## 1a) Label Accuracy

The automatic labels coincided with the human annotations for the selected three files from the most frequent categories:

* Bird Chirp (633992): 6 / 6 spans (100 % agreement)
* Dog Bark (118101): 16 / 16 spans (100 % agreement)
* Car (640611): 1 / 1 spans (100 % agreement)

In every case the cyan label trace aligns with clear peaks in the spectrogram during the annotated time ranges. Events are unambiguously audible where they’re labeled.

\begin{table}[ht]

\centering

\begin{tabular}{|c|l|c|c|c|p{6cm}|}

\hline

\textbf{file\\_id} & \textbf{class} & \textbf{n\\_annots} & \textbf{onset\\_range\\_s} & \textbf{offset\\_range\\_s} & \textbf{comment} \\ \hline

633992 & Bird Chirp & 2 & 6.21 & 6.36 & Two distinct chirps; label jumps to 1.0 exactly over both spans; no false positives. \\ \hline

118101 & Dog Bark & 8 & 16.78 & 22.33 & All eight barks captured; slight 0→0.5 dips at edges reflect minor annotator boundary shifts. \\ \hline

640611 & Car & 1 & 0.00 & 0.00 & Single continuous honk; label matches exactly over the entire region; no disagreement. \\ \hline

\end{tabular}

\caption{Label accuracy summary per file and class.}

\end{table}|

A close-up of a purple and blue gradient

AI-generated content may be incorrect.

(keep the first only if we run outta space bro)

A close-up of a graph

AI-generated content may be incorrect.

A close-up of a graph

AI-generated content may be incorrect.

Boundary stats computed via

```

boundary\_stats('633992','Bird Chirp') # → {'n\_annots':2,'onset\_range\_s':6.2108,'offset\_range\_s':6.3613}

boundary\_stats('118101','Dog Bark') # → {'n\_annots':8,'onset\_range\_s':16.7758,'offset\_range\_s':22.3322}

```

Labels perfectly cover every human-annotated span with zero false positives. Minor boundary shifts reflect individual annotator differences, which could be smoothed in post-processing if exact edge alignment is required.

## 1 b) Useful Features

A diagram of a diagram

AI-generated content may be incorrect.

ZCR and MFCC analysis over all annotated segments for the same three categories shows clear separation:

Zerocrossing Rate (ZCR):

Bird Chirp shows the highest median ZCR, consistent with its rapid, staccato sound. Dog Bark is intermediate, and Car has the lowest ZCR, indicating its smoother low-frequency profile. Therefore, ZCR strongly discriminates Bird Chirp from the other classes.

Mean MFCC (MFCC\_mean):

Bird Chirp also shows the highest median MFCC\_mean, due to its high-frequency content. Dog Bark occupies the middle range, while Car has the lowest median MFCC\_mean. Therefore, MFCC\_mean effectively separates Car from the animal sounds and refines the distinction among all three.

## 1 c) Class Clusters

A diagram of a graph

AI-generated content may be incorrect.

The classificational effectiveness of MFCCs and ZCR was assessed using two-dimensional PCA:

Bird Chirp has a distinctive combination of high ZCR and MFCC content, shown by its narrow, vertically oriented cloud at negative–to–moderate PC1 and consistently high PC2 values.

Dog Bark samples cluster in the lower‐left quadrant of the plot (low PC1, low PC2), reflecting their intermediate ZCR and MFCC characteristics relative to the other two classes.

Car instances occupy a broad region on the right side (high PC1) with moderate PC2, consistent with their low ZCR and low MFCC\_mean. Although Car overlaps partially with Dog Bark along PC2, the two remain mostly separable along PC1.

Overall, each class forms a visually coherent cluster with minimal overlap. Bird Chirp is especially well‐separated, while Dog Bark and Car exhibit only slight intermixing. These results confirm that MFCC\_mean and ZCR together produce feature spaces in which samples of the same class naturally group tightly, supporting their use as powerful discriminative features for downstream classification.