

Phase 1

Training Files Names

```
filename1/yes  
filename2/no  
_silence_  
...  
filename_n/stop
```

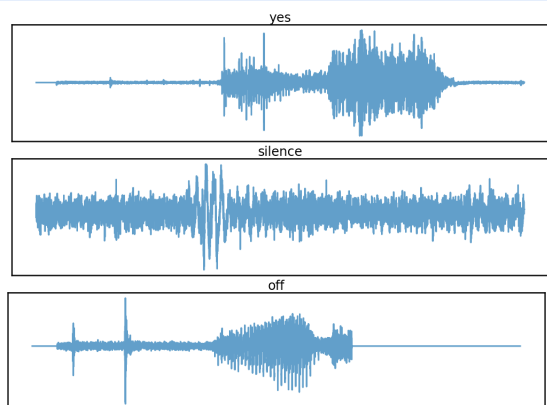
1. Decode .wav files
2. `_silence_` $\rightarrow [0, \dots, 0]$

CACHE

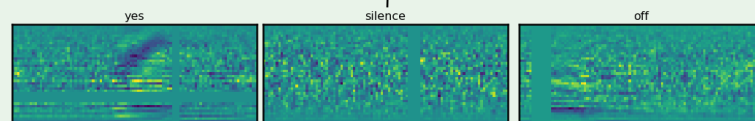
1. Random shift between $U(-0.1, 0.1)$
2. $[0, \dots, 0] \rightarrow$ random noise crop

**batch and
prefetch**

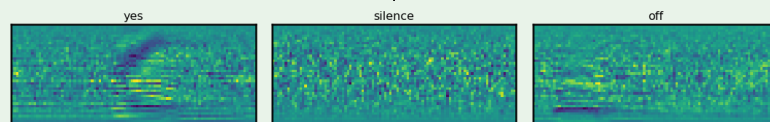
Data Batch



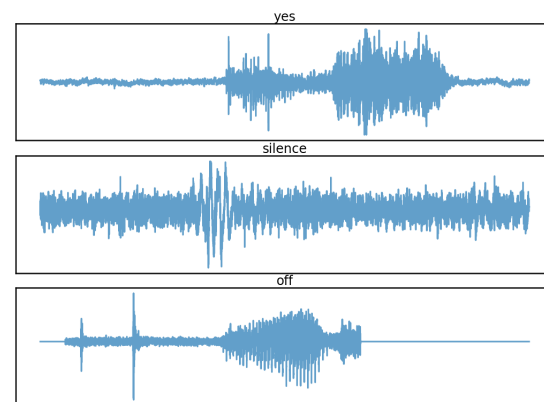
Model



SpecAugment Layer



MFCC Layer



RandomNoiseAugment Layer

Phase 2