

بسم الله الرحمن الرحيم

شرکت مهندسی نرم افزاری هلو

گزارش شناسایی پخش موسیقی در یک فایل صوتی

کاری از امیرعلی نسیمی

شبکه عصبی SampleCNN

مدل SampleCNN یک شبکه عصبی کانولوشنی (Convolutional Neural Network یا CNN) است که برای طبقه‌بندی فایل‌های صوتی طراحی شده است. این مدل از چندین لایه کانولوشنی تشکیل شده است که در ادامه توضیح داده می‌شوند:

جریان کار پردازش صوت

- main():
 - تعیین مسیر فایل صوتی `test.wav`.
 - ایجاد نمونه‌ای از کلاس `Music_Detection`.
 - اندازه‌گیری زمان شروع و پایان تحلیل فایل صوتی.
 - چاپ زمان اجرا و درصد استفاده از حافظه.

Music_Detection.py

1. load_model():
 - بارگذاری پارامترهای مدل از فایل `model-gztan-speech-music-20000.pth`.
 - تنظیم مدل به حالت ارزیابی (evaluation).
2. analyze():
 - تبدیل فایل صوتی به فرمت ۱۶ کیلوهرتز و مونو با استفاده از ffmpeg.
 - بارگذاری فایل تبدیل شده به صورت بلوک‌های زمانی با استفاده از librosa.
 - پردازش هر بلوک با مدل `SampleCNN` و چاپ نتیجه (موسیقی یا گفتار) به همراه احتمال.

پردازش صوت

۱. تبدیل فایل صوتی:

- فایل ورودی با استفاده از ffmpeg به فرمت ۱۶kHz و مونو تبدیل می شود.

۲. بارگذاری بلوک های صوتی:

- فایل تبدیل شده به بلوک های زمانی با طول ۵۹۰۴۹ نمونه و گام ۱۶۰۰۰ نمونه تقسیم می شود.
- هر بلوک به عنوان یک ورودی برای مدل استفاده می شود.

۳. پردازش بلوک های صوتی:

- هر بلوک به مدل `SampleCNN` وارد می شود.
- مدل احتمال اینکه بلوک حاوی موسیقی یا گفتار باشد را محاسبه می کند.
- نتایج همراه با زمان هر بلوک چاپ می شود.

خلاصه جریان کار

۱. اسکریپت `run.py` اجرا می شود.
۲. فایل صوتی `test.wav` بارگذاری و مسیر آن تعیین می شود.
۳. نمونه ای از کلاس `Music_Detection` ایجاد می شود.
۴. فایل صوتی به فرمت مناسب تبدیل می شود.
۵. فایل تبدیل شده به بلوک های زمانی تقسیم می شود.
۶. هر بلوک توسط مدل `SampleCNN` پردازش می شود.
۷. نتایج پردازش (موسیقی یا گفتار) همراه با احتمال مربوطه چاپ می شود.

۸. زمان اجرای تحلیل و درصد استفاده از حافظه چاپ می شود.

نتیجه گیری

این سیستم یک مدل یادگیری عمیق برای تشخیص موسیقی و گفتار از فایل های صوتی است که از چندین لایه کانولوشنی برای استخراج ویژگی ها و یک لایه کاملاً متصل برای طبقه بندی استفاده می کند. جریان کار شامل تبدیل فایل صوتی به فرمت مناسب، تقسیم آن به بلوک های زمانی و پردازش هر بلوک با مدل `SampleCNN` است.