Abstract

Classifying epitopes is essential since they can be applied in various fields, including therapeutics, diagnostics, and peptide-based vaccines. To determine the epitope or peptide against an antibody, epitope mapping with peptides is the most extensively used method. However, this method is more time-consuming and inefficient than using present methods. The ability to retrieve data on protein sequences through laboratory procedures has led to the development of computational models that predict epitope binding based on machine learning and deep learning. It has also evolved to become a crucial part of developing effective cancer immunotherapies. This paper proposes an architecture to generalize this case since various research strives to solve a low-performance classification problem. A proposed deep learning model is the fusion architecture which combines two architectures: Transformer architecture and Convolutional Neural Network (CNN), called MITNet and MITNet-Fusion. Combining these two architectures enriches feature space to correlate epitope labels with the binary classification method. The selected epitope-TCR interactions are GILG, GLCT, and NLVP, acquired from three databases: IEDB, VDJdb, and McPAS-TCR. The previous input data was extracted using Amino Acid Composition (AAC), Dipeptide Composition (DIP), Spectrum Descriptor, and the combination of all those features called AADIP composition to encode the input data to deep learning architecture. For ensuring consistency, five cross-validations were performed using the Area Under Curve (AUC) metric. Results showed that GILG, GLCT, and NLVP received scores of 0.81, 0.69, and 0.73, respectively. Those results were compared to prior architecture and outperformed four deep learning models.

Keywords: Deep Learning, Epitope Classification, Fusion Architecture, Transformer, CNN