

Sterowanie Ruchem w Grach

Damian Rogalski

Maj 2024

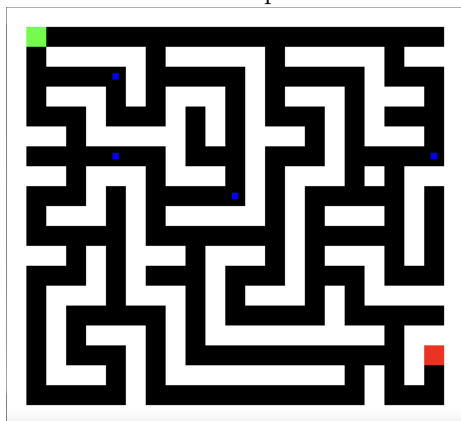
1 Wstęp

Celem mojego projektu jest sterowanie postacią w grach przy użyciu gestów dłoni. Projekt podzielony jest na trzy części, z których każda ma swoją własną demonstracyjną grę, aby zwizualizować zaimplementowane rozwiązanie dla konkretnego problemu.

2 Gry

2.1 Maze

Maze to prosta gra logiczna, w której gracz za pomocą gestów dłoni (lewo, prawo, góra, dół) wskazuje kierunek, w którym ma poruszać się zielony kwadrat. Celem gry jest dotarcie do czerwonego kwadratu, który pełni rolę mety. Gracz ma do dyspozycji trzy mapy. Podczas rozgrywki gracz może zbierać opcjonalne niebieskie punkty. Gra ilustruje detekcję położenia dłoni w przestrzeni przy użyciu biblioteki MediaPipe.



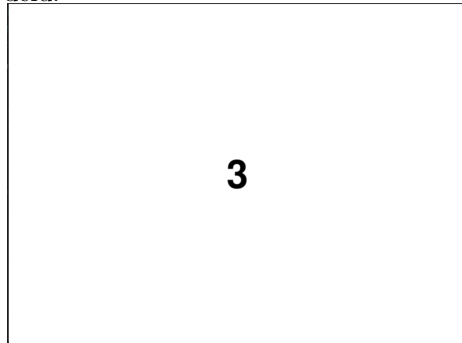
2.2 Cursor Game

Cursor Game to prosta gra, w której gracz może kontrolować kursor na ekranie za pomocą wskazującego palca, zadaniem jest najechanie na losowo generowane czerwone kwadraty. Za każdy trafiony obiekt gracz zdobywa punkt. Gracz ma ograniczony czas na trafienie w każdy obiekt. Co dziesięć punktów, kwadraty zmniejszają swoją wielkość, aby zwiększyć trudność gry. Gra demonstruje detekcję palca wskazującego i obliczanie jego pozycji w przestrzeni. Dodatkowo program oblicza odległość pomiędzy palcem wskazującym a kciukiem, ponieważ złączenie tych palców resetuje grę.



2.3 Numbers Game

Gra z Liczbami wyświetla na ekranie liczbe od 1 do 4, a zadaniem gracza jest wskazanie losowo wygenerowanej liczby palcami. Po poprawnym wskazaniu, gracz otrzymuje informacje "Poprawnie!" i po dwóch sekundach generowana jest kolejna liczba. Gra ma na celu pokazanie działania predykcji mojego wytrenowanego modelu.



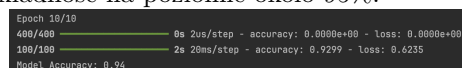
3 Model Klasyfikujący

3.1 Dane

Projekt zawiera własny model klasyfikacyjny. Do treningu i testowania zebrano 1000 obrazów przedstawiających wskazującą reke z liczbami od 1 do 4 (po 250 na klasę). Do przygotowania obrazów stworzono program, który wykrywa reke przy użyciu biblioteki cvzone, przycina obraz do określonych wymiarów i zapisuje go.

3.2 Trenowanie

Model trenowano na danych podzielonych na 80% zbioru treningowego i 20% zbioru testowego. Już po 10 epokach model uczy się bardzo szybko, osiągając dokładność na poziomie około 95%.



3.3 Testowanie i Optymalizacja

Testowałem różne funkcje aktywacyjne i ogólne osiągi najlepiej sprawdzały się w przypadku relu. Funkcja tanh drastycznie obniżała dokładność i skupiała wyniki tylko na jednej klasie. Zmniejszanie batch size wydłużało czas trenowania, lecz nie wywoływało to większych zmian. Całość trenowania trwa około 2 minut.

4 Testowanie Gier

4.1 Maze

Podczas testowania gry Labirynt szybko pojawiły się ograniczenia związane z precyzją ruchu. Aby maksymalizować komfort gry, zaleca się zachowanie odpowiedniego odległości (najlepiej około 1 metra) od kamery, aby ułatwić detekcję położenia dłoni. Przy odpowiednich warunkach oświetleniowych gra działa bez większych problemów.

4.2 Cursor Game

Ograniczenia biblioteki MediaPipe i sprzętu są najbardziej widoczne w tej grze. Ruch kursora na ekranie jest skokowy, co może być nieco niewygodne. Jednakże rozgrywka nadal dostarcza dużo frajdy i zabawy przy dobrym oświetleniu.

4.3 Numbers Game

Ta gra przyciągnęła największą uwagę, ponieważ wykorzystuje mój własny model klasyfikacyjny. Podczas testów okazało się, że program często błędnie klasyfikuje "1" jako "2". Dodatkowo, ze względu na ograniczoną liczbę danych, program ma trudności z detekcją gestów lewej ręki i wykazuje mieszane wyniki detekcji gestów u innych osób. Oświetlenie również odgrywa istotną rolę w predykcji gestów. Niemniej jednak, podczas moich własnych testów program działa poprawnie w większości przypadków.

4.4 Wnioski z eksperymentów

Testowanie funkcjonalności przynosiło sporo frajdy, jednakże również niejednokrotnie problemy natury technicznej bądź samych narzędzi wymuszały na mnie zmiany w kodzie, żeby przy kolejnym uruchomieniu programu bez poprawek (bądź delikatnej zmianie innych czynników jak oświetlenie) działać zupełnie poprawnie. Napisanie programów nie stanowiło dużego problemu wykorzystując dokumentację bibliotek jak i załączone poradniki, natomiast największą trudność stanowi zmienianie strategii bądź koncepcji rozwiązania konkretnych problemów (na przykład zmiana gestu na inny, bardziej odznaczający się od reszty) z uwagi na różne niezależne od programisty czynniki.

5 Podsumowanie

Program ciekawie demonstruje działanie różnych narzędzi do detekcji dłoni. Wyniki czasami są początkowo skromne, ale szybko stają się satysfakcjonujące. Wiele problemów wynika z ograniczeń sprzętowych, czasowych lub samych narzędzi. Poważny problem stanowi również zbyt niska augmentacja zdjęć treningowych. Niemniej jednak, osoby testujące były zadowolone i dobrze się bawiły podczas rozwoju programu.