

Bio-Medical Data Classification Approaches with Limited Annotation

原田, 翔太

<https://hdl.handle.net/2324/4784435>

出版情報 : Kyushu University, 2021, 博士 (工学) , 課程博士
バージョン :
権利関係 :

氏 名 : 原田 翔太

論 文 名 : Bio-Medical Data Classification Approaches with Limited Annotation
(アノテーションが限定された生物・医学データのための分類手法の提案)

区 分 : 甲

論 文 内 容 の 要 旨

This thesis aims to solve the problem of limited annotation in several bio-medical data analysis tasks. Specifically, data augmentation, group-based labeling utilizing constrained clustering, and semi-supervised learning are proposed as the approaches. First, as data augmentation, I proposed a time-series generation method based on Generative Adversarial Networks and applied it to a biosignal classification task. It generates various time-series from data with limited annotations and contributes to provide more training samples for a classifier. Second, I proposed a new constrained clustering method, where a user attaches annotations to several sample pairs. Annotations are two types: cannot-link and must-link. The pair with cannot-link should not belong to the same cluster, whereas the pair with must-link should belong. These annotations are useful especially for medical data, because medical experts can have a more expected clustering result by a small number of annotations. Moreover, those annotations are treated as soft-constraints and therefore medical experts can attach them without extreme carefulness. Finally, I proposed order-guided disentangled representation learning, which is semi-supervised learning for bio-medical data classification. This method performs disentangled representation learning with prior knowledge that is effective for learning bio-medical data classification tasks. This method could improve classification performance even with limited annotation by effectively utilizing the prior knowledge through disentangled representation learning.