Updated 14 Apr 18

	Params						Results (MSE)		
	Filters	Hidden units	Channels	Normalizer	Batch size	Epoch	Valence	Arousal	Average
1a b	3 x [(32,4,2)]	32	2	True	32	2	5.59/ 5.75 (5.67)	5.22/ 4.73 (4.98)	5.32
2	3 x [(32,4,2)]	32	2	True	64	2	8.19	6.89	7.54
3a b	3 x [(32,4,2)]	32	2	True	128	2	13.93/ 6.80 (10.37)	7.15/ 6.87 (7.01)	8.69
4a b	3 x [(32,4,2)]	32	2	True	32	1	6.44/ 10.51 (8.48)	12.40/ 8.77 (10.59)	9.53
5a b	3 x [(32,4,2)]	32	2	True	32	3	5.51/ 4.55 (5.03)	5.57/ 4.48 (5.03)	5.03
6a b	3 x [(32,4,2)]	32	2	True	32	4	11.36/ 5.60 (8.48)	6.54/ 8.95 (7.75)	8.11
7a b	1 x [(32,4,2)]	32	2	True	32	2	31.19/ 11.12 (21.16)	36.81/ 29.72 (33.27)	27.21
8a b	2 x [(32,4,2)]	32	2	True	32	2	11.68/ 6.70 (9.19)	9.94/ 5.57 (7.76)	8.47
9a b	4 x [(32,4,2)]	32	2	True	32	2	5.16/ 5.11 (5.14)	4.89/ 4.48 (4.68)	4.91
10a b	4 x [(64,8,2)]	32	2	True	32	2	6.25/ 6.67 (6.46)	6.19/ 7.90 (7.04)	6.75
11a b	3 x [(32,4,2)]	24	2	True	32	2	8.07/ 5.11 (6.59)	6.48/ 5.62 (6.05)	6.32
12a b	3 x [(32,4,2)]	16	2	True	32	2	5.72/ 4.94 (5.33	6.94/ 4.57 (5.76)	5.54
13a b	3 x [(32,4,2)]	32	1	True	32	2	5.85/ 5.67 (5.76)	5.43/4.44 (4.94)	5.35
14a b	3 x [(32,4,2)]	32	2	False	32	2	~~53.92~~/ 7.12/ 8.14/ 8.31	~~338.78~~/ 6.38/ 11.43/ 8.14	8.25

c d							(7.86)	(8.65)	
15a b c	4 x [(32,4,2)]	32	2	True	32	3	8.99/ 5.35/ 7.78 (7.37)	8.20/ 5.03/ 8.68 (7.30)	7.36
16a b	3 x [(32,4,2)]	32	4	False	64	3	19.04/ 8.01	15.69/ 11.19	13.48
17a b	3 x [(32,4,2)]	32	4	True	64	3	7.86/ 6.38	6.95/ 6.34	6.88
18a b	3 x [(32,4,2)]	32	3	True	64	3	7.36/ 10.63	5.94/ 5.67	7.40
19a b	3 x [(32,4,2)]	32	3	True	64	2	9.58/ 6.50	6.10/ 7.67	7.46
20a b	3 x [(32,4,2)]	32	4	True	64	2	6.22/ 10.05	6.21/ 6.58	7.27
21a b	3 x [(32,4,2)]	32	4	True	32	2	7.04/ 5.58	5.98/ 6.10	6.18
22a b	3 x [(32,4,2)]	32	3	True	32	2	5.36/ 6.60	5.78/ 9.22	6.74
23a b	4 x [(32,4,2)]	32	3	True	32	2	5.94/ 4.97	6.35/ 5.24	5.63
24a b	4 x [(32,4,2)]	32	4	True	32	2	5.72/ 5.45	7.79/ 5.08	6.01
25a b	4 x [(32,4,2)]	32	4	True	32	3	6.46/ 6.21	6.12/ 5.26	6.01

Notes

1) Optimal number of hidden units = 32
   - Compare #1 and #11, #12
   - Decreasing it increases MSE
2) Optimal number of convolution filters = 4 x (32,4,2)

- Compare #1 and #7, #8, #9
- Decreasing it increases MSE
- Increasing filter pulse and kernel size increases MSE

3) Normaliser required
   - Compare #1 and #14 or compare #16 and #17
   - Significant increase in MSE when set to false

4) Optimal batch size = 32
   - Compare #1 and #2, #3
   - Increasing batch size increases MSE

5) Optimal number of epoch = 3 with 3 x layers but = 2 with 4 x layers
   - 3 layers: Compare #1 and #4, #5, #6 or compare #24 and #25 (same results)
   - 4 layers: Compare #9, #15
   - Decreasing it increases MSE but increasing it also results in overfitting and MSE increases

6) Optimal number of channels = 2
   - Compare #1 (GSR+BVP) and #13 (GSR), #22 (GSR+BVP+zEMG), #21 (GSR+BVP+zEMG+tEMG)
   - Using GSR + BVP = smallest MSE
   - Adding in EMG did not improve the results, though using both EMG results produced better results than just using one set of EMG results. This could be due to correlation between the 2 sets of EMG results
   - Could be because GSR & BVP produce similar features so training with both data produces good results. But EMG is very different signals and may not correlate that well. Thus, adding it in resulted in non-distinct features and affected the results.

Red fonts: parameters changed from baseline
Green row: baseline parameters
Orange row: Parameters with results comparable to baseline
Blue row: Parameters with best results