

# EVOLUČNÍ NÁVRH KONVOLUČNÍCH NEURONOVÝCH SÍTÍ

## Motivace

Návrh architektury konvolučních neuronových sítí (CNN) dnes hraje důležitou roli ve výzkumu umělé inteligence. Návrh nových struktur CNN vyžaduje velké úsilí, mnoho zkušeností a pokusů. Tento proces je tak velmi namáhavý a ne vždy efektivní. Z tohoto důvodu je nutné přemýšlet o nových způsobech tvorby těchto sítí, které by nebyly tak složité a náročné, a přitom poskytovaly výsledky minimálně srovnatelné s ručně navrženými sítěmi. Technika automatizovaného hledání struktur umělých neuronových sítí, nazývaná NAS (Neural Architecture Search), je v dnešní době velmi populární a již dala vzniknout hlubokým neuronovým sítím překonávajícím ty odborně navržené. Tato práce je zaměřena na využití **neuroevoluce**, tj. použití evolučních výpočetních technik pro návrh CNN.

## Cíl

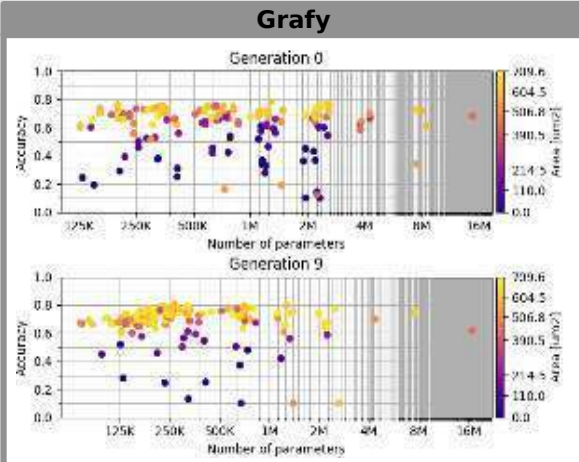
Vytvoření programu pro automatizovaný návrh CNN s možností optimalizace navrhované sítě z pohledu přesnosti, počtu parametrů, počtu vrstev sítě nebo celkového počtu násobení či množství spotřebované energie (v konvolučních vrstvách) při inferenci.

## Výsledky experimentu na datové sadě CIFAR-10

Výsledky experimentu, který byl zaměřen na evoluční návrh struktur CNN s ohledem na klasifikační přesnost a počet parametrů sítě.

Při evoluci je možné v konvolučních vrstvách použít 8-bitové aproximační násobičky z knihovny EvoApprox8b [1]. Tyto násobičky, na úkor přesnosti výsledku násobení, snižují prostorové a energetické nároky.

Během evoluce se tak mimo architektury sítě (propojení funkčních uzlů) měnila i použitá aproximační násobička. Každý jedinec (CNN model) tedy používal jinou aproximační násobičku ve všech svých konvolučních vrstvách.



V uvedených grafech jsou vyobrazeni jedinci prvních a posledních generací, z 10 běhů programu, s populací velikostí 8. Jednotlivé barvy odlišují jedince s různými aproximačními násobičkami, které zabírají různou plochu na čípu.

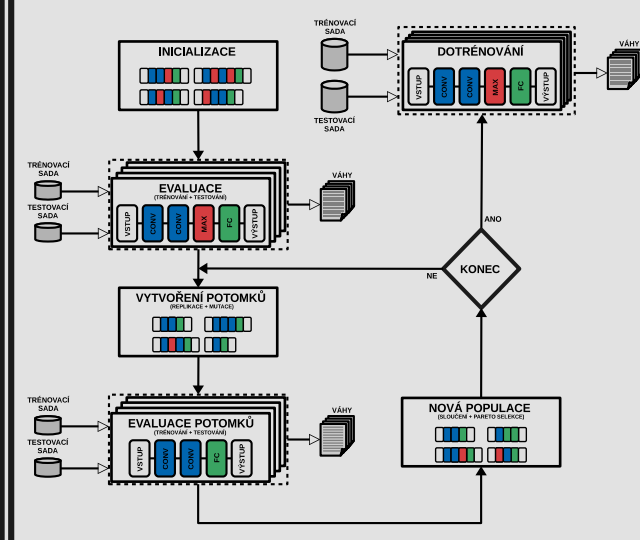
## Řešení

Pro návrh CNN byla zvolena technika kartézského genetického programování (CGP), ve spojení s vícekritériálním optimalizačním genetickým algoritmem NSGA-II (Non-dominated Sorting Genetic Algorithm II).

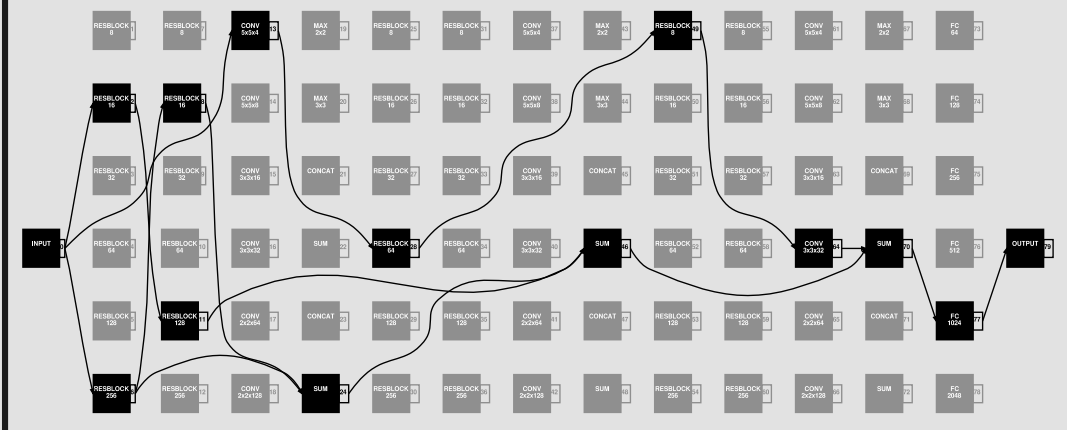
Program využívá:

- **CGP**: pro reprezentaci jedinců ve formě acyklického dopředného grafu (DAG), přičemž mutace jedince probíhá na úrovni hran grafů.
- **NSGA-II**: pro tvorbu nové populace, na základě relace Pareto dominance, umožňující při optimalizaci zohlednit více parametrů, jako je klasifikační přesnost, počet parametrů, spotřeba energie, ...

## Diagram algoritmu evoluce



## Ukázka nalezené CNN



## Reference

[1] V. Mrazek, R. Hrbacek, Z. Vasicek and L. Sekanina, EvoApprox8b: "Library of approximate adders and multipliers for circuit design and benchmarking of approximation methods" in DATE'17, pp. 258-261.