Uczenie maszynowe w animacjach

Wytyczne do projektu

Szymon Datko, Henryk Maciejewski, Marek Woda, Tomasz Zamojski

semestr letni 2021/2022

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego
Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)”
Porozumienie nr POPC.03.02.00-00-0001/20-00

1 Autonomiczne sterowanie obiektem w grze komputerowej

Opracowanie prostej gry komputerowej, wykorzystującej algorytmy uczenia ze wzmocnieniem, do nauczenia agenta sterowania ruchem/zachowaniem obiektu/postaci w określonym środowisku.

1.1 Preferowana liczba wykonawców projektu

3–4 osoby

1.2 Wprowadzenie

Celem projektu jest poznanie metod uczenia ze wzmocnieniem (ang. *reinforcement learning*, RL), które można wykorzystać do nauczenia agenta sterującego poruszaniem/zachowaniem obiektu, na przykład w prostej grze komputerowej.

1.3 Główne wytyczne i wymagania

Zadanie polega na opracowaniu szkieletu prostej gry, w którymś z popularnych silników gier – na przykład Unity albo Godot; a następnie na wykorzystaniu modułu RL silnika – Unity ML Agents albo Godot RL Agents; do nauczenia agenta jak sterować postacią, czy innym określonym obiektem w grze.

Przykładem takiej gry może być wyścig samochodowy po torze – agent ma zapewnić przejazd pojazdu do mety, z uniknięciem kolizji, w możliwie krótkim czasie. Uczenie polega na wykorzystaniu interakcji ze środowiskiem, gdzie w stanie $s(t)$ agent wykonuje akcję $a(t)$, otrzymuje za to informację zwrotną $r(t)$ od środowiska (tak zwaną nagrodę lub karę) i przechodzi do kolejnego stanu $s(t+1)$.

Na postawie tych danym uczymy polityki agenta, czyli funkcji $Q(s, a)$, tak, żeby zmaksymalizować nagrodę za epizod w grze (czyli na przykład przejazd trasy).

Fundusze Europejskie Polska Cyfrowa
Rzeczpospolita Polska
Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego
Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

1.4 Zadania do wykonania

1. Opracować prostą grę w środowisku Unity albo Godot.
2. Oprogramować środowisko do uczenia agenta w grze:
   (a) określić wektor stanu w grze (np. położenie, prędkość, warto również wyposażyć obiekt w sensory, pozwalające na uwzględnienie w wektorze $s(t)$ odległości od przeszkód);
   (b) określić przestrzeń akcji $a(t)$;
   (c) określić funkcję $r(t)$.
3. Wytrenować agenta – z wykorzystanie wybranych algorytmów, np. *Proximal Policy Optimization*, *Behavioral Cloning*, *GAIL*.
   (a) Zbadać wpływ funkcji $r(t)$ na zachowanie agenta.
4. Przygotować raport dotyczący skuteczności agenta i czasu uczenia różnymi metodami.

1.5 Literatura

1. Repozytorium środowiska Godot RL Agents, `https://github.com/edbeeching/godot_rl_agents`
2. Unity ML-Agents Toolkit Documentation, `https://github.com/Unity-Technologies/ml-agents/tree/main/docs`
3. Arthur Juliani et al., *Unity: A General Platform for Intelligent Agents*, `https://arxiv.org/abs/1809.02627`

1.6 Sposób dokumentacji/potwierdzenia zrealizowania projektu

1. Sprawozdanie – dokument opisujący szczegóły z realizacji projektu.
2. Slajdy – prezentacja multimedialna wyników przed grupą.

Fundusze Europejskie Polska Cyfrowa
Rzeczpospolita Polska
Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego
Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

2 Nadpróbkowanie niskorozdzielczych obrazów cyfrowych

Wykorzystanie technik uczenia maszynowego, takich jak sztuczne sieci neuronowe, czy algorytmy generatywne, do poprawy jakości obrazów o niskiej rozdzielczości, po zwiększeniu ich rozdzielczości, w procesie tak zwanego skalowania w górę (ang. *upscaling*).

2.1 Preferowana liczba wykonawców projektu

3–4 osoby

2.2 Wprowadzenie

Klasyczne metody przetwarzania obrazu mające na celu poprawę jakości obrazu (ostrość/czytelność) często nie przynoszą oczekiwanych rezultatów. Powiększenie obrazu o niewielkiej rozdzielczości prowadzi do utraty detali i rozmycia całego obrazu.

Algorytmy uczenia maszynowego, jak głębokie sieci neuronowe, czy modele generatywne, na przykład *Generative Adversarial Networks* (GANs), sprawiają, że to zadanie staje się możliwe. Po wyuczeniu na zbiorze treningowym, umożliwiają one generowanie obrazów o wysokiej rozdzielczości na podstawie plików wejściowych o niskiej rozdzielczości. W praktycznych aplikacjach pozwala to zaoszczędzić moc obliczeniową, albo obniżyć wymagania sprzętowe gry komputerowej, ponieważ dla uzyskania rezultatów wysokiej jakości, wystarczy wytworzyć wtedy tylko obrazy przybliżone dla każdej sceny, a wyuczony model uzupełni je szczegółami.

2.3 Główne wytyczne i wymagania

Projekt ma obejmować zapoznanie się z wybraną biblioteką, na przykład Tensorflow/Keras, oraz środowiskiem obliczeniowym, na przykład Google Colab, aby poznać podstawy algorytmów generatywnych. Należy wykorzystać sieci współzawodniczące, lub inny algorytm, i wytrenować model do sprawnego zwiększania rozdzielczości zadanych obrazów wejściowych o niskiej rozdzielczości. Należy również porównać skuteczność różnych algorytmów w podanym zadaniu.

Fundusze Europejskie Polska Cyfrowa
Rzeczpospolita Polska
Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego
Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

2.4 Zadania do wykonania

1. Przygotować kolekcję obrazów oraz różne ich wersje w niższej rozdzielczości.
2. Opracować narzędzie, odtwarzające obrazy w oryginalnej rozdzielczości.
   (a) Odszukać gotowe modele uczenia maszynowego oraz wytrenować własne algorytmy.
   (b) Wykorzystać także metody prostej interpolacji (*nearnest*, *bilinear*, ...).
3. Porównać jakość obrazów po transformacji (*upscalingu*) względem plików oryginalnych.
   (a) Obliczyć na przykład średni błąd kwadratowy (MSE) dla różnych algorytmów.

2.5 Literatura

1. Deep Learning for Image Super-Resolution
   `https://www.analyticsvidhya.com/blog/2021/05/deep-learning-for-image-super-resolution/`
2. How to use the UpSampling2D and Conv2DTranspose Layers in Keras
   `https://machinelearningmastery.com/upsampling-and-transpose-convolution-layers-for-gener`
3. Jason Brownlee, *Generative adversarial networks with python: deep learning generative models for image synthesis and image translation.*
   `https://books.google.pl/books?id=YBimDwAAQBAJ`
4. Machine Learning Mastery - 9 Books on Generative Adversarial Networks (GANs)
   `https://machinelearningmastery.com/books-on-generative-adversarial-networks-gans/`

2.6 Sposób dokumentacji/potwierdzenia zrealizowania projektu

1. Sprawozdanie – dokument opisujący szczegóły z realizacji projektu.
2. Slajdy – prezentacja multimedialna wyników przed grupą.

Fundusze Europejskie Polska Cyfrowa
Rzeczpospolita Polska
Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego
Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

3 Segmentacja elementów ludzkiego ciała

Zastosowanie rozwiązań uczenia maszynowego, przy wykorzystaniu narzędzi takich jak TensorFlow i MediaPipe [1] w aplikacji multimedialnej, do realizacji bezmarkerowego (*markerless*) systemu *motion capture* – mapowania pozy przy użyciu komputera, smartfona (przy użyciu klasycznych kamer, zamontowanych w urządzeniach).

3.1 Preferowana liczba wykonawców projektu

3–4 osoby

3.2 Wprowadzenie

Kluczowym wyzwaniem przy zbudowaniu modelu części naszego ciała jest uzyskanie realistycznych danych 3D w środowisku naturalnym. W przeciwieństwie do 2D, które można uzyskać za pomocą ludzkich adnotacji, dokładna manualna adnotacja 3D staje się wyjątkowo trudnym zadaniem. Wymaga konfiguracji laboratoryjnej lub specjalistycznego sprzętu z czujnikami głębokości do skanowania 3D – co wprowadza dodatkowe wyzwania w celu zachowania dobrego poziomu różnorodności ludzi i środowiska w zbiorze danych. Kolejna alternatywa, którą wybiera wielu badaczy to zbudowanie całkowicie syntetycznego zbioru danych, co wprowadza kolejne wyzwanie adaptacji domeny do rzeczywistych obrazów.

Inne podejście opiera się na statystycznym modelu 3D ludzkiego ciała o nazwie GHUM, który jest zbudowany z dużego korpusu ludzkich kształtów i ruchów. Aby uzyskać rzetelną pozycję 3D ludzkiego ciała, model GHUM został dopasowywany do istniejącego zestawu danych 2D pozy i rozszerzony o współrzędne punktów kluczowych 3D w świecie rzeczywistym w przestrzeni metrycznej. Podczas procesu dopasowania zmienne kształtu i pozy GHUM zostały zoptymalizowane tak, aby zrekonstruowany model był zgodny z dowodem obrazu. Obejmuje to wyrównanie semantycznej segmentacji punktu kluczowego i sylwetki 2D, a także dopasowania kształtu i jego ułożenia [3].

Fundusze Europejskie Polska Cyfrowa
Rzeczpospolita Polska
Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego
Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

3.3 Główne wytyczne i wymagania

Projekt ma obejmować zapoznanie się z platformą Media Pipe i TensorFlow.js oraz środowiskiem Google Colab – z ich podstawową strukturą, założeniami i koncepcjami. Następnie instalację platform i środowiska, zaznajomienie się z modelami i użytkowaniem bibliotek (m.in. Body Segmentation API, Pose Detection API) do przetworzenia danych wejściowych. Na koniec: zaprezentowanie działającego prototypu systemu z implementacją rozwiązania do wybranego środowiska 3D.

3.4 Zadania do wykonania

1. Zapoznać się z Platformą Media Pipe i przykładami kodu w ramach Google Colab.
2. Zarejestrować obraz przy pomocy kamery do detekcji twarzy, pozy ciała, ułożenia dłoni, itp.
3. Wygenerować animowaną siatkę (*mesh*) w wybranym środowisku 3D.
4. Zaprezentować opracowane narzędzie w porównaniu z innymi rozwiązaniami.

3.5 Literatura

1. `https://google.github.io/mediapipe/`
2. `https://cgpress.org/archives/blendarmocap-markerless-mocap-for-blender.html`
3. `https://blog.tensorflow.org/2021/08/3d-pose-detection-with-mediapipe-blazepose-ghum-tfjs.html`
4. `https://www.marktechpost.com/2022/02/02/tensorflow-team-introduce-blazepose-ghum-posture`

3.6 Sposób dokumentacji/potwierdzenia zrealizowania projektu

1. Sprawozdanie – dokument opisujący szczegóły z realizacji projektu.
2. Slajdy – prezentacja multimedialna wyników przed grupą.

Fundusze Europejskie Polska Cyfrowa
Rzeczpospolita Polska
Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego
Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

4 Zautomatyzowanie tworzenia modeli do gier lub animacji

Eksploracja możliwości wykorzystania algorytmów generatywnych w zadaniu budowania modeli 2- i 3-wymiarowych, do wykorzystania na potrzeby gier lub animacji komputerowych.

4.1 Preferowana liczba wykonawców projektu

3–4 osoby

4.2 Wprowadzenie

Jednym z wyzwań podczas budowania światów wirtualnych, w szczególności tych na potrzeby animacji lub gier komputerowych o dużej skali, jest dostępność różnorodnych modeli. Ich mnogość, a przy tym jakość wykonania, to kluczowe czynniki, które należy wziąć pod uwagę przy definiowaniu przekonującej iluzji rzeczywistości.

Klasyczne podejścia związane z tak zwanym generowaniem proceduralnym terenu i obiektów sprawdza się, pod warunkiem przygotowania odpowiednio wielu elastycznych części modeli i zadbania o poprawne ich dopasowanie. Z drugiej strony, wykorzystanie algorytmów generatywnych może pozwolić na uzyskanie równie zadowalających efektów, przy niższym koszcie roboczym w stosunku do liczby różnorodnych modeli wynikowych.

4.3 Główne wytyczne i wymagania

Celem projektu jest zapoznanie się z działaniem algorytmów generatywnych, takich jak *Generative Adversarial Networks*, *Variational Autoencoders*, czy *Flow-based generative model*. Algorytmy te mogą posłużyć do zautomatyzowanego tworzenia obrazów obiektów, które mogłyby zostać zastosowane w charakterze 2-wymiarowej tekstury dla 3-wymiarowego obiektu, albo bezpośrednio do stworzenia gotowej 3-wymiarowej siatki. Należy zbadać możliwości wybranego algorytmu w zadaniu generacji obiektów do późniejszego wykorzystania w środowisku 3D.

Fundusze Europejskie Polska Cyfrowa
Rzeczpospolita Polska
Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego
Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

4.4 Zadania do wykonania

1. Zapoznać się z dostępnymi modelami algorytmów generatywnych.
2. Wykorzystać gotowy model generujący jakiś obiekt lub teksturę, albo wyuczyć własny.
3. Stworzyć siatkę modelu i zwizualizować go w wybranym środowisku 3D.
4. Zaprezentować działanie opracowanego narzędzia i efekty jego pracy.

4.5 Literatura

1. Analyzing and Improving the Image Quality of StyleGAN
`https://arxiv.org/abs/1912.04958`
2. StyleGAN2 — Official TensorFlow Implementation
`https://github.com/NVlabs/stylegan2`
3. Simple StyleGan2 for Pytorch
`https://github.com/lucidrains/stylegan2-pytorch`
4. 'Lightweight' GAN
`https://github.com/lucidrains/lightweight-gan`
5. This person does not exist / This cat does not exist
`https://thispersondoesnotexist.com` / `https://thiscatdoesnotexist.com`
6. Stylegan2-Ada-Colab-Starter
`https://colab.research.google.com/github/Hephyrius/Stylegan2-Ada-Google-Colab-Starter-No`
`blob/main/Stylegan2_Ada_Colab_Starter.ipynb`
7. Face detection with OpenCV and deep learning
`https://www.pyimagesearch.com/2018/02/26/face-detection-with-opencv-and-deep-learning/`

4.6 Sposób dokumentacji/potwierdzenia zrealizowania projektu

1. Sprawozdanie – dokument opisujący szczegóły z realizacji projektu.
2. Slajdy – prezentacja multimedialna wyników przed grupą.

Fundusze Europejskie Polska Cyfrowa
Rzeczpospolita Polska
Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego
Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)"
Porozumienie nr POPC.03.02.00-00-0001/20-00

5 Inne pomysły

Wyżej przedstawione propozycje projektów obejmują jedynie zgrubny zarys możliwych do realizacji tematów i nie wyczerpują wszystkich możliwości. Dopuszczalne są własne wariacje zaproponowanych tematów, w szczególności w zakresie zaproponowanych bibliotek i technologii w nich stosowanych. Możliwe są także wszystkie inne projekty, uzgodnione z prowadzącymi, o ile ich tematyka będzie związana z zagadnieniem zastosowań uczenia maszynowego w animacjach.

Kilka innych przykładów, dla inspiracji:

1. Selektywne zamazywanie postaci/obiektów na obrazach.
2. Optymalizacja obliczeń związanych z dynamiką płynów.

Fundusze Europejskie Polska Cyfrowa
Rzeczpospolita Polska
Unia Europejska Europejski Fundusz Rozwoju Regionalnego

Projekt współfinansowany ze środków Unii Europejskiej w ramach Europejskiego Funduszu Rozwoju Regionalnego
Program Operacyjny Polska Cyfrowa na lata 2014-2020,
Oś Priorytetowa nr 3 "Cyfrowe kompetencje społeczeństwa" Działanie nr 3.2 "Innowacyjne rozwiązania na rzecz aktywizacji cyfrowej"
Tytuł projektu: „Akademia Innowacyjnych Zastosowań Technologii Cyfrowych (AI Tech)”
Porozumienie nr POPC.03.02.00-00-0001/20-00