

Umělá inteligence pro běžnou praxi.

Nejmodernější „chytré“ přístroje v sobě nesou prvky umělé inteligence (AI). Je to ale ta umělá inteligence, před kterou varují známé osobnosti? Jistě, že ne. Jedná se sice o poměrně složité, rafinované, ale stále jen pevné počítačové programy, které nám přinášejí pokroky do běžného života. Pomineme-li zneužívané systémy pro rozpoznávání lidských tváří, či nepopulární automatické čtečky poznávacích značek automobilů, můžeme uvítat šikovné rozpoznávání textu či hlasu a mnoho dalších užitečných aplikací.

Například parní generátor žehličího systému nejmenovaného výrobce dokáže sám optimalizovat nastavování teploty. Pomocí zabudované kamery a umělé inteligence jako údajně jediný generátor na světě pozná, co právě žehlíte a přizpůsobí látce své nastavení.

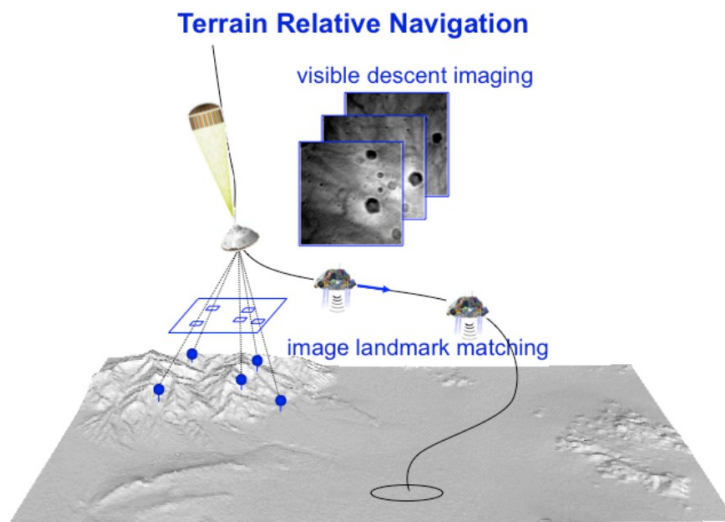
Automatickému řidiči bychom v běžném provozu zřejmě ještě nedůvěřovali, ale ten se na rozdíl od nás nikdy neunaví. Jak spolehlivě se asi dokáže umělá inteligence orientovat v „terénu“ si můžete simulovat sami. Stačí mít nainstalovanou některou mobilní aplikaci využívající předučený detektor objektů YOLO (poslední verze 3) a namířit vestavěný fotoaparát chytrého telefonu na vaše bezprostřední okolí. V ulicích New Yorku se vám na displeji objeví něco podobného jako na obrázku níže. Využitelnost pro řídicí systémy autonomních dopravních prostředků je nasnadě.



Rozpoznávání objektů pomocí algoritmu YOLO v3 (zdroj YouTube).

Umělá inteligence se může starat rovněž o naši bezpečnost. Na základě signálů ze speciálních senzorů umístěných na kolejnicích rozpoznává projíždějící vlak, na železniční svršek padající kameny nebo rozvoj defektů v kritických částech výhybek. Může vyhodnocovat kvalitu a závadovost různých produktů přímo na běžících výrobních linkách a to na základě mnoha senzorů a detekčních metod, které nekompromisně předčí lidské smysly.

Tyto technologie se mohly ještě nedávno zdát jako „z Marsu“. Však i tam nachází strojové učení své uplatnění. Kvůli příliš dlouhé době cesty řídicího signálu ze Země byla sonda Perseverance odkázána sama na sebe a její přistání bylo navigováno automatickým vizuálním systémem LVS (Lander Vision System), který v reálném čase porovnával data ze své kamery s palubní mapou v podobě snímků pořízených s vysokým rozlišením předchozími orbitálními sondami. Systém dokázal rozpoznat charakteristické geologické útvary a vybrat optimální bezpečné místo pro přistání v předem vytyčené oblasti kráteru Jezero.



*Ilustrace „relativní navigace podle terénu“ při přistání sondy Perseverance na Marsu.
(zdroj http://mepag.jpl.nasa.gov/reports/MEP/Mars_2020_SDT_Report_Final.pdf)*

V současnosti existuje mnoho známých algoritmů založených na modelech tzv. umělých neuronových sítí a někteří výrobci hardwaru rozšiřují své procesory o jádra speciálně určená pro strojové učení (viz např. Neural Engine revolučního procesoru M1 společnosti Apple). Jak ale probíhá teoretický vývoj zmiňovaných softwarových technologií a jak to vypadá pod jejich „kapotou“?

Na začátku všeho byl v polovině minulého století tzv. jednoduchý perceptron, který se mohl naučit pouze základní logické operace, či oddělování dat pomocí lineární hranice. Po vyřešení matematického problému učení celého systému obdobných jednotek (propojených a uspořádaných do vrstev) nastal zlom v teoretických i reálných aproximačních schopnostech a tzv. vrstevnaté sítě našly mnohé uplatnění v praxi. Objevily se paralelně i jiné architektury, koncipované pro rozpoznávání obrazu nebo klasifikaci dat. Výkon tehdejšího hardwaru (přelom století) a teoreticky malá informační kapacita známých modelů sítí ale neumožňovaly tak široké nasazení, jako je tomu dnes.

Kýžený boom nastal s návrhem tzv. konvolučních neuronových sítí (CNN), které mají aktuálně až nepřehledné množství variant. Bez hlubšího komentáře ponechme typické schéma vybrané architektury GoogLeNet na obrázku níže. Všimněme si jen, že má poslední výpočetní bloky převzaté od klasických vrstevnatých sítí a vstupní jednotky koncipované pro zpracování obrazových dat, což záměrně předesílá její využití. Lze ještě dodat, že problémem současnosti již není ani výpočetní výkon, ale dostupnost dostatečně rozsáhlé sady relevantních tréninkových dat.

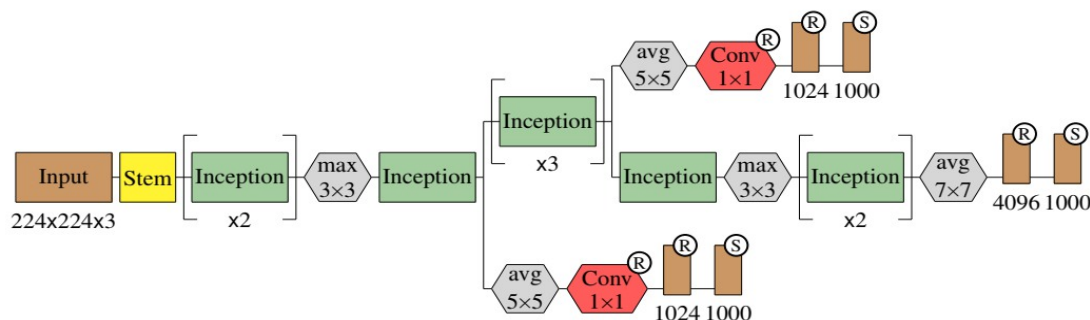


Schéma architektury GoogLeNet.

Klíčovou fází návrhu systému využívajícího strojové učení je totiž označení, či pojmenování objektů pro neuronovou síť k zapamatování. Jednou z nejrozsáhlejších takových databází je známý ImageNet dataset, který obsahuje aktuálně přes 14 milionů slovně popsaných fotografií z 21841 kategorií různých předmětů, objektů, živočichů, rostlin atd. Od roku 2010 pak pravidelně probíhá soutěž výzkumných týmů v rozpoznávání těchto fotografií. Po natrénování sítě se prověřuje jejich úspěšnost na základě obrázků, které nebyly síti dosud předloženy. Srovnání některých známých architektur sítě na redukované databázi mající „pouze“ 1000 kategorií můžeme vidět na obrázku. Vodorovná osa představuje počet milionů parametrů sítě, které musejí být ve fázi učení optimalizovány. Svislá osa pak uvádí procentuální úspěšnost správné klasifikace předložené fotografie.



Vztah velikosti a úspěšnosti vyvíjených architektur na testovací databázi ImageNet.

Z obrázku je patrné jak se např. mění přístupy vývojářů architektur v průběhu let. Na začátku vývoje byly testovány architektury, ve kterých se střídaly konvoluční (filtrační) a poolingové (obraz „zmenšovací“) vrstvy. Od dob AlexNetu (není na obrázku, protože má příliš nízkou přesnost) se začalo experimentovat s použitím více konvolučních vrstev za sebou. Hlavní argument spočíval v tom, že vlivem více nelineárních transformací získává síť více možností pro učení a generalizaci dat. V tomto duchu se nesly i modely VGG-16 a VGG-19, které obsahovaly velké množství parametrů, a byly proto hůře upočítatelné. Výhodou však bylo výrazné zpřesnění predikce modelů.

Jinou cestou se vydali tvůrci sítě s názvem GoogLeNet (Inception-v1), kteří přišli s paralelizací vláken v rámci jednoho bloku sítě. CNN síť se potom tvořila z takto definovaných bloků a výsledkem bylo nepatrné zvýšení přesnosti, ovšem s výrazným snížením počtu parametrů a tím pádem ke zjednodušení trénovacího procesu.

Zajímavým milníkem bylo použití residuálního učení (residual learning), které do modelu přináší jakési zkratky v podobě přičtení vstupu jednoho bloku k jeho výstupu. Autoři tohoto konceptu experimentálně dokázali možnou užitečnost tohoto nápadu, který se následně začal objevovat ve většině nejlepších modelů. Vhodnost tohoto postupu se snažil dokázat mimo jiné i tým pana Szegedy, který vyvinul architektury Inception-v4 a Inception-ResNet-v2, kde residual learning použil pouze u druhé jmenované a snažil se je porovnávat mezi sebou.

Další pokrok přišel s modelem NASNet, pro nějž byl zadefinován matematický prostor, na kterém se hledají a porovnávají architektury. Výběr struktur se tímto omezil a tudíž i zjednodušil. Tohoto kroku následně využila přelomová síť AmoebaNet, v jejímž případě byly jednotlivé strukturní bloky modifikovány pomocí evolučního algoritmu. V poslední době byl vyvinut koncept EfficientNetu, který díky vhodnému poměru mezi počtem vrstev a šířkou rozvětvení modelu dokázal zvýšit, či přinejmenším zachovat kvalitu modelu za výrazného snížení počtu parametrů, které je při učení potřeba optimalizovat (viz samostatná vývojová linie B0-B7 na obrázku). Vzhledem k tomu, že se CNN sítě často využívají i v mobilních zařízeních, nezáleží pouze na přesnosti, ale také na numerické náročnosti modelů. Kvůli tomu vzniklo několik verzí EfficientNetu, které se hodí na různé typy úloh a zařízení. S postupem doby lze pochopitelně očekávat další rozvoj v této oblasti, jelikož nějaké zásadní technické či teoretické limity zatím nejsou na dohled.

Akademičtí členové konsorcia B4I mají dlouhodobou zkušenost se zpracováním signálů z metod nedestruktivního testování (NDT), které jsou jakýmsi ekvivalentem smyslových vjemů u lidské inteligence. Velmi dobře se osvědčily klasické neuronové sítě při lokalizaci poruch na složitých tělesech nebo konvoluční neuronové sítě pro rozpoznávání spojitých emisních signálů z převodových ústrojí. Velký potenciál má jejich nasazení při výrobních procesech, konkrétně sledování stavu opotřebených obráběcích nástrojů. Poslední experimenty rovněž dokazují možnost detekce počátečních fází plastizace kovových materiálů a tudíž dostatečně včasného varování před kolapsem zatížené konstrukce.

Přestože může akademická komunita vidět aplikační potenciál propojení NDT s AI především prostřednictvím projektů, které zavála praxe do jejich laboratoří, vnímá současnou vyjímečně příznivou konstalaci připravenosti teoretických matematických modelů a všeobecných technických možností. „Třetí“ vlna rozvoje v oblasti umělé inteligence, ruku v ruce s dostupností výkonného hardwaru pro NDT metody, poskytují značný inovační prostor nejen pro velké hráče, vyvíjející avantgardní novinky na poli umělé inteligence. Nově vzniklé centrum B4I propojuje akademické výzkumné laboratoře se širokou platformou malých a středních podniků. Takto zamýšlená symbióza má za úkol přinést řadu užitečných aplikací, zvýšit konkurenceschopnost průmyslu a orientovat vědu věrohodnými progresivními směry.

Milan Chlada, Martin Kovanda 2021