

УДК 546.06

**ТЕХТ-ТО-IMAGE ПРЕДСКАЗАТЕЛЬНАЯ СИСТЕМА МОРФОЛОГИИ
НАНОМАТЕРИАЛОВ, ОСНОВАННАЯ НА ПРОЦЕДУРАХ СИНТЕЗОВ
НАНОЧАСТИЦ, ПОЛУЧАЕМЫХ МЕТОДАМИ РАСТВОРНОЙ ХИМИИ, И МОДЕЛИ
ОБРАБОТКИ ЕСТЕСТВЕННОГО ЯЗЫКА**

**Дубровский И.С. (Университет ИТМО), Серов Н.С. (Университет ИТМО)
Научный руководитель – профессор, доктор химических наук, Виноградов В.В.
(Университет ИТМО)**

Введение. На сегодняшний день, наноматериалы используются в различных областях, таких как катализ, оптика, медицина, в частности при разработке фармацевтических препаратов, диагностических систем и систем адресной доставки лекарств. При этом часто на эффективность наноматериалов в данных областях влияет их морфология, а именно линейные размеры, форма, структура поверхности. Например, геометрия наноматериалов влияет на их эндоцитоз клетками иммунитета [1]. Ввиду этого часто возникает необходимость контролировать данные параметры в процессе синтеза наноматериалов. При этом синтез наноматериалов включает в себя наличие большого количества взаимозависимых параметров, таких как концентрации реагентов, температуры, скорости перемешивания и так далее. Это делает предсказание результата конкретного синтеза нетривиальной задачей, вследствие чего для создания наноматериала с желаемой морфологией необходимо проводить большое количество экспериментов, тратить время и другие ресурсы.

Глубокое обучение потенциально может позволить справиться с данной задачей, однако пока что разработаны системы лишь для качественного предсказания геометрических форм и предсказания размеров по одной из осей [2]; при этом, характер отображения данных принципиально не позволяет расширить подобные подходы на большее количество химических систем.

Основная часть. Разработанная нами система способна генерировать изображения наноматериалов, получаемые со сканирующего электронного микроскопа, основываясь на процедурах синтеза данных наноматериалов. Это достигается за счет использования предобученной модели обработки естественного языка BERT, которая переводит тексты синтеза в более сжатое векторное представление, а также за счет генеративной системы, которая работает по принципу вариационного автоэнкодера и позволяет создавать изображения наноматериалов из полученного векторного представления текстов синтезов. Наша система основана на собранном вручную датасете из 215 синтезов с изображениями со сканирующего электронного микроскопа и структурированной информации о синтезе. Для того, чтобы получить тексты синтезов из структурированной информации (например значения температур, концентраций и названий реагентов) нами были использованы специальные шаблоны, которые представляли собой тексты синтезов с пропусками для определенных значений. Подобная система позволит исследователям тратить значительно меньше времени на эксперименты благодаря возможности предсказания результата синтеза.

Выводы. Охватывая все детали морфологии наноматериала, разработанный подход также позволит значительно сократить затраты на их разработку. При этом данный подход является универсальным, так как любой наноматериал потенциально может быть отображен в виде процедуры синтеза. Таким образом, становится возможным обратный дизайн наноматериалов с заранее заданными морфологическими свойствами.

Список использованных источников:

1. Champion J. A., Mitragotri S. Role of target geometry in phagocytosis //Proceedings of the National Academy of Sciences. 2006 Mar 28;103(13):4930-4.

2. Khayati G. R. A predictive model on size of silver nanoparticles prepared by green synthesis method using hybrid artificial neural network-particle swarm optimization algorithm //Measurement. 2020 Feb 1;151:107199.