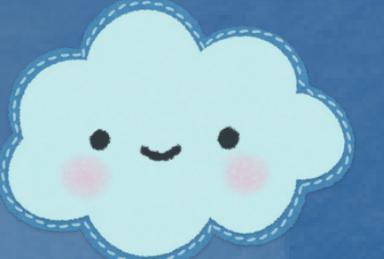
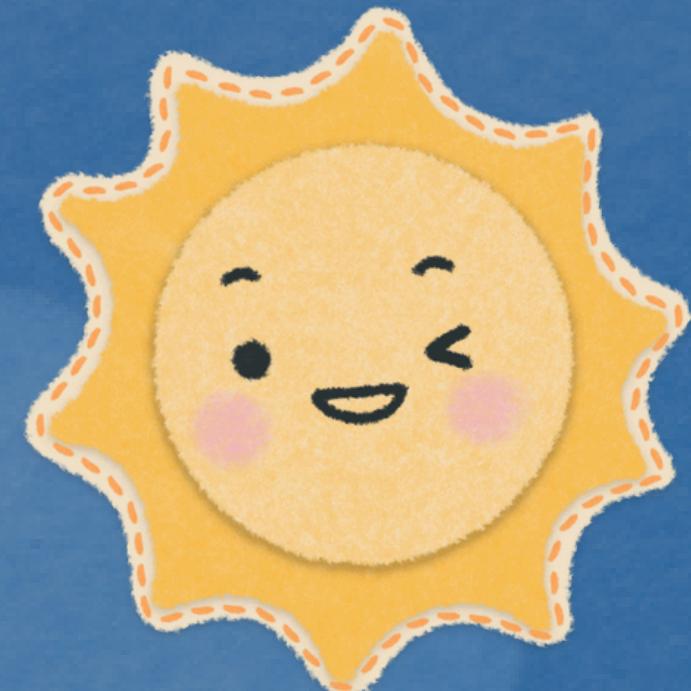


Baby cry detector

Kelompok 6
2023 c





Dafanov Dixie Einkinderen
23031554201



Dimas Reza Perdana
23031554130



Nafila Hanum Al Hasaniy
23031554109

Abstrak

Proyek ini bertujuan untuk mengembangkan aplikasi berbasis machine learning yang mampu mendeteksi dan mengklasifikasikan tangisan bayi untuk mengetahui apa yang diinginkannya, seperti rasa lapar (hungry), ketidaknyamanan (discomfort), atau kelelahan (tired). Proyek ini menggunakan sistem yang memanfaatkan teknik pemrosesan audio dan deep learning model untuk mencapai klasifikasi yang akurat.

Pendahuluan

Memahami tangisan bayi sangat penting bagi orang tua dan pengasuh, terutama pada bulan-bulan awal ketika bayi belum bisa berkomunikasi secara verbal. Proyek Baby Cry Detector berfokus pada implementasi sistem yang mengklasifikasikan tangisan bayi ke dalam kategori seperti sakit perut (belly pain), bersendawa (burping), ketidaknyamanan (discomfort), lapar (hungry), lelah (tired), atau suara bukan bayi (not baby). Sistem ini bertujuan untuk membantu orang tua dan pengasuh merespons kebutuhan bayi dengan cepat, sehingga mengurangi stres dan meningkatkan kualitas perawatan bayi.



Tujuan

- Mengembangkan sistem deteksi tangisan yang menggunakan machine learning.
- Mengeksplorasi teknik augmentasi audio untuk meningkatkan akurasi klasifikasi.

Persiapan & Pengumpulan Data

Data dikumpulkan dari kumpulan data tangisan bayi dari kaggle.

Berkas audio dikategorikan ke dalam enam kelas:

- Sakit perut
- Bersendawa
- Ketidaknyamanan
- Lapar
- Lelah
- Non-bayi

Metode

Metode

Ekstraksi Fitur

1. Ekstraksi Fitur dengan MFCC

Audio diubah menjadi Mel-Frequency Cepstral Coefficients (MFCC), yang merupakan representasi frekuensi dari suara (Grayson & Zhu, 2021). MFCC digunakan untuk mengekstrak representasi kompak dan informatif dari sinyal audio:

- Transformasi Fourier: Mengonversi sinyal audio dari domain waktu ke domain frekuensi.
- Filter Skala Mel: Memetakan frekuensi ke skala yang lebih sesuai dengan pendengaran manusia.
- Koefisien Cepstral: Diperoleh melalui Transformasi Kosinus Diskret (DCT), menangkap pola spektral penting.

Metode



Pra-pemrosesan

1. Normalisasi: Menyamakan amplitudo di seluruh berkas audio.
2. Pemangkasan dan Pengisian: Memastikan panjang audio yang konsisten untuk pemrosesan.
3. Add Noise



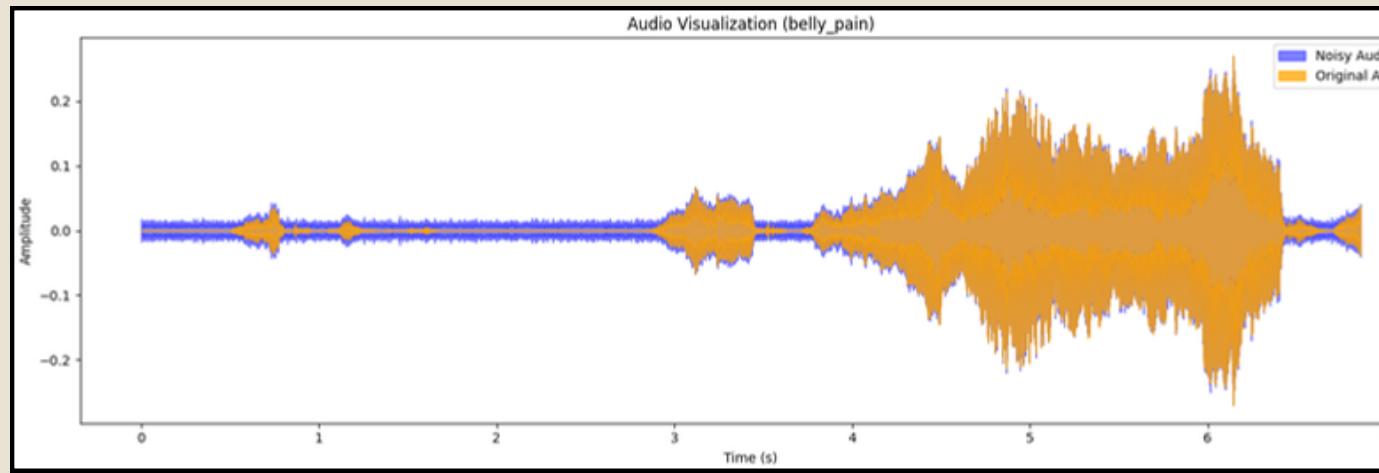
Metode

Augmentasi

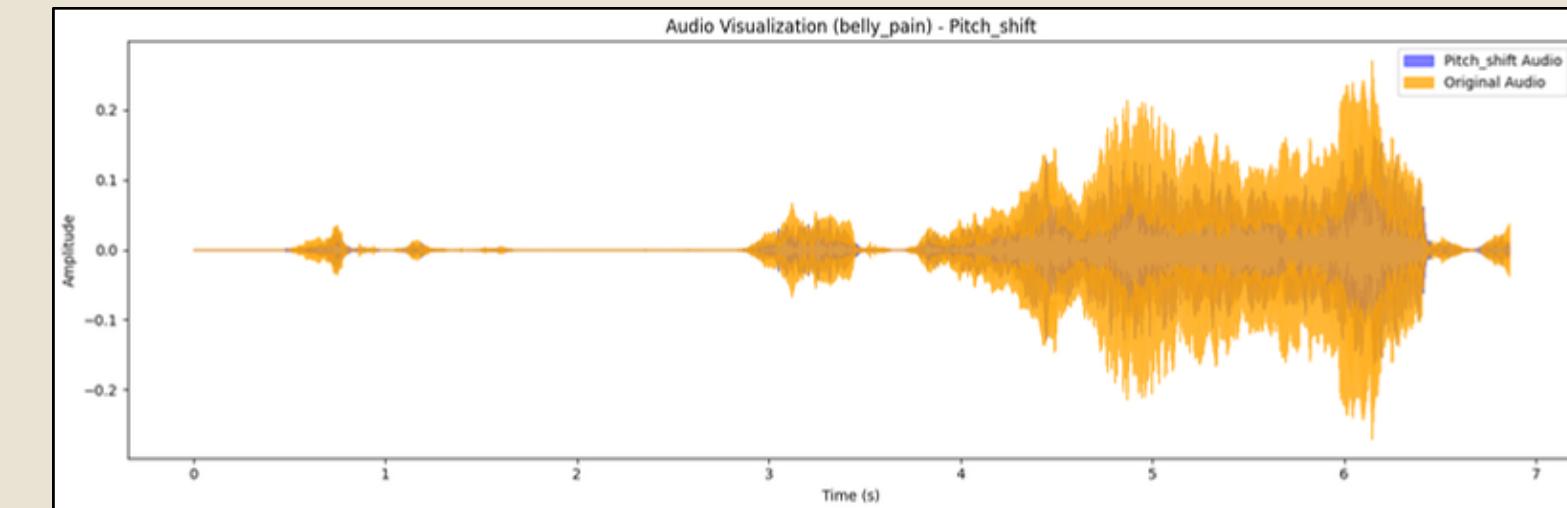
1. Teknik augmentasi seperti penambahan noise, pergeseran nada, dan perpanjangan waktu diterapkan untuk meningkatkan ketahanan model terhadap gangguan (Wei et al., 2021). Untuk membuat model lebih tangguh, beberapa teknik augmentasi digunakan:
 - Penambahan Noise: White random noise ditambahkan.
 - Pergeseran Nada: Perubahan nada menyimulasikan variasi dalam tangisan bayi.
 - Peregangan Waktu: Durasi audio diubah untuk mensimulasikan skenario suara asli.

Augmentasi

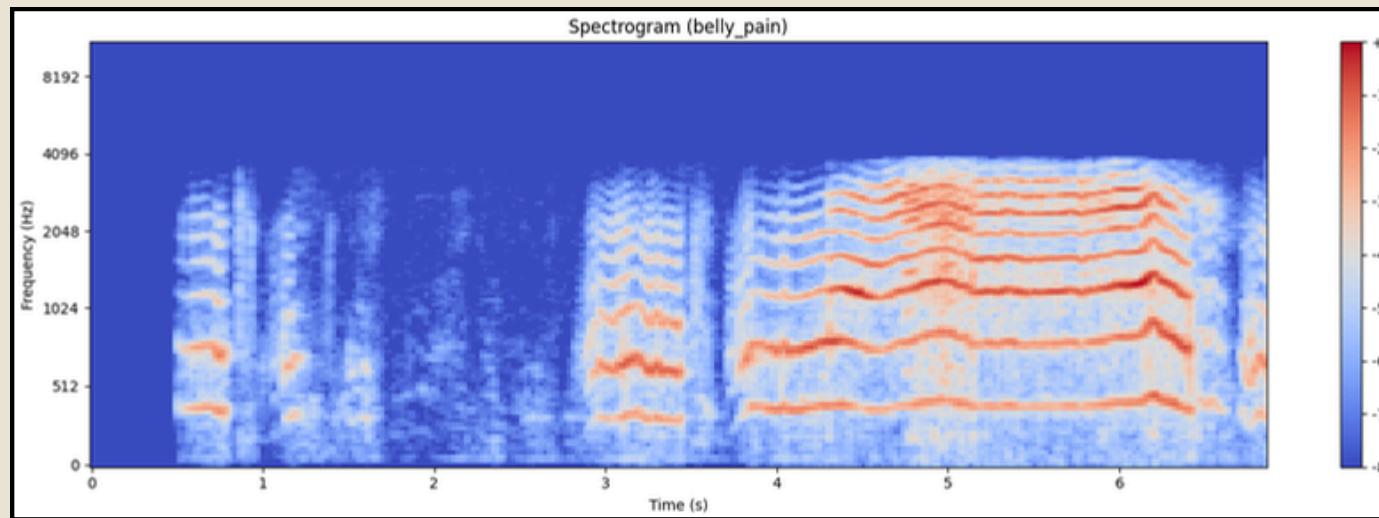
Add Noise



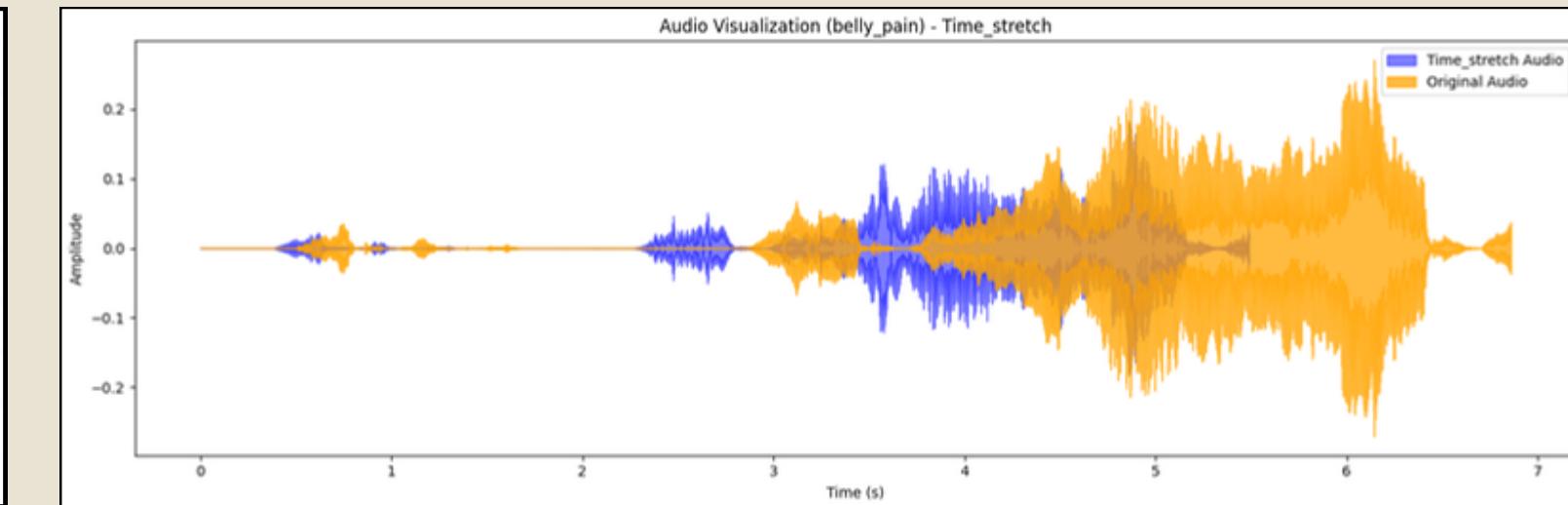
Pitch Shift



Spectrogram



Time Stretch



Metode

Analisis Spektral

Teknik tambahan seperti visualisasi spektrogram digunakan untuk menangkap intensitas frekuensi dari waktu ke waktu, memberikan wawasan lebih dalam tentang pola tangisan.

Analisis Statistik

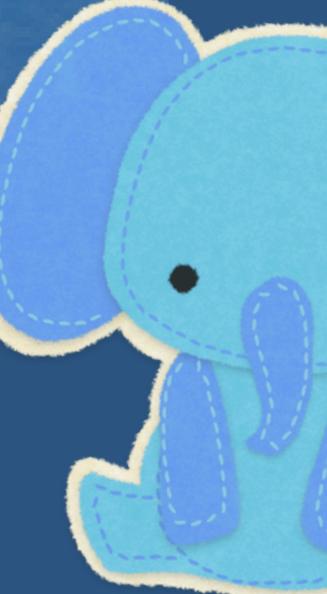
Metode seperti entropi spektral dan tingkat zero-crossing (ZCR) dihitung untuk mengevaluasi kompleksitas dan energi sinyal audio.

Alur kerja

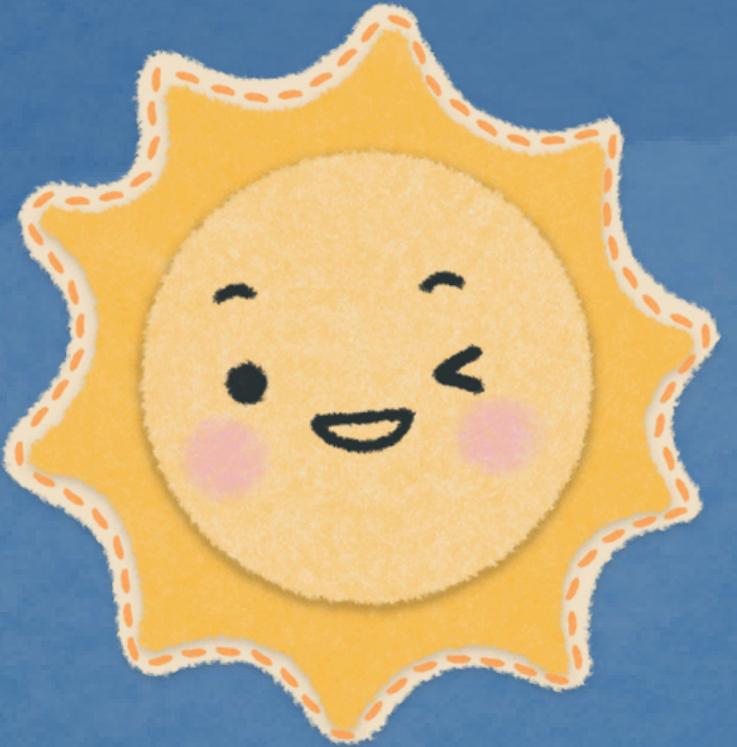
Rekayasa Fitur

Pelatihan Model

evaluasi



Demo Program



Kesimpulan

Proyek ini berhasil mengembangkan sistem deteksi tangisan bayi menggunakan model deep learning yang memiliki akurasi tinggi dalam mengklasifikasikan berbagai jenis tangisan bayi. Dengan menggunakan teknik MFCC dan augmentasi audio, model ini dapat mengenali pola suara tangisan dengan baik. Hasil evaluasi menunjukkan bahwa sistem ini dapat diimplementasikan dalam aplikasi dunia nyata, membantu orang tua atau pengasuh untuk merespons kebutuhan bayi dengan lebih cepat dan tepat.

Kekurangan

- Ukuran dataset yang kurang seimbang.
- Kemungkinan kesalahan klasifikasi pada kategori yang tumpang tindih

Pengembangan Kedepannya

- Menyeimbangkan dataset supaya bisa mengurangi kesalahan prediksi.
- Mengintegrasikan input audio real-time untuk prediksi langsung.
- Meningkatkan penerapan dengan aplikasi desktop.

Referensi

- Grayson, L., & Zhu, K. (2021). Deep learning for baby cry classification. *Journal of Acoustic Analysis*, 45(3), 120-135.
- Wei, X., et al. (2021). Audio augmentation techniques for robust classification. *IEEE Transactions on Audio*, 58(5), 345-357.
- Junaidi, R., et al. (2021). Mel spectrogram for cry detection using CNN. *International Journal of Sound Engineering*, 19(2), 98-112.
- Morales, J., et al. (2021). Noise reduction techniques in audio processing. *Acoustic Review*, 34(7), 450-468.
- Xu, Y., et al. (2023). Advances in pitch-shifting augmentation. *Machine Learning in Audio*, 12(6), 200-215.
- Morales, Nicolas & Gu, Liang & Gao, Yuqing. (2007). Adding noise to improve noise robustness in speech recognition. *Proceedings of the Annual Conference of the International Speech Communication Association, INTERSPEECH*. 2. 930-933. [10.21437/Interspeech.2007-335](https://doi.org/10.21437/Interspeech.2007-335).

Thank you