World Academy of Science, Engineering and Technology
International Journal of Biomedical and Biological Engineering
Vol:17, No:01, 2023

Recognition of Gene Names from Gene Pathway Figures Using Siamese Network

Authors : Muhammad Azam, Micheal Olaolu Arowolo, Fei He, Mihail Popescu, Dong Xu

Abstract : The number of biological papers is growing quickly, which means that the number of biological pathway figures in those papers is also increasing quickly. Each pathway figure shows extensive biological information, like the names of genes and how the genes are related. However, manually annotating pathway figures takes a lot of time and work. Even though using advanced image understanding models could speed up the process of curation, these models still need to be made more accurate. To improve gene name recognition from pathway figures, we applied a Siamese network to map image segments to a library of pictures containing known genes in a similar way to person recognition from photos in many photo applications. We used a triple loss function and a triplet spatial pyramid pooling network by combining the triplet convolution neural network and the spatial pyramid pooling (TSPP-Net). We compared VGG19 and VGG16 as the Siamese network model. VGG16 achieved better performance with an accuracy of 93%, which is much higher than OCR results.

Keywords : biological pathway, image understanding, gene name recognition, object detection, Siamese network, VGG
Conference Title : ICBCSB 2023 : International Conference on Bioinformatics and Computational Systems Biology
Conference Location : New York, United States
Conference Dates : January 30-31, 2023