

Střední průmyslová škola strojní   
a elektrotechnická a Vyšší odborná škola, Liberec 1, Masarykova 3

Využití Stable Diffusion pro vytváření ilustrační grafiky

Ročníková práce

Autor **Natálie Ryšavá**

Obor **Technické lyceum**

Vedoucí práce **Ing. Tomáš Kazda, DiS.**

Školní rok **2022/2023**

Anotace

Práce se zabývá …

Vychází z …

Přináší …

Summary

This work …

Čestné prohlášení

Prohlašuji, že jsem předkládanou maturitní práci vypracoval sám a uvedl jsem veškerou použitou literaturu a bibliografické citace.

V Liberci dne 22.05.2023

Natálie Ryšavá

Obsah

[Úvod 1](#_Toc130646050)

[1 Co je to neuronová síť? 2](#_Toc130646051)

[1.1 Využití neuronové sítě pro strojové učení 2](#_Toc130646052)

[1.1.1 Příklady aplikací strojového učení 2](#_Toc130646053)

[2 CLIP a difuzní modely 3](#_Toc130646054)

[2.1 Dall-E 2 3](#_Toc130646055)

[2.2 Imagen 3](#_Toc130646056)

[2.3 Stable Diffusion 3](#_Toc130646057)

[2.4 Midjourney 3](#_Toc130646058)

[3 Prostředí Stable Diffusion web UI 4](#_Toc130646059)

[4 Proces tvorby a optimalizace ilustrační grafiky 5](#_Toc130646060)

[5 Srovnání samplovacích metod 6](#_Toc130646061)

[6 Tvorba návodu 7](#_Toc130646062)

[Závěr 8](#_Toc130646063)

[Seznam zkratek a odborných výrazů 9](#_Toc130646064)

[Seznam obrázků 10](#_Toc130646065)

[Použité zdroje 11](#_Toc130646066)

[A. Seznam přiložených souborů I](#_Toc130646067)

Úvod

Autorka využila příležitosti více se seznámit s novou průlomovou technologií umělé inteligence Stable Diffusion.

Vývoj umělé inteligence se totiž v posledních letech velmi rychle rozvíjí a může být v budoucnu klíčovým faktorem v mnoha oblastech. Výzkumníci se snaží vylepšovat algoritmy, hardware a využívat nové technologie jako například kvantové výpočty a neuronové sítě. V budoucnu se očekává, že umělá inteligence bude mít ještě větší vliv na lidský život a bude schopna řešit složitější problémy.

Stable Diffusion je otevřený text-to-image model založený na metodě učení CLIP, který dokáže generovat podrobné obrázky z libovolného textového vstupu. Byl vyvinut společností StabilityAI ve spolupráci s EleutherAI a LAION a vydán v roce 2022. Model je primárně používán ke generování fotorealistických obrázků z latentního prostoru, upravování obrázků podle textu nebo doplňování chybějících detailů do obrázků s nízkým rozlišením nebo nedostatečnou kvalitou. Byl natrénován na miliardách obrázků a dosahuje výsledků srovnatelných s DALL-E 2 nebo MidJourney.

Stable Diffusion představuje významný krok vpřed v oblasti strojového učení. Tento model nabízí nové možnosti v oblasti umělé inteligence a může být využit mnoha způsoby, například v oblasti tvorby filmových efektů, reklamních kampaní nebo v herním průmyslu.

Cílem je využít tuto práci jako učební pomůcku neboli návod pro studenty na to, jak Stable Diffusion vůbec používat a implementovat ji do svých projektů.

[Stable Diffusion 2.0 Release — Stability AI](https://stability.ai/blog/stable-diffusion-v2-release)

Nastavte **název** dokumentu a **autora** v nabídce Soubor/Informace.

Pro vkládání zdrojů použijte Reference/Spravovat prameny.

# Co je to neuronová síť?

Neuronová síť (ANN neboli artificial neural network) je výpočetní model používaný v umělé inteligenci. Používá se k modelování vztahu mezi **vícerozměrnou vstupní proměnnou x** a **vícerozměrnou výstupní proměnnou y**. Obecně se jedná o nelineární regresní model, který lze vyjádřit síťovou strukturou.

Neuronové sítě můžeme definovat jako sadu algoritmů vytvořených podle fungování lidského mozku a nervového systému – pracují zkrátka na stejném principu jako trénovaný mozek dítěte. Skládají se z umělých (formálních) neuronů, což je jakási matematická funkce, která přijímá libovolný počet vstupů a poté je klasifikuje podle použitého algoritmu právě na jeden výstup. Každý neuron je tedy složen ze **vstupní vrstvy**, **více skrytých vrstev** a **výstupní vrstvy**. Ty jsou vzájemně propojeny synaptickými vazbami vytvářenými a upravenými pomocí strojového učení, což umožňuje neuronovým sítím adaptovat se na data a zlepšovat svou schopnost provádět úlohy, jako je klasifikace či predikce. Vrstvy si navzájem předávají signály a transformují je pomocí aktivačních přenosových funkcí.

## Učení neuronové sítě

Pro učení neuronové sítě se musí nejprve vytvořit struktura sítě, která slouží jako základ pro její následné trénování. Nicméně, sama o sobě tato struktura neumí v podstatě nic, dokud není trénována na konkrétní úlohu.

Trénování neuronové sítě spočívá v přijímání obrovského množství dat a postupném zlepšování její schopnosti interpretovat a chápat tato data. Přičemž během tohoto procesu se síti podávají různé podněty, které ji vedou právě k vhodnému porozumění předložených informací. V průběhu trénování si síť vytváří model, což je v podstatě abstraktní reprezentace vnitřních parametrů a vah, který se poté používá ke klasifikaci dat nových. Po dokončení trénování se sítě stává aktivní, kdy může být využita pro řešení zadaných úloh.

Existují různé výukové metody neuronových sítí, které se liší v závislosti na typu úlohy, datového souboru a požadovaném výstupu. Klasickými jsou dvě paradigmata typická pro strojové učení: **učení s učitelem** a **učení bez učitele**.

### Učení s učitelem

Učení s učitelem, známé také jako „supervised learning“, představuje třídu postupů trénování dat obsahujících vstupní objekty a odpovídající výstupní ohodnocení. To umožňuje strojům učit se z dat, která jsou již kategorizována nebo anotována lidmi pomocí klíčových slov, a tyto vzory pak slouží jako referenční bod. Proces probíhá tak, že stroj obdrží trénovací data, porovná je a snaží se najít funkci, která nejlépe popisuje jejich závislost mezi vstupem a výstupem. Na základě tohoto tréninku by měl s velkou přesností umět definovat data nová.

Například při trénování obrazové klasifikace by se síti mohly prezentovat různé obrázky spolu s příslušnými texty, které specifikují, co se na obrázku nachází. Síť pak adaptivně upravuje své vnitřní parametry tak, aby příště dokázala vstupní data správně vyhodnotit.

### Učení bez učitele

Učení bez učitele neboli „unsupervised learning“ se od učení s učitelem liší tím, že trénovací data neobsahují výstupní ohodnocení. To znamená, že funkce musí najít vztahy v datech sama metodou pokus-omyl. Využívá shlukování dat, aby objevila podobnosti mezi různými objekty, přiřadila jim vlastní označení a zařadila je do skupin s podobnými vlastnostmi.

## Využití neuronové sítě pro strojové učení

V dnešní době existuje mnoho známých nástrojů strojového učení, které přinášejí inovativní a užitečné řešení pro různé oblasti. Například to je rozpoznávání obrazů a řeči, klasifikace a segmentace dat, předpovídání vývoje časových řad, analýzy psaného textu nebo filtrování spamu. V lékařství se používá k diagnostice onemocnění a v řízení pro podporu rozhodování.

### CLIP a difúzní modely

CLIP (Contrastive Language-Image Pre-training) představuje model neuronové sítě vyvinutý společností OpenAI, který kombinuje obrazovou a jazykovou reprezentaci. Je předtrénován na více než 400 milionech dvojic textu a obrázku, a to například pomocí **difúzních modelů**. Tradičně bylo nutné pro kategorizaci shromáždit velké množství označených obrázků a na nich natrénovat klasifikační model. CLIP je však zero-shot model, což znamená, že dokáže identifikovat nespočet věcí, které nikdy předtím neviděl.

Difuzní modely jsou pokročilé strojově učené algoritmy pro generování dat, které využívají princip difuze k šíření informací a provádění změn v datových strukturách. Zpracovávají obrazy (případně i videa) s cílem odhadnout a předpovědět neznámé hodnoty na základě těch známých.

Během trénování se difuzní modely učí odstraňovat šum, který je zaveden do obrázků. Tento proces obnovy šumu je pak využíván při generování obrazů zcela nových, kde model vytváří realistické vizuální obsahy z původních nejasností. Díky principu difuze tyto modely dokáží efektivně šířit a propagovat informace skrze datové struktury, což přináší vylepšené výsledky při práci s obrazovými a video daty.

### Příklady aplikací strojového učení

Mezi nejpopulárnější aplikace strojového učení patří **generátory obrázků** nebo **chatovací nástroje**. Mohou být buď zdarma a volně dostupné pro veřejnost, nebo komerční a vyžadovat placenou licenci. Výběr konkrétní aplikace závisí na individuálních potřebách a požadavcích uživatele. Mezi ty nejznámější generátory obrázků patří Dall-E 2, Imagen, Midjourney a Stable Diffusion. K významným chatovacím nástrojům bychom naopak mohli přiřadit chatbot ChatGPT.

#### Dall-E 2

DALL-E 2 je nástroj umělé inteligence pro generování realistických obrázků a umění na základě textového popisu. Byl vyvinut společností OpenAI, na jejichž webu je dostupný veřejnosti s omezeným množstvím bezplatných obrázků a možností zakoupit další. Od dubna 2022, kdy byl odhalen, poskytuje ještě realističtější obrázky s 4× vyšším rozlišením než jeho původní verze DALL-E 1. Umí podávat výstupy v jakémkoli stylu, třeba i podle konkrétního umělce.

#### Imagen

Imagen reprezentuje jakousi verzi DALL-E od OpenAI vytvořenou společností Google, ale zatím není dostupná pro veřejnost – přístup mají pouze zaregistrovaní beta testeři. Je totiž trénována na obrovských souborech dat stažených z webu, které nejsou upravovány, což přináší řadu problémů.

Využívá umělou inteligenci pro zpracování obrazu, chlubí se nevídaným stupněm fotorealismu a hlubokým porozuměním jazyka. Dokáže také automaticky rozpoznat obsah obrázků či videí a přiřadit jim metadata. Staví na síle velkých transformačních jazykových modelů.

#### Midjourney

[Stable Diffusion vs. MidJourney vs. Dall-E-2 (buildspace.so)](https://buildspace.so/notes/stable-diffusion-vs-midjourney-vs-dalle-2)

#### Stable Diffusion

# Prostředí Stable Diffusion web UI

V této kapitole je bnfvj

## Technologie generování obrázků

### Text-to-Image

### Image-to-Image

#### Inpaint

[Diffusion Models: A Practical Guide | Scale AI](https://scale.com/guides/diffusion-models-guide#diffusion-model-prompt-engineering)

### Extras

## Prompt

[Diffusion Models: A Practical Guide | Scale AI](https://scale.com/guides/diffusion-models-guide#diffusion-model-prompt-engineering)

## Negative prompt

## Sampling Steps

## Sampling Method

### Srovnání samplovacích metod

Bylo vybráno dpm2, Euler a, lms karras.

## Rozměry

## Batch count

## Batch size

# Proces tvorby a optimalizace ilustrační grafiky

Zaměřila jsem se na realistické fotografie…. Atd (dopíši úvod)

## Člověk

## Zvíře

## Krajina

## Architektura

Závěr

Tak jsem se dostal až na konec.

Seznam zkratek a odborných výrazů

HTML

HyperText Markup Language – značkovací jazyk používaný pro tvorbu webových stránek.

Seznam obrázků

[Obrázek 1 Úplně bez legrace, mě tohle kotě docela děsí. 3](#_Toc86059903)

[Obrázek 2 Modré borůvky 3](#_Toc86059904)

Použité zdroje

1. **Stehlík, Michal.** *Návod k maturitním pracím 2020.* Liberec : Albatros, 2020.

1. Seznam přiložených souborů

Na přiloženém datovém nosiči se nacházejí následující soubory a složky:

* **MP2010-Novák-Jan-L4-Tepelné\_čerpadlo.docx** – editovatelná verze dokumentace maturitní práce
* **MP2010-Novák-Jan-L4-Tepelné\_čerpadlo.pdf** – tisknutelná verze dokumentace maturitní práce
* **Výkresy** – kompletní výkresová dokumentace
* **Aplikace** – zdrojové kódy