

Prepoznavanje pesme na osnovu isečka u igri SongPop2

Nikola Papić SW4/2017, Dušan Stojančević SW11/2017, Vera Kovačević SW19/2017

Univerzitet u Novom Sadu, Fakultet tehničkih nauka, Novi Sad

Uvod

Cilj aplikacije je automatizovano prepoznavanje pesme u video igri *SongPop2*. Igra pušta isečak pesme 10 sekundi i za to vreme je potrebno odabrati jedan od 4 ponuđena odgovora koji predstavljaju ili numeru ili izvođača. Aplikacija se sastoji iz tri dela:

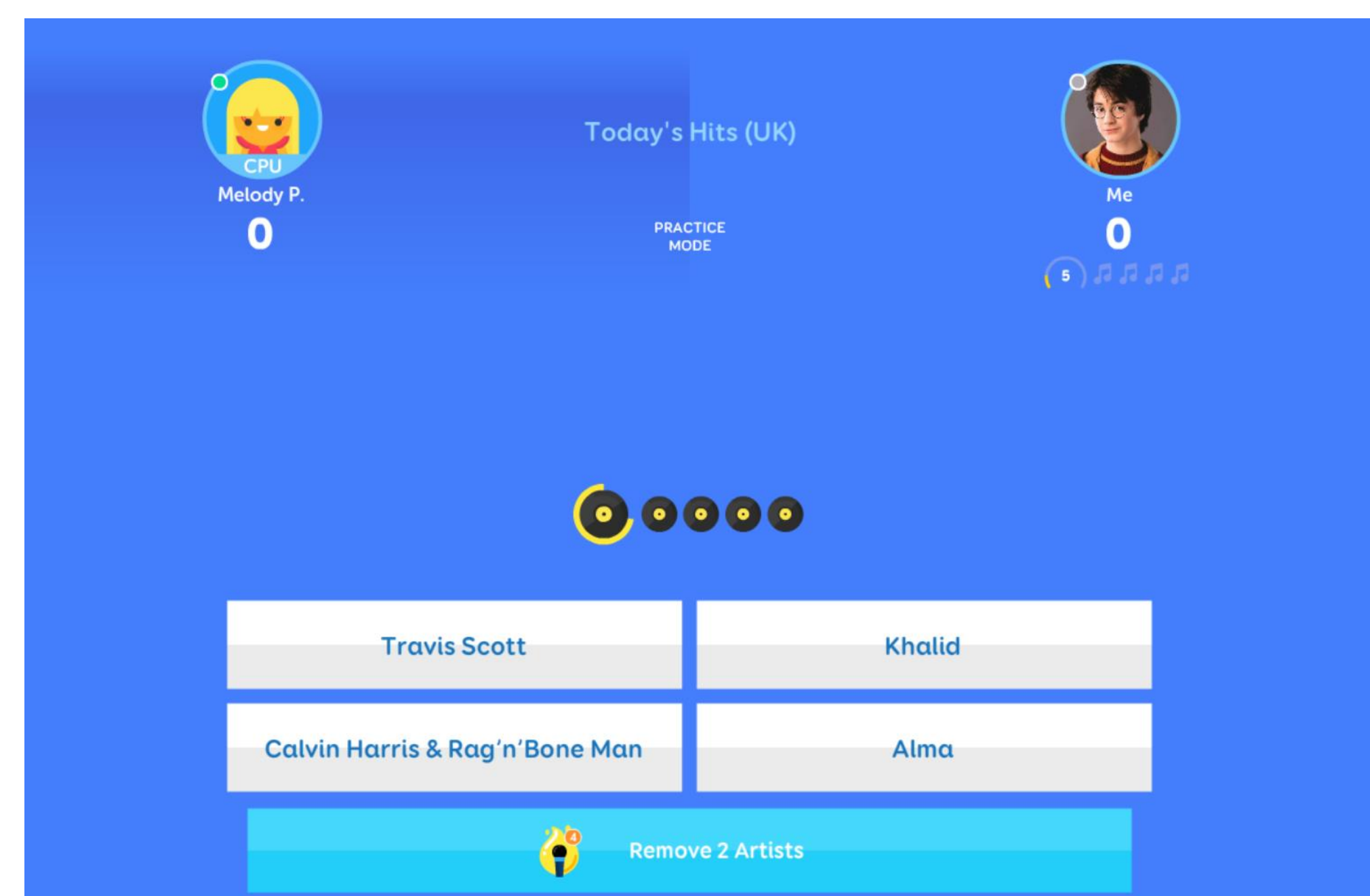
- Automatsko upravljanje aplikacijom
- Prepoznavanje pesme
- Prepoznavanje ponuđenih odgovora

Skup podataka

Kao skup podataka korištena je jedna od plejlista dostupnih u igri (*Today's Hits UK*). Generisani su isecci za većinu pesama (118). Za prepoznavanje ponuđenih odgovora i automatsko upravljanje aplikacijom korištene su slike ekrana igre.

Automatsko upravljanje aplikacijom

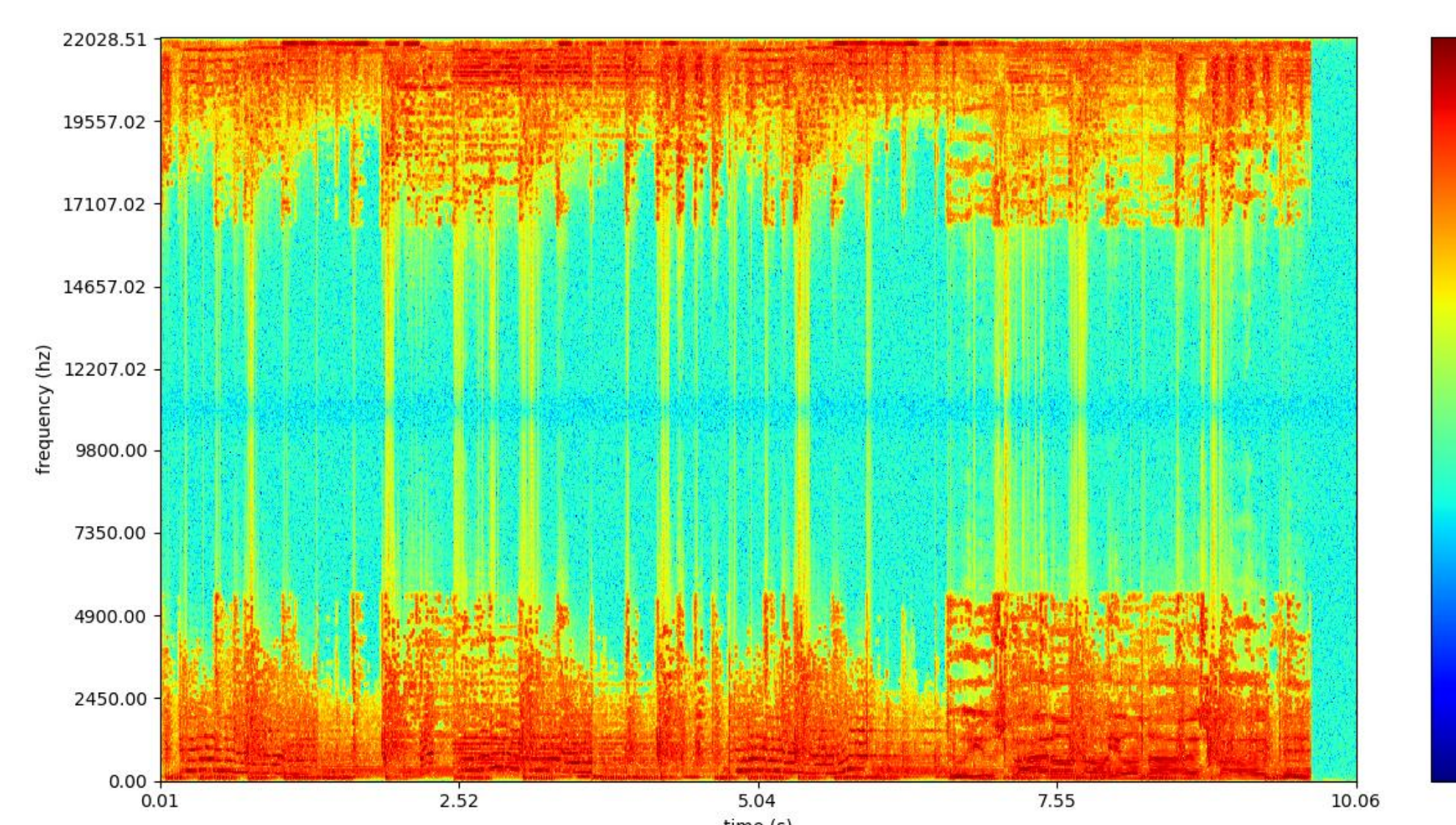
Na ekranu igre prepoznaju se akcije na osnovu odgovarajućih sličica (npr. dugme za novu igru). Njihova lokacija na ekranu se detektuje metodom *"matchTemplate"* biblioteke *"opencv"* koja koristi konvoluciju, a samo upravljanje odrađeno je pomoću biblioteke *"pyautogui"*. Međutim da bi detektovali kada je igrice pustila pesmu korištena je konvoluciona mreža. Mreža bi posmatrala grafičke elemente koji ukazuju na početak pesme i njen kraj.



Primer jedne partije u igri

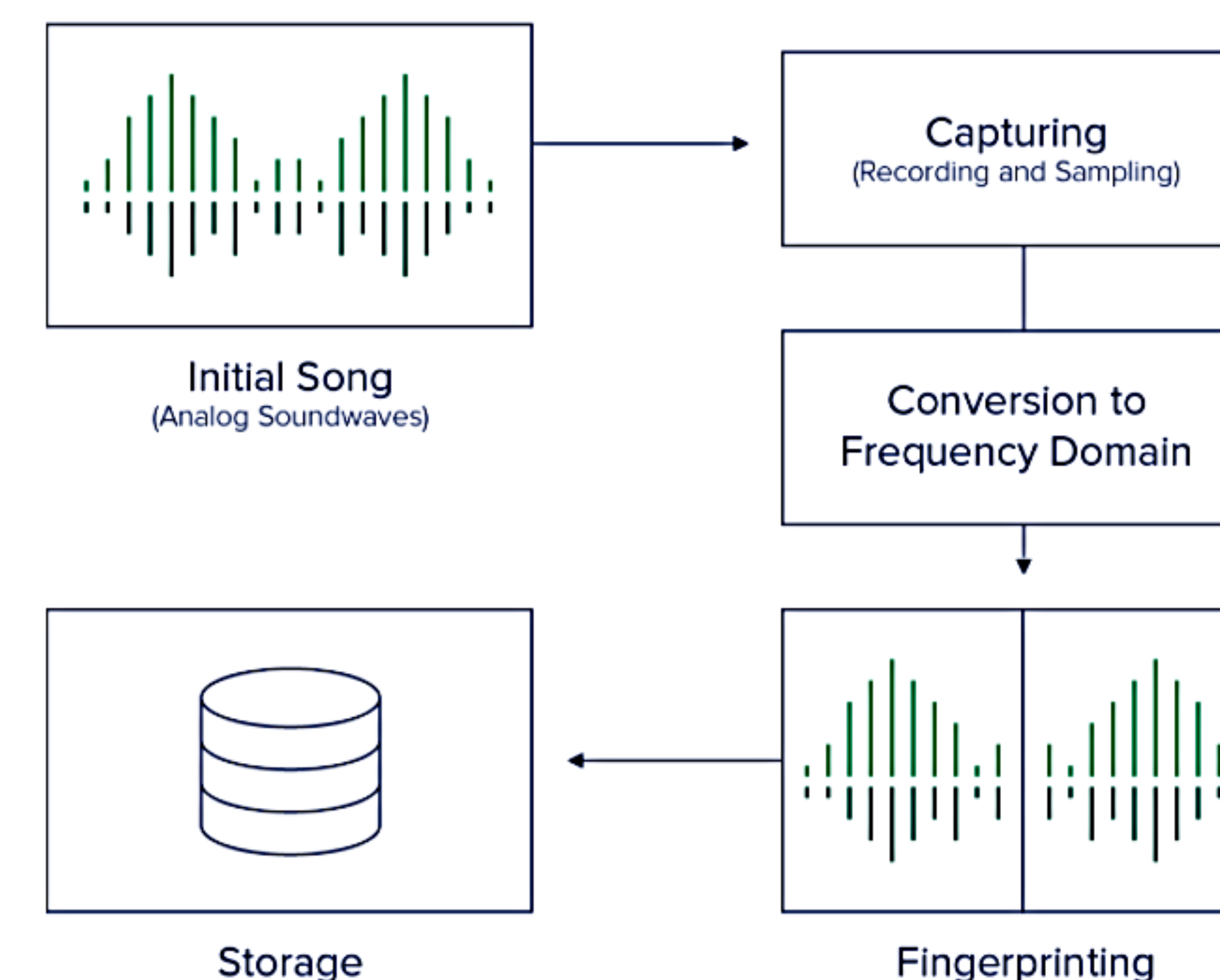
Prepoznavanje pesme

Za prepoznavanje pesme korišten je *"An Industrial-Strength Audio Search Algorithm"*, koji je osnova za aplikaciju Shazam. Algoritam pravi *"otisak"* (*fingerprint*) audio fajla na osnovu spektrograma gde y-osa predstavlja frekvenciju, x-osa predstavlja vreme a boja predstavlja amplitudu (metoda *specgram* biblioteke *matplotlib.mlab*). Izdvajaju se najveće amplitude (*peaks*) za svaki frekvencijski interval. Na osnovu njih se formiraju heševi koji se čuvaju u bazi (*sqlite3*), zajedno sa identifikatorom pesme (naziv pesme i ime izvođača).



Primer spektrograma

Kada program detektuje početak isečka pesme u igri, taj isečak se snima i na osnovu njega se pravi novi fingerprint. Dobijeni fingerprint se poredi sa podacima u bazi, i kao rezultat se uzima pesma čiji fingerprint iz baze se najviše podudara sa novim fingerprintom.

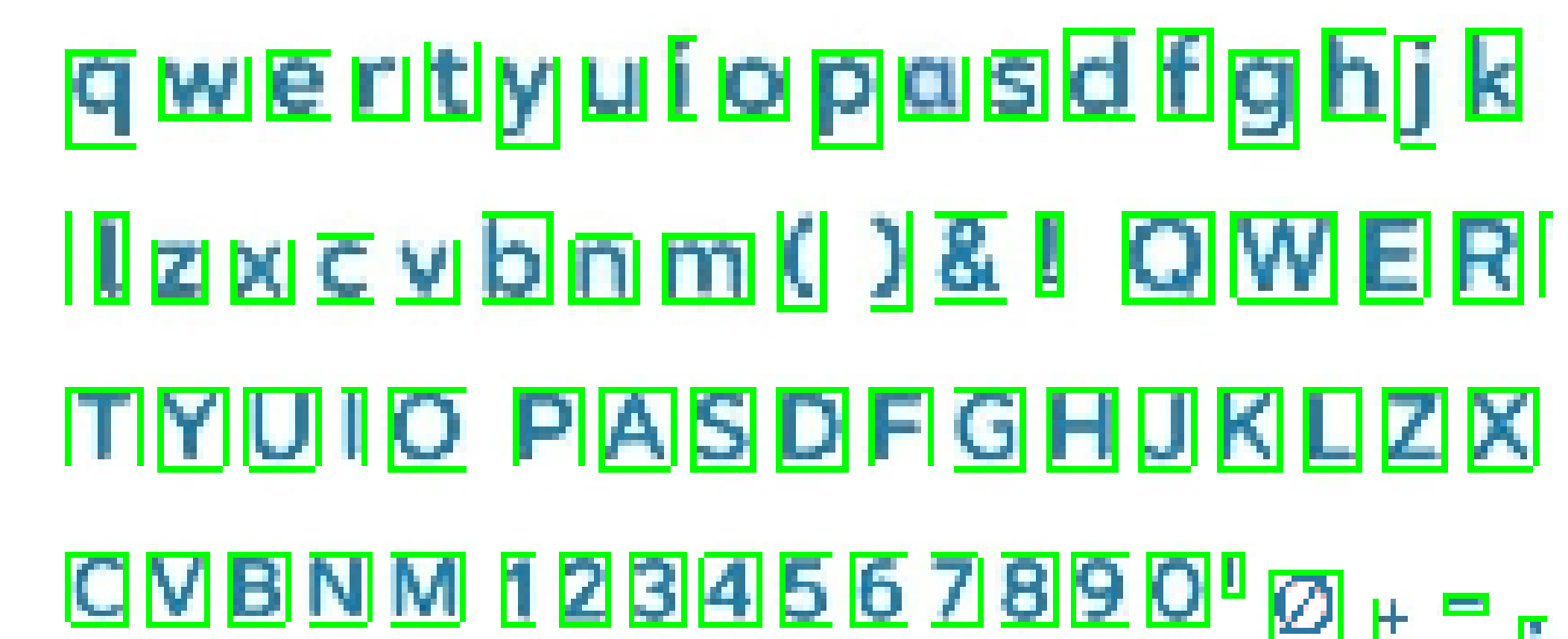


Postupak popunjavanja baze

Prepoznavanje ponuđenih odgovora

Nakon što algoritam prepozna pesmu iz baze, potrebno je odrediti koje dugme na ekranu predstavlja tu pesmu. Za čitanje teksta sa ekrana korištena je metoda OCR (*Optical Character Recognition*). Neuronska mreža je trenirana na osnovu znakova koji se pojavljuju u igri (slova, brojevi, neki specijalni znakovi). Prilikom prepoznavanja, za svaki od 4 ponuđena odgovora izdvajaju se konture i šalju se neuronskoj mreži koja prepoznaje karaktere. Od 4 dobijena rezultata bira se tekst koji je najbliži nazivu ili imenu

izvođača prepoznate pesme, program odabira taj odgovor i nastavlja dalje sa igrom.



Karakteristi za treniranje neuronske mreže

Rezultati

Na osnovu 20 odigranih partija sa po 5 pogađanja, program je prosečno prepoznao 4 od 5 pesama, pa možemo zaključiti da je preciznost (*accuracy*) 80%.

Zaključak

Program radi prilično dobro za većinu slučajeva, odnosno za slučajeve kada tražena pesma postoji u bazi. Mogao bi se unaprediti proširivanjem baze novim pesmama koje se pojavljuju u igri (npr. ukoliko ne prepozna pesmu, da se nova pesma doda u bazu). Takođe, tačnost bi se mogla povećati dužim snimanjem, ali bi se program tada igrao sporije i dobijao bi manje poena u igri.