

Д. М. Синюгина, Е. В. Комракова
(ГГТУ имени П. О. Сухого, Гомель)

ИСПОЛЬЗОВАНИЕ ГЕНЕРАТИВНО-СОСТЯЗАТЕЛЬНЫХ СЕТЕЙ ДЛЯ СИНТЕЗА И АУГМЕНТАЦИИ ДАННЫХ С ЦЕЛЬЮ УЛУЧШЕНИЯ СЕГМЕНТАЦИИ РЕДКИХ МИКРООРГАНИЗМОВ

Генеративно-состязательные сети (*GAN*) – это подход машинного обучения, который может использоваться для синтеза и аугментации данных с целью улучшения обучения моделей нейронных сетей. Особенно *GAN* полезны в случаях, когда имеется ограниченное количество данных.

Традиционные подходы к классификации и сегментации изображений требуют большого количества размеченных данных. Однако для редких микроорганизмов часто имеется ограниченное количество исходных изображений. Это приводит к переобучению моделей.

GAN позволяют генерировать новые изображения на основе набора реальных данных. Это достигается за счет одновременного обучения двух нейронных сетей – генератора и дискриминатора. Генератор создает новые изображения, а дискриминатор пытается отличить настоящие изображения от сгенерированных.

Сгенерированные изображения можно использовать для увеличения обучающей выборки. Это позволяет избежать переобучения и улучшить способность моделей обобщать на новых данных.

Помимо генерации новых изображений, *GAN* можно использовать для аугментации – внесения небольших искажений в исходные данные. Это увеличивает разнообразие данных для обучения моделей.

Было разработано программное обеспечение, использующее генеративно-состязательные сети для улучшения сегментации бактерий на микробиологических снимках. Использование *GAN* позволило добиться более высокой точности сегментации.

Генеративно-состязательные сети являются многообещающим инструментом для решения проблемы нехватки данных при сегментации. Дальнейшие исследования в этом направлении могут способствовать более широкому практическому применению *GAN*.