

YSCMR 2021

Proceedings of the
**Young Scientists' Conference on
Multidisciplinary Research - 2021**

Virtual International Conference

21st October 2021

“Multidisciplinary Research for Tomorrow’s Challenges”



ISSN 2815-0260

Proceedings of the Young Scientists' Conference on Multidisciplinary Research - 2021

21st October 2021

The material in this publication has been supplied by the authors, and only minor copy editing has been done by YSCMR 2021 editorial committee. The views expressed in the abstracts in this publication remain the responsibility of the named authors and do not necessarily reflect those of the National Institute of Fundamental Studies (NIFS) or the NIFS-Young Scientists' Association.

This document is made freely available on the public internet, permitting any users to read, download, copy, distribute, print, search, or link to the full texts of research articles, crawl them for indexing, pass them as data to software, or use them for any other lawful purpose, without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. The copyright is shared by authors and YSCMR 2021 to control over the integrity of their work and the right to be properly acknowledged and cited.

Published by:

Young Scientists' Association (NIFS-YSA),
National Institute of Fundamental Studies,
Hanthana Road, Kandy,
Sri Lanka

www.nifs.ac.lk

Tel: +94 (0) 812 232 002

Email: ysa@nifs.ac.lk

Cover page design & Page setup

Buddhika Karunaratne, Mahesh Senarathna, Hiran Kankanamge

An accurate deep learning model for classification of cancer types

G.G.H.T. Deshappriya, C.T. Wannige*

University of Ruhuna, Matara, Sri Lanka

**wannige@dcs.ruh.ac.lk*

Background: Cancer is one of the most threatening diseases in the world. According to the molecular structure, cancers differ from one type to another. Gene expression data can be used for cancer type classification which is publicly available in TCGA (The Cancer Genome Atlas). However, gene expression data is high-dimensional. Recently machine learning and deep learning models use those data for cancer type classification. However, their accuracies were not enough. This research proposes how transfer learning can be used to perform the cancer classification process effectively.

Objectives: The objective of this research is to implement a deep learning model to perform the classification of cancer types and subtypes using gene expression data more accurately.

Methods: Gene expression data from TCGA was used for this research. Gene expression data related to Kidney Renal Cell Carcinoma (KIRC), Breast Invasive Carcinoma (BRCA), Lung Squamous Cell Carcinoma (LUSC), Lung Adenocarcinoma (LUAD), and Uterine Corpus Endometrial Carcinoma (UCEC) was used in this research based on the availability of the data. Before inputting it into a deep learning model, the dataset was pre-processed using an equation. Each sample consists of 20533 genes. Gene expression data related to each sample was then represented in a 3D space. Using available different Convolutional Neural Network architectures, cancer types were classified. Among the Resnet50, Resnet101, VGG16, VGG19 and Densenet121 CNN architectures Densenet121 architecture showed the best classification results.

Results: The proposed Deep learning model has obtained higher accuracies than the other models. Training Accuracy is 99.948% and validation accuracy is 98.926%.

Conclusion: Densenet121 can be used for cancer type classification using transfer learning. In future work, other CNN models can experiment with this problem.

Keywords: *Gene expression data, Cancer classification*