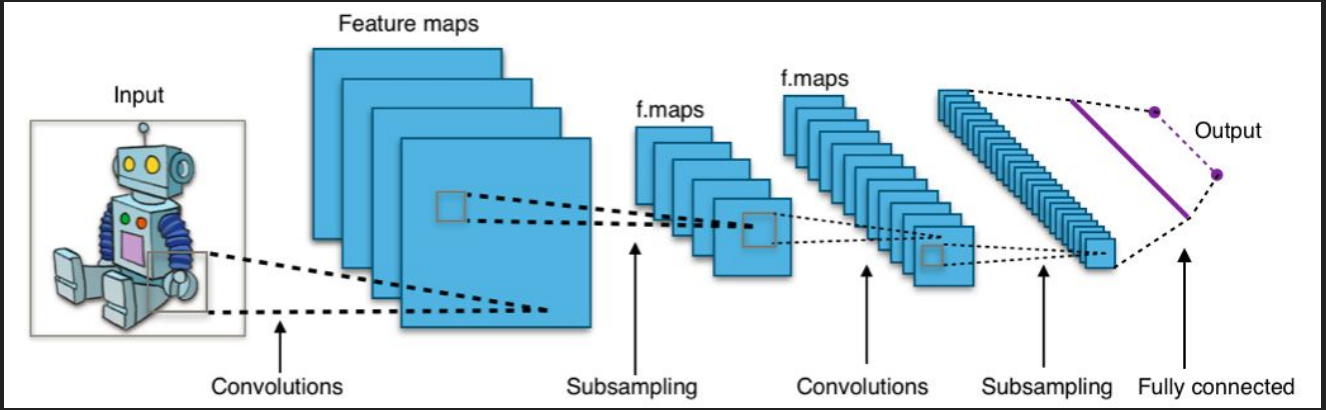
Towards a Time-efficient CNN Design with Cartesian Genetic Programming

Team7 Final Project Presentation

Introduction

Constructing CNN Architecture ...



... is too rough physical labor.

Introduction

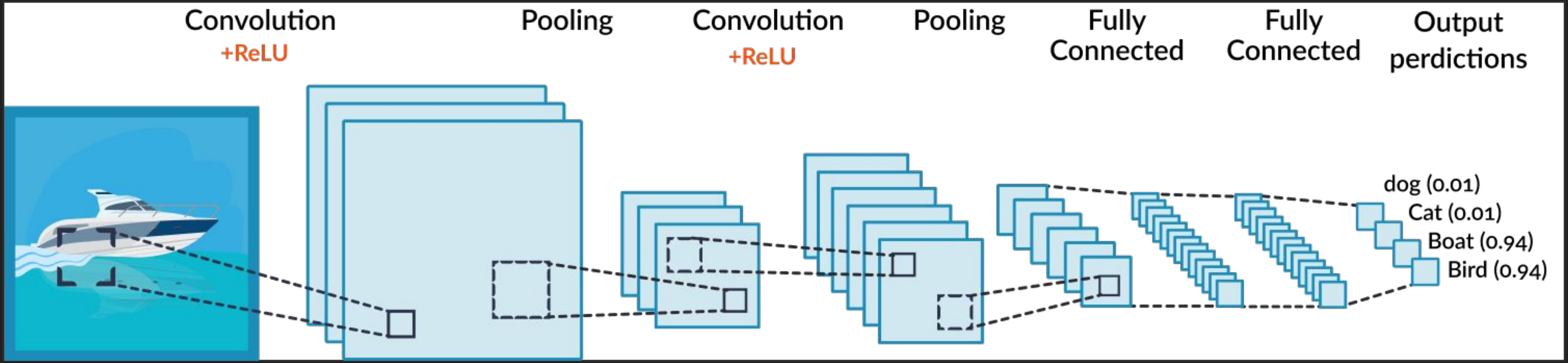
What if we can make these architecture automatically using genetic algorithm?

But finding the "best" solution takes too much time...

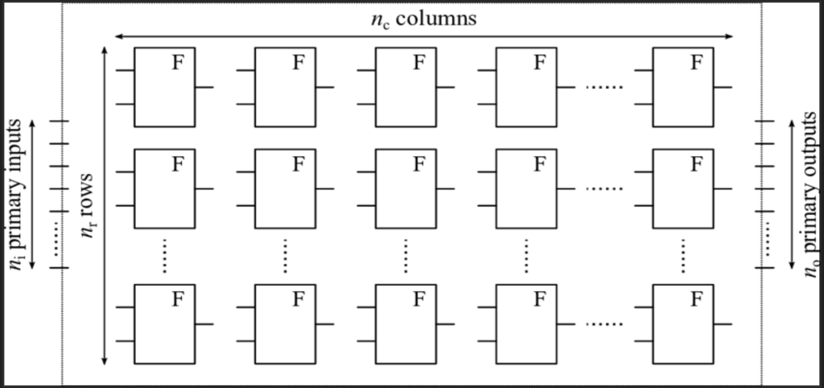
...

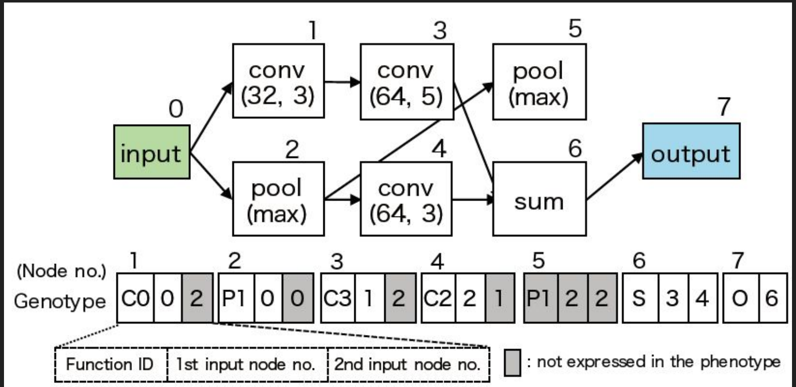
OK. Find the best species in limited time budget (# of evaluation)

Background :: Convolutional Neural Network



Background :: Cartesian Genetic Programming





Background :: Previous Research

A Genetic Programming Approach to Designing Convolutional Neural Network Architectures

Masanori Suganuma
Yokohama National University
79-7 Tokiwadai Hodogaya-ku
Yokohama, Japan 240-8501
suganuma-masanori-hf@ynu.jp

Shinichi Shirakawa
Yokohama National University
79-7 Tokiwadai Hodogaya-ku
Yokohama, Japan 240-8501
shirakawa-shinichi-bg@ynu.ac.jp

Tomoharu Nagao
Yokohama National University
79-7 Tokiwadai Hodogaya-ku
Yokohama, Japan 240-8501
nagao@ynu.ac.jp

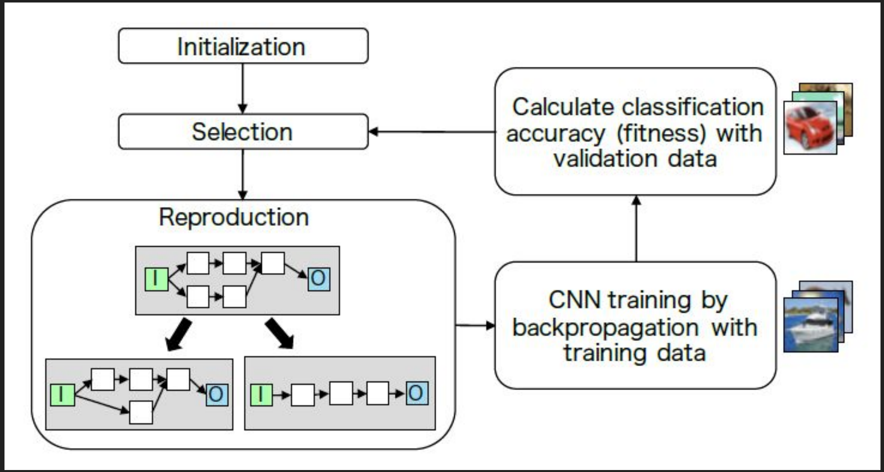
ABSTRACT

The convolutional neural network (CNN), which is one of the deep learning models, has seen much success in a variety of computer vision tasks. However, designing CNN architectures still requires expert knowledge and a lot of trial and error. In this paper, we attempt to automatically construct CNN architectures for an image classification task based on Cartesian genetic programming (CGP). In our method, we adopt highly functional modules, such as convolutional blocks and tensor concatenation, as the node functions in CGP. The CNN structure and connectivity represented by the CGP encoding method are optimized to maximize the validation accuracy. To evaluate the proposed method, we constructed a CNN architecture for the image classification task with the CIFAR-10 dataset. The experimental result shows that the proposed method can be used to automatically find the competitive CNN architecture compared with state-of-the-art models.

have seen huge success in image recognition tasks in the past few years and are applied to various computer vision applications [37, 38]. A commonly used CNN architecture consists mostly of several convolutions, pooling, and fully connected layers. Several recent studies focus on developing a novel CNN architecture that achieves higher classification accuracy, e.g., GoogleNet [33], ResNet [10], and DensNet [12]. Despite their success, designing CNN architectures is still a difficult task because many design parameters exist, such as the depth of a network, the type and parameters of each layer, and the connectivity of the layers. State-of-the-art CNN architectures have become deep and complex, which suggests that a significant number of design parameters should be tuned to realize the best performance for a specific dataset. Therefore, trial-and-error or expert knowledge is required when users construct suitable architectures for their target datasets. In light of this situation, automatic design methods for CNN architectures are highly beneficial.

SUGANUMA, Masanori; SHIRAKAWA, Shinichi; NAGAO, Tomoharu. A genetic programming approach to designing convolutional neural network architectures. In: Proceedings of the Genetic and Evolutionary Computation Conference. ACM, 2017. p. 497-504.

Background :: Previous Research (Cont.)



We use the modified $(1 + \lambda)$ evolutionary strategy (with $\lambda = 2$ in our experiments) in the above artifice. The procedure of our modified algorithm is as follows:

(1) Generate an initial individual at random as parent P, and train the CNN represented by P followed by assigning the validation accuracy as the fitness.
(2) Generate a set of λ offsprings C by applying the forced mutation to P.
(3) Train the λ CNNs represented by offsprings C in parallel, and assign the validation accuracies as the fitness.
(4) Apply the neutral mutation to parent P.
(5) Select an elite individual from the set of P and C, and then replace P with the elite individual.
(6) Return to step 2 until a stopping criterion is satisfied.

Limitations...?

SUGANUMA, Masanori; SHIRAKAWA, Shinichi; NAGAO, Tomoharu. A genetic programming approach to designing convolutional neural network architectures. In: Proceedings of the Genetic and Evolutionary Computation Conference. ACM, 2017. p. 497-504.

We focus on...

The purpose is "finding the best species in limited time budget (# of evaluation)."

Focus on <u>initialisation</u> & <u>search space</u>

Methods & Experimental Setup

- Initialisation based on Statistical Analysis
- Strong Neutral Mutation (SNM)
- Enlarging Parent Population Size

Code based on PyTorch with CUDA

Two RTX 2080 machines with Intel i7-6700K CPU and 15GB RAM

Methods :: Initialisation based on Statistical Analysis

Problem & Motivation

- Currently, CGP is initialised by the random initialisation.
- <u>Initialise with architectures sampled by CGP</u>
- <u>CGP is time-consuming. Limit the # of fitness evaluations for initialisation.</u>

Methods :: Initialisation based on Statistical Analysis

Approach

- Fix the number of total fitness evaluations (N), and calculate the expectation of maximum fitness value among M trials of each N/M generations.

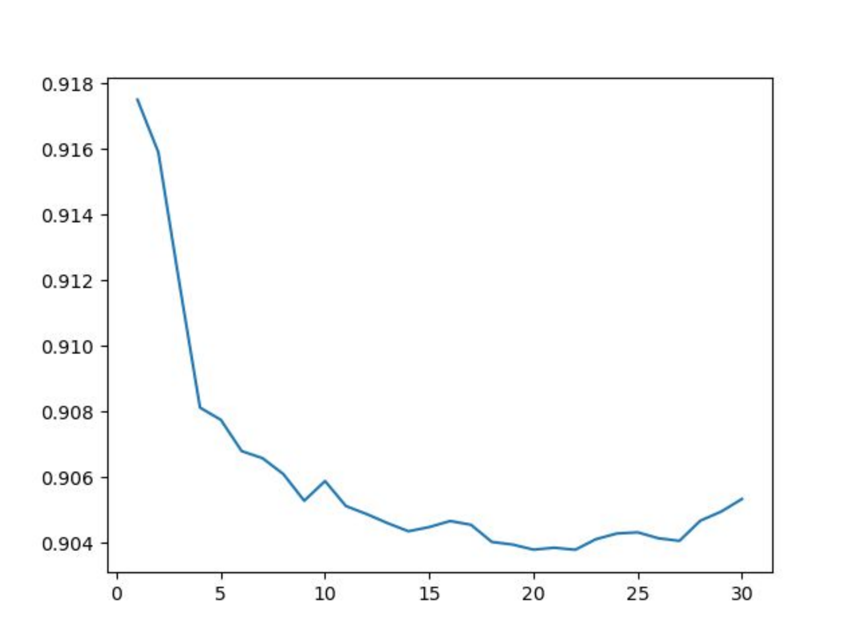
Example (N = 30)

- M = 5 → 5 trials of each 6 generations. $E[\max(X_{i,\ generation=6}),\ i=1..5]$
- M = 15 → 15 trials of each 2 generations. $E[\max(X_{i,\ generation=15}),\ i=1..15]$

Methods :: Initialisation based on Statistical Analysis

Result

- 30 trials of 1 generation shows the maximum expectation of fitness.
- Shows slow convergence of CGP-based CNN generation.



Methods :: Strong Neutral Mutation (SNM)

Problem & Motivation

- Original method performs neutral mutation on the parent for each generation.
- However, CGP shows very low rate of convergence.

Methods :: Strong Neutral Mutation (SNM)

Hypothesis

- Neutral mutation is not strong enough to make difference on offsprings.

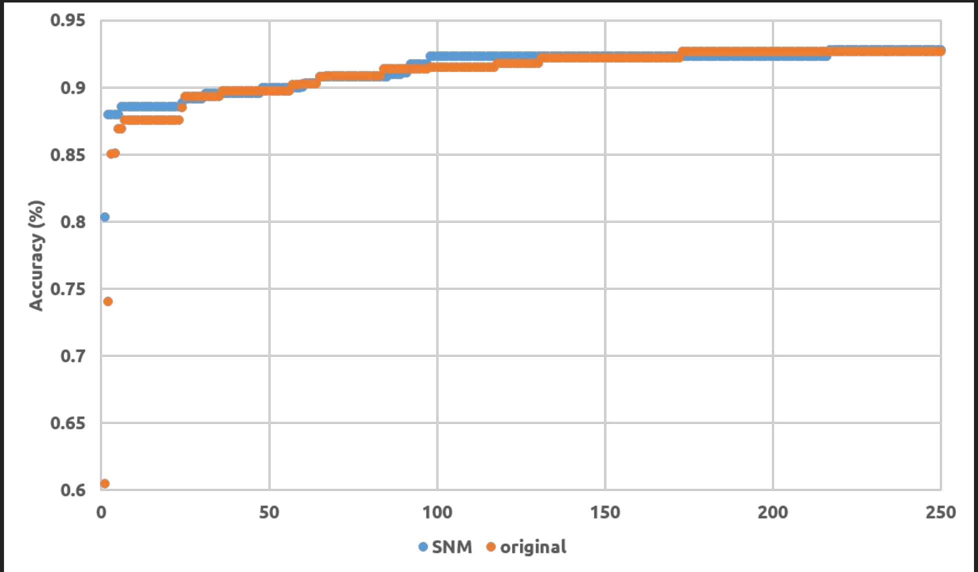
Approach

- Perform multiple neutral mutations simultaneously if the parent does not change for certain generations.

Methods :: Strong Neutral Mutation (SNM)

Result

- SNM failed to show better accuracy / fast convergence than the original one.



Methods :: Enlarging Parent Population Size

Problem & Motivation

- Original CGP-based approach only uses one parent and two offsprings.
- No reason to limit the parent population size to one.

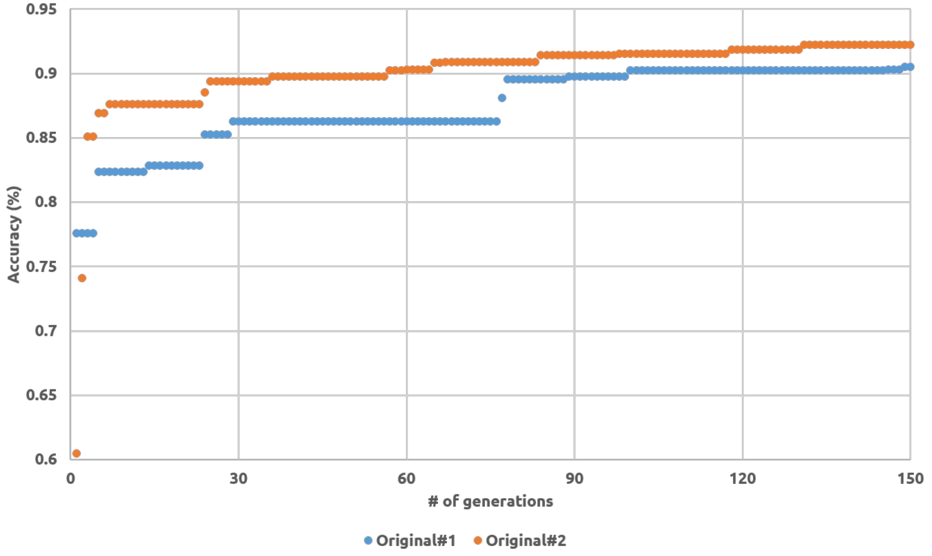
Hypothesis & Approach

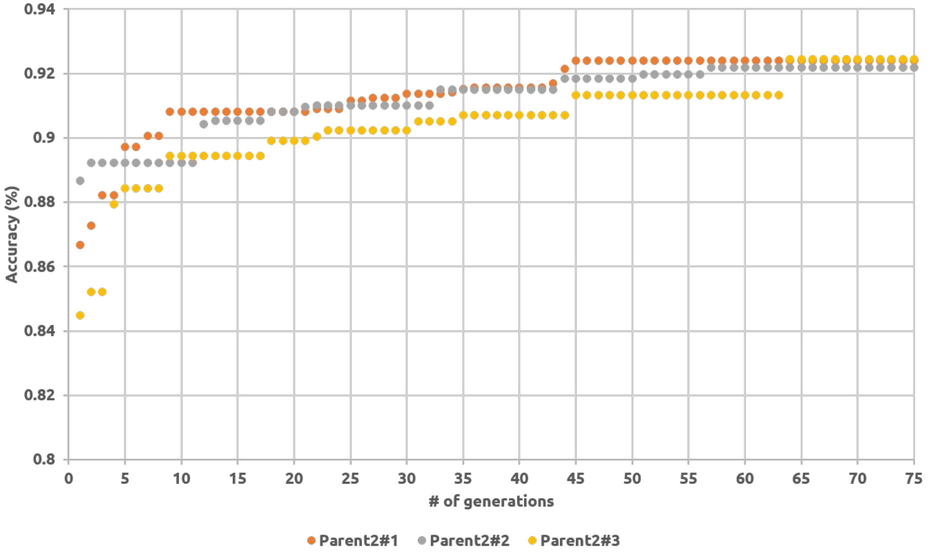
- <u>Larger parent population size will give an opportunity of the fast convergence.</u>

Methods :: Enlarging Parent Population Size

Result

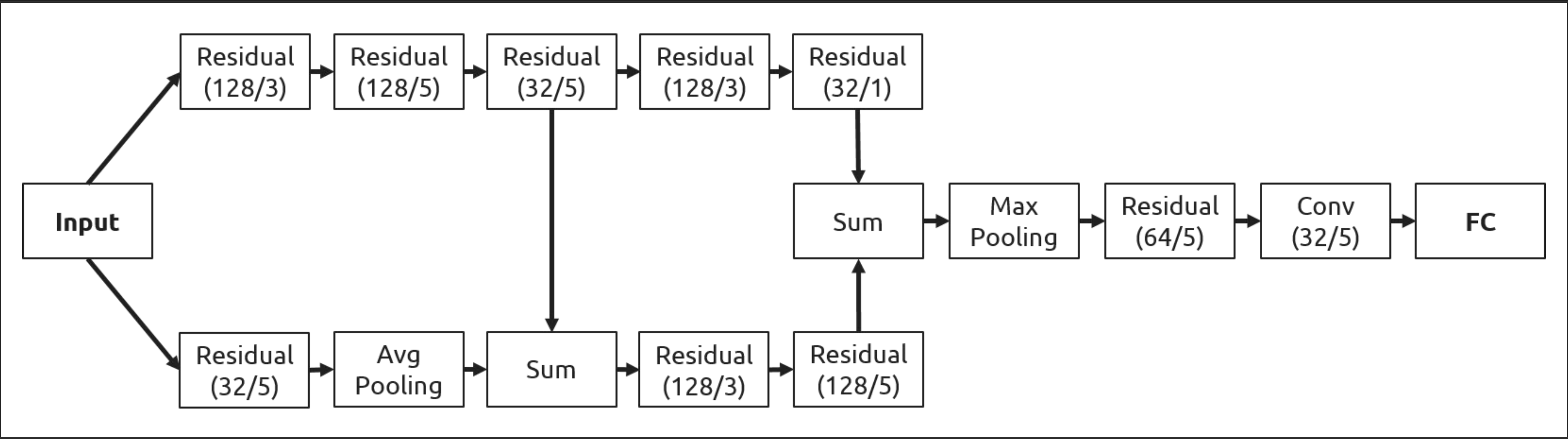
- Smooth & fast conversion (better fitness values at earlier steps).





Result

Maximum accuracy 92.44% resulted in our Enlarged Parent method.



Conclusion

We suggest three approaches to improve existing CGP-based CNN architecture generation:

- Initialisation based on Statistical Analysis
- Strong Neutral Mutation (SNM)
- Enlarging Parent Population Size

Limitations & Future Works

- SNM-based approach need to be improved to perform more mutations.
- Due to the heavy computation, we failed to repeat experiments more.
- Overall, CGP showed slow convergence; need fundamental improvements.