

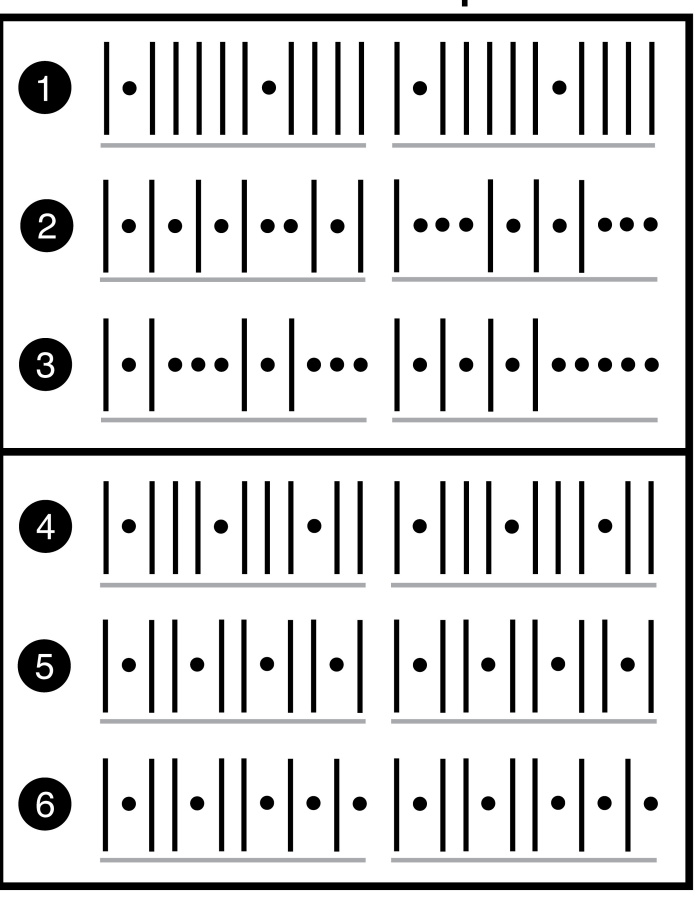
# Using Convolutional Neural Networks to Recognize Rhythm Stimuli from Electroencephalography Recordings

Sebastian Stober, Daniel J. Cameron, Jessica A. Grahn  
Brain and Mind Institute – Department of Psychology – Western University, London, ON, Canada

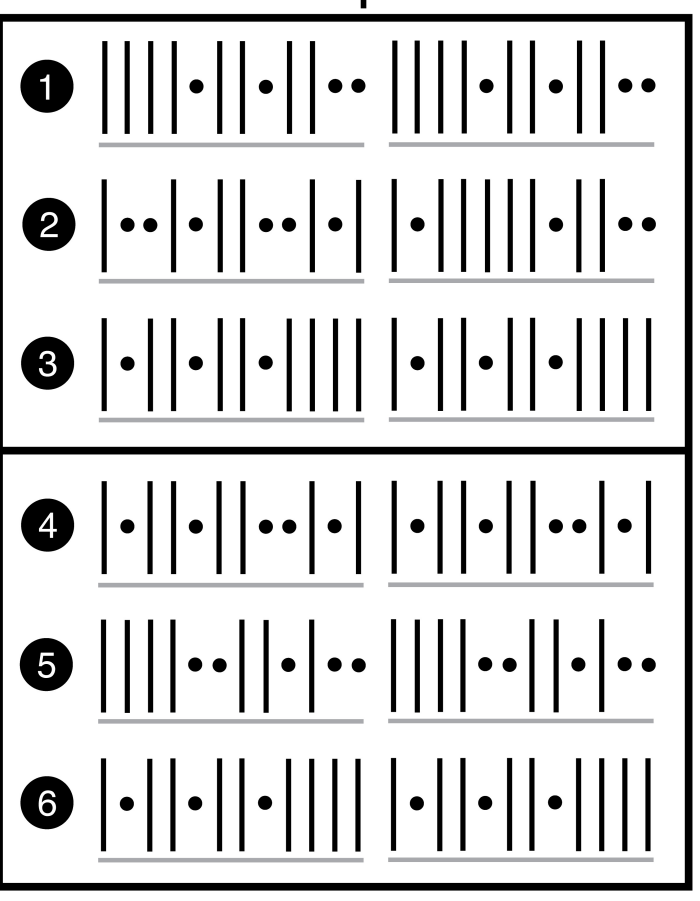
### Dataset: Rhythm Perception Study

- recorded 2013 in Kigali, Rwanda
- 13 participants (mean age 21)
- 14 EEG channels @ 400Hz
- 12 East African/12 Western rhythms
  - constructed from 4 groups of 3 sequences
  - 3 different pairs per group (12 pairs in total)
  - for each pair within each group, one played at 375Hz and the other at 500Hz (perfect fourth)
  - 24 stimuli = 4 groups \* 3 pairs \* 2 pitches
- presented for 32s each, 4s break
- same tempo within participants
- block design (randomized within)

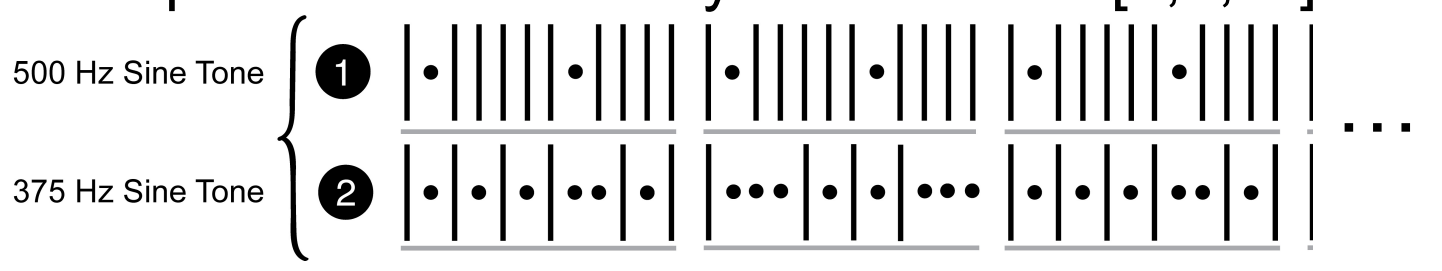
#### East African Sequences



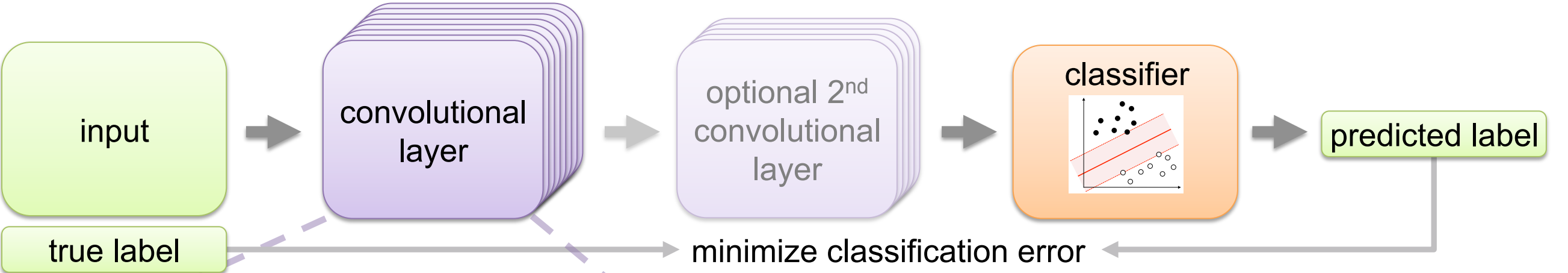
#### Western Sequences



#### Example: East African rhythm stimulus [1,2,'a']



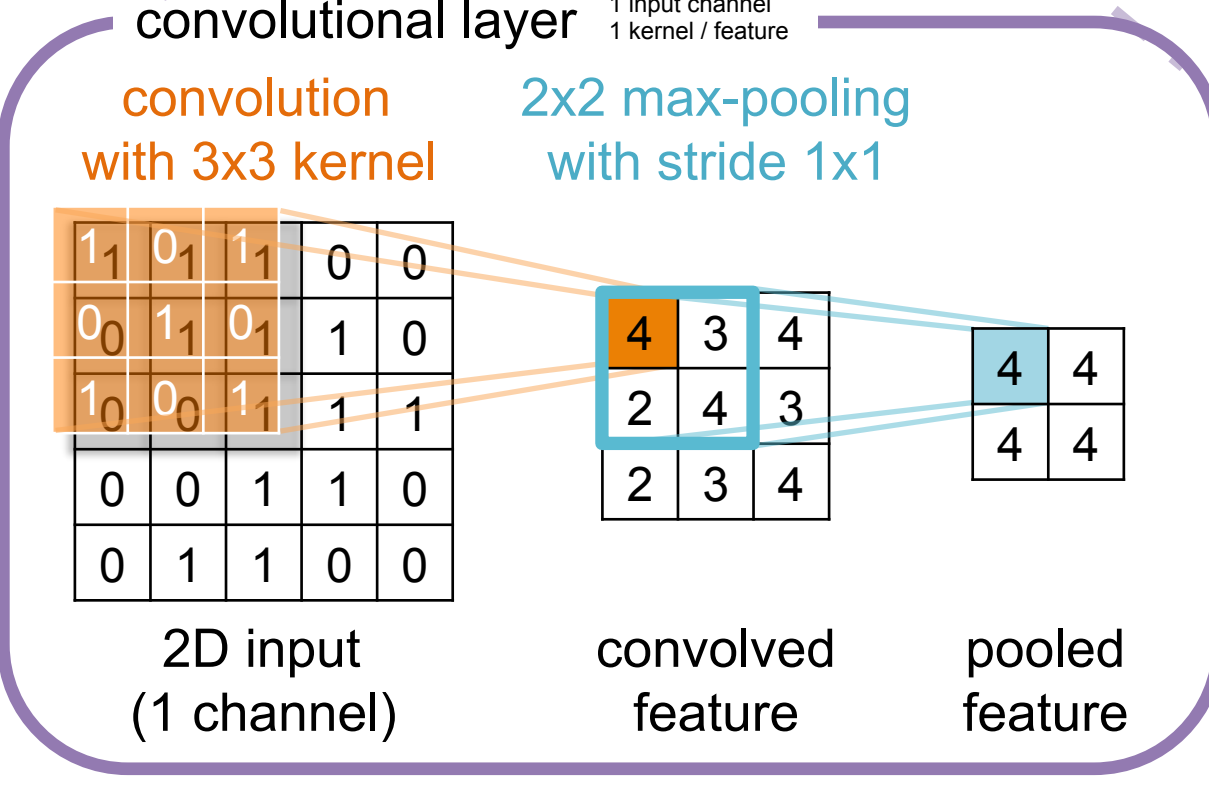
### Convolutional Neural Networks (CNNs)



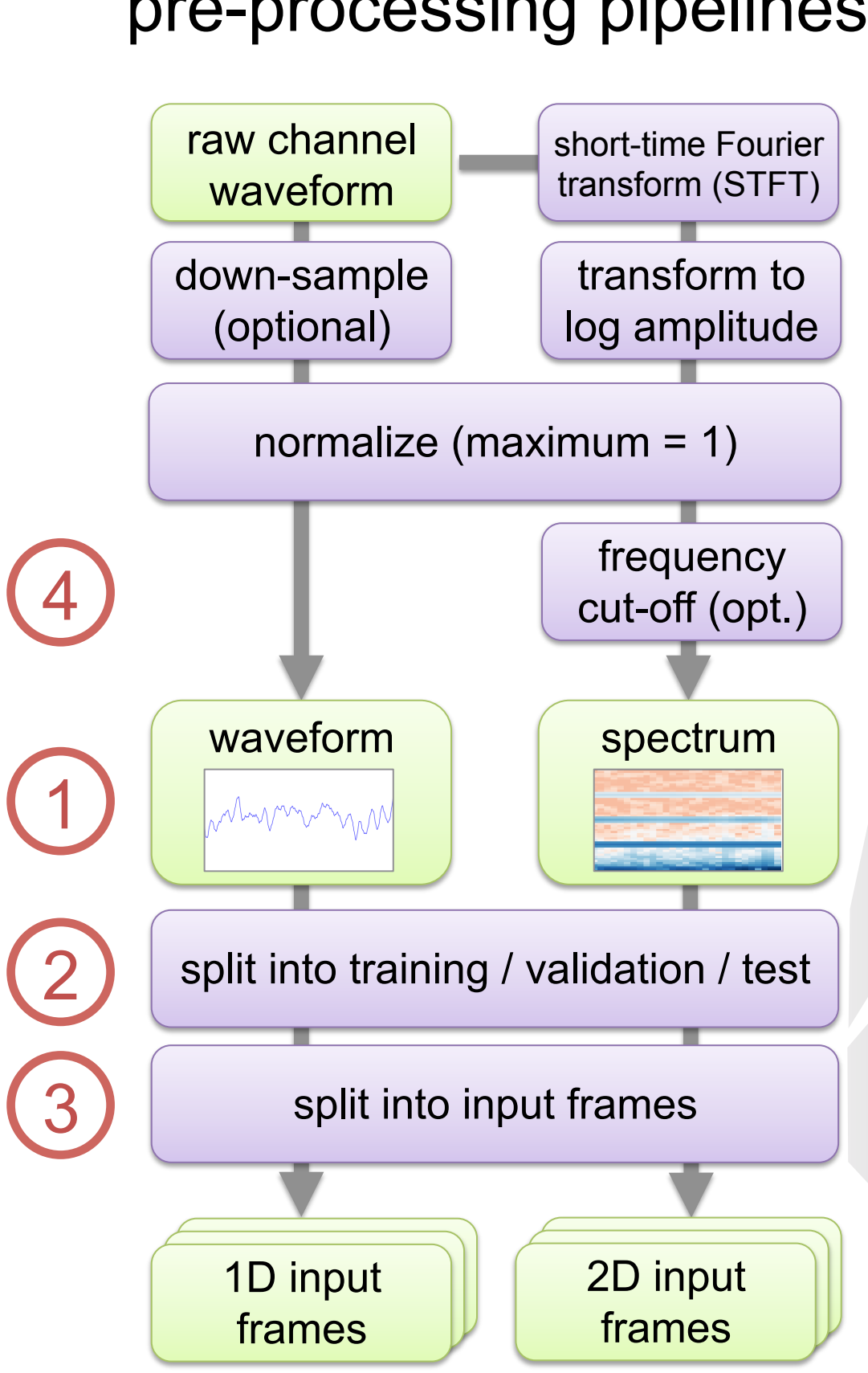
#### implementation details:

- DLSVM classifier (hinge loss)
- Rectified Linear Units (ReLU)
- Dropout regularization
- based on Pylean2 / Theano
- Bayesian optimization of hyper-parameters with Spearmint

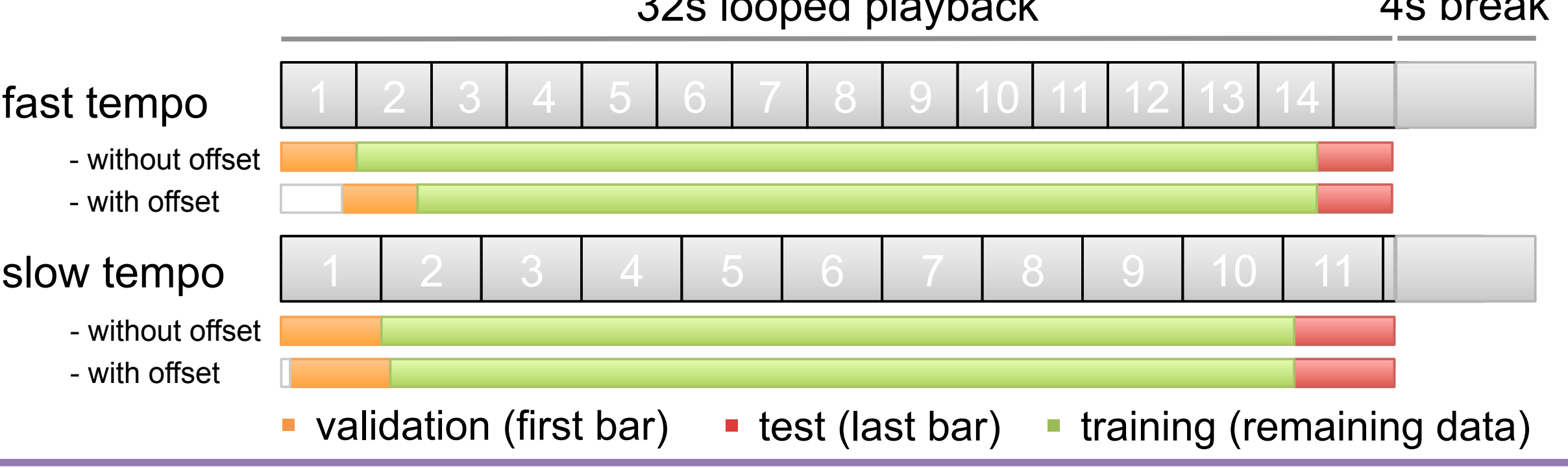
#### convolutional layer



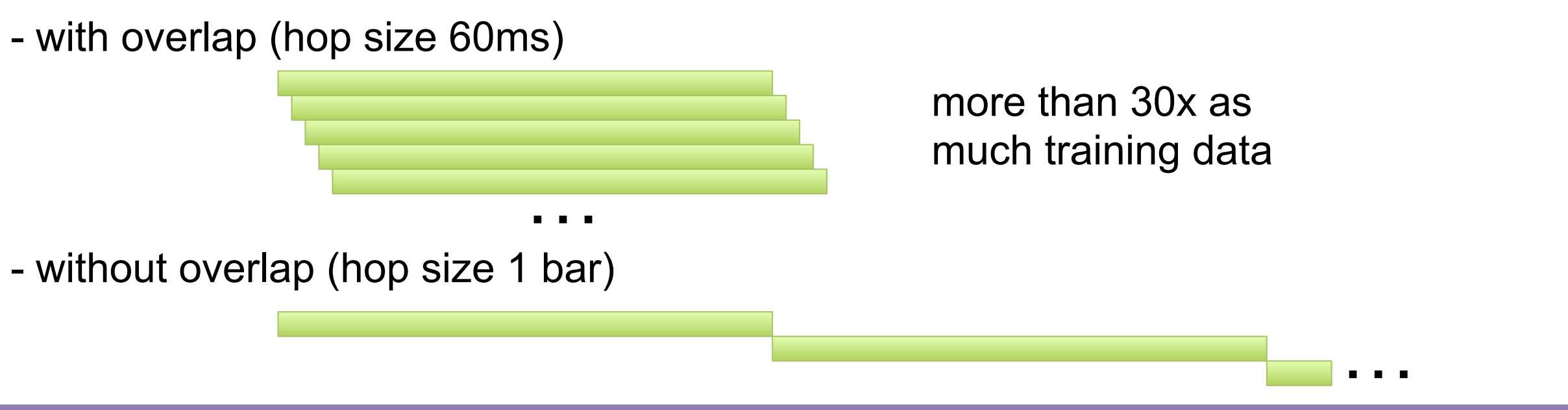
### Impact of Pre-Processing



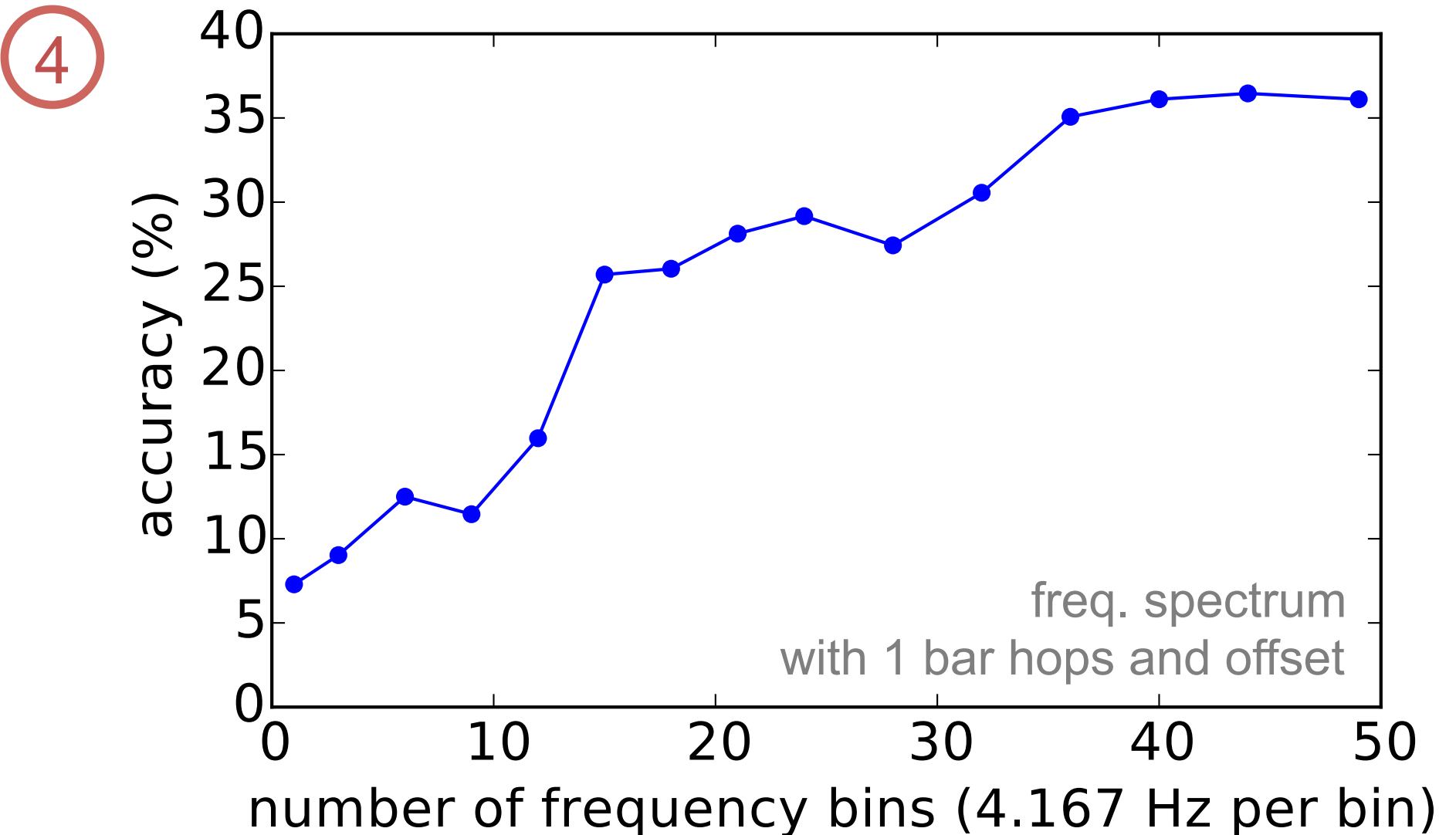
#### splitting each trial into training, validation and test data (with or without offset)



#### splitting the (training) data into input frames



#### accuracy (%) vs number of frequency bins (4.167 Hz per bin)

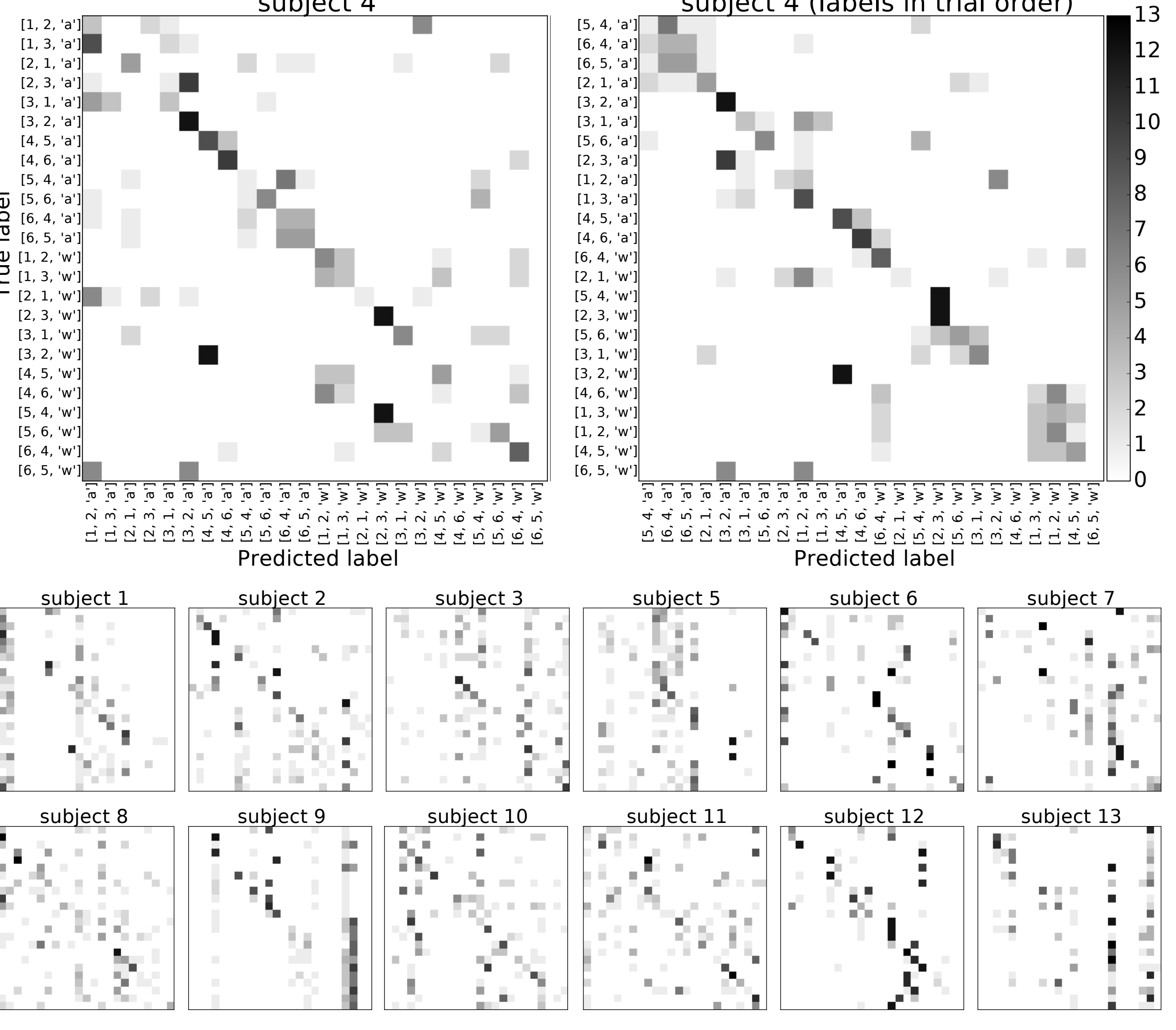


	hop size	offset	waveform	freq. spectrum
1	60ms	no	33.3%	233.7s
2	(60 runs)	yes	34.8%	119.5s
3	1 bar	no	33.0%	12.7s
4	(300 runs)	yes	24.7%	5.3s

### Individual and Aggregated Classification Results

subj.	input	network structure		single channel mean			aggregated trial accuracy			
		1st layer	2nd layer	24 classes accuracy	24 classes prec.@3	24 classes MRR	24 classes (stimuli)	12 classes (pairs)	4 classes (groups)	2 classes (types)
1*	33x49	[5x49]/3x16	[16x1]/5x12	19.1%	36.1%	0.34	25.0%	29.2%	58.3%	79.2%
2*	33x49	[10x49]/1x22		27.1%	46.5%	0.42	37.5%	37.5%	50.0%	87.5%
3*	33x49	[17x49]/1x30		21.9%	38.2%	0.36	20.8%	25.0%	45.8%	66.7%
4	45x49	[35x49]/1x30		36.1%	63.5%	0.55	50.0%	62.5%	75.0%	83.3%
5	45x49	[40x49]/2x30		18.1%	34.7%	0.33	16.7%	25.0%	41.7%	70.8%
6	45x49	[26x49]/5x30	[1x1]/10x30	29.5%	48.1%	0.45	37.5%	41.7%	54.2%	75.0%
7*	33x49	[15x49]/1x13		23.1%	43.9%	0.40	33.3%	45.8%	54.2%	66.7%
8*	33x49	[5x49]/2x21	[2x1]/2x24	24.0%	44.2%	0.41	41.7%	41.7%	58.3%	91.7%
9*	33x49	[13x49]/2x21	[6x1]/4x30	21.8%	33.7%	0.36	25.0%	29.2%	58.3%	91.7%
10	45x49	[7x49]/1x30		26.6%	51.0%	0.44	33.3%	33.3%	45.8%	66.7%
11	45x49	[27x49]/1x30		26.6%	55.1%	0.45	33.3%	37.5%	41.7%	75.0%
12	45x49	[5x49]/5x30	[5x1]/10x30	32.1%	60.9%	0.51	29.2%	33.3%	54.2%	83.3%
13	45x49	[18x49]/10x21	[1x1]/6x30	20.2%	37.2%	0.36	25.0%	29.2%	50.0%	70.8%
mean (1 convolutional layer)				24.4%	46.4%	0.41	30.8%	36.5%	51.6%	74.7%
mean (2 convolutional layers)				24.4%	44.2%	0.40	29.5%	34.0%	52.2%	77.2%
fast*	33x49	[8x49]/1x22		9.7%	22.1%	0.23	10.4%	16.7%	35.4%	66.7%
slow	45x49	[31x49]/1x30		9.9%	22.9%	0.24	10.7%	13.7%	32.7%	56.5%
all	33x49	[1x49]/1x30		7.3%	19.0%	0.21	7.7%	12.2%	29.2%	57.1%

#### confusion mostly between similar stimuli and consecutive trials



contact: [ssstober@uwo.ca](mailto:ssstober@uwo.ca)  
website: <http://www.uwo.ca/bmi/>  
supplementary material & code: <http://dx.doi.org/10.6084/m9.figshare.1213903>