

AUTOMATSKO PREPOZNAVANJE ŽANRA PESME POMOĆU KONVOLUCIONIH NEURONSKIH MREŽA

Motivacija

- Ogromna količina muzike
- Težak zadatak za ljude
- Sistemi za preporuku muzike
- Velike kompanije



Skup podataka - GTZAN

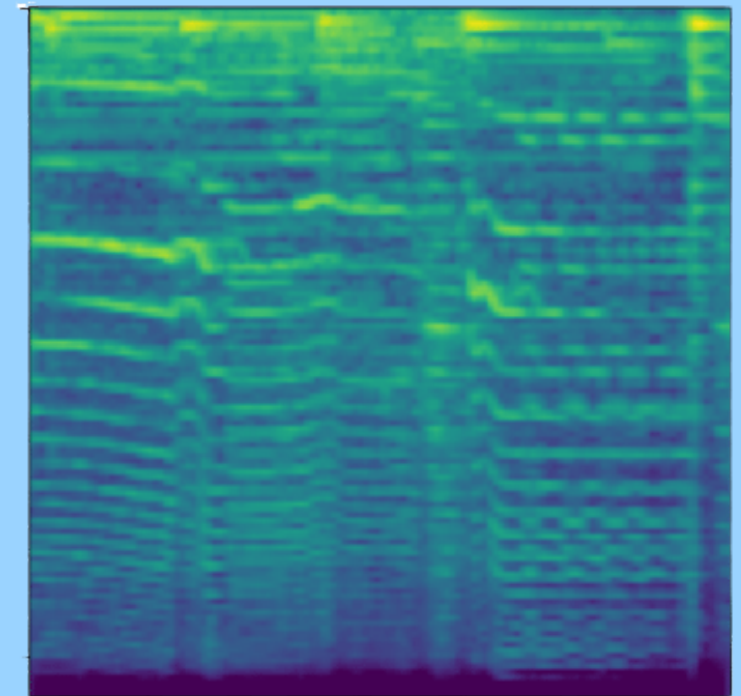
Najčešće korišćen skup podataka u ovoj oblasti je GTZAN skup podataka



- Bluz
- Klasična muzika
- Kantri
- Disko
- Pop
- Hiphop
- Metal
- Rege
- Rok
- Džez

Spektrogram

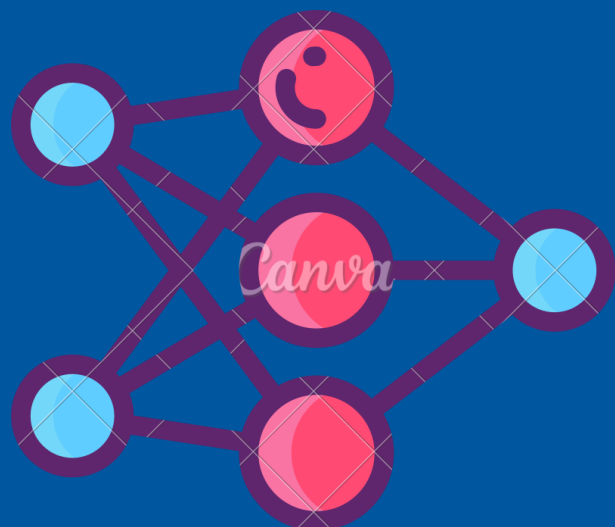
Spektrogram je vizuelna reprezentacija zvuka. On je, takođe, očekivani ulaz u konvolucionu neuronsku mrežu



CNN Model

Konvlucione neuronske mreže se sastoje od niza slojeva. Arhitektura modela je sledeća:

- Konvolucioni sloj
- Pooling sloj
- Soft max aktivaciona funkcija
- Potpuno povezan sloj



Treniranje

Ulazni skup podataka se deli tako što se 80% koristi za treniranje neuronske mreže a preostalih 20% za validaciju. Potrebno je 20 epoha da se stigne do maksimalne testne i trenažne preciznosti.



Rezultati

Rezultati se ogledaju u vidu preciznosti. Preciznost modela predstavlja odnos broja tačnih klasifikacija i ukupnog broja klasifikacija. Model je ostvario preciznost od 80%.

