

В отличие от языковых моделей, подобных ChatGPT, хороших универсальных моделей для робототехники пока нет. Octo позволяет роботам выполнять широкий спектр задач, например, поднимать предметы, закрывать ящики или вытирать столы.

Модель основана на том же типе нейронных сетей, что и ChatGPT, — трансформаторах. Ее главное отличие — масштаб данных для обучения и гибкость. Ее обучали на самом большом наборе траекторий манипуляций роботов из открытого набора данных Open X-Embodiment.

Исследователи проверили Octo на девяти разных роботах, разработанных в их университетах. Модель успешно управляла роботами, позволяя им выполнять различные задачи, даже если во время обучения она не видела данных с сенсоров этих роботов.