

عنوان مقاله:

A hybrid multi-scale CNN-LSTM deep learning model for the identification of protein-coding regions in DNA sequences

محل انتشار:

فصلنامه مهندسی برق دانشگاه تبریز، دوره 52، شماره 2 (سال: 1401)

تعداد صفحات اصل مقاله: 10

نویسندگان:

عباس درویش - گروه بیوالکتریک، دانشکده مهندسی پزشکی، دانشگاه صنعتی سهند، تبریز، ایران

سینا شامخی - گروه بیوالکتریک، دانشکده مهندسی پزشکی، دانشگاه صنعتی سهند، تبریز، ایران

خلاصه مقاله:

Identification of the exact location of an exon in a DNA sequence is an important research area of bioinformatics. The main issues of the previous signal processing techniques are accuracy and robustness for the exact locating of exons. To address the mentioned issues, in this study, a method has been proposed based on deep learning. The proposed method includes a new preprocessing, a new mapping method, and a multi-scale modified and hybrid deep neural network. The proposed preprocessing method enriches the network to accept and encode genes at any length in a new mapping method. The proposed multi-scale deep neural network uses a combination of an embedding layer, a modified CNN, and an LSTM network. In this study, HMR195, BG570, and F56F11.4 datasets have been used to compare this work with previous studies. The accuracies of the proposed method have been 0.982, 0.966, and 0.965 on HMR195, BG570, and F56F11.4 databases, respectively. The results reveal the superiority and effectiveness of the proposed hybrid multi-scale CNN-LSTM network.

کلمات کلیدی:

Deep learning, DNA sequences, CNN, LSTM, Multi-scale, Protein coding region

لینک ثابت مقاله در پایگاه سیویلیکا:

<https://civilica.com/doc/1545030>

