Midterm Progress Report

AudioT - Team EUT

An audio classification approach to assess bird transportation welfare

Georgia Tech

Table of contents

Georgia Tech

1. Introduction - The chicken industry

Every year, the USA ranks in the **top 3 countries for chicken consumption** per capita[2].

Concurrently, **animal welfare** is a growing consumer, industry and public administration concern[3].

Transportation stress is known to have an impact on **meat quality**, with duration and conditions impacting pH and cortisol levels, glycogen in muscles and the development of PSE (Pale, Soft, and Exudative) meat, undesirable in the industry[4,5,6,7,8].

Georgia Tech

1. Introduction - AudioT

AudioT specializes in leveraging **acoustics**, **machine learning** and **signal processing** expertise to **monitor poultry farms**.

AudioT solutions enable industry actors to **detect anomalies** including intrusion, sickness, equipment failures, and changes in chicken behavior, ranging from hatching to transport to the processing facilities.

2. Project details - Objectives

The objective of the Eurotruck project (EUT) is to build upon AudioT's knowledge and insights regarding vocalizations from late-stage broiler flocks, and **apply it to the birds during transport** in order to **detect** and measure **distress events**.

The underlying idea is to detect, measure and analyze **flock activity**, assuming that **a quiet flock is a good flock**.

2. Project details - Recordings

Just like for its poultry farms monitoring, AudioT relies on microphones to generate its primary source of data: **recordings of bird vocalizations**.

Recordings can then be analyzed, labeled and used to train models able to distinguish between **different types of vocalizations** and alert in case of distress.

Microphones were therefore placed in trucks to record flock activity for future analysis and processing.

2. Project details - Recording process

Two different types of trucks were used:

- **Regular trailer** (the most common ones in the US)
- **Aircool trailer**

Six trips were recorded in total, and **five different microphones** were used. Times provided are in CET.

- **EUT1** - Trip 1 (single trip on Aug 30, 6:30 to 8:10 AM)
  - Mic 2
  - Mic 9
- **EUT2** - Trip 2 (single trip on Aug 30, 12:48 to 17:37)
  - Mic 2
  - Mic 8
- **EUT3** - Trip 3 (out and back on Nov 29, 8:25 to 15:28)
  - Mic 2 (failed)

- **EUT1** - Trip 1 (single trip on Oct 27, 7:30 to 10:49 AM)
  - Mic 2
  - Mic 10
- **EUT2** - Trip 2 (single trip on Oct 27, 10:50 to 13:48)
  - Mic 2
  - Mic 10
- **EUT4** - Trip 3 (out and back twice on Nov 29, 8:23 to 16:23)
  - Mic 2

Georgia Tech

2. Project details - Recording process

Regular trailer

Aircool truck

3. State of data

The data consists of **one minute audio segments** of the recordings mentioned previously. These segments are available in the form of `.flac` files (free of compression).

The data comes from:

- Different **trips**
- Different **trucks**
- Different **microphones**
- Different **microphone placements** in trucks

These are the source of **variability** in the data quality.

Georgia Tech

3. State of data

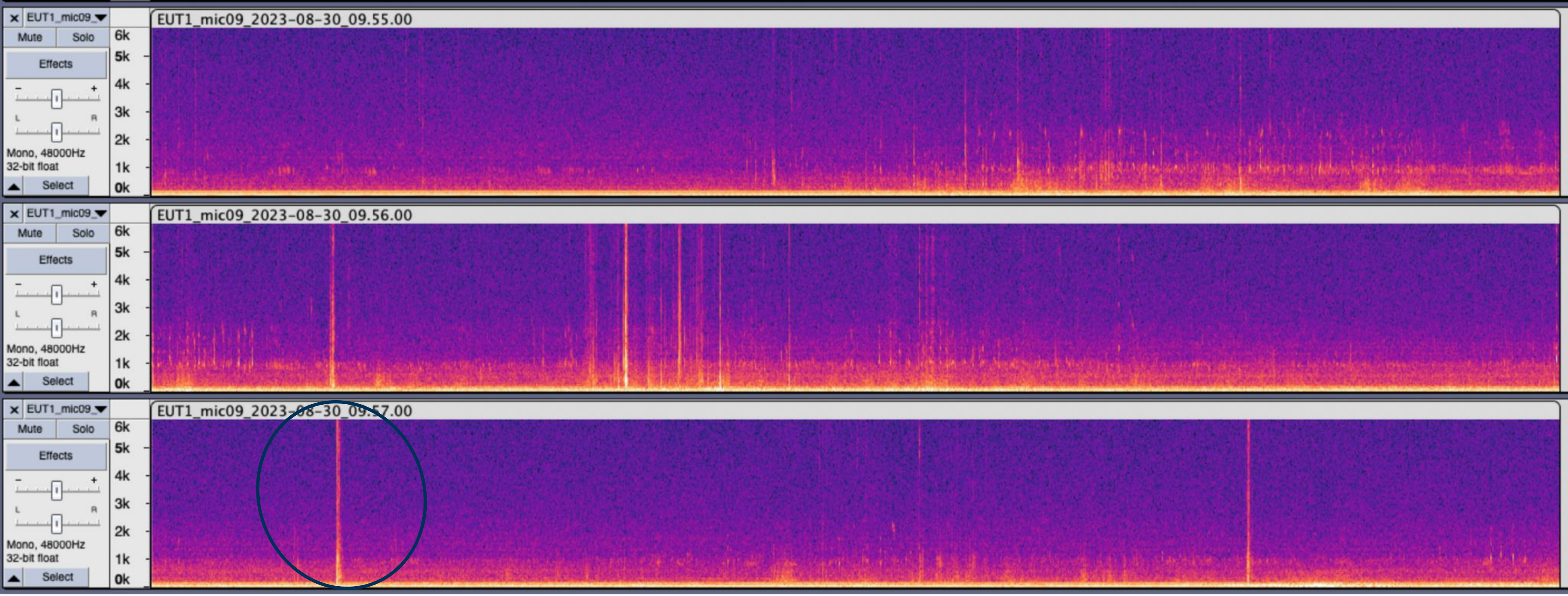
The `.flac` files provided have a standard naming convention.

- **Example: `EUT1_mic02_2023-08-30_07.00.00.flac`**
- **Decoding the naming structure:**
  - `EUT1` matches the legend on slide 7, indicating this is from trip EUT1, an aircool trailer trip on 8/30/2023.
  - `mic02` indicates that the audio is from Mic 2
  - `2023-08-30` indicates the date, though this is redundant with information provided on the legend
  - `07.00.00` indicates the minute in GMT, so an hour needs to be added for it to match the legend, which is in CET (i.e., 8:00 CET.)
    - This is particularly important, as there are audio files of the trucks when chickens are not present.

3. State of data - Processing

We rely on **audio and visual analysis** (spectrogram, wavelength) to **label the data** in Audacity. The final result, on top of the features contained in the audio file itself, are a timestamp and corresponding labels, and the presence or absence of labels in itself.

We need to identify interesting segments, label them accurately and enough of them to have a solid training dataset. Audacity lets us listen and visually analyze several files at the same time.



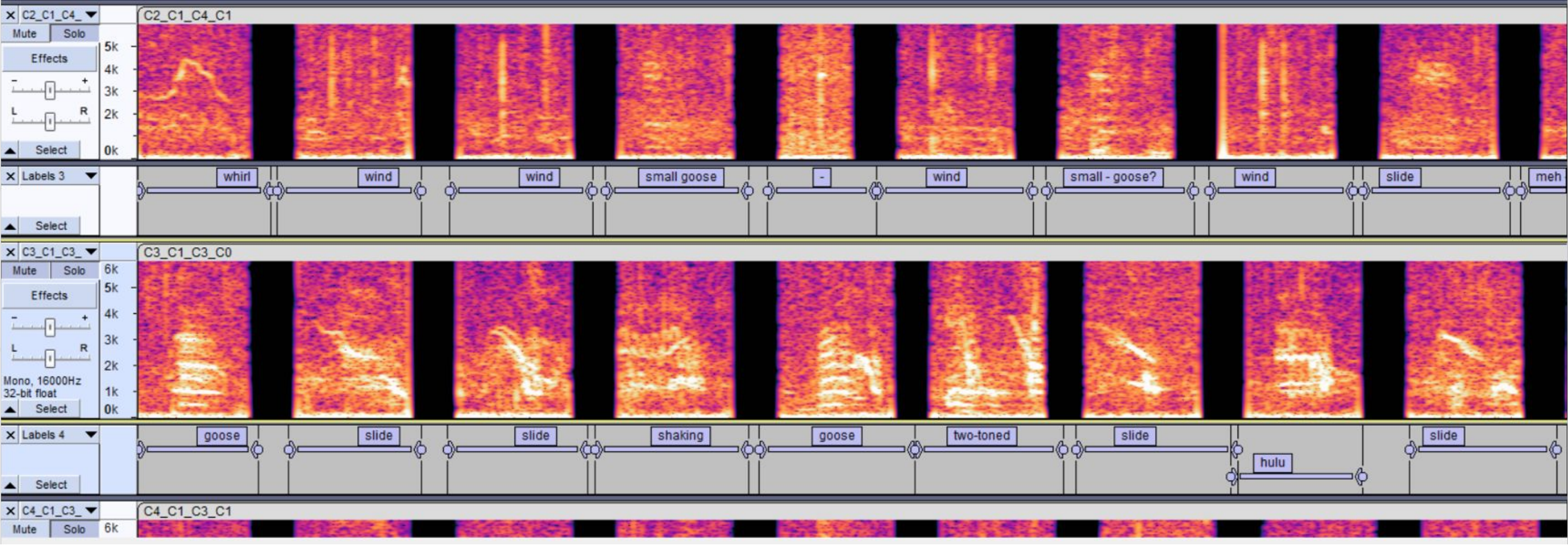
The 3 audio data files here shows **EUT truck 1 (Aircooled) on 08/30/2023 from 9:55 AM - 9:57 AM CET** using **microphone #9**. Visually, the warmer coloured tones (yellow/orange/red) shows audio captured through the microphone. When identifying bird vocalization, they are generally identified to be up to the 3000 Hz frequencies. For any above that range - as identified with the circle above, that audio is most likely a “truck” noise (rustling, clanking, brushing of the microphone).

3. State of data - Processing

Bird vocalizations can be labeled in different fashions, and we **collaborate** with another AudioT team (Team AW6) working on **late stage vocalization** (right before they get transported). Such vocalizations can be labeled as cooing, cackling, chirping, trilling, coughing…

In transport, **a lot of other noises are recorded** as well, but shouldn’t be mixed up for bird activity. Such noises are air vents, crates rattling or rustling, wind blowing, road bumping, etc.

Georgia Tech



An example of existing AudioT taxonomy, presenting a dictionary matching specific sound patterns and audio cues to visual patterns on the corresponding spectrogram. Our initial effort will focus on binary classification (detecting flock activity vs none). However, such a taxonomy can help bring variance to the training dataset and improve model accuracy in a multiclass classification approach (more on this point in Part 6).

4. Data preparation challenges

- Labeling sounds accurately on the **timeframe**
- Labeling sounds accurately (our interpretation of the sound is **subjective**)
- Finding one-minute segment with **enough activity to learn from**
- **Balancing segments** with quiet flock, segments with active flocks and segments without flock
- Potential **imbalance** (distress events are hopefully rare)

4. Data preparation challenges

- Selecting quality segments **without biasing** the training set
- Working with microphones in **real-life conditions** (close to an air vent, on the exterior of the truck, obstructed by crates)
- Agreeing and deciding on the **label dictionary** (binary - activity / no activity, or more precise for multi-classification, bringing more variance to learn from but potentially more data to train on)
- Dealing with **some low quality recordings** (we can't just increase the sound, because noise also increases)
- Noise reduction vs filtering lower frequencies

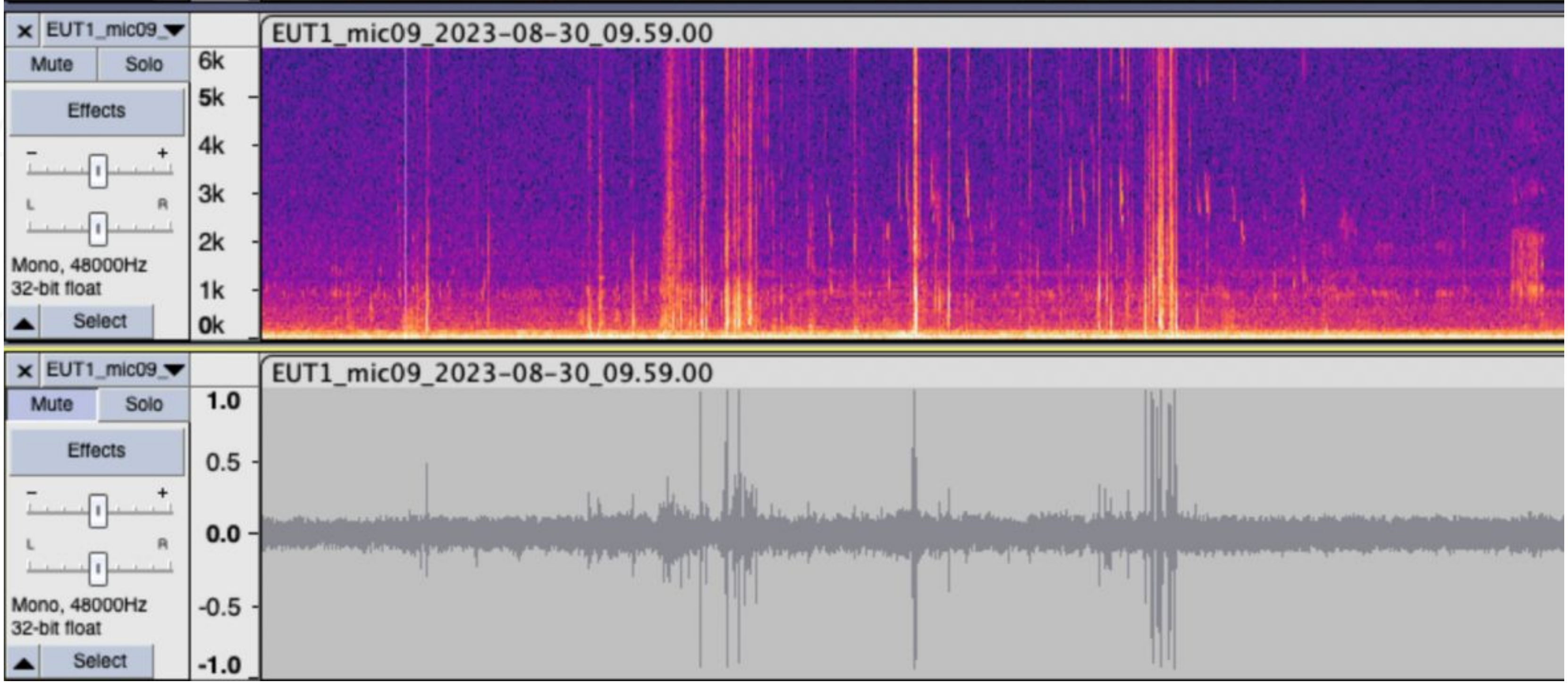
4. Data preparation challenges

Due to fan noise, road noise, and other external factors, extra noise in the data is a challenge. Two potential ways of addressing this are **noise reduction** and **filtering**.

- **Noise Reduction** involves identifying a noise profile of the noise to remove and then applying it to the audio file
- **Filtering** involves attenuating (reducing) specific frequencies:
  - **Low-pass filters** attenuate frequencies above a specific cutoff
  - **High-pass filters** attenuate frequencies below a specific cutoff
- For our project, a filtering approach is more appropriate, as noise reduction leads to loss of bird vocalizations due to shared frequencies

5. Data features

- Truck
- Trip number
- Microphone used
- Timestamp
- Presence or absence of labels
- Statistical summaries of mean and variance of spectrogram coefficients from Mel spectrogram (frequency) and MCFF (timber)
- Spectral contrast
- Chroma features

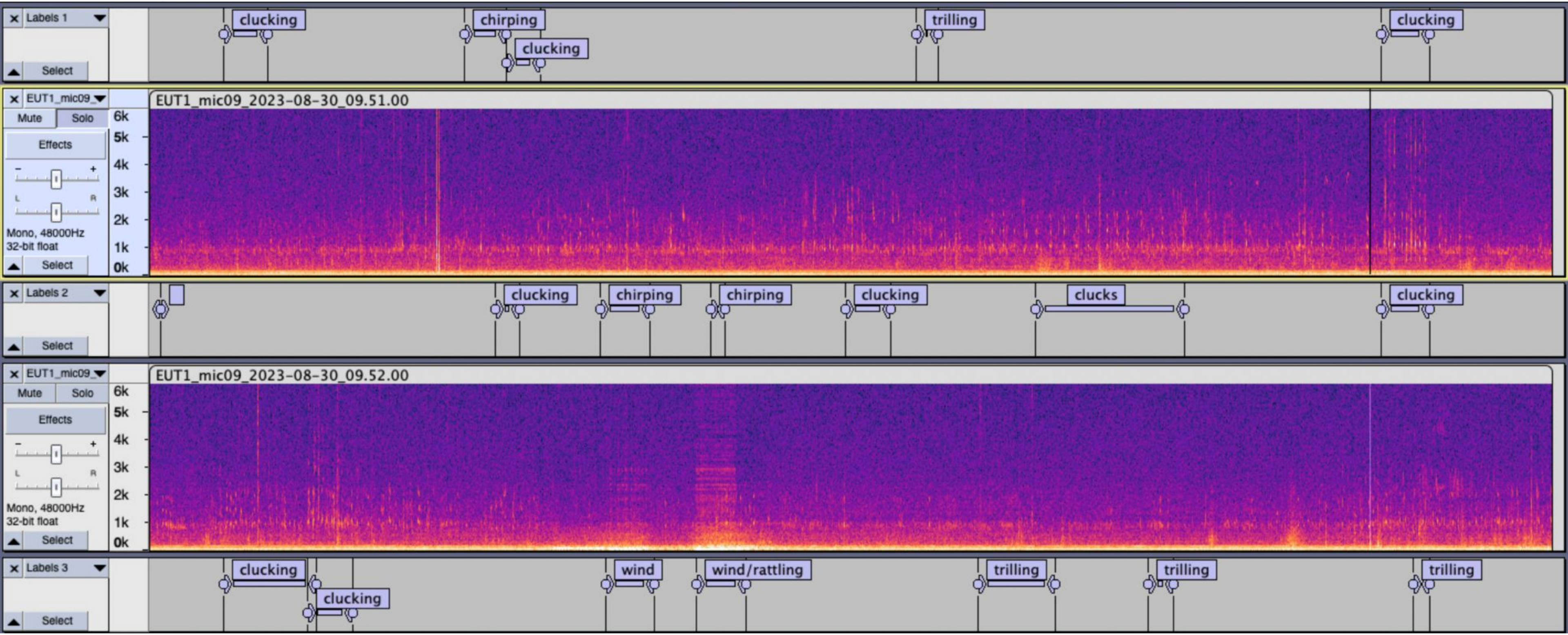


The same audio file in spectrogram and waveform views. "Spectrogram view provides a visual indication of how the energy in different frequency bands changes over time" - here set with a minimum frequency of 0 Hz and a maximum frequency of 6000 Hz. Below the waveform view provides a "linear vertical scale running from -1.0 (negative values) to +1.0 (positive values), centered on zero." The +1 denotes the maximum possible loudness when positive and the -1 denotes the maximum possible when negative. Within waveform view, there is also a decibel scale available.

5. Data features

In terms of labeling, **three different approaches** are possible:

- **Binary** labeling (flock activity vs everything else)
- **Multiclass** labeling (classifying flock and other noises more accurately, e.g. the taxonomy dictionary mentioned in Part 3)
- **Multilabel** labeling (each team member labels each sound based on a common taxonomy)



The 3 audio files **EUT truck 1 (Aircooled) on 08/30/2023 from 9:55 AM - 9:57 AM CET** using **microphone #9**, with labels. It is important to note that the audio files from the transportation of the birds have much lower amounts of bird activity. This is a positive outcome as there are no sounds of distress. From the data that has been labelled, we have found only positive bird activity including clucks, trilling/coos, and chirping.

6. Feature selection

- Mel spectrogram for feature selection
  - Emphasizes **frequency** content
  - Matches the human auditory system's sensitivity, especially in the lower frequency range
  - Helps identify frequency **patterns specific to certain sounds**
- Mel-Frequency Cepstral Coefficients (MCFF) for feature selection
  - Emphasizes **timber**
  - Helps **distinguish between sounds** with different textures despite **having similar frequencies**
- Random Forest, CNN or RNN for frame-based approach

6. Feature selection

→ *A **Mel Spectrogram** can help us identify segments of high activity*

→*A **MCFF** approach can complement this by helping distinguish between sounds within this segment of high activity*

7. Modeling approaches

- **Frame-based** approach
  - Audio is analyzed in short, overlapping frames
  - Efficient in detecting short-term events and sound characteristics
  - Increases computational complexity
  - Sensitive to noise (requires noise reduction preprocessing)
- **Segment-based** approach
  - Better suited for capturing the overall activity level or the presence of sustained sounds
  - Reduces computational complexity
  - Lower temporal resolution (may miss short-lived sounds or rapide changes)
  - Choice of segment length impacts specific event detection

7. Modeling approaches

→ *A **segment-based** approach seems more appropriate to our **flock activity detection** objective*

→*A **frame-based** approach could complement the segment-based approach, intervening only within active segments:*

- *To **ensure activity is indeed flock activity** and not noise: a current issue at AudioT is some models detecting crates rattling in empty trucks before loading or after unloading as flock activity*
- *To assess the **nature** of the flock activity: are we hearing **comfort** chirping or **distress** chirping?*

Georgia Tech

7. Modeling approaches

- ***Logistic regression, SVM and Random Forest*** *could help in binary classification*
- ***CNN, RNN and gradient boosting machines*** *could help in multiclass classification*
- ***Binary relevance*** *(with Logistic Regression, SVMs, or Random Forests),* ***classifier trains and neural networks with multiple outputs layers*** *could help in multilabel classification*

Tools used: Audacity for labeling, scikit learn and librosa for modeling

7. Modeling approaches - simple to complex

- Establish **baselines** using simpler approaches (binary classification, segment-based only)
- Increase **model complexity** with **classification complexity** (move to multiclass and multilabel, merge segment and frame-based approaches)
- Revise feature engineering and hyperparameter tuning as needed based on **continuous learning**
- Consider **ensemble models** if we discover several well-performing models, potentially weighing them on their respective accuracy on flock vs non-flock detection

Sources

1. Title slide picture: Chicken by Cassie, Flickr, https://flic.kr/p/9uUefi, Licence CC BY-NC-SA 2.0 DEED, changed tone to sepia.
2. Faostat, Helgi - https://www.helgilibrary.com/indicators/poultry-meat-consumption-per-capita/
3. National Chicken Council - https://www.nationalchickencouncil.org/policy/animal-welfare/
4. Grandin, T. (1997). "Assessment of Stress during Handling and Transport." Journal of Animal Science, 75(1), 249-257.
5. Warriss, P. D. (2000). "The welfare of slaughter pigs during transport." Animal Welfare, 9(4), 365-381.
6. Knowles, T. G., et al. (1999). "Effects of road transport on indices of stress in sheep." Veterinary Record, 144(15), 418-422.
7. Fazio, E., & Ferlazzo, A. (2003). "Evaluation of stress during transport." Veterinary Research Communications, 27(1), 519-524.
8. Swanson, J. C., & Morrow-Tesch, J. (2001). "Cattle transport: Historical, research, and future perspectives." Journal of Animal Science, 79(E-Suppl), E102-E109.

Georgia Tech