

Kök Nota Analizi II

1. Alperen Ovalı
231307086
Kocaeli Üniversitesi
Kocaeli, Türkiye
alperen.oval@gmail.com

2. Tolga Boz
211307036
Kocaeli Üniversitesi
Kocaeli, Türkiye
tolgaboz350@gmail.com

Abstract—Kök nota analizi üzerine yapılan bu çalışmada, ses verisi sınıflandırma modelleri karşılaştırılmıştır. MiniWav2Vec, HuBERT, Wav2Vec2, YAMNet ve AudioSpectrogramTransformer (Lite) modelleri kullanılarak ses verileri üzerinde doğruluk, duyarlılık, özgüllük, precision, recall, F-Score, AUC değerleri hesaplanmış, her modelin karmaşıklık matrisi ve ROC eğrileri oluşturulmuştur. Ayrıca, her modelin eğitim ve test veri kümeleri için epoch sayısı ve kayıp fonksiyonu (loss) grafikleri çıkarılmış, eğitim ve çıkarım süreleri değerlendirilmiştir.

Index Terms—Kök Nota, Ses Analizi, MiniWav2Vec, HuBERT, Wav2Vec2, YAMNet, AudioSpectrogramTransformer, ROC Eğrisi, Eğitim Zamanı, Çıkarım Zamanı, AUC

I. TANITIM

- Problem Tanımlaması: Kök nota analizi, müzik teorisinde bir şarkının temel notasının tespiti ile ilgilidir. Bu problem, ses verisi üzerinde yapılan sınıflandırma görevlerinden biridir.
- Çalışma Amacı: Bu çalışmanın amacı, ses verisi sınıflandırma modellerini karşılaştırarak kök nota tespitinde en iyi performans gösteren modelin belirlenmesidir.
- Literatür Taraması: Ses analizi ve kök nota tespiti üzerine yapılmış önceki çalışmaların kısa bir özeti. Farklı ses işleme yöntemleri ve modellerin kullanımı.

II. MATERYALLER VE YÖNTEMLER

- Veri Seti: Projeye ait verilerin nasıl düzenlendiği, her klasörde bulunan ses dosyaları ve bunların etiketlenmesi. (Örneğin: "A", "B", "F", "G" gibi notalar)
- Model Seçimi: Kullandığınız modellerin kısa açıklamaları:
 - MiniWav2Vec2: Özellik çıkarımı için kullanılan bir model.
 - HuBERT: Özellikle ses temsilleri için kullanılan derin öğrenme tabanlı bir model.
 - Wav2Vec2: Küçük model boyutu ile ses sınıflandırması yapan bir model.
 - YAMNet: Yüksek düzeyde ses tanıma modelidir.
 - AudioSpectrogramTransformer (Lite): Ses verilerinin dönüştürülmesini sağlayan transformer tabanlı model.

- Veri Ön İşleme: Verilerin normalizasyonu, ses dosyalarının spectrogram'a dönüştürülmesi gibi işlemler.
- Model Eğitimi: Her modelin eğitim süreci, kullanılan optimizasyon teknikleri ve parametreler (örneğin epoch sayısı, batch boyutu, learning rate).
- Değerlendirme Metrikleri: Model performansını ölçmek için kullanılan metrikler (Accuracy, Precision, Recall, Sensitivity, Specificity, F-Score, AUC, vb.).
- Çalışma Ortamı ve Donanım: Google Colab üzerinde yapılan çalışmalar ve kullanılan donanım (örneğin GPU veya TPU kullanımı).

III. SONUÇLAR

- Performans Metrikleri: Her bir modelin sonuçları, her metrik için ayrı bir tablo veya grafik.
 - Accuracy, Precision, Recall, Sensitivity, Specificity, F-Score ve AUC değerleri her model için gösterilecek.
- Karmaşıklık Matrisi (Confusion Matrix): Her model için karmaşıklık matrislerinin görsel gösterimi.
- ROC Eğrileri: Her modelin ROC eğrisinin çizimi, AUC değerleri ile birlikte.
- Epoch ve Loss Grafikleri: Her modelin eğitim ve test verileri için epoch sayısı ve kayıp fonksiyonu (loss) grafikleri. Tüm modellerin grafikleri aynı grafik üzerinde karşılaştırılabilir.
- Eğitim ve Çıkarım Zamanı: Her modelin eğitim ve çıkarım sürelerinin karşılaştırılması.

A. Tartışmalar

- Model Performansının Karşılaştırılması: Her modelin başarısını metrikler açısından kıyaslama. Hangi modelin daha iyi sonuç verdiği ve nedenleri.
- Model Karmaşıklığı ve Zaman Performansı: Model karmaşıklığı ve çıkarım sürelerinin tartışılması.
- Çalışmanın Sınırlamaları: Veri seti, model yapılandırılmaları, ve potansiyel hata kaynakları üzerine bir değerlendirme.

B. Sonuçlar

- Sonuçlar: En iyi performansı gösteren modelin belirlenmesi ve elde edilen sonuçların genel bir özeti.

- Gelecek Çalışmalar: Bu araştırmanın sonuçlarına dayanarak yapılabilecek iyileştirmeler veya gelecekteki araştırma alanları.

REFERENCES

- [1] YAMNet Model Linki : <https://tfhub.dev/google/yamnet/1>
- [2] HuBERT Model Linki : <https://huggingface.co/facebook/hubert-large-ls960-ft>
- [3] Wav2Vec' Model Linki : <https://huggingface.co/facebook/wav2vec2-base-960h>
- [4] MiniWav2Vec2 Model Linki : <https://huggingface.co/facebook/wav2vec2-base-960h>
- [5] AudioSpectrogramTransformer(Lite) : <https://huggingface.co/MIT/ast-finetuned-audioset-10-10-0.4593>
- [6] Google Drive Linki : <https://drive.google.com/drive/folders/1Akw0NPcKhm8qKdkgR7yNawxMQzYM3pSj?usp=sharing>
- [7] Github Linki : <https://github.com/alperenvali/Root-Note-Analyzer-II>