



Bird Species Classification Using Audio Signal Processing and Convolutional Neural Networks

Lukas Bertsch and Tom Troska

Machine Learning Lecture

Result Presentation | 18 Jul 2024

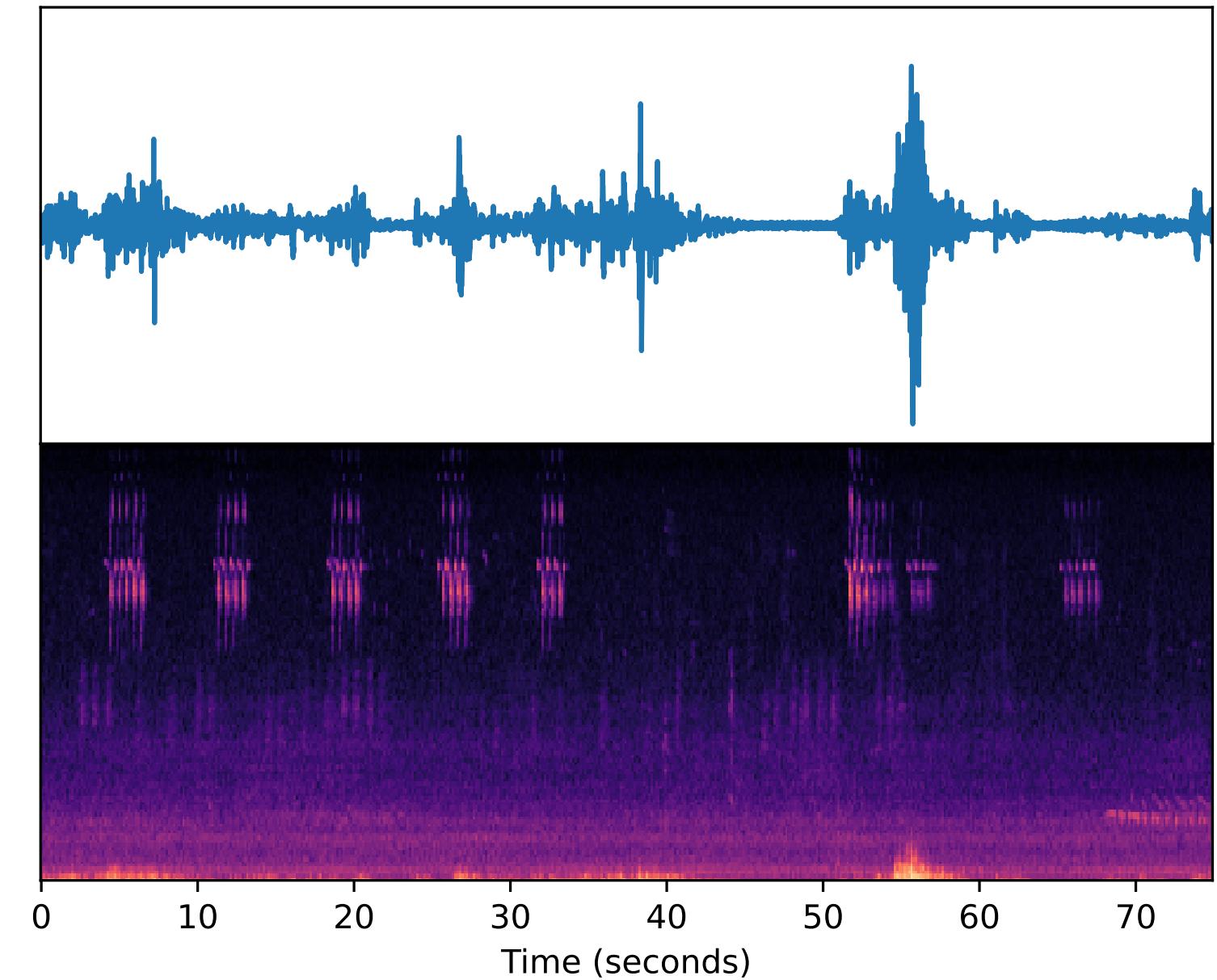
tu technische universität
dortmund

Problem

How effectively can we classify bird species by analyzing sound files using a convolutional neural network (CNN)?

Motivation

- Classifying and observing different species of birds is of great interest for wildlife research
- But only hear bird calls without seeing them
- Differentiating bird species from sound recordings requires expert knowledge and is very time consuming



→ CNN: automatically identify species by patterns in spectrograms of audio recordings

Dataset

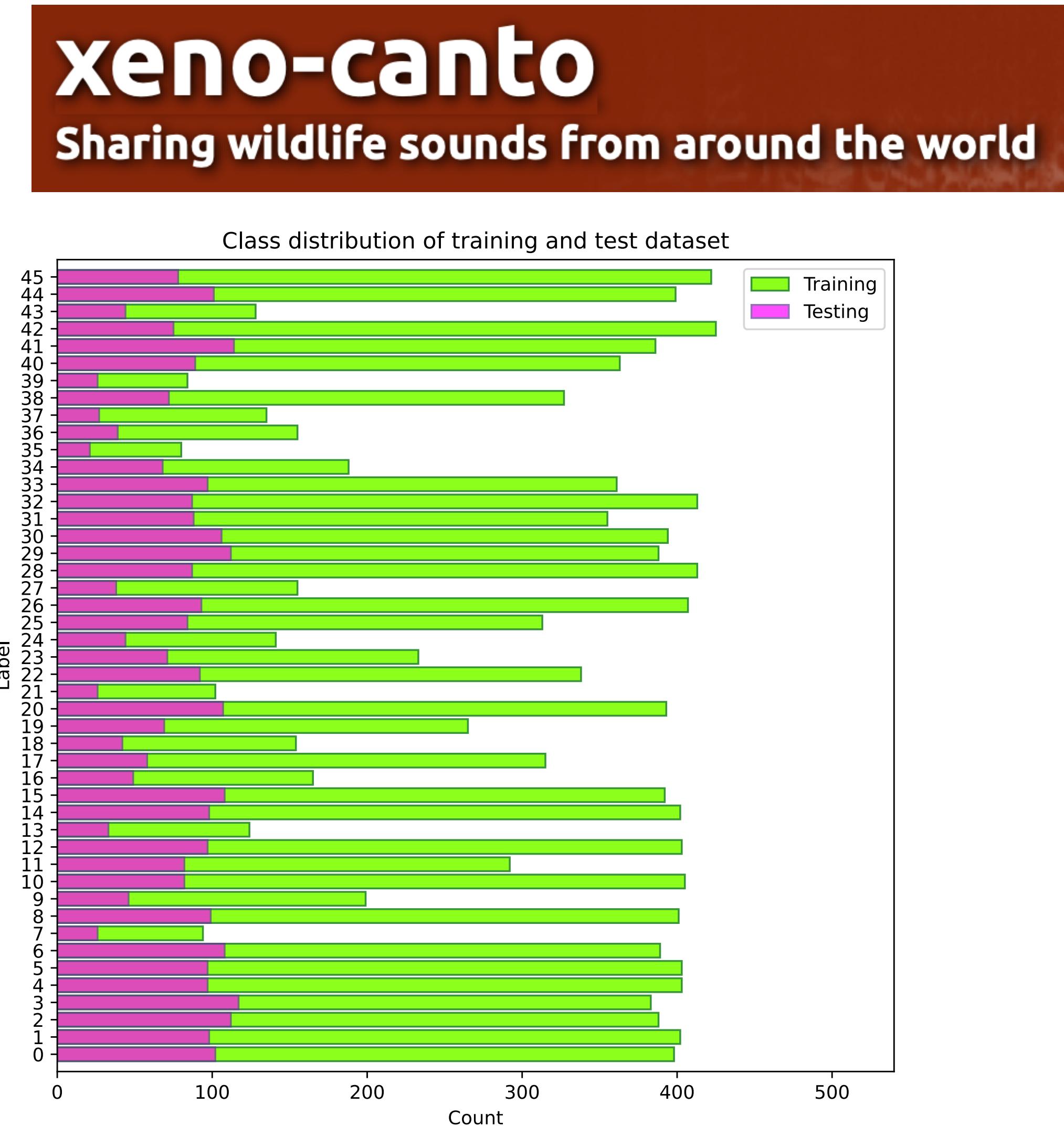
- Website: xeno-canto.org
- Platform, where animal sounds can be shared
- Data labeled by users
- Download audio files via API, individual licenses, all CC
- Total of 46 bird species with at least 100 recordings each
- Minimal duration of audio file: 5 s



| dataset_raw | | | | | | | | | | | | | | |
|-------------|--------|--------|------------------|--------------|------------------|------|---------|---|-------|-----------|-------------|--|-----|--|
| gen | sp | ssp | en | cnt | type | sex | url | lic | time | bird-seen | length_secc | fullfilename | | |
| Turdus | merula | | Common Blackbird | Germany | song | | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 19:34 | yes | 299 |/data/dataset/Turdus merula_Germany_19:34.mp3 | 299 |/data/dataset/Turdus merula_Germany_19:34.mp3 |
| Turdus | merula | | Common Blackbird | Poland | song | male | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 17:35 | yes | 297 |/data/dataset/Turdus merula_Poland_17:35.mp3 | 297 |/data/dataset/Turdus merula_Poland_17:35.mp3 |
| Turdus | merula | merula | Common Blackbird | Germany | call, song | | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 06:45 | no | 295 |/data/dataset/Turdus merula_Germany_06:45.mp3 | 295 |/data/dataset/Turdus merula_Germany_06:45.mp3 |
| Turdus | merula | | Common Blackbird | Norway | song | | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 17:04 | no | 291 |/data/dataset/Turdus merula_Norway_17:04.mp3 | 291 |/data/dataset/Turdus merula_Norway_17:04.mp3 |
| Turdus | merula | | Common Blackbird | Portugal | song | | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 05:30 | no | 290 |/data/dataset/Turdus merula_Portugal_05:30.mp3 | 290 |/data/dataset/Turdus merula_Portugal_05:30.mp3 |
| Turdus | merula | | Common Blackbird | Netherlands | alarm call, song | male | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 22:41 | yes | 289 |/data/dataset/Turdus merula_Netherlands_22:41.mp3 | 289 |/data/dataset/Turdus merula_Netherlands_22:41.mp3 |
| Turdus | merula | merula | Common Blackbird | Denmark | song | | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | xx:xx | unknown | 288 |/data/dataset/Turdus merula_Denmark_xx:xx.mp3 | 288 |/data/dataset/Turdus merula_Denmark_xx:xx.mp3 |
| Turdus | merula | | Common Blackbird | France | song | male | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 06:00 | yes | 288 |/data/dataset/Turdus merula_France_06:00.mp3 | 288 |/data/dataset/Turdus merula_France_06:00.mp3 |
| Turdus | merula | | Common Blackbird | Ireland | song | male | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 18:30 | no | 283 |/data/dataset/Turdus merula_Ireland_18:30.mp3 | 283 |/data/dataset/Turdus merula_Ireland_18:30.mp3 |
| Turdus | merula | | Common Blackbird | Germany | song | | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 05:42 | yes | 280 |/data/dataset/Turdus merula_Germany_05:42.mp3 | 280 |/data/dataset/Turdus merula_Germany_05:42.mp3 |
| Turdus | merula | | Common Blackbird | Germany | song | | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 05:04 | yes | 279 |/data/dataset/Turdus merula_Germany_05:04.mp3 | 279 |/data/dataset/Turdus merula_Germany_05:04.mp3 |
| Turdus | merula | | Common Blackbird | Germany | song | | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 16:30 | yes | 279 |/data/dataset/Turdus merula_Germany_16:30.mp3 | 279 |/data/dataset/Turdus merula_Germany_16:30.mp3 |
| Turdus | merula | | Common Blackbird | Poland | song | male | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 10:50 | yes | 277 |/data/dataset/Turdus merula_Poland_10:50.mp3 | 277 |/data/dataset/Turdus merula_Poland_10:50.mp3 |
| Turdus | merula | | Common Blackbird | Portugal | aberrant, song | | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 04:42 | no | 271 |/data/dataset/Turdus merula_Portugal_04:42.mp3 | 271 |/data/dataset/Turdus merula_Portugal_04:42.mp3 |
| Turdus | merula | | Common Blackbird | Germany | song | | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 06:24 | yes | 270 |/data/dataset/Turdus merula_Germany_06:24.mp3 | 270 |/data/dataset/Turdus merula_Germany_06:24.mp3 |
| Turdus | merula | | Common Blackbird | Portugal | alarm call, song | | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 03:54 | no | 267 |/data/dataset/Turdus merula_Portugal_03:54.mp3 | 267 |/data/dataset/Turdus merula_Portugal_03:54.mp3 |
| Turdus | merula | | Common Blackbird | Russian Fede | song | male | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 17:30 | yes | 266 |/data/dataset/Turdus merula_Russian Fede_17:30.mp3 | 266 |/data/dataset/Turdus merula_Russian Fede_17:30.mp3 |
| Turdus | merula | merula | Common Blackbird | Netherlands | song | | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 04:20 | no | 265 |/data/dataset/Turdus merula_Netherlands_04:20.mp3 | 265 |/data/dataset/Turdus merula_Netherlands_04:20.mp3 |
| Turdus | merula | | Common Blackbird | Germany | song | | //xeno- | //creativecommons.org/licenses/by-nd/4.0/ | 01:20 | no | 263 |/data/dataset/Turdus merula_Germany_01:20.mp3 | 263 |/data/dataset/Turdus merula_Germany_01:20.mp3 |

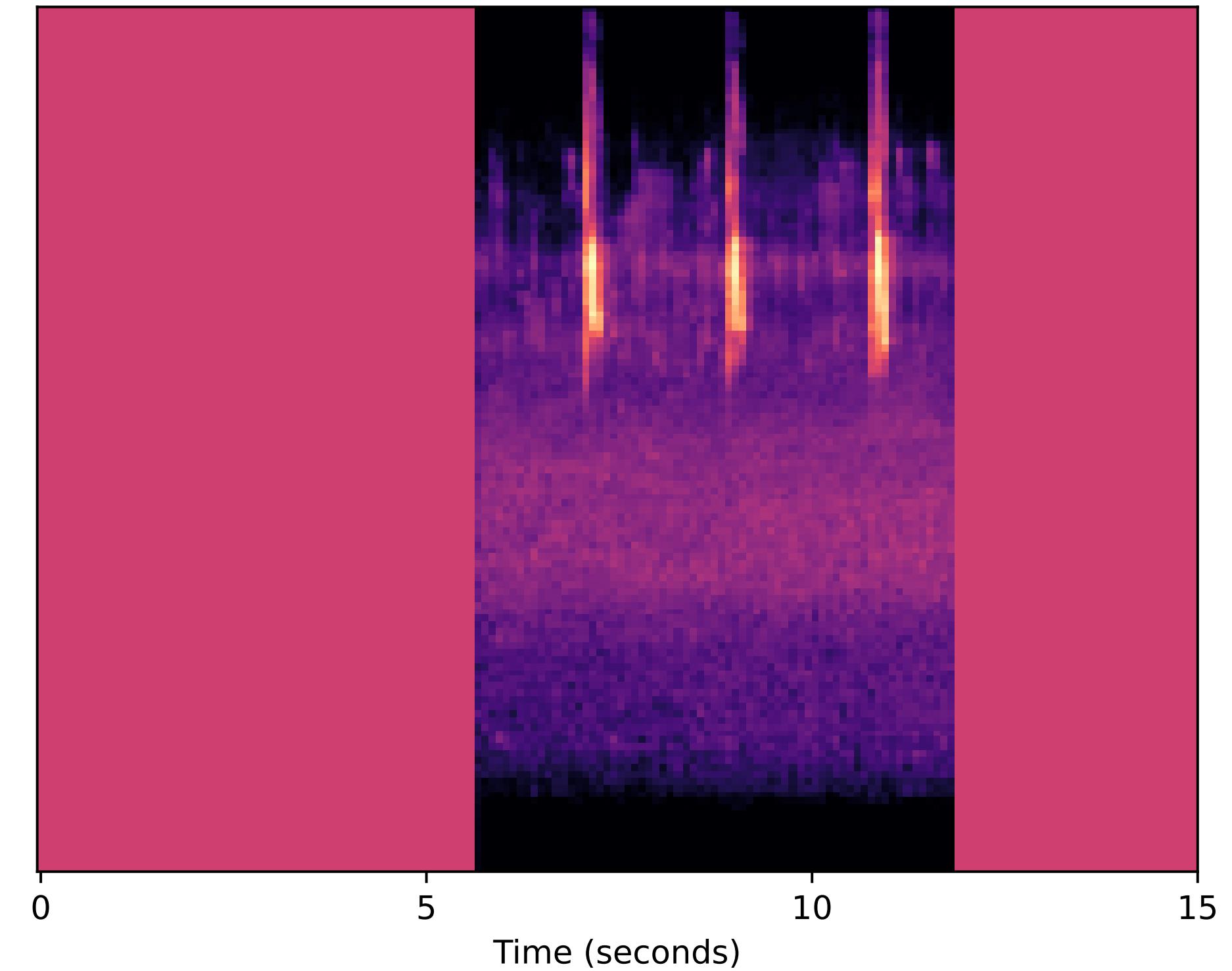
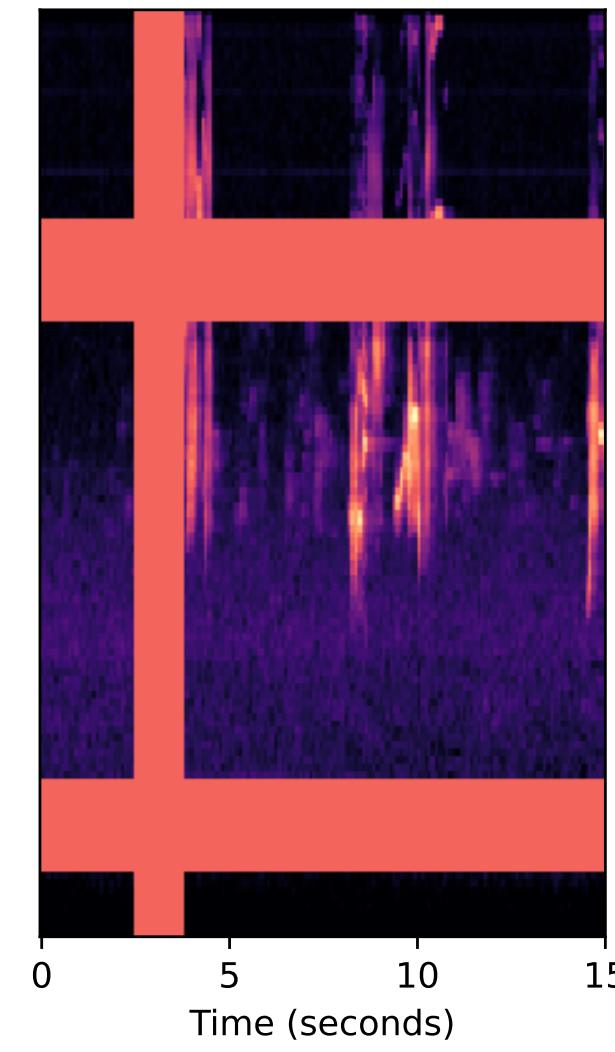
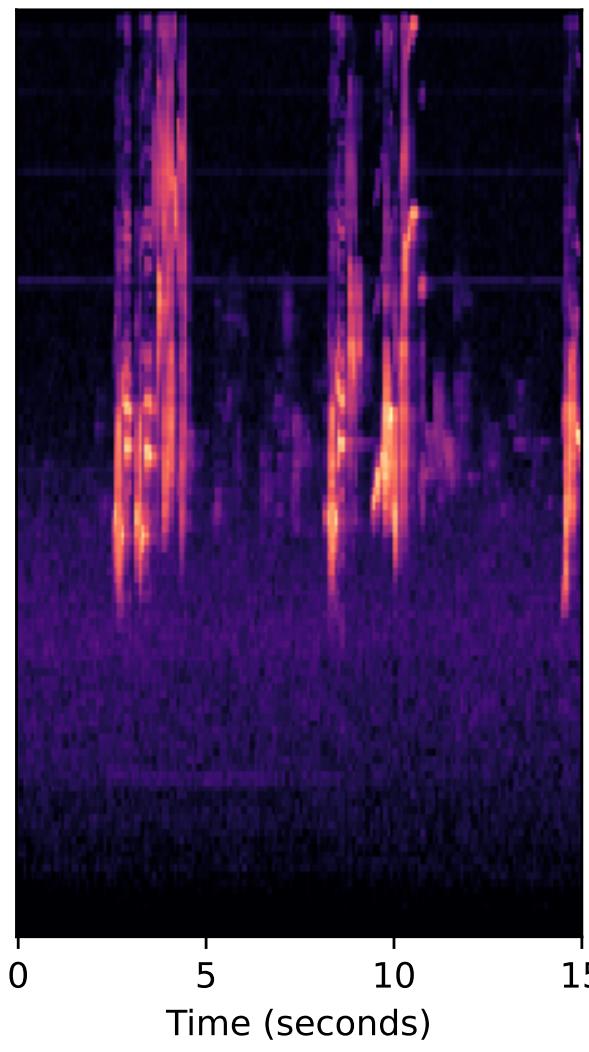
Dataset

- Website: xeno-canto.org
- Platform, where animal sounds can be shared
- Data labeled by users
- Download audio files via API,
individual licenses, all CC
- Total of 46 bird species with
at least 100 recordings each
- Minimal duration of audio file: 5 s

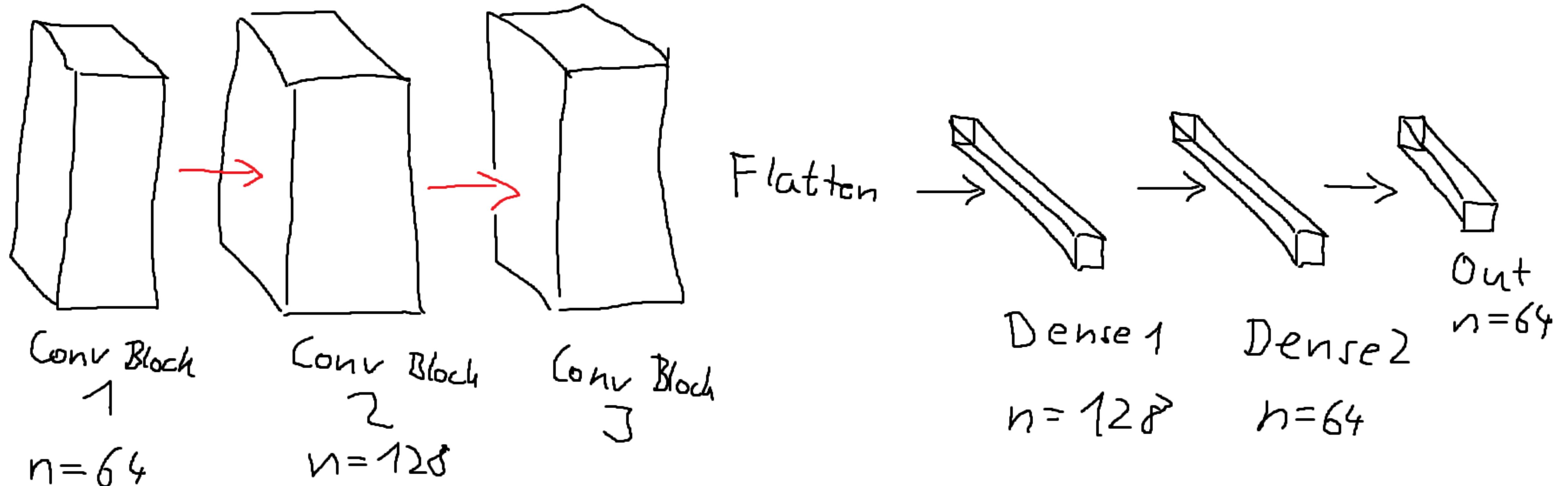


Details on Method

- Compute spectrograms for each audio file
- While training: draw new random 15 s clips from spectrogram every epoch
- Predict audio file: split spectrogram into 15 s clips → take average prediction
- If clip shorter than 15 s: pad spectrogram
- Data augmentation: time-frequency masking

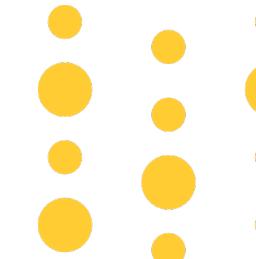


Network Architecture

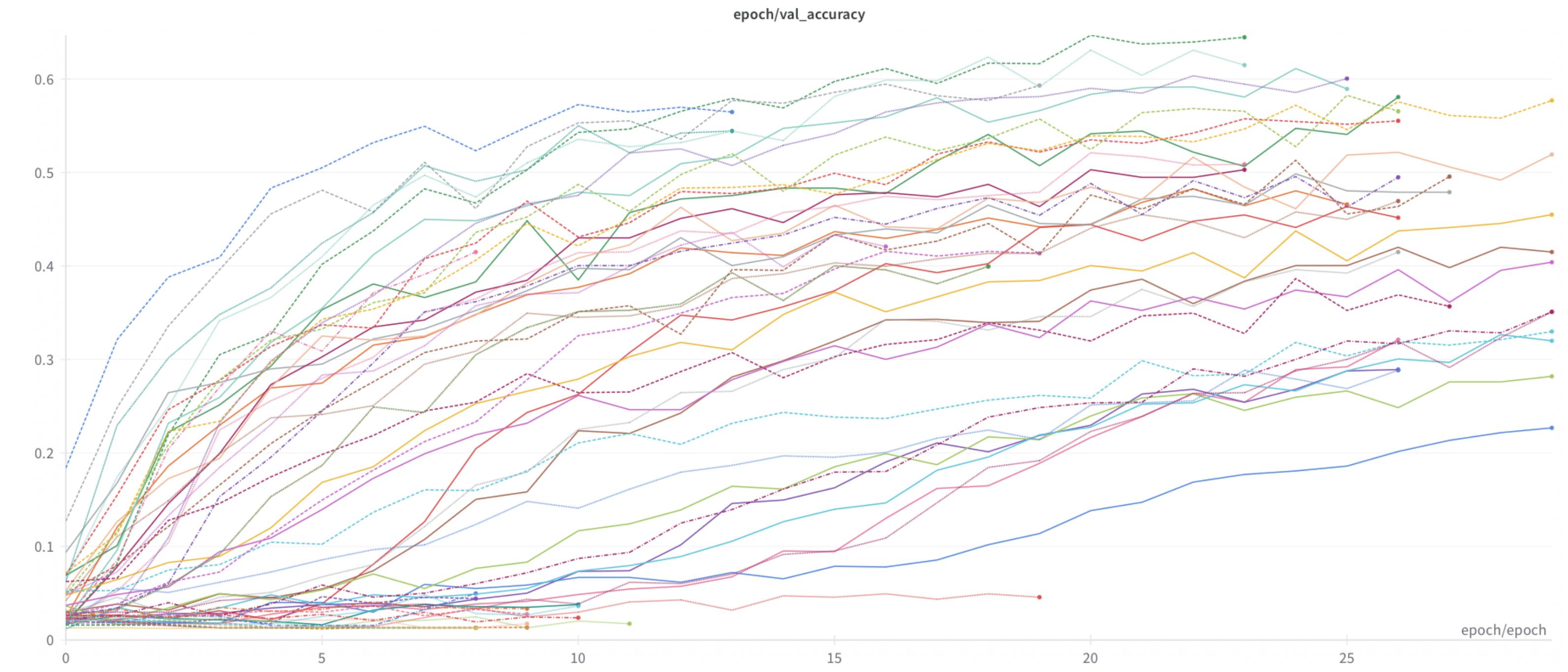


Hyperparameter Optimization

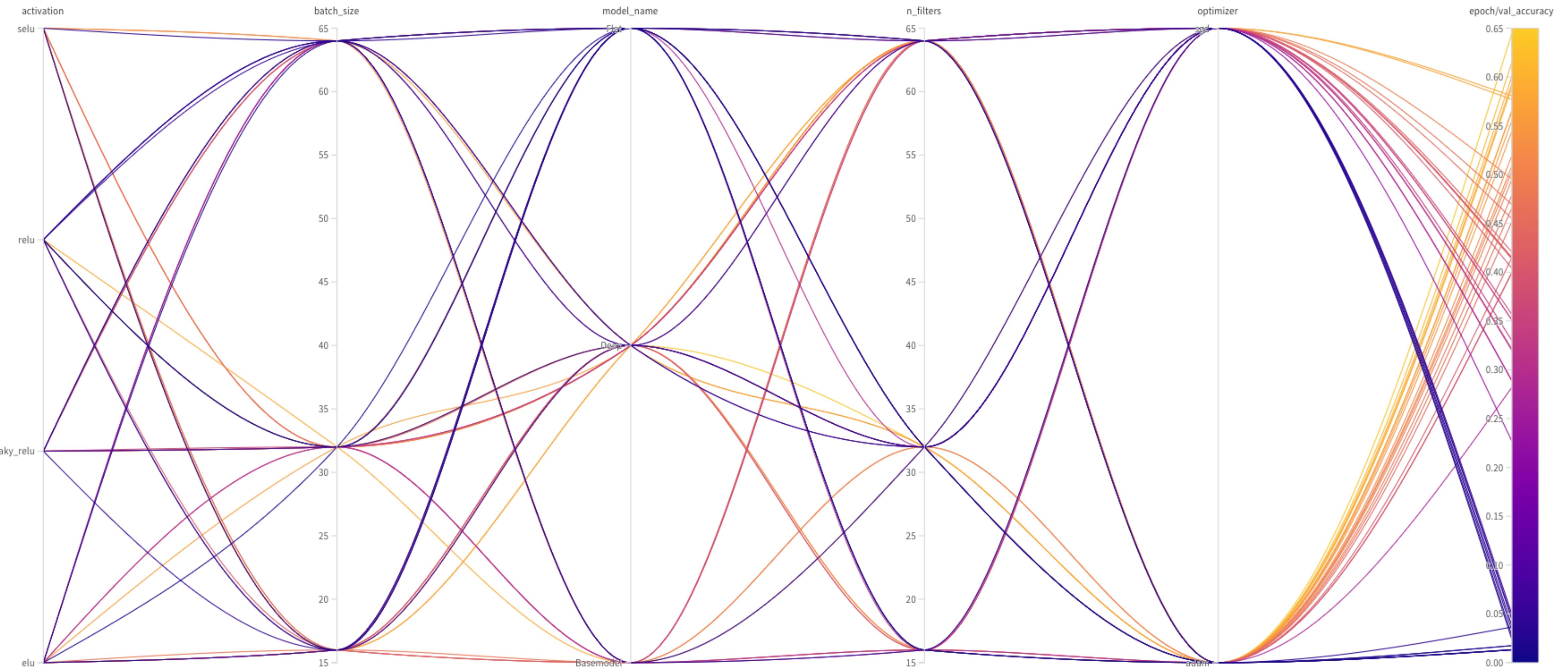
- Random search with wandb.ai sweep function
- Optimizing parameters
 - Model architecture
 - Optimizer
 - Activation function
 - N filter
 - Dropout %
 - Clip duration



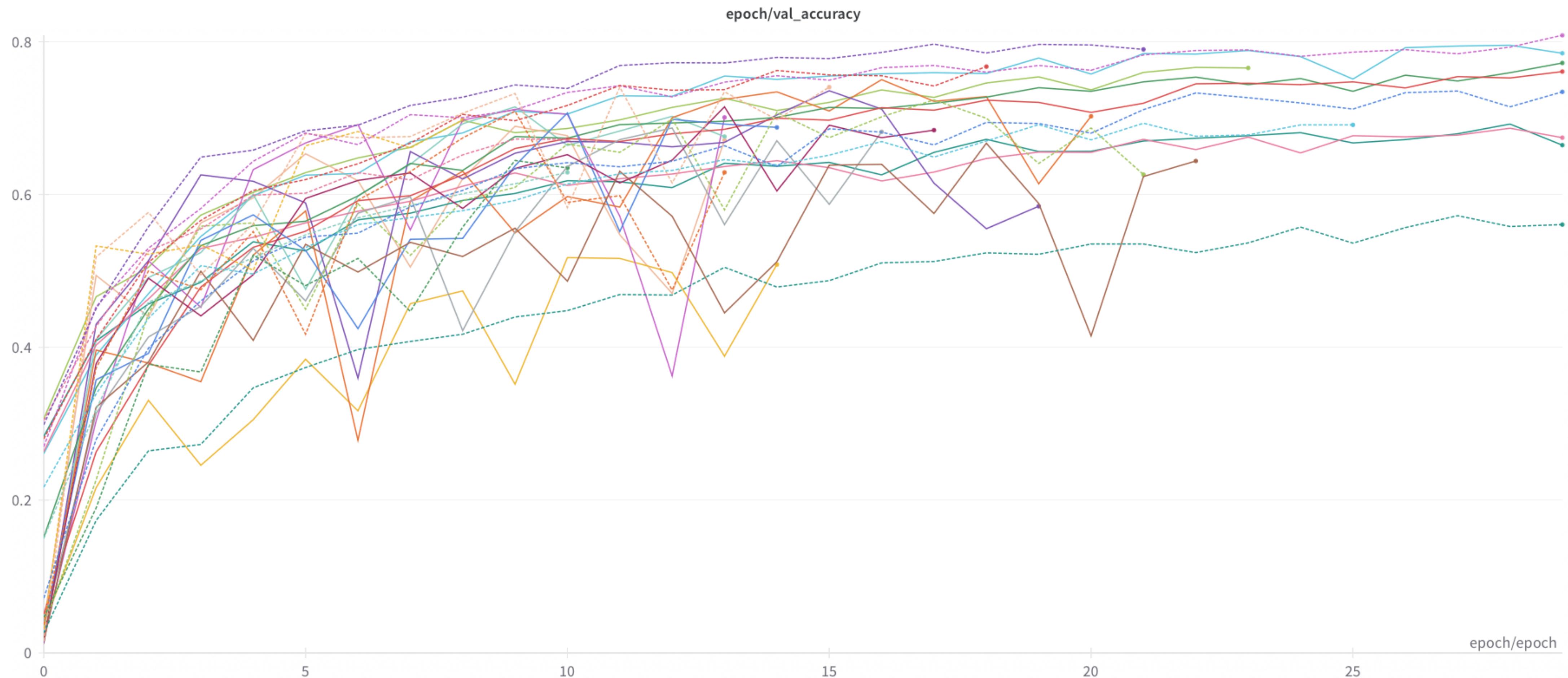
Weights & Biases



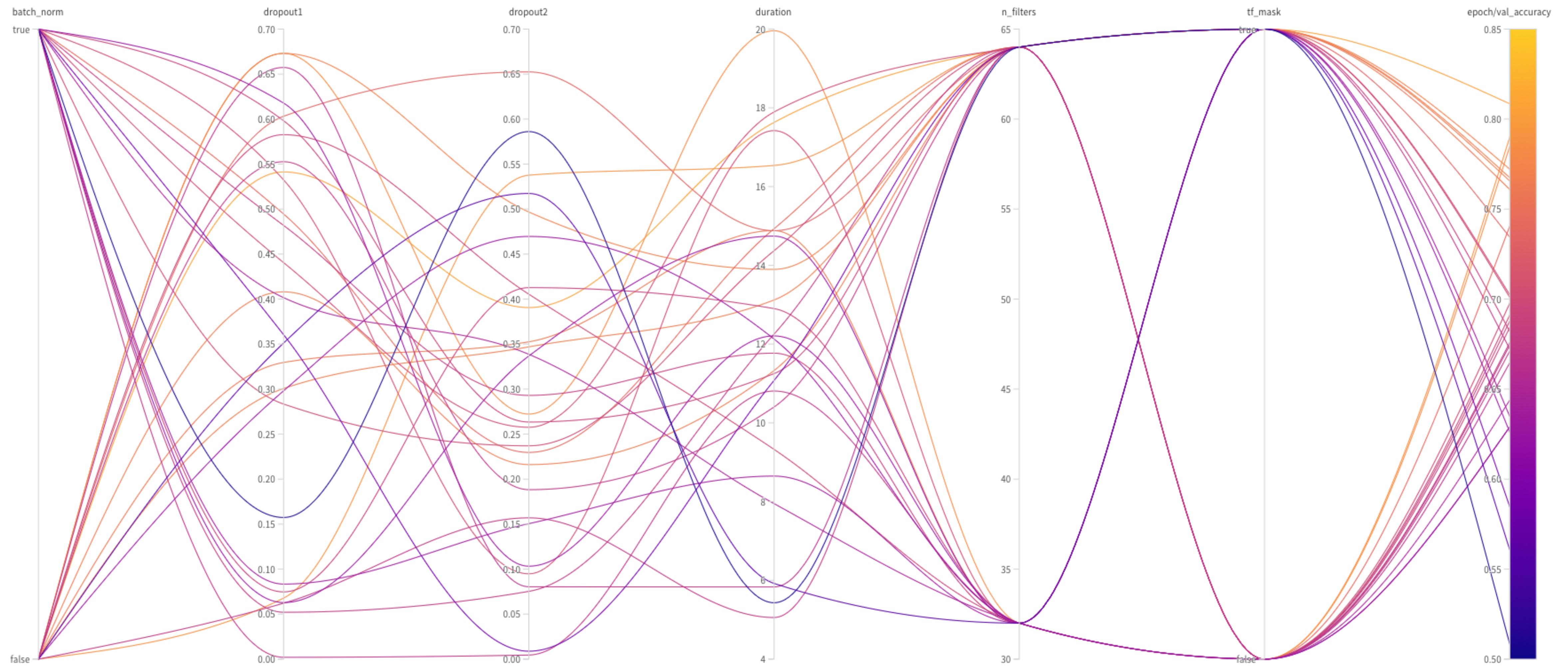
Hyperparameter Optimization Sweep #1



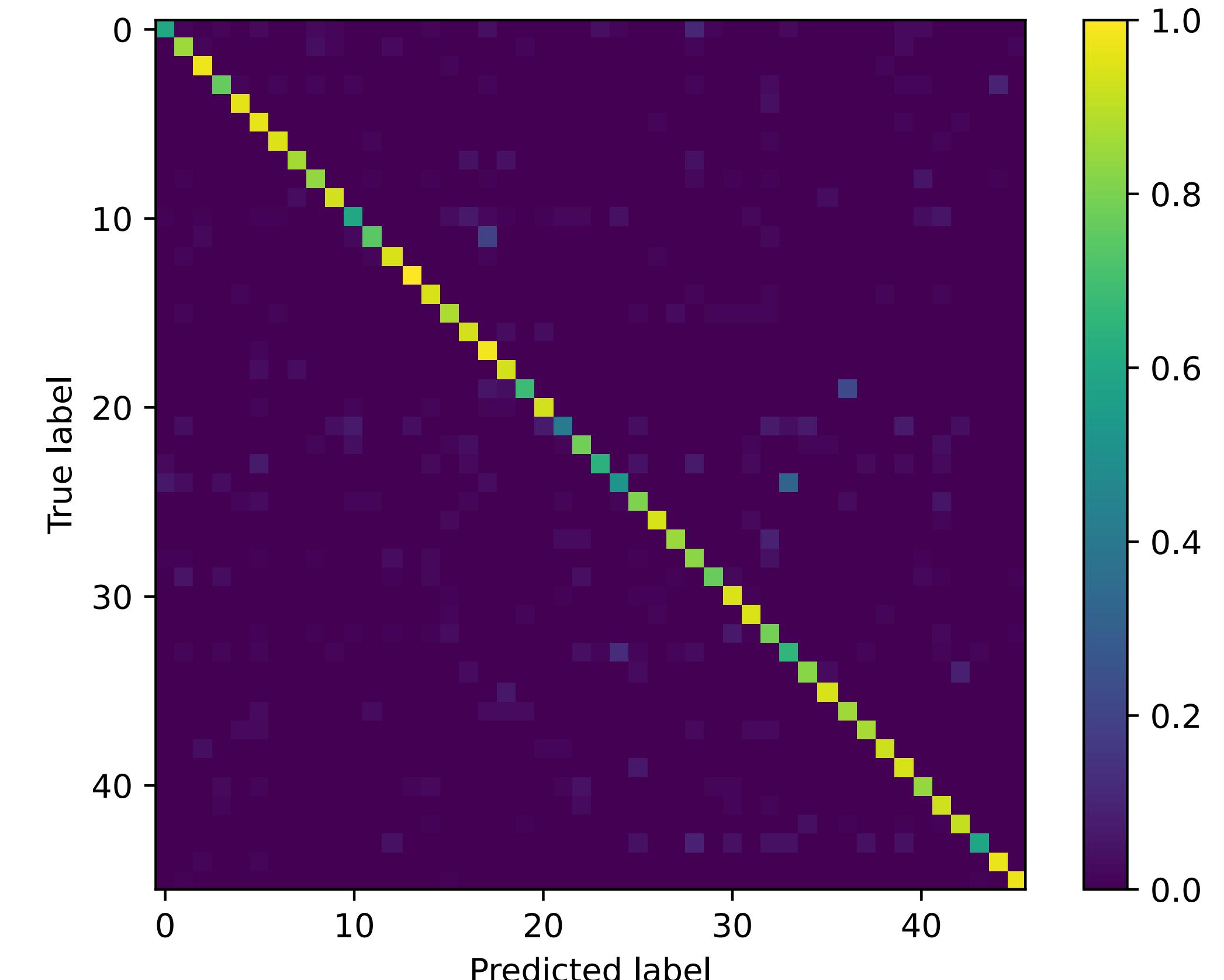
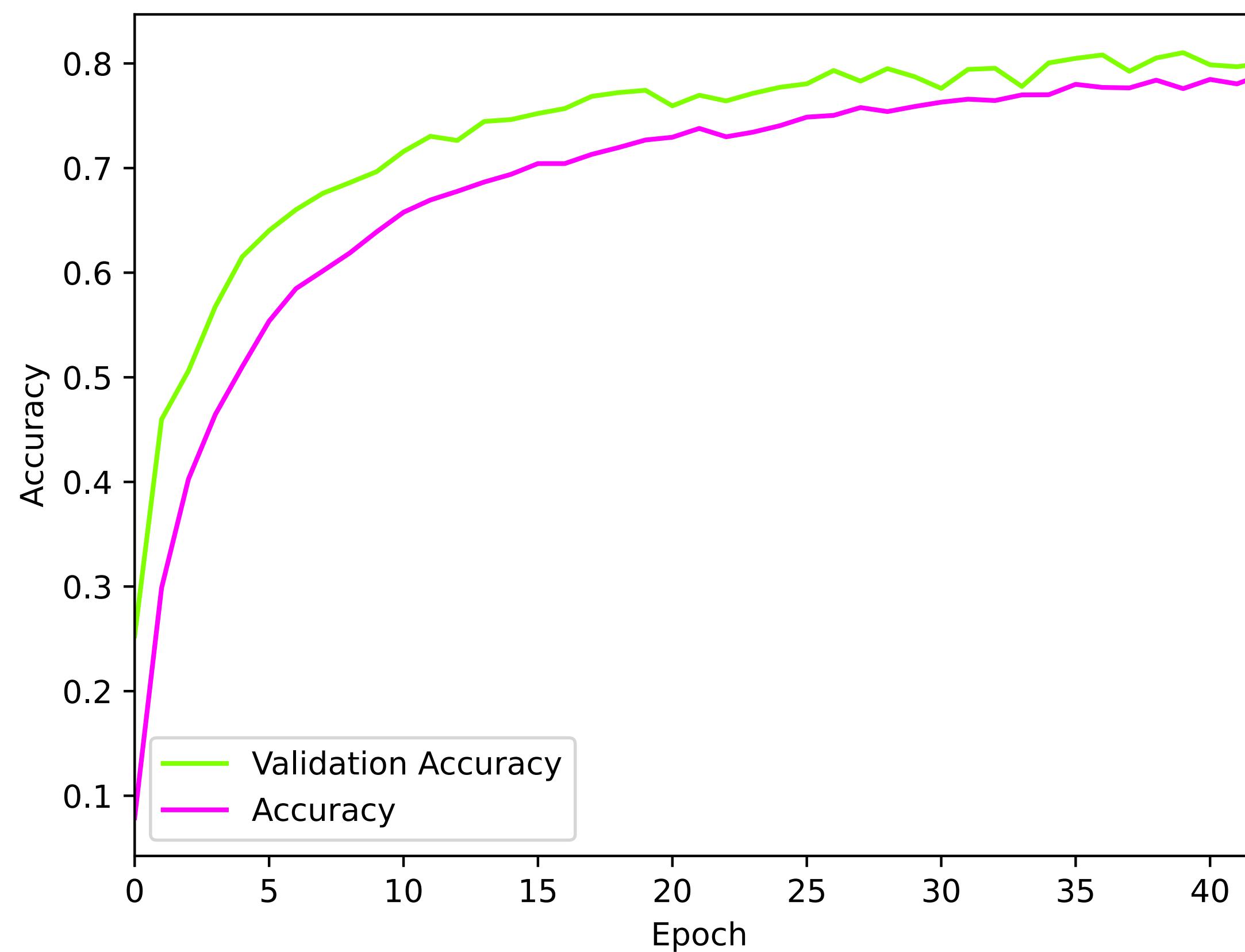
Hyperparameter Optimization Sweep #2

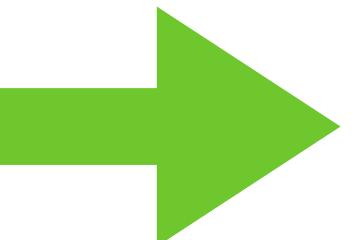


Hyperparameter Optimization Sweep #2



Results on Validation Data



 **Total Accuracy: 85.05 %**

Baseline: 3.57%

Alternative Method

- CNN most obvious and promising approach, kNN classifier also viable

Strategy

- Extract meaningful features from audio files
 - Using the spectrograms directly is unfeasible: $128 \times 162 > 20.000$
- Possible features
 - Zero crossing rate, most prominent frequencies

→ Compare accuracy to CNN's accuracy

