# Convolutional Neural Network Optimized by Genetic Algorithm for Detection of COVID-19 from Chest X-Ray Image

**Muhammad Ihza Mahendra, Isman Kurniawan**

Fakultas Informatika, Universitas Telkom, Bandung
ihzamuhammad@students.telkomuniversity.ac.id, ismankrn@telkomuniversity.ac.id

**Abstrak**

**Coronavirus Diseases 19 (COVID-19) kasus awalnya terjadi di kota Wuhan, China, pada penghujung tahun 2019 hingga kasusnya mencapai beberapa negara dengan kecepatan yang sangat cepat. Diagnosis dini COVID-19 diperlukan untuk mencegah penyebaran lebih lanjut, mengurangi jumlah kematian. Reverse Transcription Polymerase Chain Reaction (RT-PCR) adalah salah satu metode yang digunakan untuk mendeteksi COVID-19. Namun cara ini membutuhkan waktu yang lama sehingga diperlukan cara lain yang lebih cepat. Sebagai alternatif, foto rontgen dada dapat digunakan untuk mendeteksi gejala spesifik terkait COVID-19 dengan mengotomatiskan diagnosis. Convolutional Neural Network (CNN) banyak digunakan untuk deteksi penyakit berdasarkan citra medis berbantuan komputer. Pada penelitian ini diusulkan model klasifikasi untuk pendeteksian COVID-19 dengan menggunakan metode CNN, dimana bobot sinaptik dioptimasi oleh algoritma genetika. Hasil penelitian menunjukkan bahwa CNN yang dioptimasi oleh algoritma genetika dengan Adam optimizer memberikan performa terbaik untuk akurasi, F1-Score dan Area Under Curve (AUC) dengan nilai masing-masing 87,5%, 85,1% dan 86,8%.**

**Kata kunci: Covid-19, Rontgen Dada, Convolutional Neural Network, Algoritma Genetika.**

**Abstract**

**Coronavirus Diseases 19 (COVID-19) first appeared in the city of Wuhan, China, at the end of 2019 and spread to several countries at a very fast rate. An early diagnosis of COVID-19 is needed to prevent further spread, reduce the number of deaths. Reverse Transcription Polymerase Chain Reaction (RT-PCR) is one of the methods used to detect COVID-19. However, this method takes a long time so another faster method is required. As the alternative, Chest x-rays image can be used to detect specific symptoms related to COVID-19 by automating the diagnosis. Convolutional Neural Network (CNN) is widely used for disease detection based on computer-aided medical images. In this study, we proposed classification model for the detection of COVID-19 by using CNN method, in which the synaptic weight is optimized by genetic algorithm. The result show that, CNN optimized by genetic algorithm with Adam optimizer gives best performance for accuracy, F1-Score and Area Under Curve (AUC) with value 87.5%, 85.1% and 86.8% respectively.**

**Keywords: Covid-19, Chest X-ray image, Convolutional Neural Network, Genetic Algorithm.**

## A. Introduction

### Background

In 2020, there has been a pandemic of Coronavirus Disease 19 (COVID-19) caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) which first appeared in the city of Wuhan, China, at the end of 2019 [1]. Until June, 2021, this disease spread fast and it has caused more than 180 million people to be infected and more than three million people died worldwide.

The current detection method most commonly used is the Reverse Transcription Polymerase Chain Reaction (RT-PCR) to treat the nucleic acid form of SARS-CoV-2 [2]. However, this method took a long time, while the number of suspected host increased everyday. For another method to detect COVID-19, the radiological image obtained from the COVID-19 case contains credible information to serve as an answer to the patient's diagnosis.

Several studies using this visual diagnosis method have found differences and changes in the images of the lungs on chest X-rays and CT-scans (Computerized Tomography scans) before the onset of symptoms in COVID-19 patients. Therefore, an automatic diagnostic method is needed that can help in predicting COVID-19 from early X-rays and CT-Scan images [3].

It is a very common things in a several study for an automatic diagnostic for disease detection based on medical images and non medical images using Convolutional Neural Network (CNN) [4-7]. CNN has been successfully implemented in many classification tasks, among several studies that have been done by V. Sirish

Kaushik et al. and co-workers, 2020, with mechanism to detect pneumonia from X-ray images by using 3 layers of convolution with an accuracy value 92.31% [4]. There are also plant disease detection done by E. Harte et al. and co-workers, 2020, using ResNet34 and managed to achieve an accuracy value of 97.2% [5].

Another study for cancer disease detection by S. Dabeer et al. and co-workers, 2019, with 93.45% accuracy [6]. Malignancy Detection by S. P B et al. and co-worker, 2019, reach the highest accuracy value 99.97% for 2 class problems [7]. Also Mammogram-based image for breast cancer detection by Gardezi et al. and co-workers, 2019, with more then 90% accuracy for the performance metrics [8]. And the other like skin cancer detection using CNN done by Hasan et al. and co-workers, 2019, with 83% of precision [9].

In this study, we aim to develop classification model for prediction of COVID-19 based on chest X-ray image database as a credible dataset by using CNN optimized by genetic algorithm [10]. Further more, data augmentation techniques used in the pre-processing section to extract features so that it can improve model's prediction[11]. This methods of optimize synaptic weight with GA.

**Topics and Limits**

The topic of this research is the process of detecting the COVID-19 disease by using Deep Learning method using CNN (Convolutional Neural Network). Which will be optimized using genetic algorithm for synaptic weight from CNN. The input data to be used is a chest X-Ray image that is confirmed positive and normal. The image will classification was carried out to determine whether the chest X-Ray image was positive for COVID-19 or normal. In this study, the data used is Chest X-Ray data which consists of two classes, namely Chest X-Ray positive for COVID-19 and Chest X-Ray normal with a total dataset of 3616 for normal and 3636 for the COVID-19 positive class. The model can only carry out the classification process to detect disease COVID-19 or normal.

**Purpose**

The purpose of this research is to optimize the CNN synaptic weight using the algorithm genetics to obtain an optimal model so that it can carry out the process COVID-19 disease classification based on chest X-Ray images and to find out the best fitness value can produce a model with a maximum accuracy value.

**B. Literature Review**

**1. Genetic Algorithm**

Genetic algorithm is a type of algorithm that offers optimum solution inspired with natural selction. Genetic algorithms generally used to create optimum solution for optimization and search problems by relying on natural selection operators such as mutation, crossover and selection [12]. In the early 1970s, these algorithms became famous by the work of John Holland in his book which is Adaptation in Natural and Artificial Systems (1975) [13]. Example of Genetic Algorithm is used in data mining optimization with the idea of recombination where the probability given as follow:

$$F(x_i) = \frac{f(x_i)}{\sum_{i=1}^{N_{ind}} f(x_i)} \qquad (1)$$

Where $f$(xi) states the score of the evaluation for each solution xi and F(xi) state the rate of probability of that individual being selected. In the case of an optimization problem, these are the standard Genetic Algorithm life cycle procedures used in the optimization is by start and initial population begin by initialize the number of population and other validation parameter. It start each generation to be calculated through evaluation of its fitness function. Then, evaluate fitness function by calculate the fitness f(x) of each individual in the population at current generation by fitness function evaluation.

After that, parent selection is to a couple of parents from the the most latest population, the probability rate increase the fitness function. The individual competition will go through process where it will be replaced which leads to the same parents could be selected more than once. Crossover is performed after the best parent is selected based on the crossover rate, perform the process for the parents at a randomly chosen point to create a couple of new individual. If there is no crossover performed, the new couple individual will just be the same with its predecessor. Lastly is mutation takes place based on mutation rate, and the resulting strings in the new population will be placed then replace the latest generation with the new generation and back to evaluate fitness