

ABSTRAK

Tangisan bayi adalah bentuk komunikasi utama yang digunakan untuk menyampaikan kebutuhan mereka. Tangisan bayi dapat menimbulkan stres pada orang tua, terutama jika mereka kesulitan mengidentifikasi penyebabnya. Selain itu, tangisan bayi dapat menyebabkan ketidaknyamanan berkepanjangan yang mengganggu tidur, pertumbuhan, dan perkembangan bayi. Oleh karena itu, dibutuhkan sistem otomatis untuk mengidentifikasi suara tangisan bayi. Penelitian ini menggunakan metode *Mel-Frequency Cepstral Coefficients* (MFCC) untuk mengekstraksi fitur dan *Convolutional Neural Networks* dengan *Long Short Term Memory* (CNN-LSTM) untuk mengenali pola dalam suara tangisan bayi. Dataset yang digunakan adalah data suara tangisan bayi berdasarkan teori *Dunstan Baby Language*, yang membagi data menjadi lima kategori yaitu lapar, lelah, bersendawa, tidak nyaman, dan lelah. Penelitian ini mengevaluasi beberapa skenario, termasuk variasi fitur seperti MFCC, MFCC Delta, dan MFCC Delta-Delta, serta jumlah epoch dan nilai *k-fold cross-validation*, dengan teknik *balancing* data menggunakan SMOTE dan Tomek Links. Berdasarkan hasil pengujian, akurasi yang dicapai sebesar 89% dengan menggunakan 20 epoch, fitur MFCC, nilai k-fold 10, dan teknik *balancing* data SMOTE. Peningkatan jumlah epoch menjadi 50 dengan fitur MFCC, nilai k-fold 10, dan teknik *balancing* SMOTE menghasilkan akurasi tertinggi sebesar 95%. Sebaliknya, pada teknik *balancing* data Tomek Links, akurasi yang dicapai sebesar 78% dicapai dengan 20 epoch, fitur MFCC, dan nilai k-fold 10, sementara dengan jumlah epoch 50 pada kondisi yang sama, akurasi tertinggi mencapai 80%. Penelitian ini menunjukkan bahwa peningkatan jumlah epoch memungkinkan model untuk belajar lebih baik dari data pelatihan, sehingga meningkatkan performa klasifikasi suara tangisan bayi.

Kata Kunci : *Klasifikasi suara, Tangisan Bayi, MFCC, CNN, LSTM, Dunstan Baby Language*

ABSTRACT

Infant crying is a primary form of communication used by babies to express their needs. This crying can cause stress for parents, especially if they struggle to identify its cause. Additionally, persistent crying can lead to prolonged discomfort, disrupting the baby's sleep, growth, and development. Therefore, an automated system is needed to identify infant cries. This study employs Mel-Frequency Cepstral Coefficients (MFCC) for feature extraction and Convolutional Neural Networks with Long Short Term Memory (CNN-LSTM) to recognize patterns in infant cries. The dataset used consists of infant cry recordings based on the Dunstan Baby Language theory, categorizing the cries into five groups: hungry, tired, burping, uncomfortable, and tired. The study evaluates several scenarios, including variations in features such as MFCC, MFCC Delta, and MFCC Delta-Deltas, as well as different numbers of epochs and k-fold cross-validation values, with data balancing techniques using SMOTE and Tomek Links. The results show that the highest accuracy of 89% was achieved using 20 epochs, MFCC features, a k-fold value of 10, and the SMOTE balancing technique. Increasing the number of epochs to 50, while using MFCC features, a k-fold value of 10, and SMOTE, resulted in the highest accuracy of 95%. Conversely, with Tomek Links balancing technique, the highest accuracy of 78% was achieved with 20 epochs, MFCC features, and a k-fold value of 10, while increasing the number of epochs to 50 in the same conditions resulted in an accuracy of 80%. This study demonstrates that increasing the number of epochs allows the model to better learn from training data, thus improving the performance of infant cry classification.

Keywords: *Cry Classification, Infant Crying, MFCC, CNN, LSTM, Dunstan Baby Language*

