

## Głębokie warunkowe modele generatywne

**Głębokie modele generatywne**, takie jak GANy (generative adversarial networks) oraz VAE (variational autoencoders), osiągnęły w ostatnich latach ogromny sukces, szczególnie w takich dziedzinach jak wizja komputerowa, cheminformatyka czy modelowanie kształtów 3D. Opisując rozkład danych, możemy tworzyć nowe przykłady na podstawie istniejących danych. Jednakże skoro modele generatywne są zwykle trenowane w sposób nienadzorowany, trudno jest sterować właściwościami generowanych danych, zobacz Rysunek 1 przedstawiający nasze wstępne wyniki. W wielu sytuacjach proces generowania danych musi spełniać określone własności i ograniczenia:

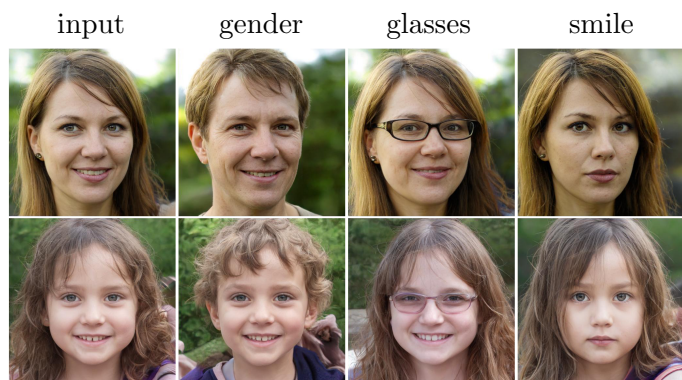
- Aby **wyjaśnić decyzje modeli sztucznej inteligencji**, chcemy odkryć atrybuty mające największy wpływ na predykcje. Można to zrealizować generując przykłady kontrfaktyczne (counterfactual examples), które przez lekkie modyfikacje interpretowanych atrybutów prowadzą do zmiany predykcji modelu.
- W **cheminformatyce**, generowanie cząstek chemicznych o określonych właściwościach jest pierwszym krokiem w konstrukcji nowych leków. Kandydat na lek powinien być aktywny na określone białka oraz nieaktywny na inne, aby nie powodował efektów ubocznych.
- W **wypełnianiu obrazów (image inpainting)**, szukamy realistycznego wypełnienia brakującego obszaru, który jest zgodny z pozostałą częścią obrazu. Jako że wiele rozwiązań spełnia wymagane ograniczenia, naturalnym jest wykorzystanie modeli generatywnych, które mogą zwracać wiele rezultatów dla danego warunku.

**Biorąc pod uwagę wagę powyższych problemów, naszym celem jest konstrukcja modeli generatywnych, które pozwalają na kontrolowanie procesu generowania.**

Budowa nowych warunkowych modeli generatywnych stanowi duże wyzwanie z koncepcyjnego oraz obliczeniowego punktu widzenia, ponieważ wymaga wyboru architektury oraz szeregu hiperparametrów, co jest czasochłonne oraz wymaga zaangażowania dużych zasobów obliczeniowych. Ponadto, zaadaptowanie wypracowanego rozwiązania do nowych danych stanowi dodatkowy koszt. Ten problem jest szczególnie ważny, gdyż nauczenie bardzo dużego modelu może być fizycznie niewykonalne bez dostępu do mocy obliczeniowej, którą dysponują globalne przedsiębiorstwa takie jak Google czy Facebook.

Aby zmierzyć się z różnorodnymi problemami warunkowymi, będziemy podążać nowym trendem stosowanym w głębokim uczeniu, który polega na **zaadaptowaniu pre-trenowanych modeli** do zadanego problemu, co znacząco zminimalizuje koszty koncepcyjne oraz obliczeniowe. Planujemy opracować **niewielkie przenośne sieci (plugin networks)**, które będą mogły pracować na reprezentacji stworzonej przez model pre-trenowany i modyfikować wybrane atrybuty przykładów. Naszym zdaniem przenośne sieci lepiej nadają się do aplikacji w różnorodnych zastosowaniach. Postawiona hipoteza będzie weryfikowana poprzez zastosowanie proponowanych modeli do generowania przykładów kontrfaktycznych, cząstek chemicznych, a także modelowania kształtów 3D.

Nasze rezultaty będą miały konsekwencje w wyjaśnianiu decyzji modeli uczenia maszynowego, co jest kluczowym wymaganiem stawianym obecnie przed systemami sztucznej inteligencji. Osiągnięte wyniki będą miały również wpływ na rozwiązanie realnych problemów, takich jak komputerowo-wspomagane tworzenie leków. Będziemy również pracować nad modelami głębokiego uczenia do rozwiązania problemów odwrotnych takich jak wypełnianie obrazów. Z uwagi na interdyscyplinarny charakter prowadzonych badań, nasz projekt przyniesie efekty poza samym uczeniem maszynowym. Aby lepiej spopularyzować nasze wyniki, planujemy publikowanie kodów implementujących nasze algorytmy na serwisach takich jak GitHub (<https://github.com/>).



Rysunek 1: Manipulacja atrybutami twarzy (płeć, okulary, uśmiech) za pomocą naszego wstępnego modelu.