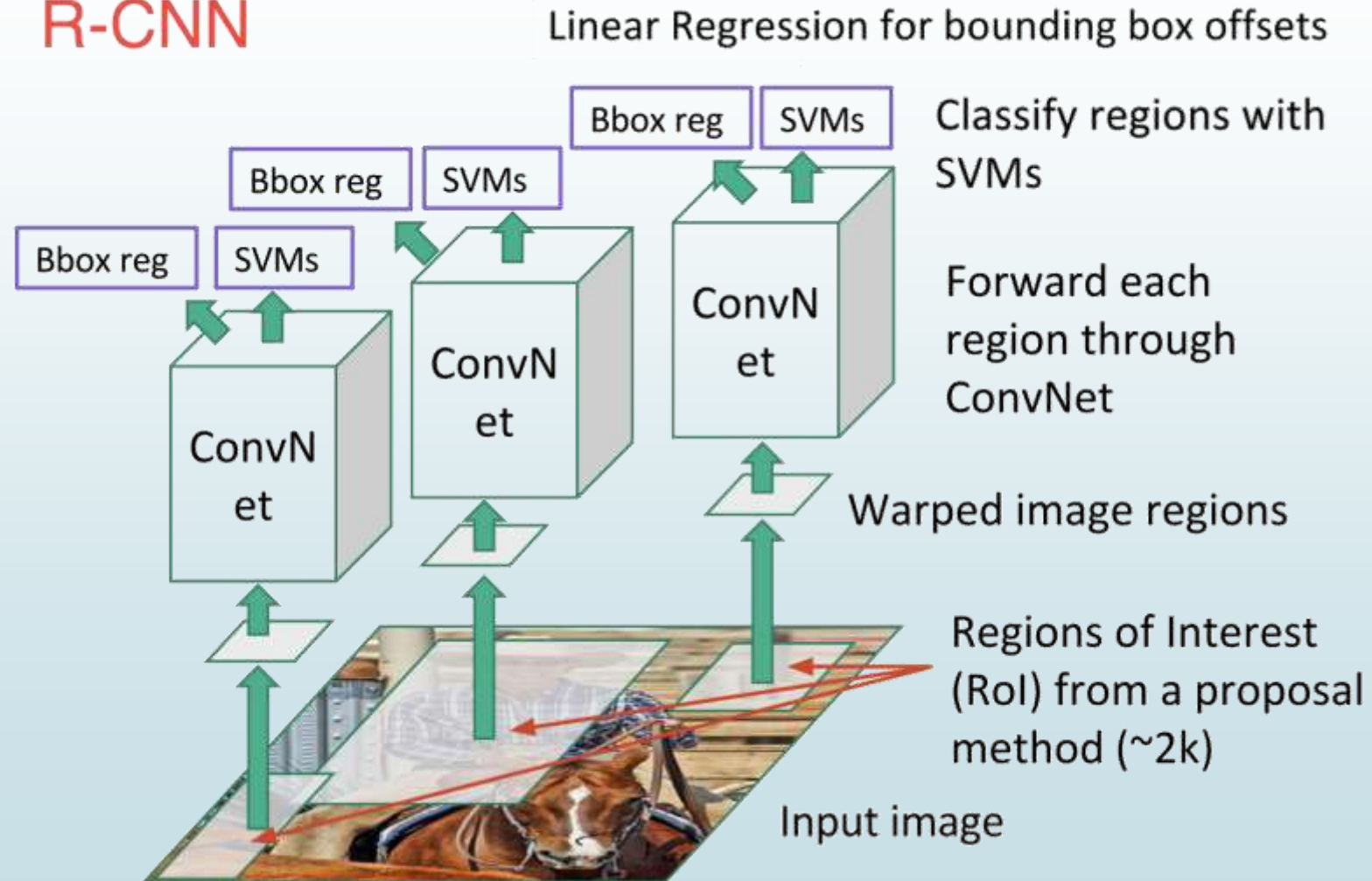


Komplexnejšie modely

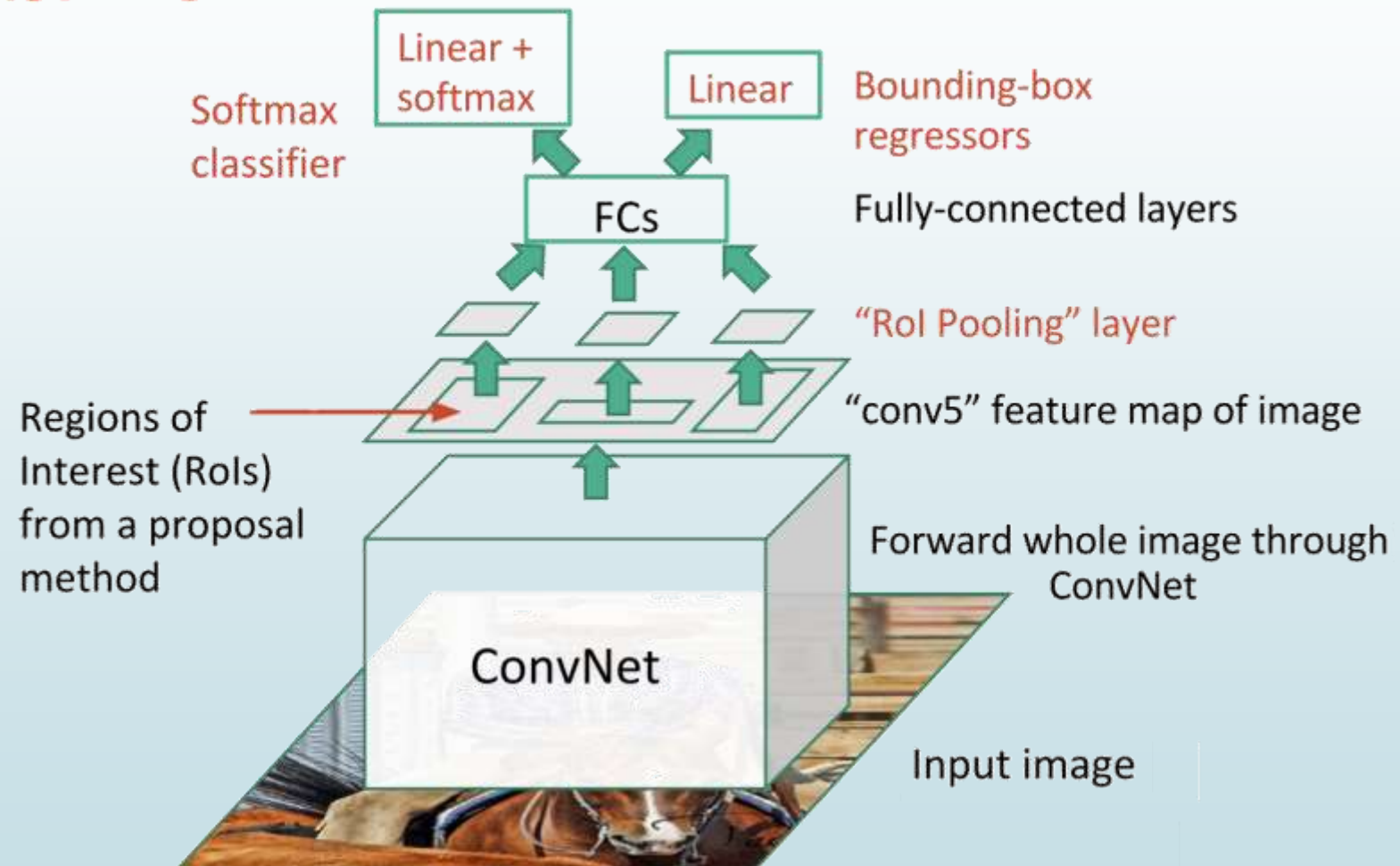
- Detektory – povedia kde a čo je na obrázku
- Segmentačné detektory – povedia kde, čo a aký to má tvar na obrázku
- Detektory s odhadom pózy – povedia kde, čo a v akej póze je



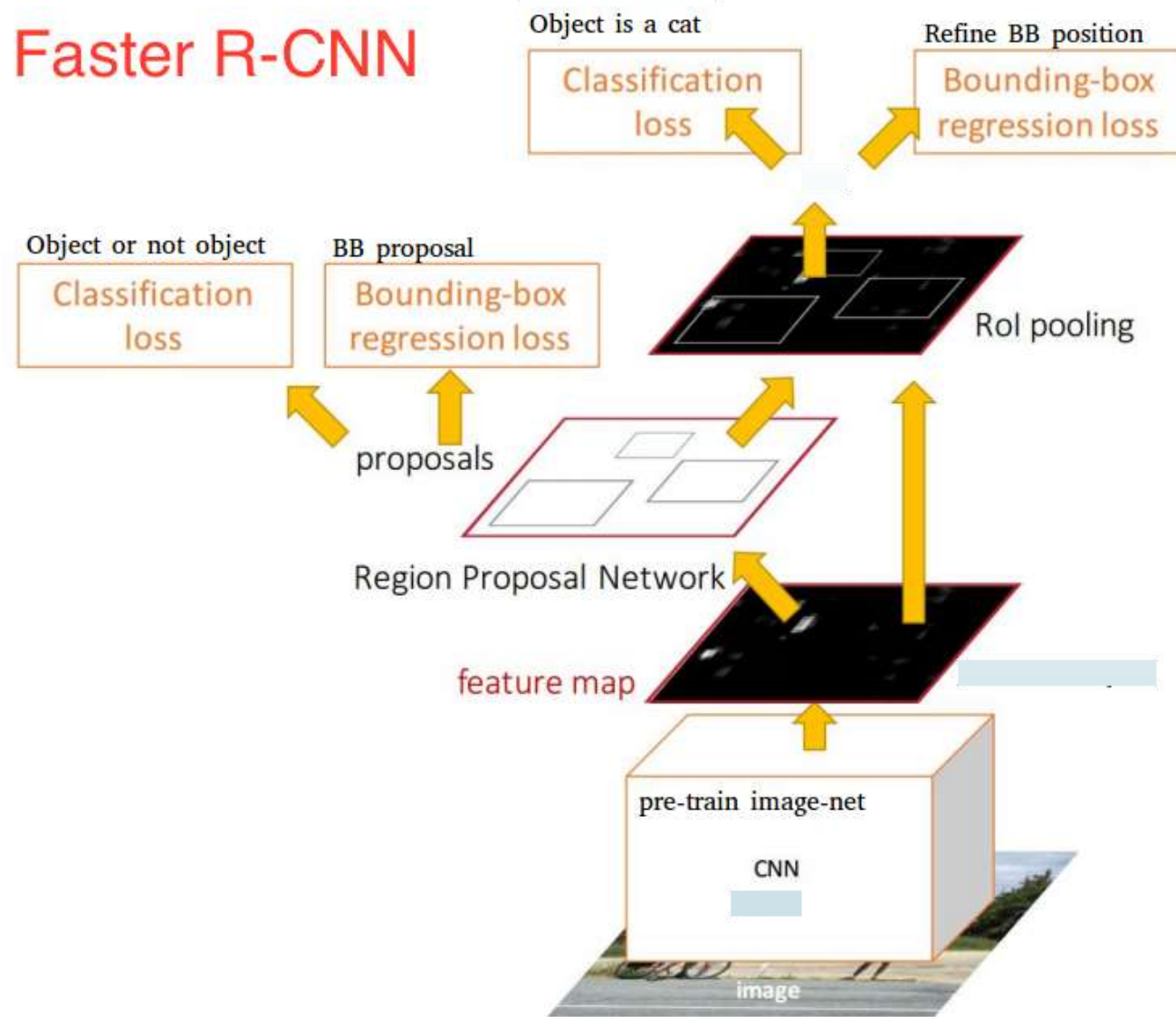
R-CNN



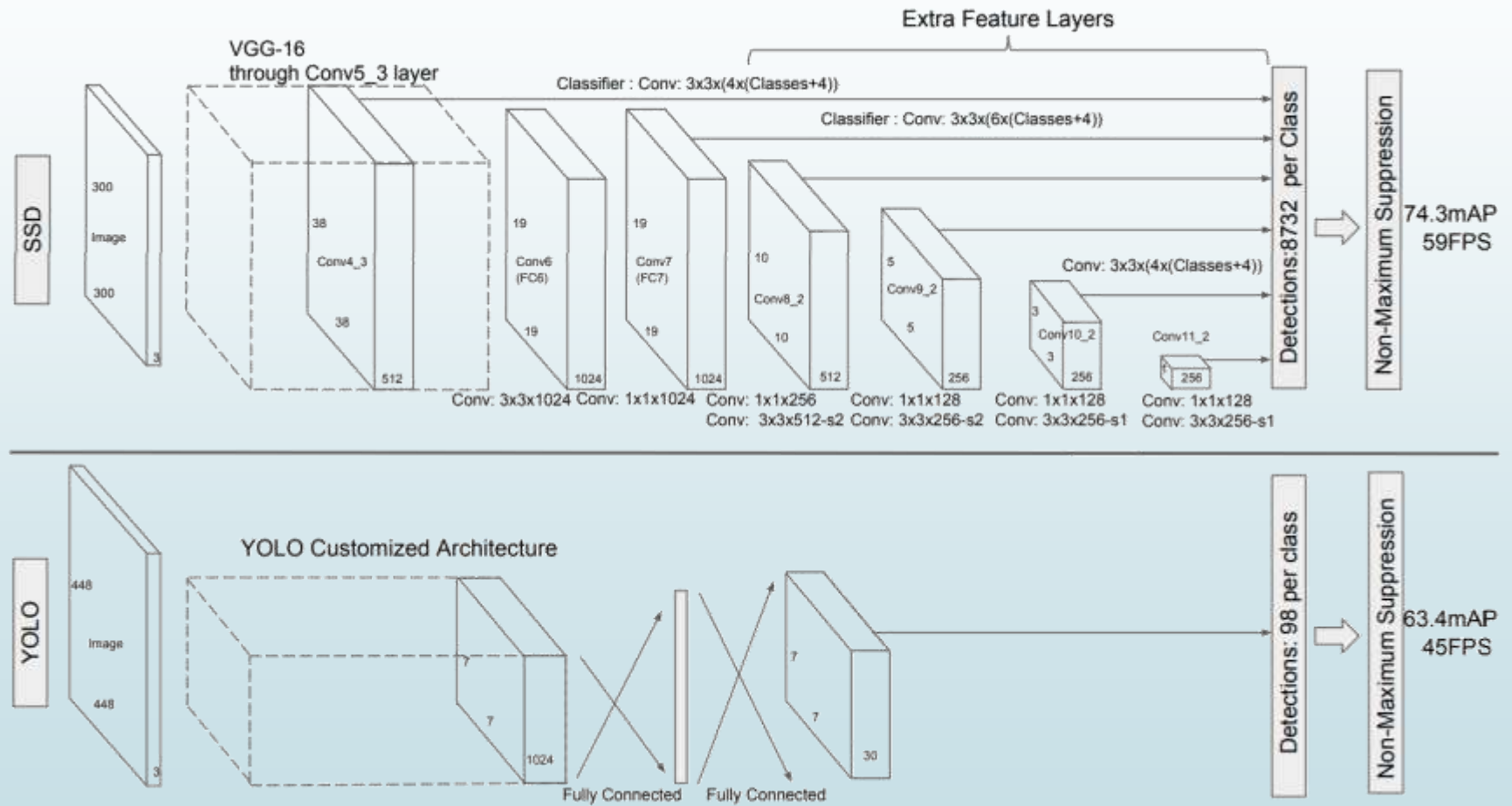
Fast R-CNN

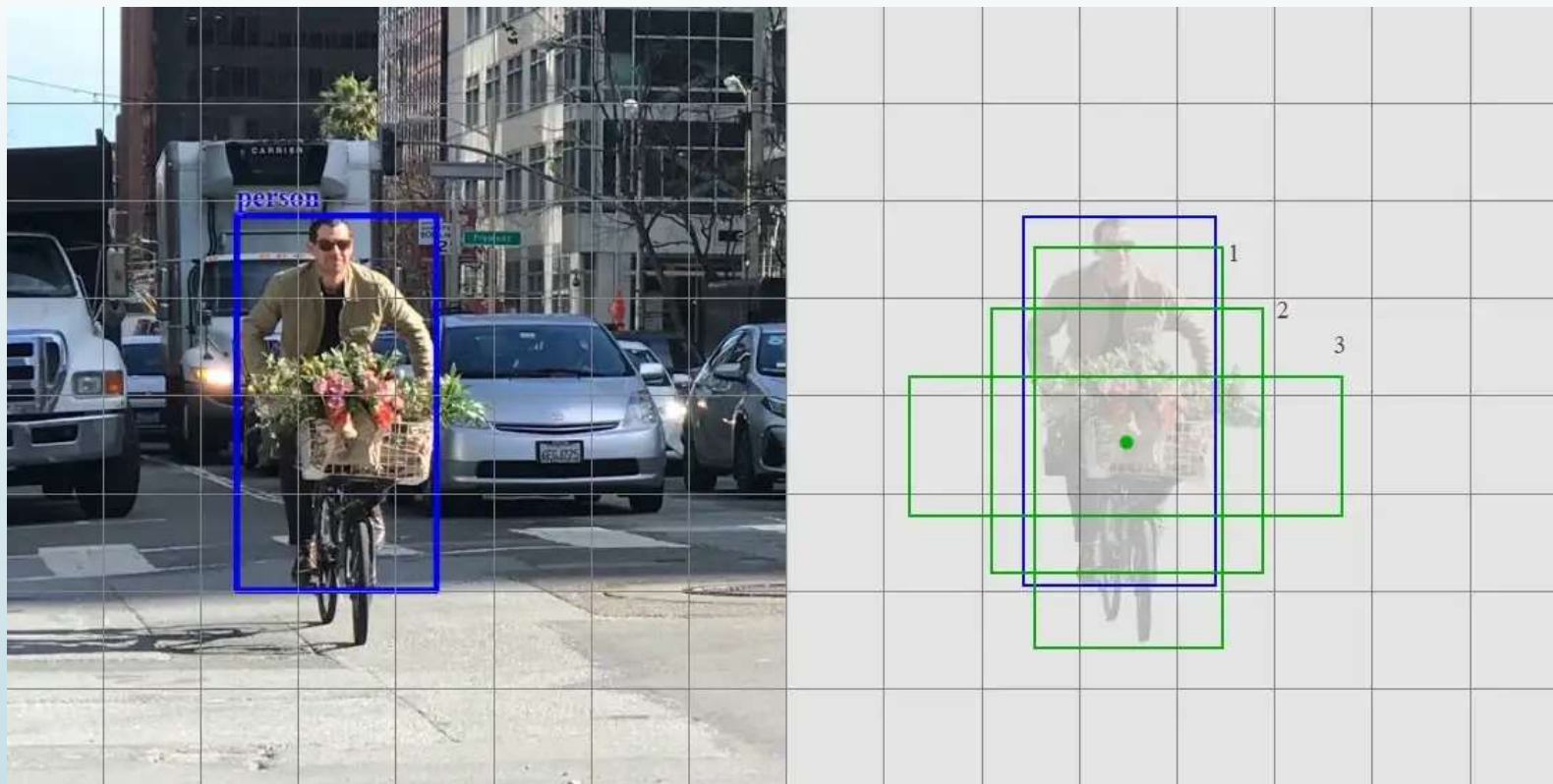


Faster R-CNN



SSD α YOLO

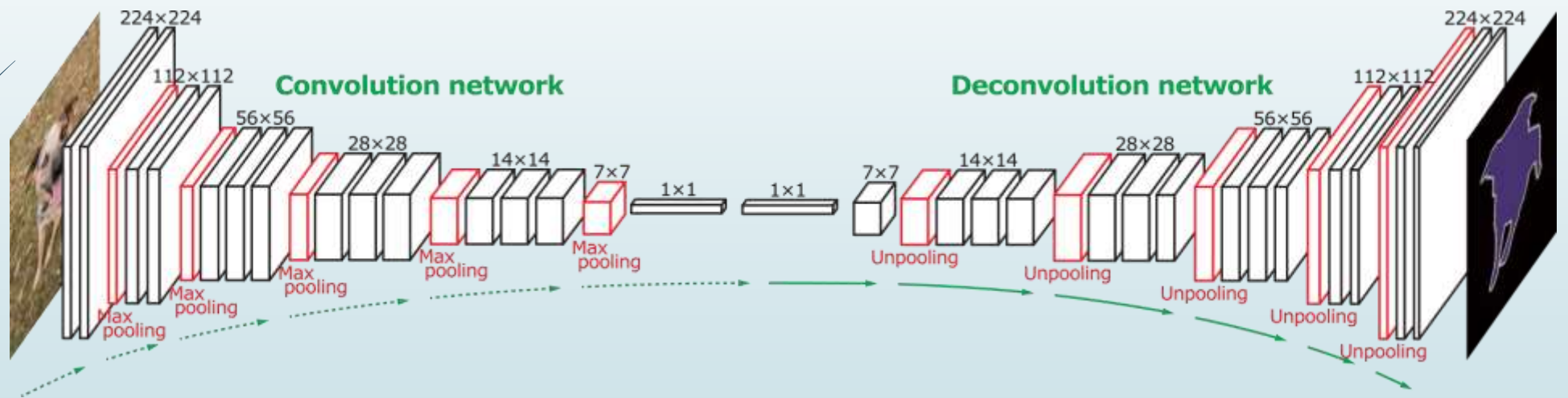


SSD α YOLO

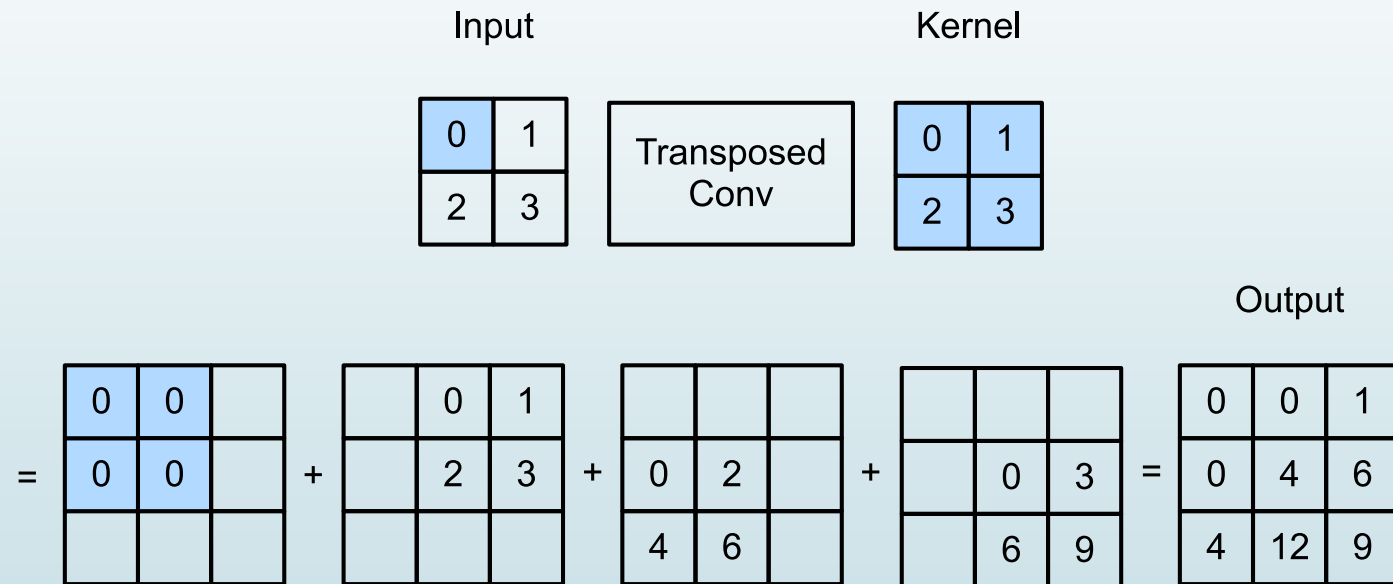
SSD a YOLO



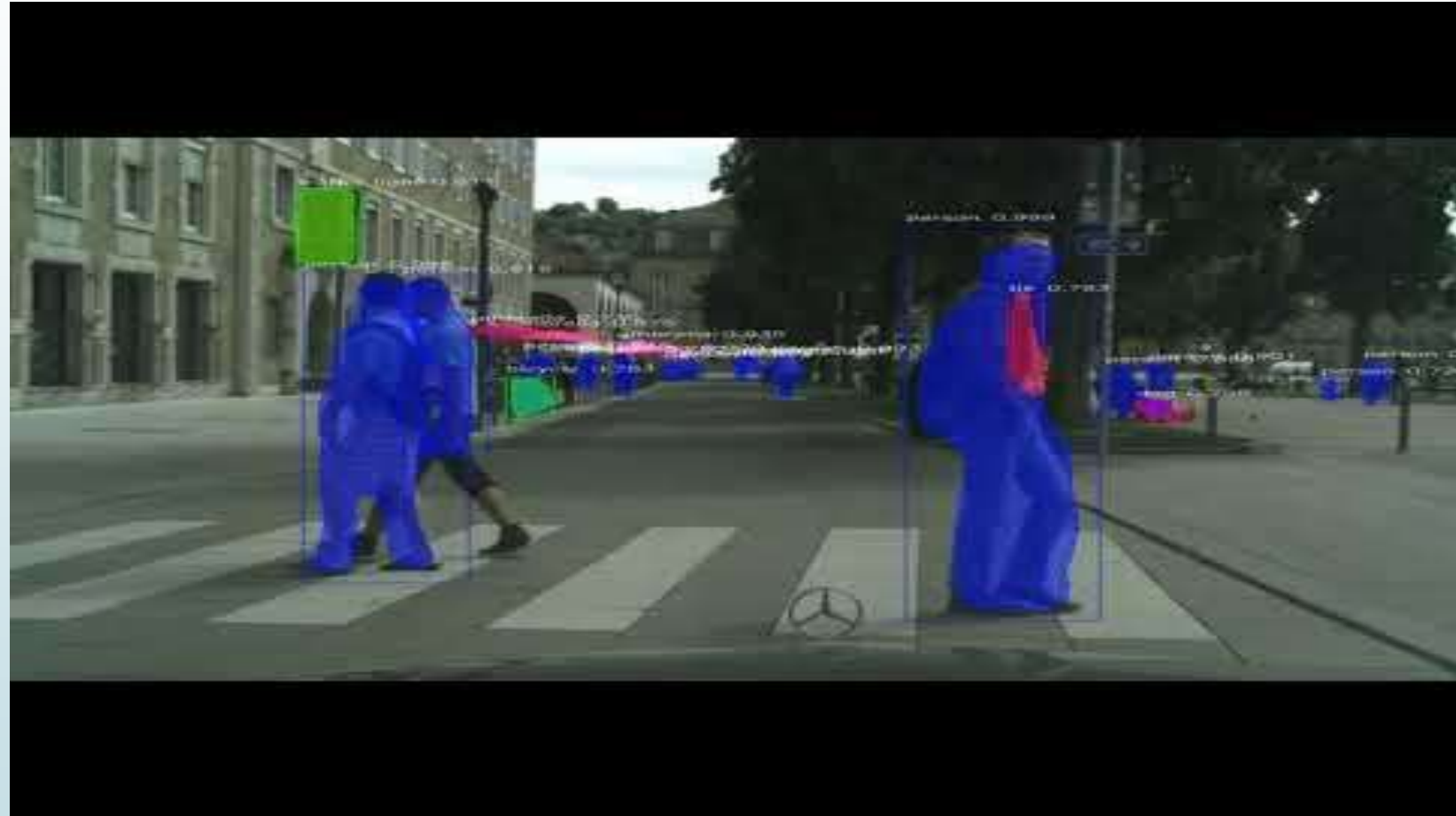
Segmentácia



Transponovaná konvolúcia



Segmentácia Mask RCNN

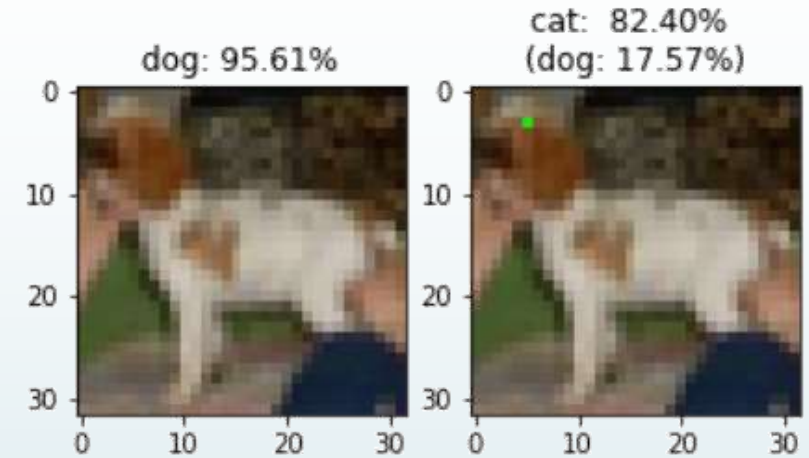


Využitie

- Autonómne autá, vlaky, drony
- Rozpoznávanie nádorov, kožných ochorení, fMRI
- Rozpoznanie aktivít, reči tela, analýza bezpečnostných záznamov
- Filmový priemysel, vylepšená realita
- Generovanie dát pre tréning robotov

One pixel attack

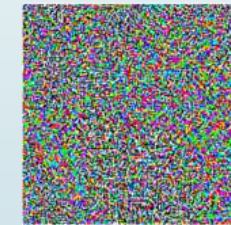
- Pridaním nevýrazného šumu je možné úplne zmeniť predikcie siete
- Spätnou analýzou sa dá nájsť 1 pixel, ktorý rozhádže pravdepodobnostné intervaly siete
- Riešením je pridávať takéto perturbácie už počas trénovania



“panda”

57.7% confidence

+ .007 ×



noise

=



“gibbon”

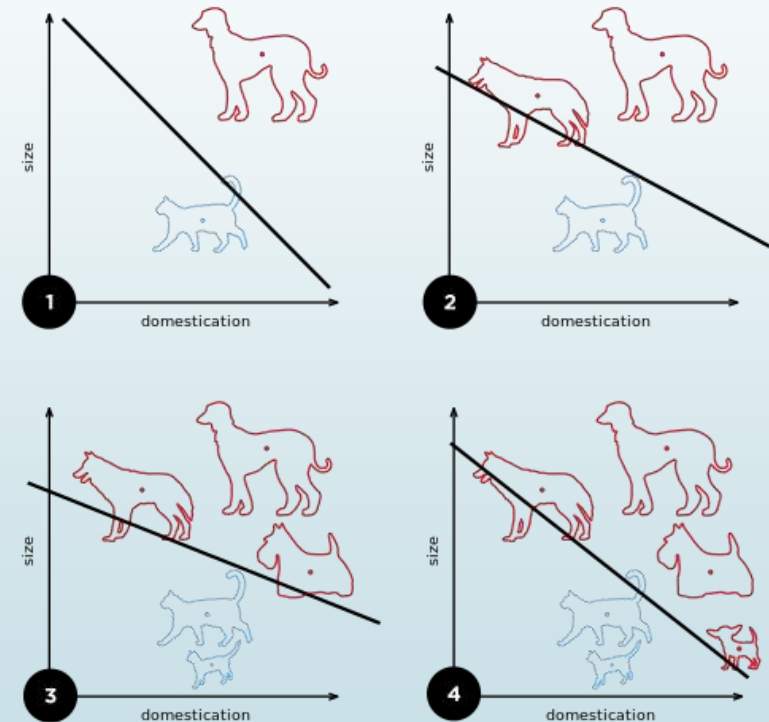
99.3% confidence

Generatívne modely



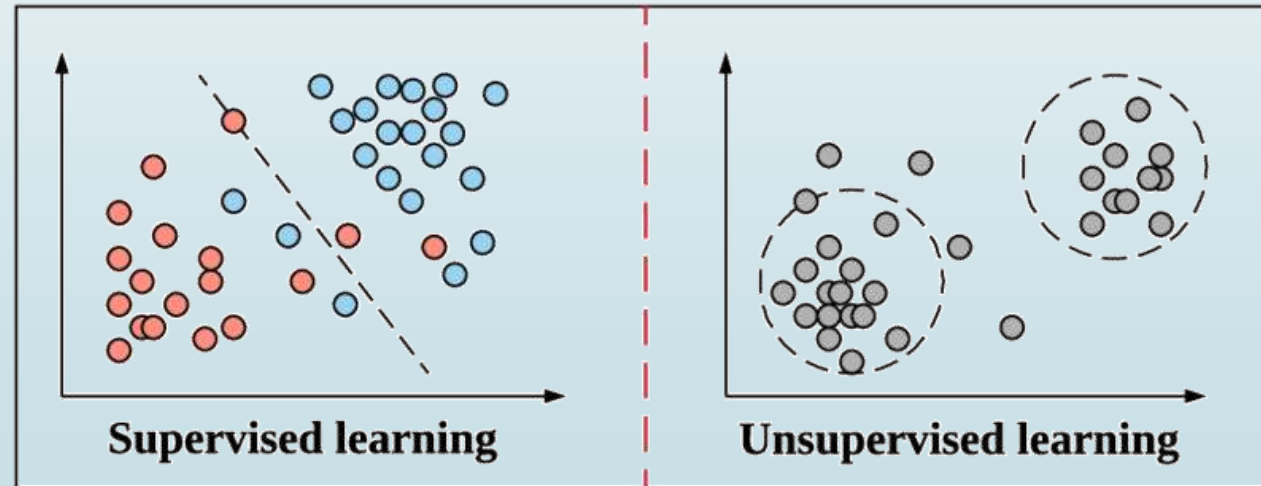
Učenie s učiteľom

- Poznáme vstup a požadovaný výstup
- Model sa učí zobrazenie (funkciu) vstupných hodnôt na výstupné
- Natrénovaný model vie potom kategorizovať vstup, ktorý nevidel alebo aproximovať neznámu funkciu



Učenie bez učiteľa

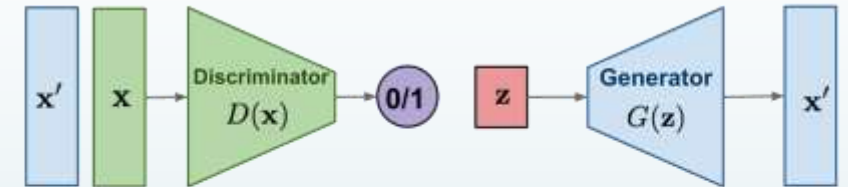
- Cieľom je naučiť sa vzorce z neoznačených dát
- Model má svoju vnútornú štruktúru , ktorá odráža štruktúru dát



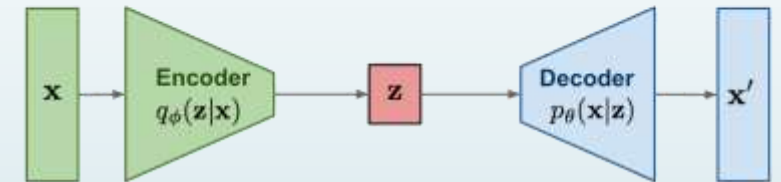
Generatívne modely

- Predstavujú jeden z kľúčových komponentov AI
- Pridávajú vlastnosť „predstavivosti“ a „pamäte“
- Umožňujú modelovanie komplexných vysoko rozmerných dát ako obraz, text či reč
- Generujú inštanície výstupných dát na základe vstupných dát

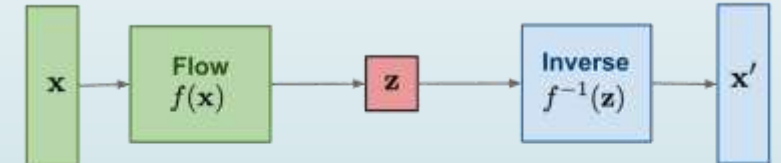
GAN: Adversarial training



VAE: maximize variational lower bound



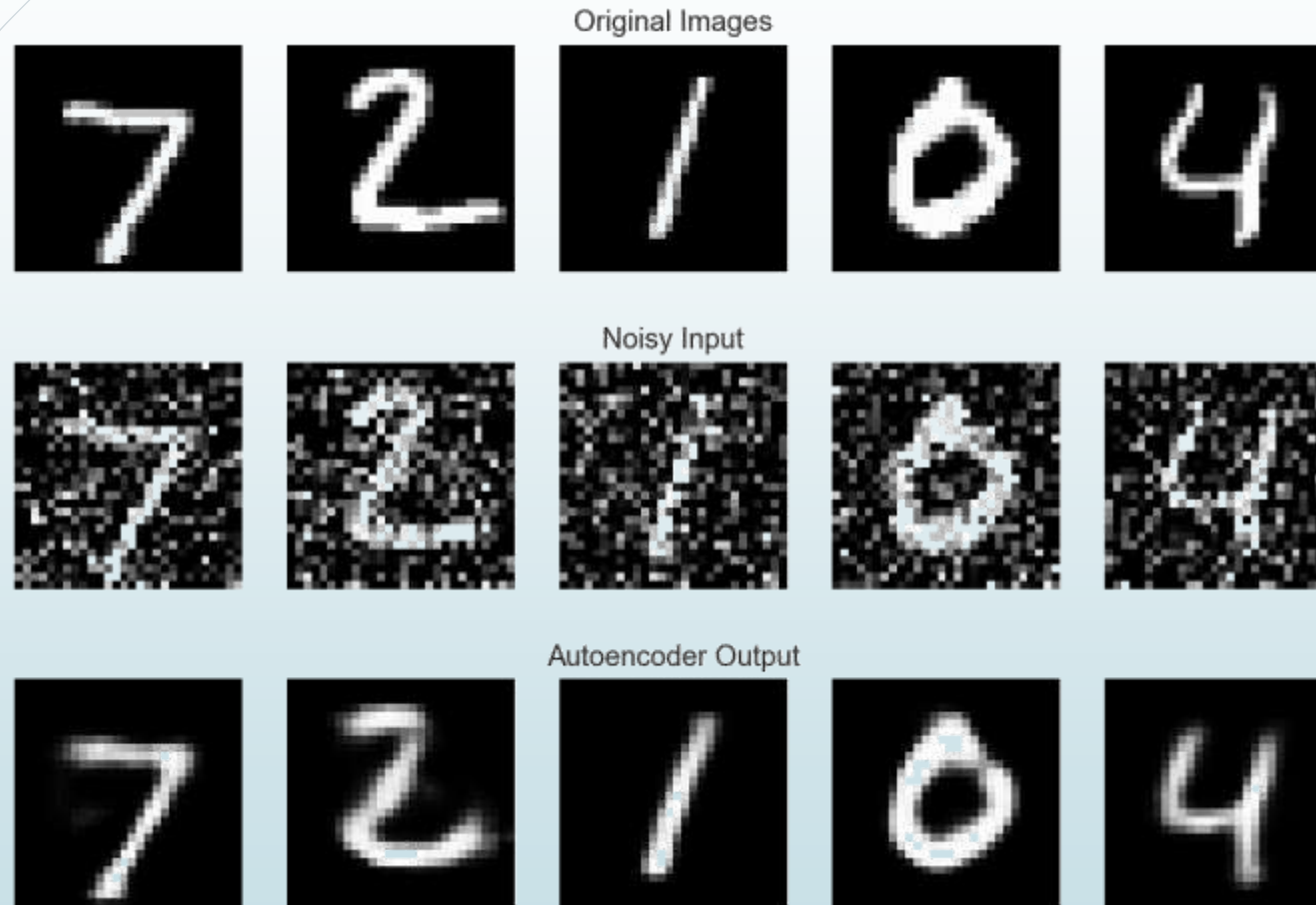
Flow-based models: Invertible transform of distributions



Diffusion models: Gradually add Gaussian noise and then reverse

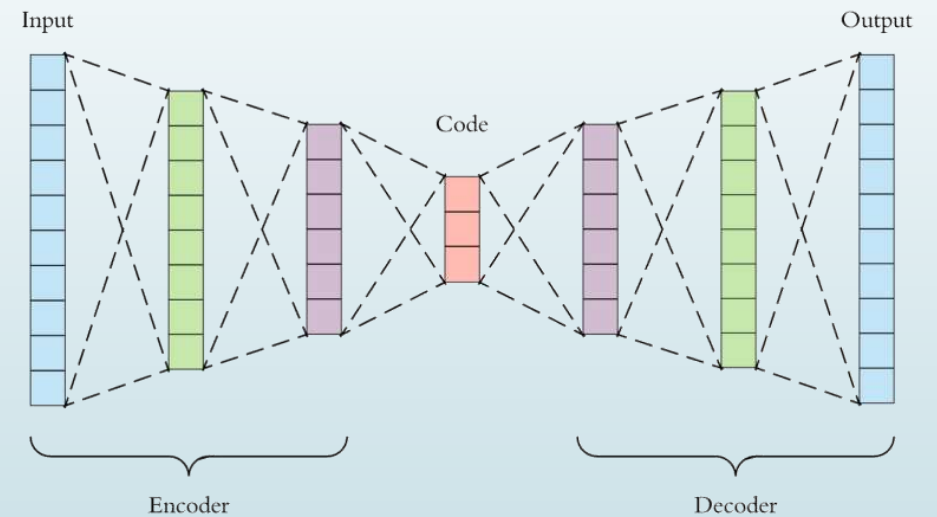


Autoenkóder – redukcia šumu



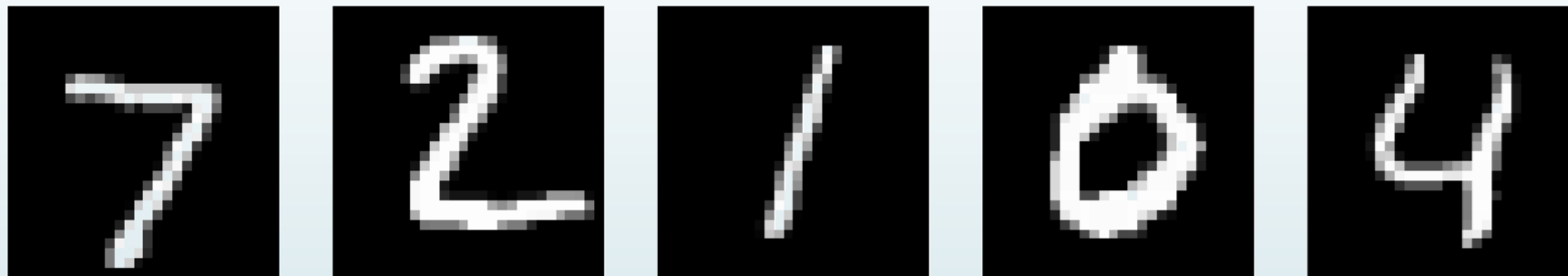
Autoenkóder

- Má dve časti: enkóder a dekóder
- Uprostred vzniká komprimovaná informácia o vstupe - reprezentácia v latentnom priestore
- Cieľom je naučiť sa preniesť vstup na výstup
- Kompresia vstupu
- Redukcia šumu

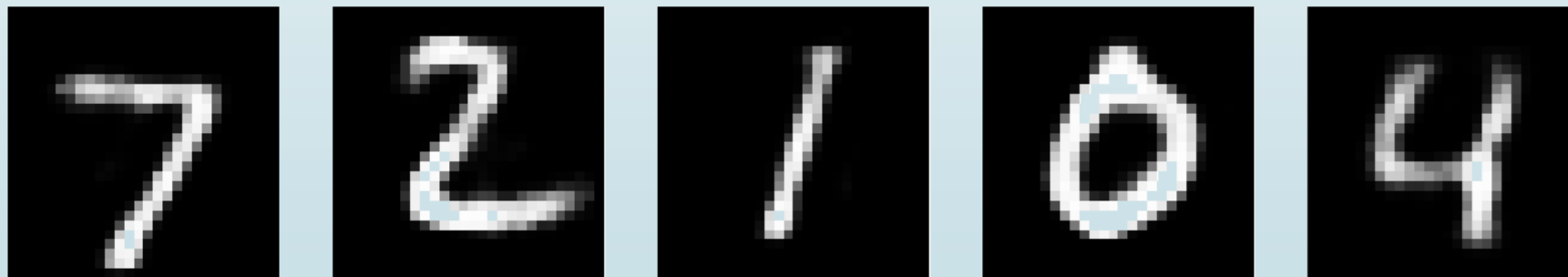


Autoenkóder - príklad

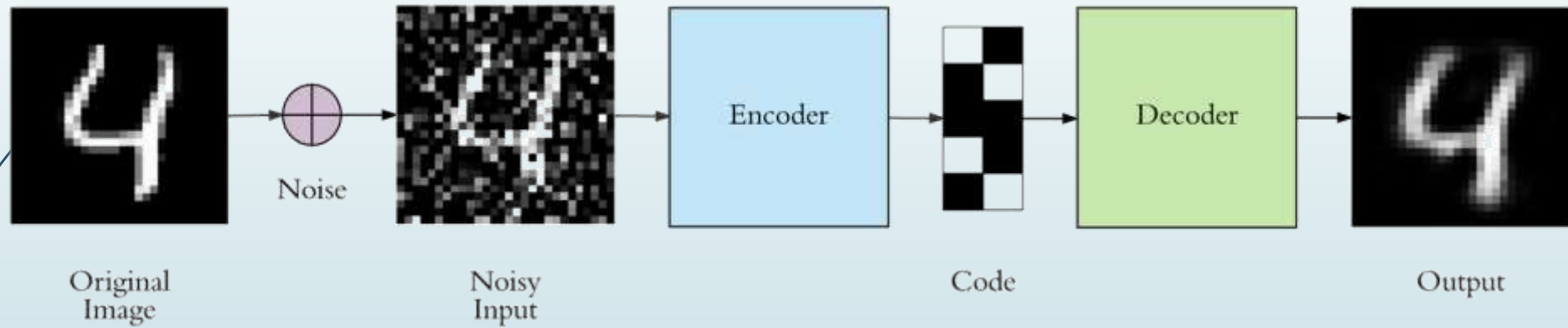
Original Images



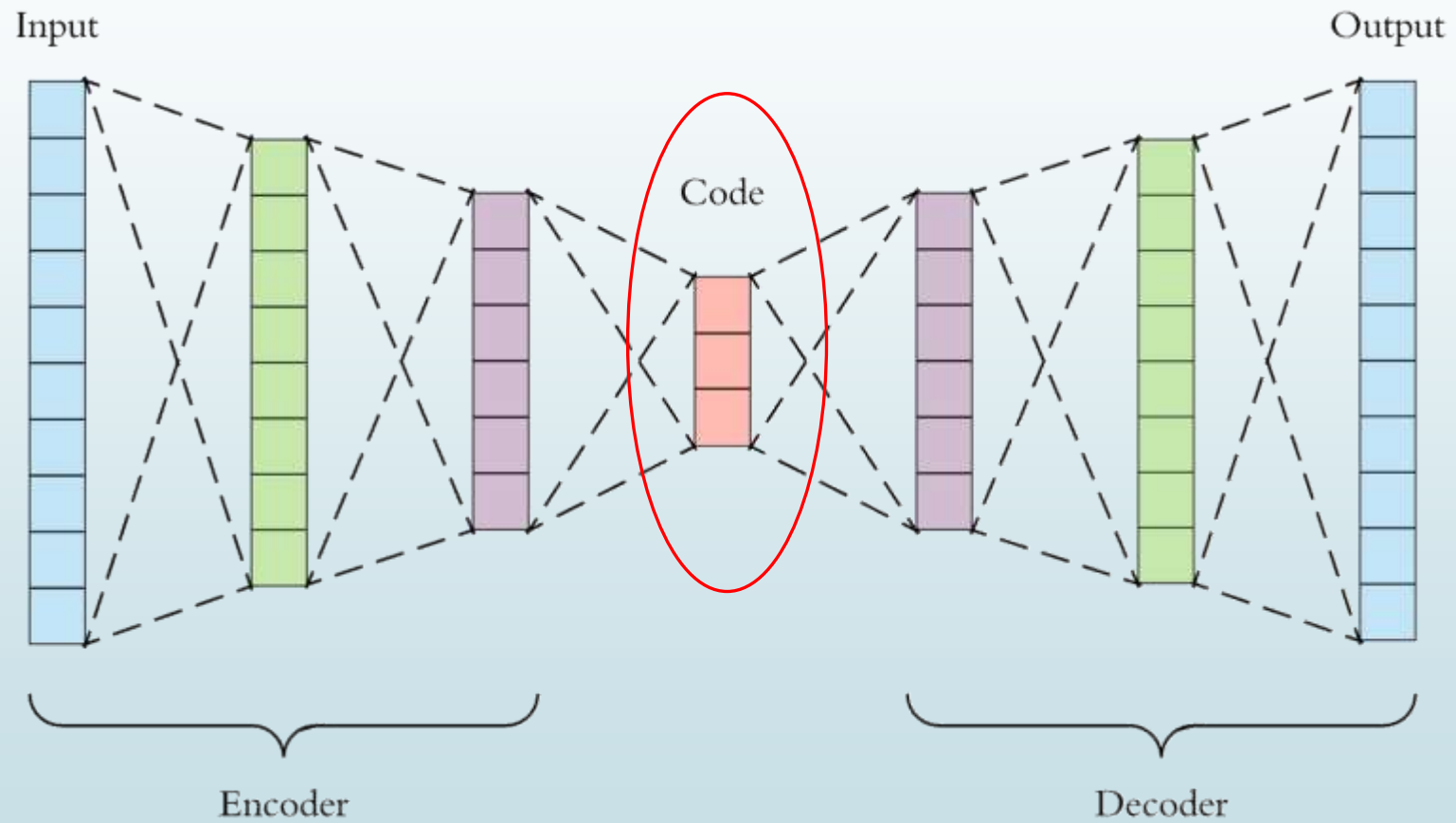
Reconstructed Images



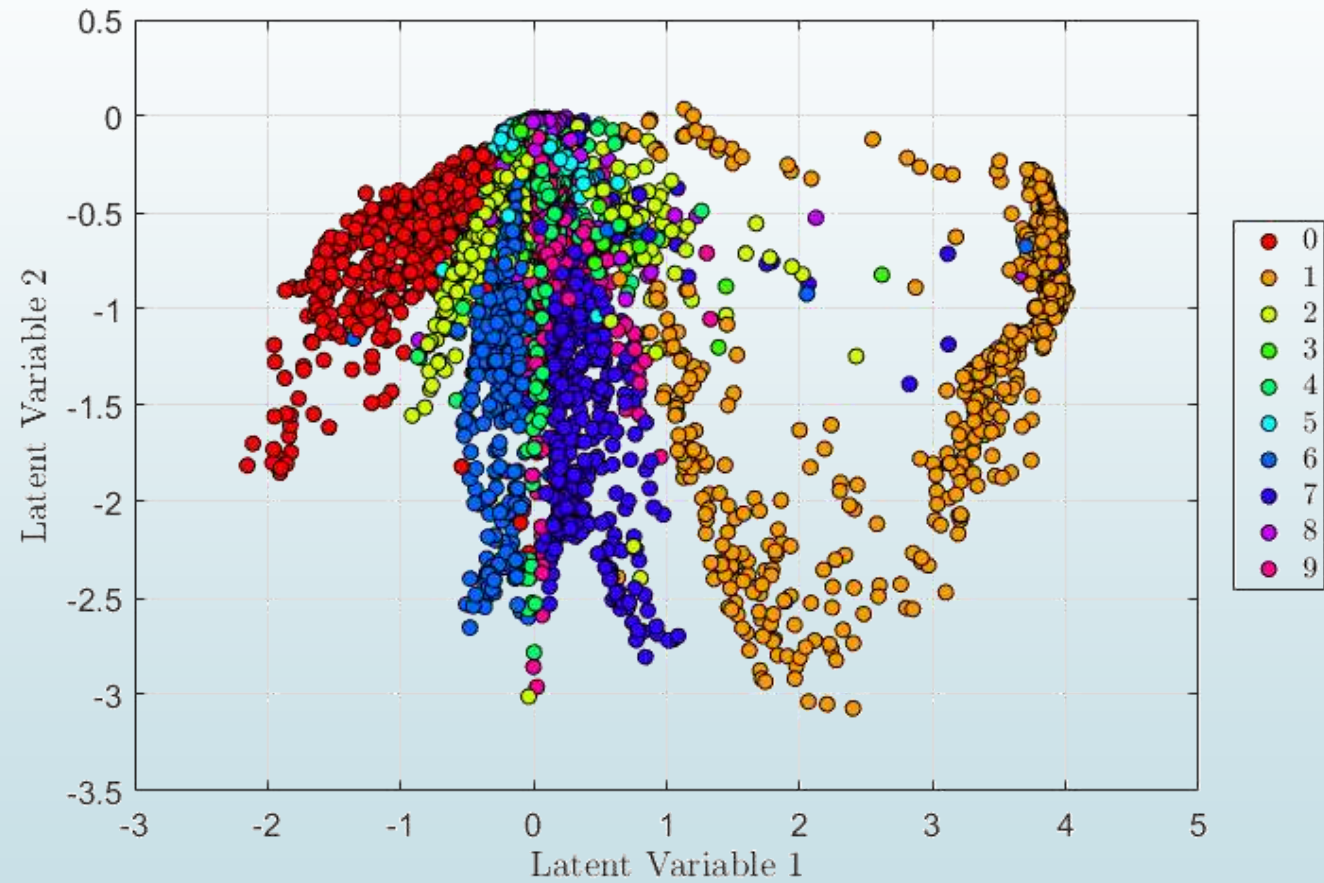
Autoenkóder – redukcia šumu



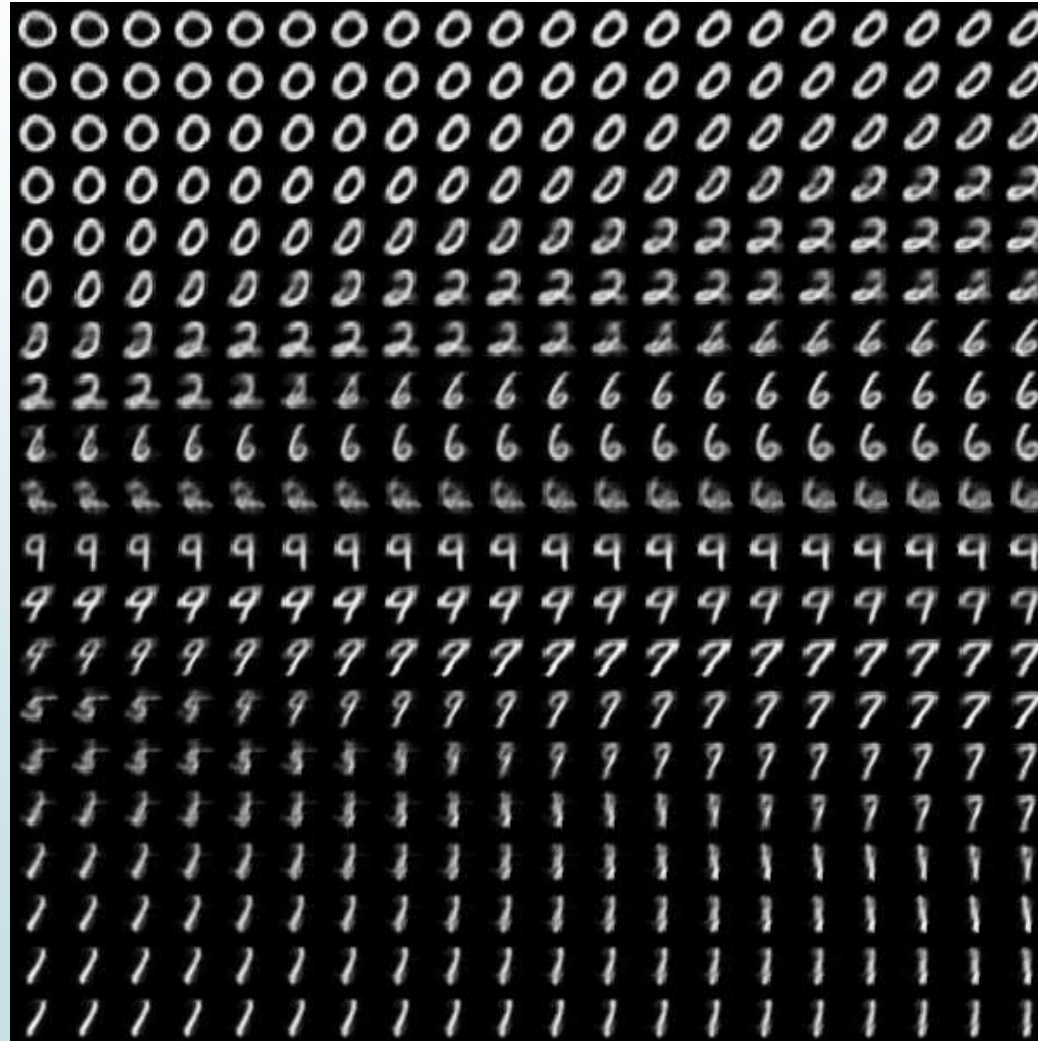
Latentný priestor



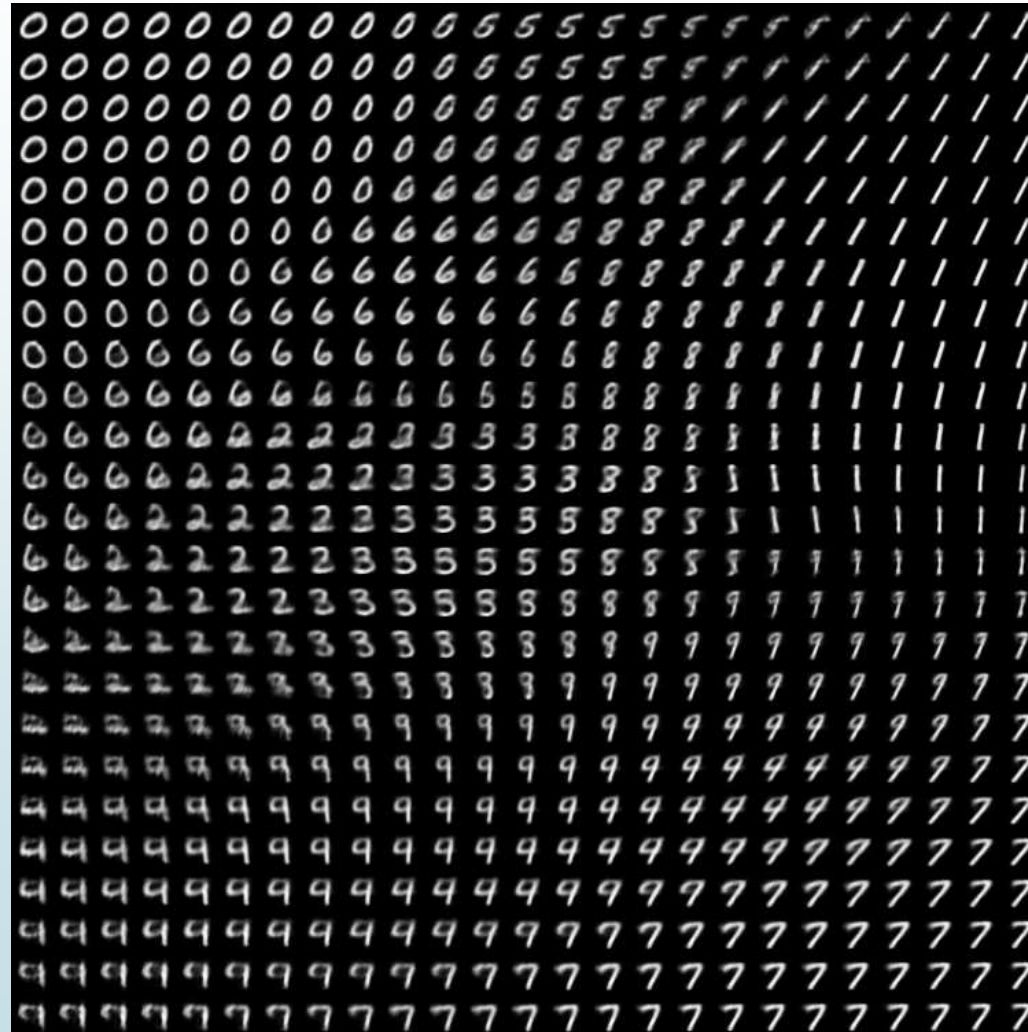
Latentný priestor



Latentný priestor bez metriky

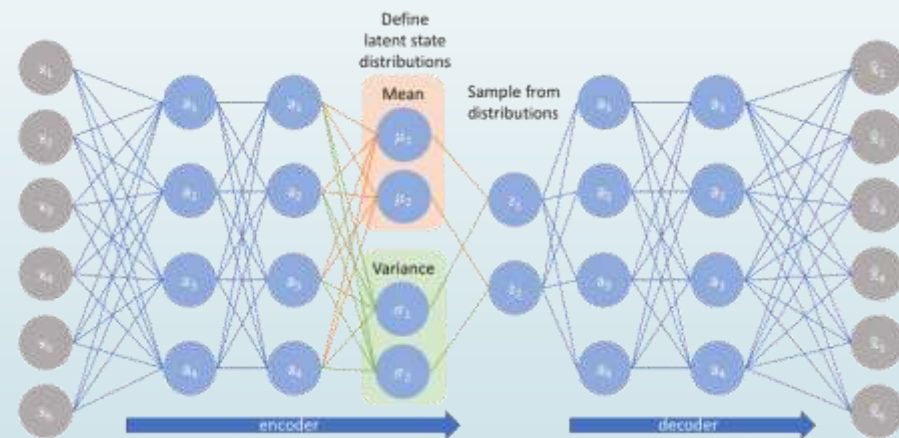


Latentný priestor s metrikou



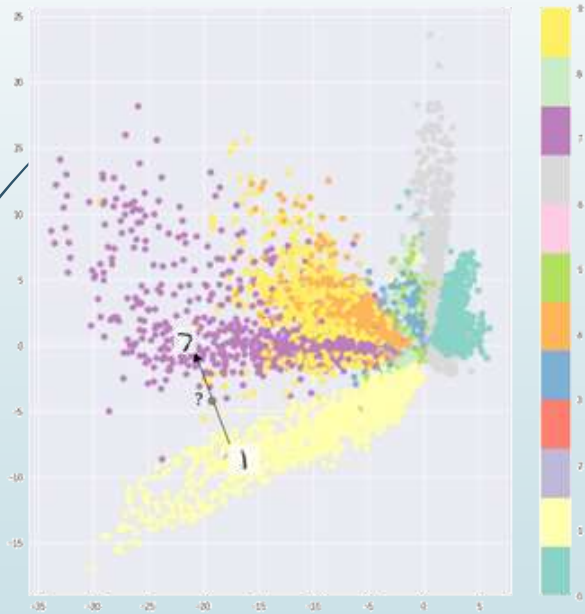
Variačný autoenkóder

- Má dve časti: enkóder a dekóder
- Cieľom je naučiť sa preniesť vstup na výstup
- VAE sa učí distribúciu vstupných dát namiesto priameho zobrazenia
- Predpokladá sa, že distribúcia je normálna (gaussovská)

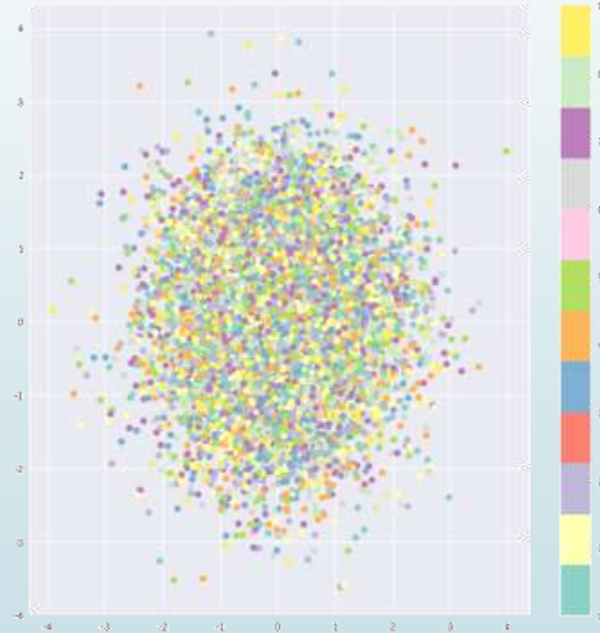


Latentný priestor VAE

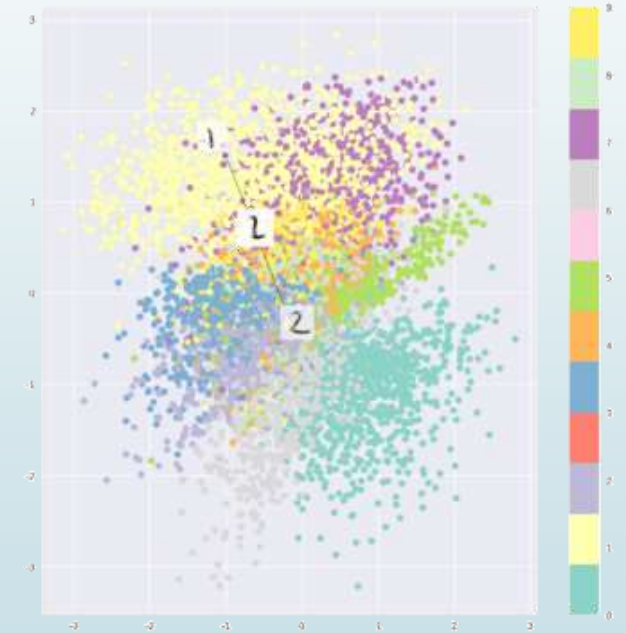
Only reconstruction loss



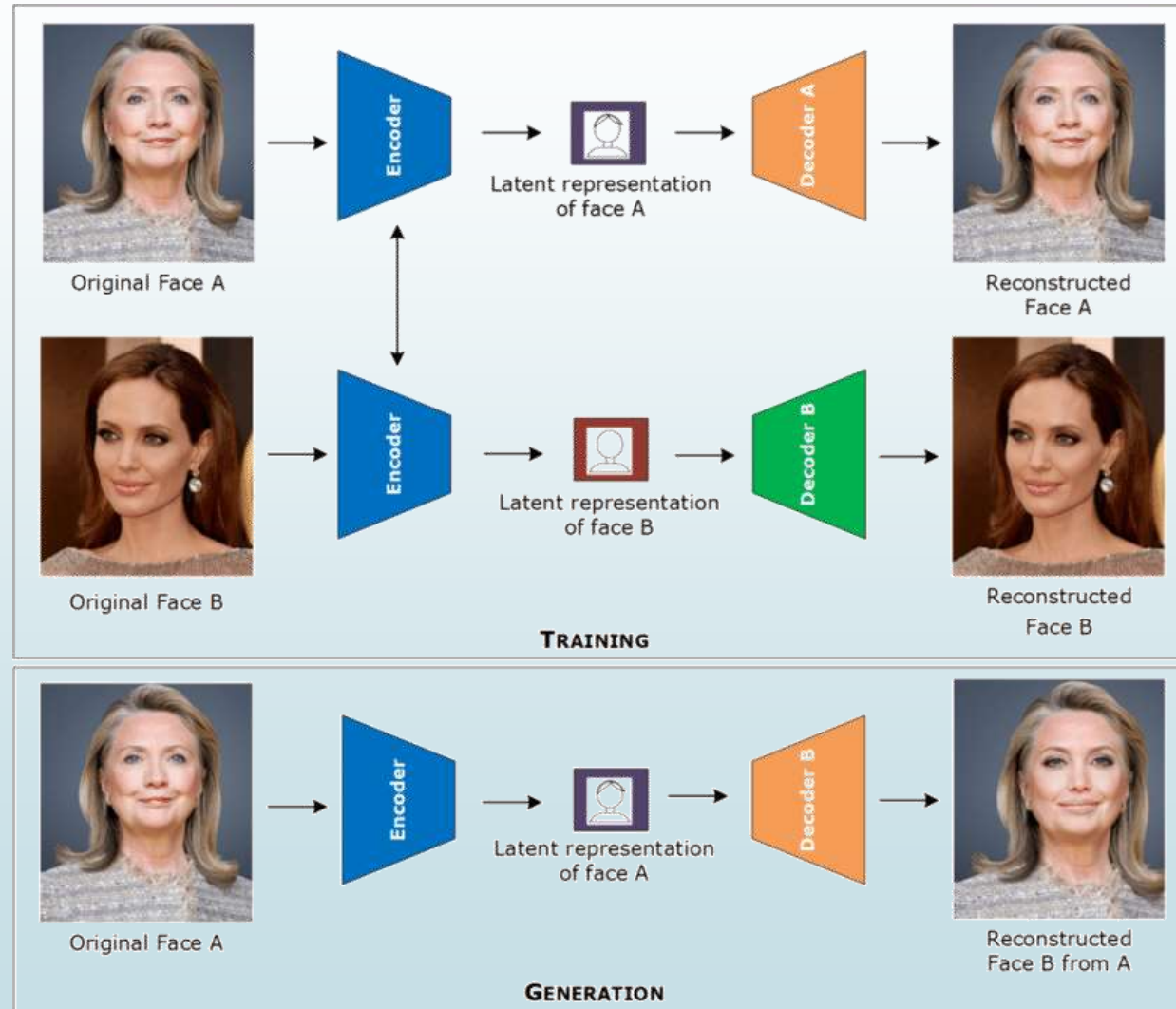
Only KL divergence



Combination

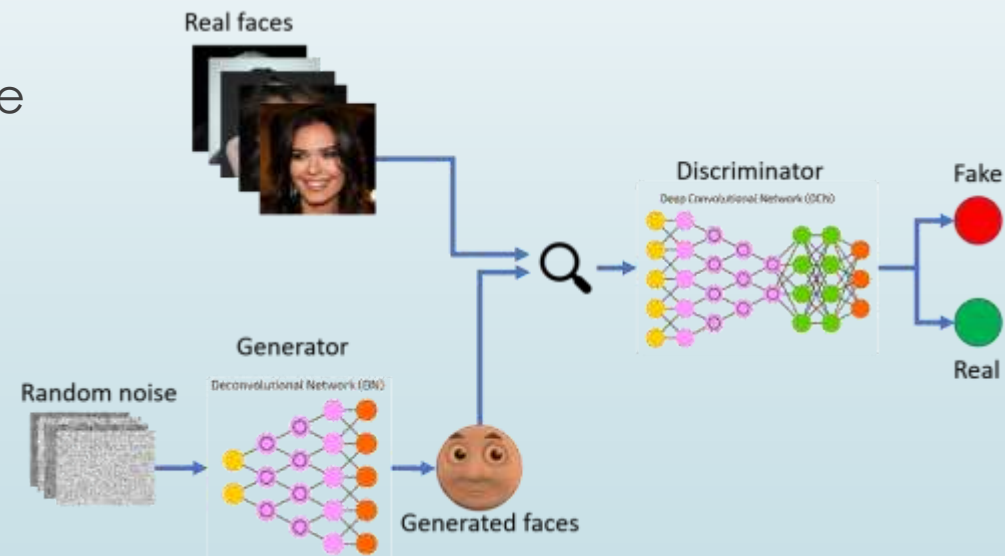


Deep Fake

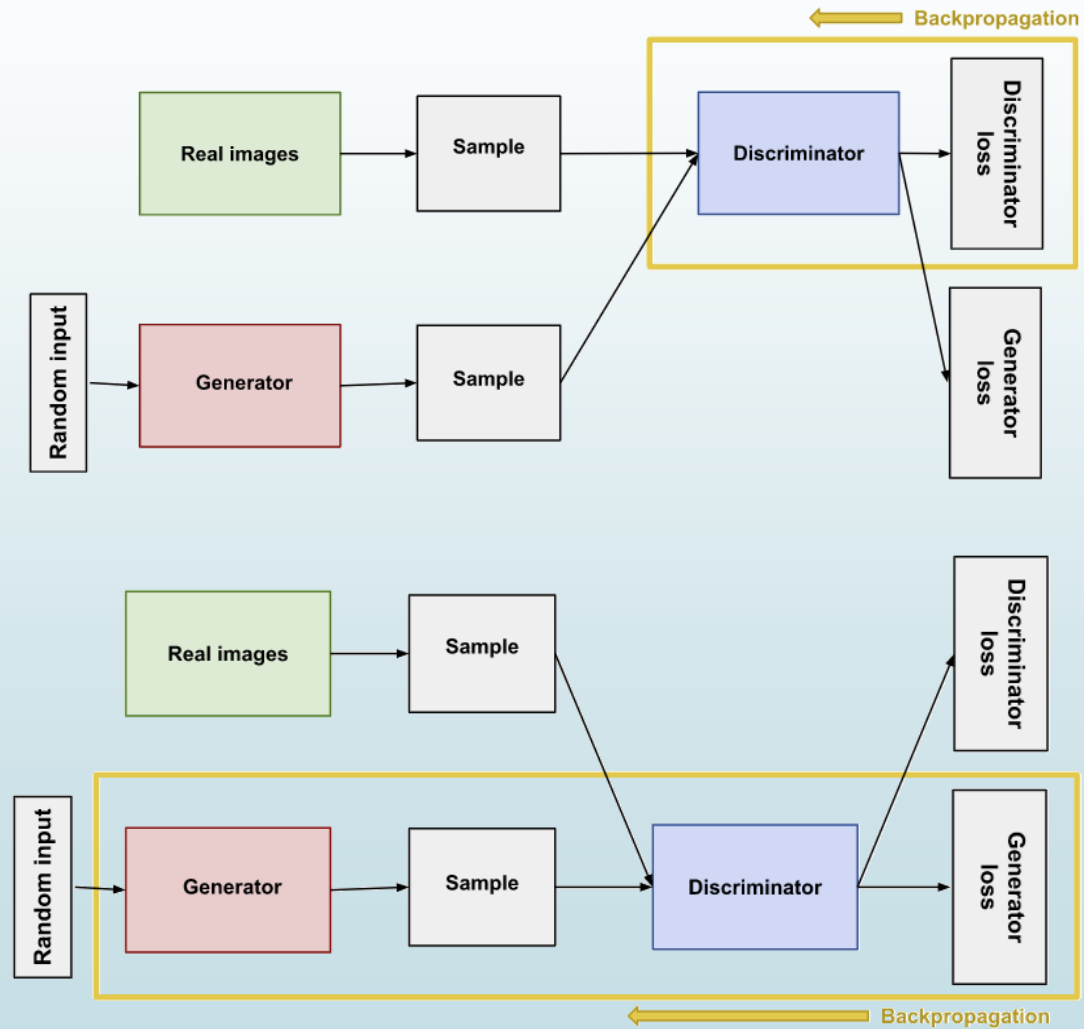


GAN – generative adversarial networks

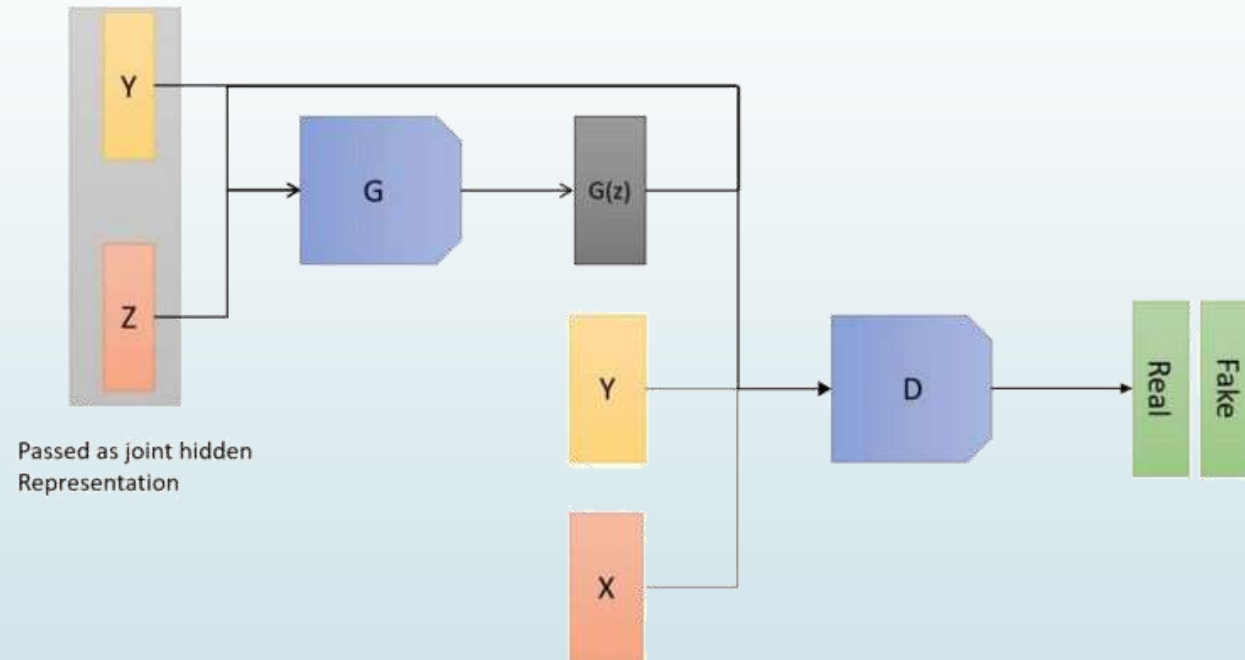
- Má dve časti: **generátor** a **diskriminátor**
 - Generátor vytvára vstup pre diskriminátor
 - Diskriminátor klasifikuje, či je vstup z reálnej alebo generovanej množiny
- GAN sa učí pomocou min-max game



GAN učenie



Podmienená GAN



Podmienená GAN

0-18

19-29

30-39

40-49

50-59

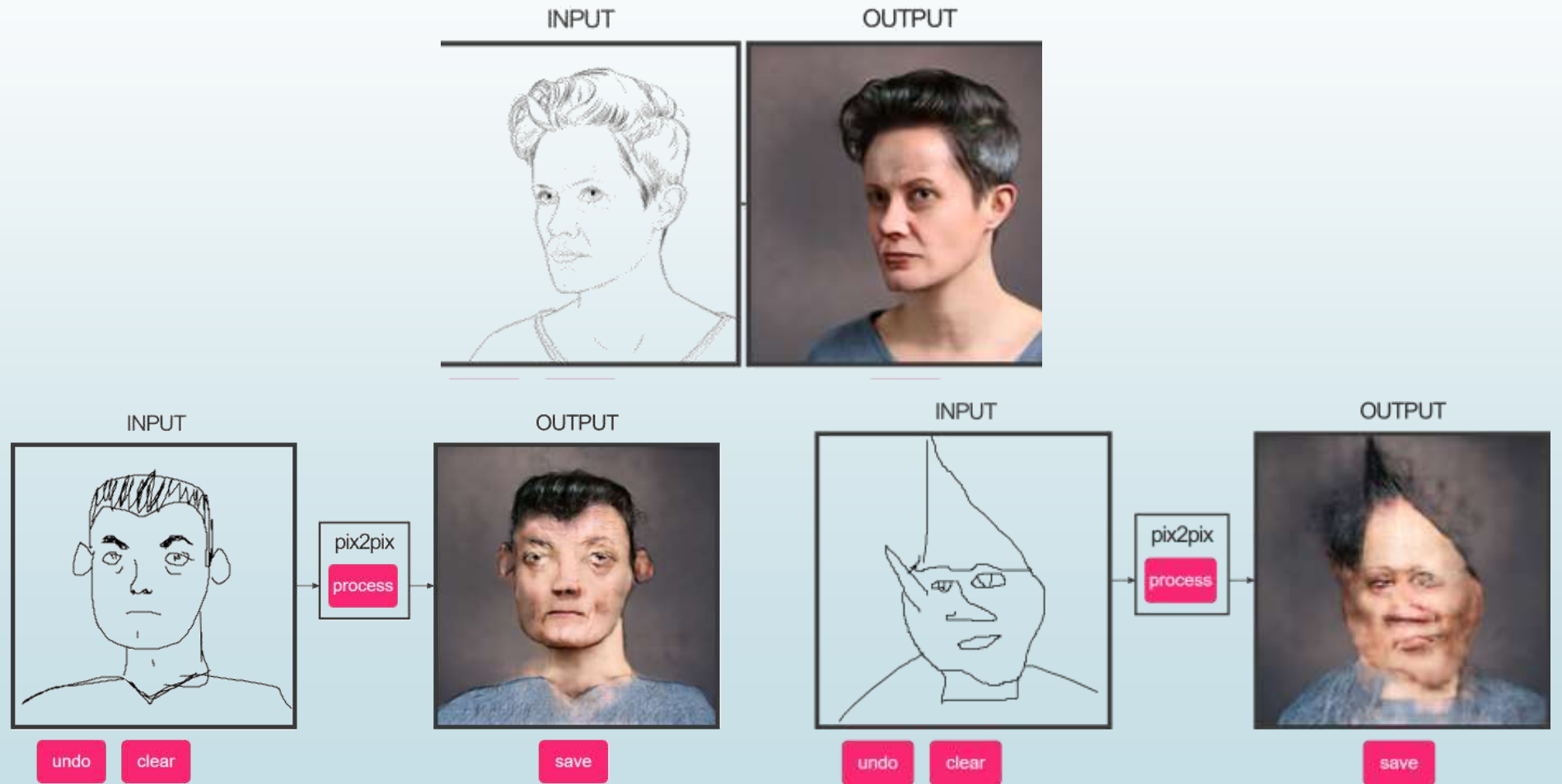
60+



GAN: zobrazenie obrázka na obrázok



<https://gan-paint-demo.mybluemix.net/>



GAN: zobrazenie textu na obrázok

Text
description

This bird is blue with white and has a very short beak

This bird has wings that are brown and has a yellow belly

A white bird with a black crown and yellow beak

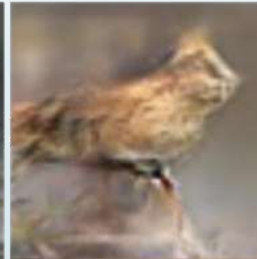
This bird is white, black, and brown in color, with a brown beak

The bird has small beak, with reddish brown crown and gray belly

This is a small, black bird with a white breast and white on the wingbars.

This bird is white black and yellow in color, with a short black beak

Stage-I
images



Stage-II
images



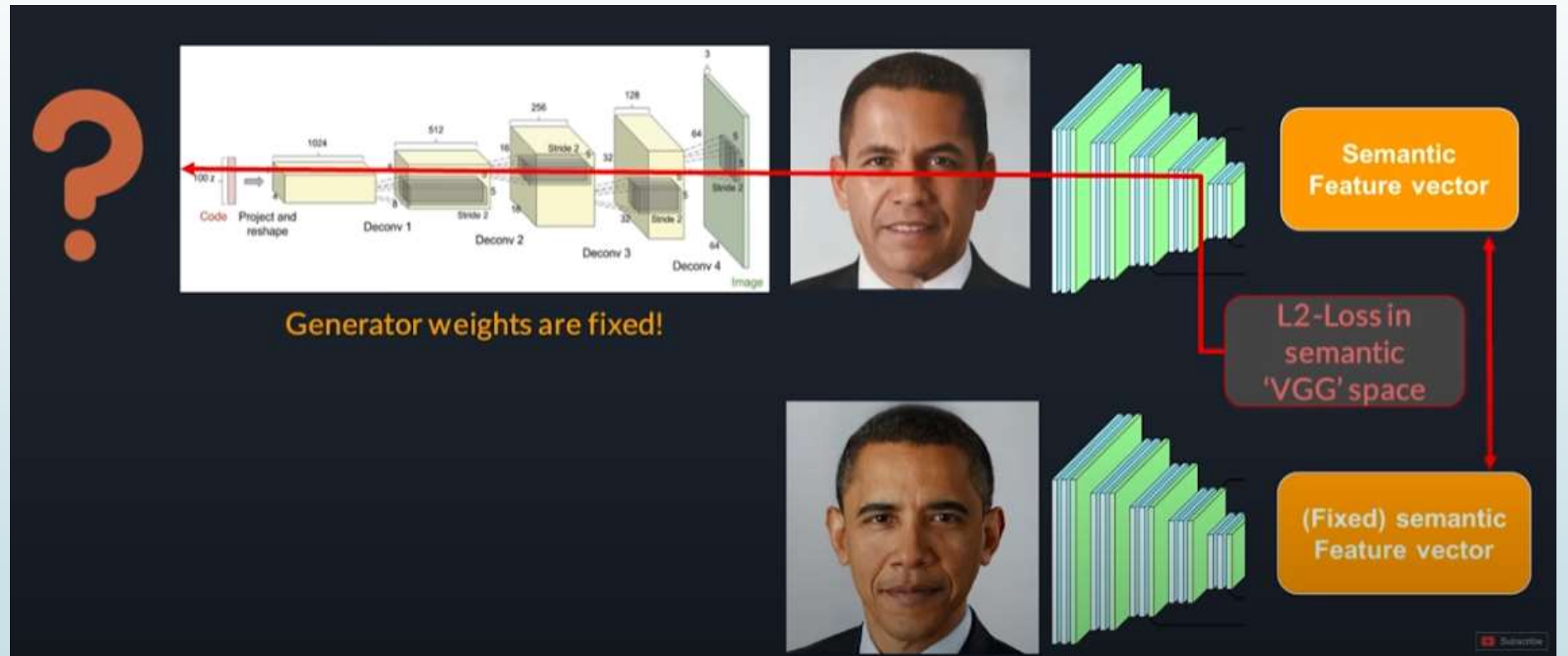
GAN: Super rozlíšenie



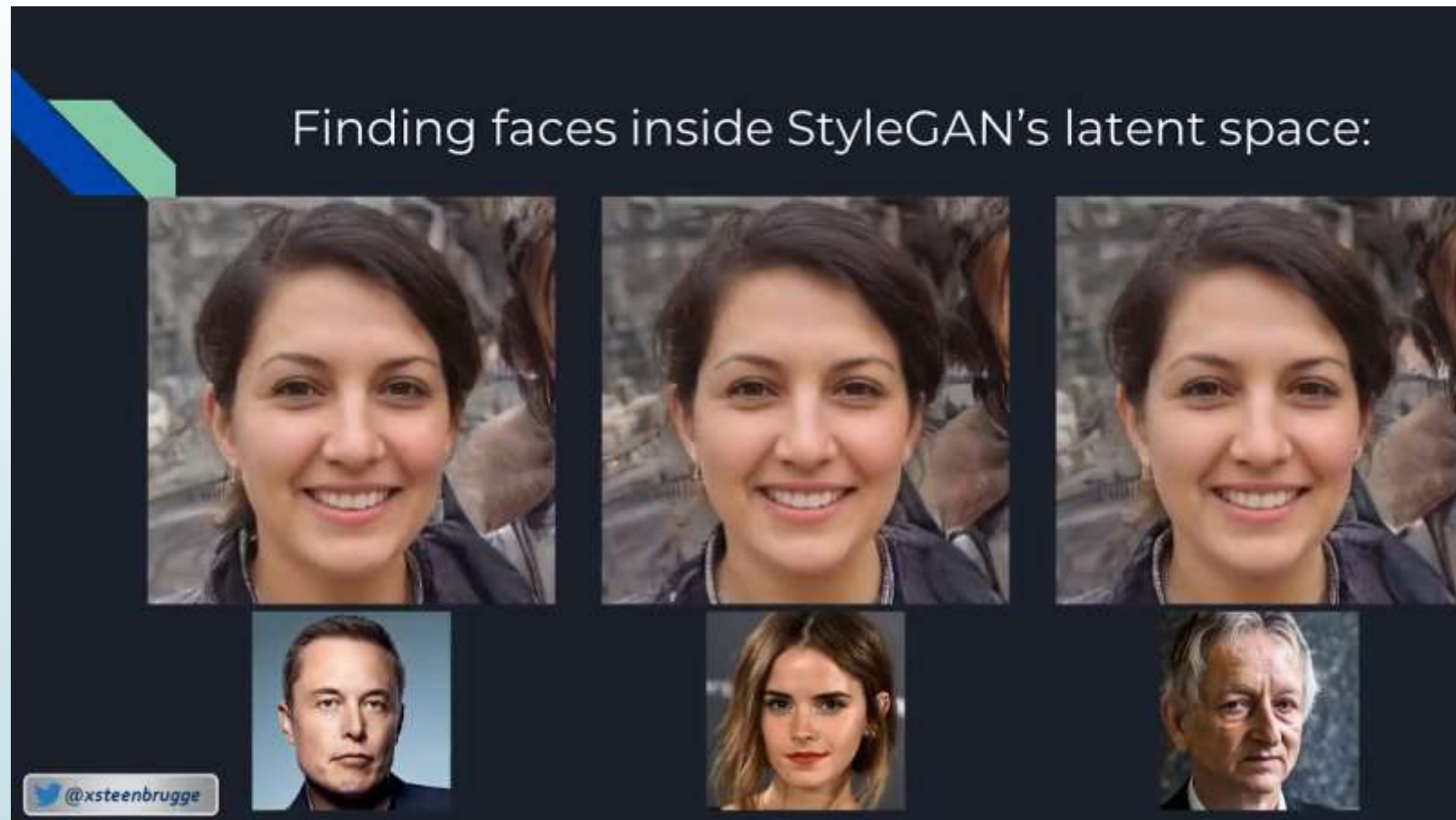
GAN: doplnenie obsahu



Ako nájsť latentný kód?

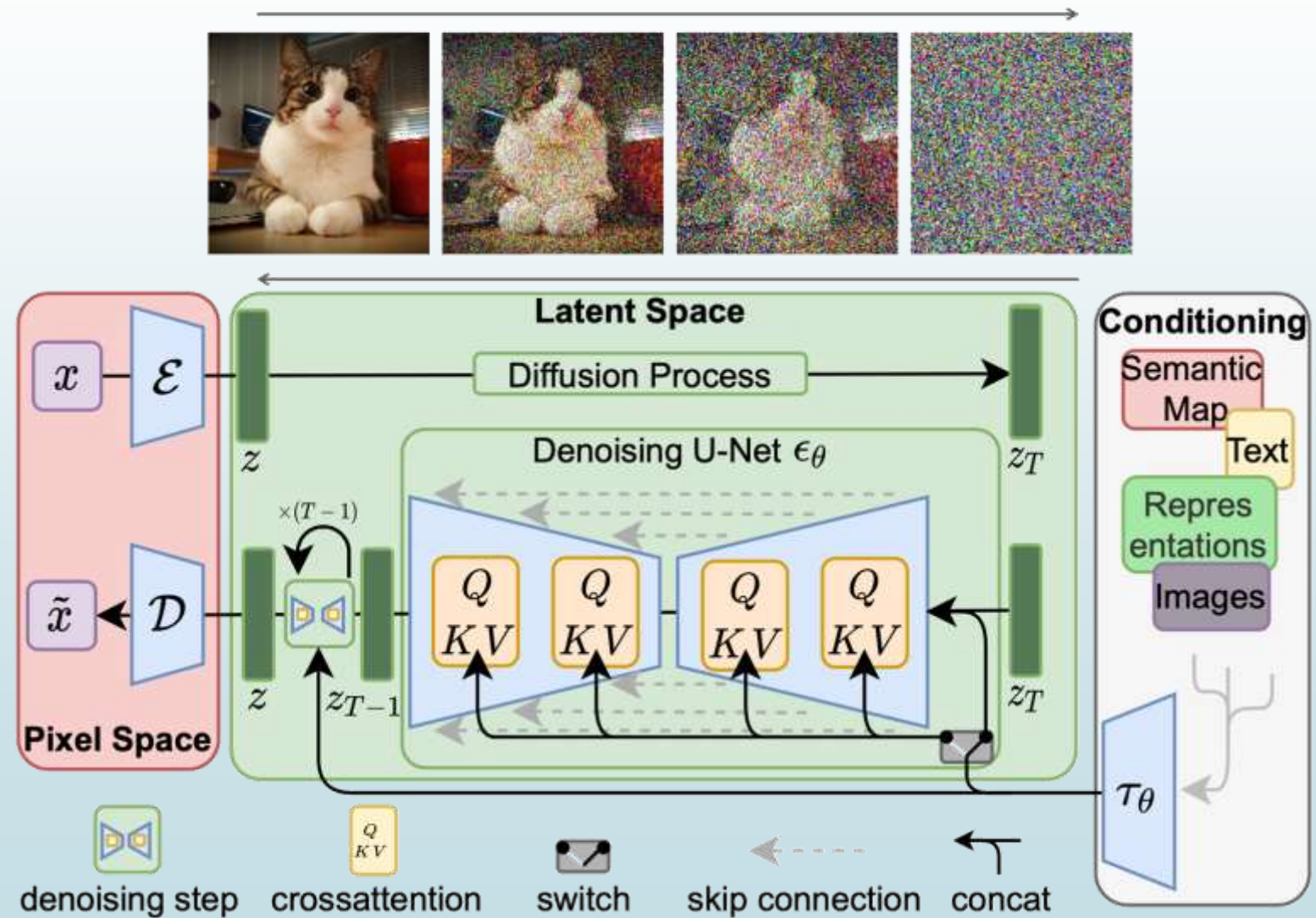


Ako nájsť latentný kód?





Difúzne modely



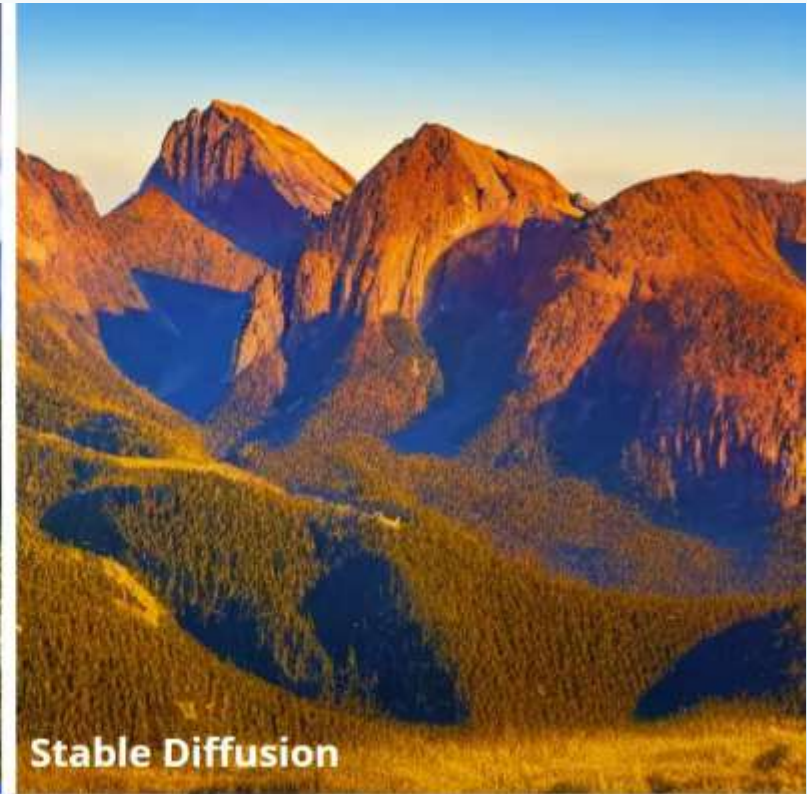
DALL-E2 vs Stable Diffusion

Portrait of a girl in a coffeeshop, reading a book, dramatic lighting



DALL-E2 vs Stable Diffusion

A realistic photo of a beautiful landscape





- ➔ <https://huggingface.co/spaces/stabilityai/stable-diffusion>

Využitie

- Generatívne modely poskytujú AI agentom „predstavivosť“
- Doplnenie obsahu na základe kontextu
- Odstránenie šumu – satelitné snímky, dáta z radarov, ...
- Vývoj liekov v medicíne
- Filmový priemysel - generovanie tvárí hercov
- Superrozlíšenie
- Editovanie fotografií

World's 1st Universal Processor for Servers / AI / HPC

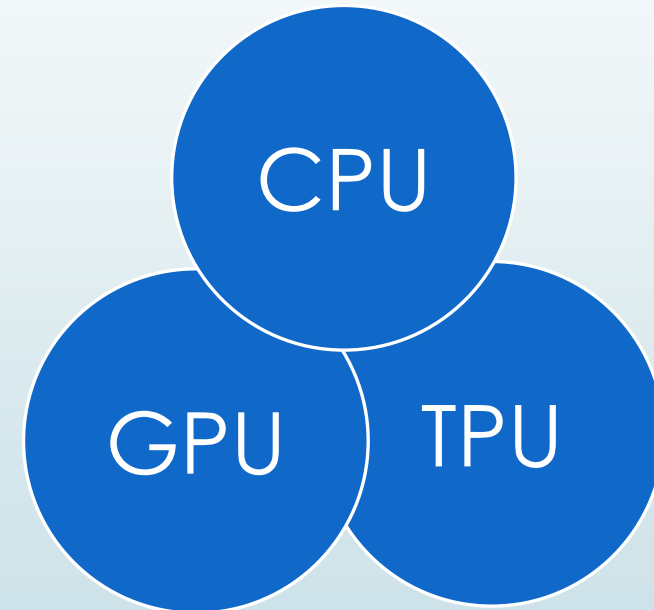
Server / Supercomputer / AI Chip

- For hyperscale datacenters

Prodigy is faster than Xeon/GPU/TPU

- Faster, 10x less power, 1/3 cost of Xeon
- Faster than NVIDIA A100 in HPC and AI

Tachyum Universal Processor is Best of



Long Vectors Matrix Operations

World's Fastest AI Supercomputer

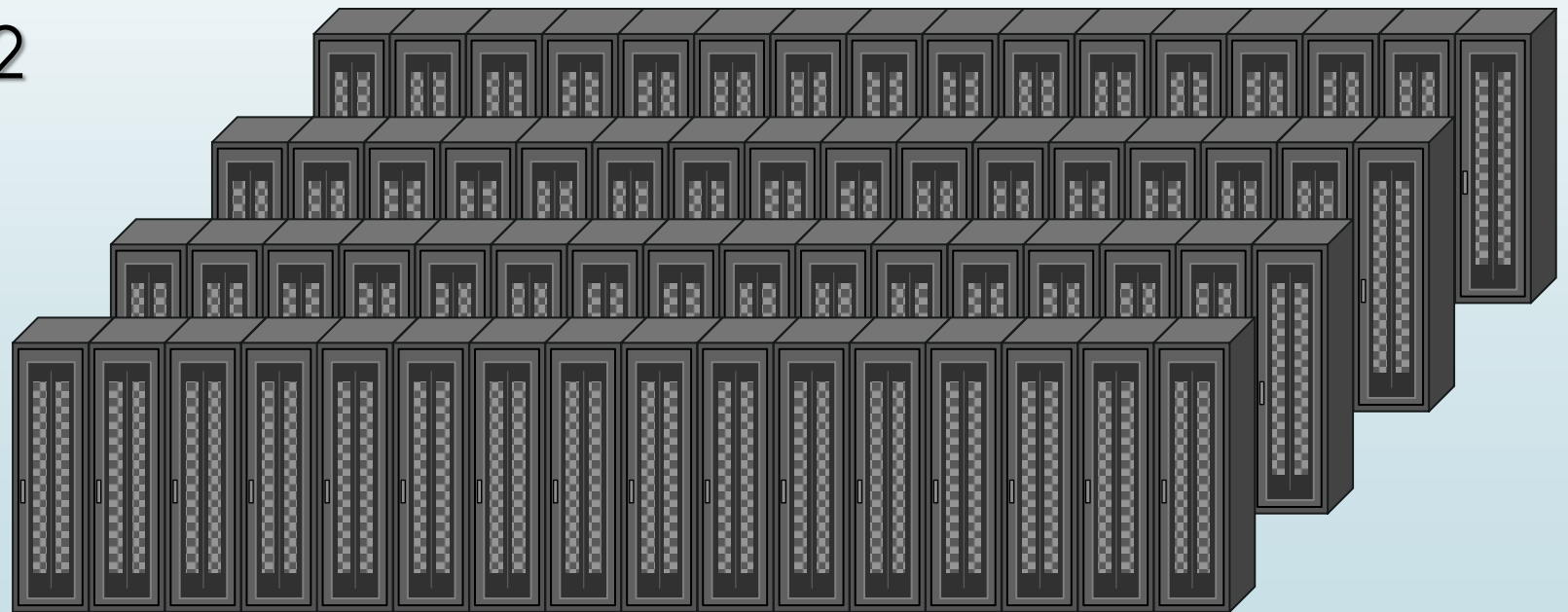


64 Compute Racks

64 AI ExaFLOPs

Operational in 2022

NSCC Slovakia Supercomputer



Prodigy-
Powered