

Głębokie modele generatywne dla uczenia reprezentacji danych trójwymiarowych

Dane trójwymiarowe, na przykład chmury punktów lub RGB-D stają się obecnie coraz bardziej popularne. Znajdują one zastosowania w nowoczesnych technologiach, gdzie wykorzystywana jest wizja przestrzenna, takich jak drony, czy autonomiczne pojazdy. Zastosowania te odnoszą się do sytuacji, gdzie decyzje podejmuje się w czasie rzeczywistym, co wymusza projektowania algorytmów przetwarzających te dane, mając na uwadze nie tylko ich skuteczność, ale również wydajność.

Jedną z najczęściej spotykanych form danych trójwymiarowych są chmury punktów. Uzyskuje się je za pomocą urządzeń typu LIDAR, które laserowo rejestruje otoczenie i mapuje je jako punkty w przestrzeni trójwymiarowej. Co istotne, taka reprezentacja danych, nie wymusza żadnej konkretnej struktury punktów (w przeciwieństwie do np. tradycyjnych obrazów, gdzie każdy piksel ma swoje stałe, niezmiennie miejsce). Taka dowolność umożliwia wierne odwzorowanie środowiska, w którym znajduje się LIDAR, jednak znacznie komplikuje procedury operujące na zebranych danych. Jako przykład można sobie wyobrazić sytuację, gdzie dana chmura punktów jest zapisana dwukrotnie, lecz w obu przypadkach punkty podane są w różnej kolejności. Ważne jest, by wykorzystywane podejście potrafiło określić, że jest to ten sam obiekt, co uchroni przed nadmierną złożonością i długim czasem pracy algorytmów. Inną trudnością związaną z danymi zapisanymi jako chmury punktów, jest rozmiar danych. Podczas jednego skanu, LIDAR potrafi stworzyć reprezentację otoczenia składającą się z od kilku tysięcy do kilku, a nawet kilkudziesięciu milionów punktów. Stanowi to kolejne wyzwanie dla projektowanych wydajnych sposobów przetwarzania danych 3D.

Jednym z problemów obszarze uczenia maszynowego jest zagadnienie generowania obiektów. Problem generowania obiektów polega na zbudowaniu modelu który będzie w stanie w budować wcześniej niezobserwowane obiekty zgodne z dystrybucją danych. W praktyce oznacza to, że model generatywny wyuczony na zdjęciach twarzy będzie generować zdjęcia twarzy ludzi nieistniejących. Istnieje szereg metod generowania obiektów które operują na obrazach. Interesującym rozwinięciem tych metod są rozwiązania stosowane dla obiektów 3D, które są przedmiotem tego projektu. Generowanie obiektów 3D może być przydatne w projektowaniu przestrzennym, w autonomicznych samochodach czy rozszerzaniu zbioru uczącego. Ponadto, modele generatywne mogą być wykorzystane do budowy zestawu cech (zwanego również *reprezentacją ukrytą*) o znacznie niższej wymiarowości, zachowującej jednak możliwie dużo informacji na temat oryginalnych danych.

Celem projektu jest opracowanie szeregu metod generatywnych dla reprezentacji trójwymiarowych. Od opracowanych metod wymaga się, by uczyły się one w efektywny sposób, generowały wysokiej jakości obiekty 3D i operowały na reprezentacjach ciągłych. Dodatkowo, opracowane metody zostaną rozszerzone o znajdowanie reprezentacji binarnych, tj. składających się tylko z wartości 0 i 1, co jeszcze bardziej usprawnia prędkość przetwarzania (na poziomie sprzętowym przetwarzanie liczb całkowitych jest zwykle o 1-2 rzędy wielkości szybsze niż liczb zmiennoprzecinkowych). Opracowane metody zostaną przeanalizowane w sposób ilościowy i jakościowy z wykorzystaniem znanych zbiorów dostępnych w internecie.