

Best video-based model on SEWA dataset	Seconds, window size	Pooling	MAE (V)	RMSE (V)	MAE (A)	RMSE (A)
	2 sec, 10	AVG Pool	0.08	0.10	0.08	0.11
	3 sec, 15	AVG Pool	0.07	0.10	0.08	0.11
	4 sec, 20	AVG Pool	0.08	0.10	0.08	0.11

Method name	Seconds, window size	Pooling	MAE (V)	RMSE (V)	MAE (A)	RMSE (A)
Video-based model on SEWA dataset	1 sec, 5	AVG Pool	0.08	0.11	0.09	0.12
		MAX Pool	0.09	0.12	0.09	0.12
		1D CNN	0.95	0.96	0.95	0.96
	2 sec, 10	AVG Pool	0.08	0.10	0.08	0.11
		MAX Pool	0.08	0.11	0.09	0.11
		1D CNN	1.04	1.05	0.95	0.96
	3 sec, 15	AVG Pool	0.07	0.10	0.08	0.11
		MAX Pool	0.08	0.11	0.08	0.11
		1D CNN	1.04	1.05	1.04	1.05
	4 sec, 20	AVG Pool	0.08	0.10	0.08	0.11

		MAX Pool	0.09	0.12	0.08	0.11
		1D CNN	0.95	0.96	0.95	0.96

Best video-based model on AFEW-VA dataset	Seconds, window size	Pooling	MAE (V)	RMSE (V)	MAE (A)	RMSE (A)
	2 sec, 10	MAX Pool	0.07	0.11	0.06	0.07
	3 sec, 15	MAX Pool	0.07	0.11	0.06	0.08
	4 sec, 20	MAX Pool	0.07	0.10	0.06	0.08

Method name	Seconds, window size	Pooling	MAE (V)	RMSE (V)	MAE (A)	RMSE (A)
Video-based model on AFEW-VA dataset	1 sec, 5	AVG Pool	0.09	0.12	0.06	0.08
		MAX Pool	0.08	0.11	0.07	0.09
		1D CNN	0.77	0.82	1.19	1.22
	2 sec, 10	AVG Pool	0.09	0.13	0.07	0.09
		MAX Pool	0.07	0.11	0.06	0.07
		1D CNN	0.77	0.82	1.20	1.23

	3 sec, 15	AVG Pool	0.08	0.12	0.07	0.09
		MAX Pool	0.07	0.11	0.06	0.08
		1D CNN	0.77	0.82	0.79	0.84
	4 sec, 20	AVG Pool	0.09	0.12	0.06	0.08
		MAX Pool	0.07	0.10	0.06	0.08
		1D CNN	1.23	1.26	0.79	0.83

Best audio-based model on SEWA dataset	Seconds	MAE (V)	RMSE (V)	MAE (A)	RMSE (A)
	1 sec	0.13	0.16	0.13	0.17

Method name	Seconds	MAE (V)	RMSE (V)	MAE (A)	RMSE (A)
Audio-based model on SEWA dataset	1 sec	0.13	0.16	0.13	0.17
	2 sec	0.13	0.16	0.13	0.18
	3 sec	0.13	0.16	0.13	0.18
	4 sec	0.13	0.16	0.13	0.18

Algorithm (SEWA Audio)	Seconds	MSE (V)	RMSE (V)	MSE (A)	RMSE (A)
------------------------	---------	---------	----------	---------	----------

SVR	1 sec	0.02	0.16	0.03	0.18
	2 sec	0.02	0.16	0.03	0.18
	3 sec	0.02	0.16	0.03	0.18
	4 sec	0.02	0.16	0.03	0.18
Decision Tree	1 sec	0.05	0.23	0.05	0.24
	2 sec	0.04	0.22	0.05	0.24
	3 sec	0.04	0.21	0.04	0.22
	4 sec	0.04	0.22	0.04	0.22

Best video-based model on SEWA dataset	Seconds, window size	Pooling	MAE (V)	RMSE (V)	MAE (A)	RMSE (A)
	2 sec, 10	AVG Pool	0.08	0.10	0.08	0.11
	3 sec, 15	AVG Pool	0.07	0.10	0.08	0.11
	4 sec, 20	AVG Pool	0.08	0.10	0.08	0.11

Best Fusion Models on test data	Seconds, window size	Pooling	MAE (V)	RMSE (V)	MAE (A)	RMSE (A)
Fusion 1	2 sec, 10	AVG Pool 2x	0.0893	0.1145	0.0930	0.1183

Fusion 2	4 sec, 20	AVG Pool	0.0902	0.1174	0.0988	0.1284
Fusion 3	1 sec, 5	AVG Pool	0.0928	0.1206	0.1025	0.1355
Fusion 4 (v2)	4 sec, 20	1D CNN, AVG Pool	0.0815	0.1035	0.0956	0.1239

Best Fusion Models on dev data	Seconds, window size	Pooling	MAE (V)	RMSE (V)	MAE (A)	RMSE (A)
Fusion 1	2 sec, 10	AVG Pool 2x	0.0606	0.0830	0.0689	0.0972
Fusion 2	4 sec, 20	AVG Pool	0.0601	0.0892	0.0639	0.0904
Fusion 3	1 sec, 5	AVG Pool	0.0606	0.0822	0.0709	0.0985
Fusion 4 (v2)	4 sec, 20	1D CNN, AVG Pool	0.0583	0.0809	0.0679	0.0940

Method name (dev)	Seconds, window size	Pooling	MAE (V)	RMSE (V)	MAE (A)	RMSE (A)
Fusion 1	1 sec, 5	AVG Pool 2x	0.0697	0.0811	0.0689	0.0958
Fusion 1	2 sec, 10	AVG Pool 2x	0.0606	0.0830	0.0689	0.0972
Fusion 1	3 sec, 15	AVG Pool 2x	0.0626	0.0857	0.0718	0.1009
Fusion 1	4 sec, 20	AVG Pool 2x	0.0642	0.0888	0.0718	0.1010

Fusion 2	1 sec, 5	AVG Pool	0.0694	0.0814	0.0666	0.0943
Fusion 2	2 sec, 10	AVG Pool	0.0602	0.0820	0.0667	0.0940
Fusion 2	3 sec, 15	AVG Pool	0.0602	0.0816	0.0661	0.0924
Fusion 2	4 sec, 20	AVG Pool	0.0601	0.0892	0.0639	0.0904
Fusion 3	1 sec, 5	AVG Pool	0.0606	0.0822	0.0709	0.0985
Fusion 3	2 sec, 10	AVG Pool	0.0619	0.0848	0.0727	0.1012
Fusion 3	3 sec, 15	AVG Pool	0.0627	0.0848	0.0712	0.0992
Fusion 3	4 sec, 20	AVG Pool	0.0641	0.0877	0.0715	0.1009
Fusion 4 (v1)	1 sec, 5	1D CNN, AVG Pool	0.0900	0.1163	0.1026	0.1341
Fusion 4 (v1)	2 sec, 10	1D CNN, AVG Pool	0.0666	0.0898	0.0735	0.1037
Fusion 4 (v1)	3 sec, 15	1D CNN, AVG Pool	0.0661	0.0881	0.0740	0.1028
Fusion 4 (v1)	4 sec, 20	1D CNN, AVG Pool	0.0663	0.0877	0.0717	0.0992
Fusion 4 (v2)	1 sec, 5	1D CNN, AVG Pool	0.0588	0.0813	0.0696	0.0987
Fusion 4 (v2)	2 sec, 10	1D CNN, AVG Pool	0.0600	0.0826	0.0683	0.0984
Fusion 4 (v2)	3 sec, 15	1D CNN, AVG Pool	0.0594	0.0822	0.0679	0.0944
Fusion 4 (v2)	4 sec, 20	1D CNN, AVG Pool	0.0583	0.0809	0.0672	0.0940
Fusion 4 (v3)	1 sec, 5	1D CNN, AVG Pool	0.0906	0.1166	0.1023	0.1337

Fusion 4 (v3)	2 sec, 10	1D CNN, AVG Pool	0.0910	0.1171	0.1001	0.1326
Fusion 4 (v3)	3 sec, 15	1D CNN, AVG Pool	0.0922	0.1195	0.1022	0.1341
Fusion 4 (v3)	4 sec, 20	1D CNN, AVG Pool	0.0908	0.1193	0.1034	0.1363

Method name (test)	Seconds, window size	Pooling	MAE (V)	RMSE (V)	MAE (A)	RMSE (A)
Fusion 1	1 sec, 5	AVG Pool 2x	0.0876	0.1113	0.0944	0.1213
Fusion 1	2 sec, 10	AVG Pool 2x	0.0893	0.1145	0.0930	0.1183
Fusion 1	3 sec, 15	AVG Pool 2x	0.0923	0.1174	0.0975	0.1246
Fusion 1	4 sec, 20	AVG Pool 2x	0.0941	0.1197	0.0999	0.1297
Fusion 2	1 sec, 5	AVG Pool	0.0940	0.1232	0.1021	0.1289
Fusion 2	2 sec, 10	AVG Pool	0.0923	0.1201	0.1033	0.1327
Fusion 2	3 sec, 15	AVG Pool	0.0914	0.1193	0.1013	0.1292
Fusion 2	4 sec, 20	AVG Pool	0.0902	0.1174	0.0988	0.1284
Fusion 3	1 sec, 5	AVG Pool	0.0928	0.1206	0.1025	0.1355
Fusion 3	2 sec, 10	AVG Pool	0.1058	0.1401	0.1034	0.1328
Fusion 3	3 sec, 15	AVG Pool	0.0925	0.1203	0.1092	0.1470

Fusion 3	4 sec, 20	AVG Pool	0.1129	0.1494	0.1057	0.1350
Fusion 4 (v1)	1 sec, 5	1D CNN, AVG Pool	0.1371	0.1835	0.1336	0.1682
Fusion 4 (v1)	2 sec, 10	1D CNN, AVG Pool	0.1151	0.1533	0.1064	0.1346
Fusion 4 (v1)	3 sec, 15	1D CNN, AVG Pool	0.1165	0.1548	0.1097	0.1391
Fusion 4 (v1)	4 sec, 20	1D CNN, AVG Pool	0.1196	0.1589	0.1042	0.1346
Fusion 4 (v2)	1 sec, 5	1D CNN, AVG Pool	0.0861	0.1106	0.0932	0.1208
Fusion 4 (v2)	2 sec, 10	1D CNN, AVG Pool	0.0867	0.1098	0.0937	0.1210
Fusion 4 (v2)	3 sec, 15	1D CNN, AVG Pool	0.0827	0.1050	0.0976	0.1261
Fusion 4 (v2)	4 sec, 20	1D CNN, AVG Pool	0.0815	0.1035	0.0956	0.1239
Fusion 4 (v3)	1 sec, 5	1D CNN, AVG Pool	0.1379	0.1852	0.1337	0.1683
Fusion 4 (v3)	2 sec, 10	1D CNN, AVG Pool	0.1385	0.1855	0.1341	0.1678
Fusion 4 (v3)	3 sec, 15	1D CNN, AVG Pool	0.1379	0.1842	0.1374	0.1713
Fusion 4 (v3)	4 sec, 20	1D CNN, AVG Pool	0.1377	0.1836	0.1368	0.1719