

# Protocol for Experiments on SEP-28k Dataset

## A. SEP-28k Dataset

In this case study, we used SEP-28k stuttering dataset [1] which consists of 28,177 speech samples from 385 podcasts. After removing non-stuttered samples, we are left with 23573 annotated speech segments. We randomly selected 80% podcasts (without mixing podcasts) for training, 10% podcasts for validation and the remaining 10% of the podcast for evaluation in a 10-fold cross validation scheme. The speaker information is missing from the SEP-28k dataset, so we divided the dataset based on podcast ids (assuming each podcast is having unique speaker). This dataset contains two different types of annotations including stuttering and non-stuttering. We considered only stuttering annotations (repetitions, blocks, interjections, prolongations, and fluent speech) and avoided the non-stuttering annotations (unsure, no speech, poor audio quality, music, unintelligible and natural pauses) in this experimental study. The experimental results mentioned are the average of 10-fold cross validation experiments, and are compared to the baseline results from [2] which is trained on MFCC features.

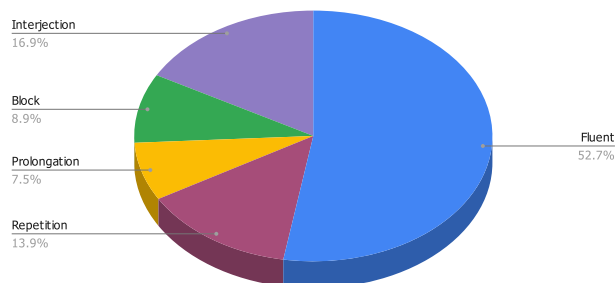


Fig. 1. SEP-28k Dataset Distribution

## REFERENCES

- [1] C. Lea *et al.*, “Sep-28k: A dataset for stuttering event detection from podcasts with people who stutter,” in *Proc. ICASSP*, 2021, pp. 6798–6802.
- [2] S. A. Sheikh *et al.*, “Stutternet: Stuttering detection using time delay neural network,” in *Proc. 29th EUSIPCO*, 2021, pp. 426–430.

TABLE I  
SEP-28K DATASET STATISTICS ( B: BLOCK , F: FLUENT , R: REPETITION , P: PROLONGATION , I: INTERJECTION, ID: SETID)

Distribution						
ID	R	P	B	I	F	Count
Train						
1	2681	1384	1726	3181	9950	18922
2	2696	1495	1758	3274	10356	19579
3	2652	1481	1761	3202	10301	19397
4	2676	1518	1753	3305	10293	19545
5	2513	1430	1679	3115	9832	18569
6	2745	1457	1716	3119	9806	18843
7	2805	1437	1718	3265	10174	19399
8	2491	1439	1666	3085	9914	18595
9	2409	1412	1701	3179	10091	18792
10	2556	1348	1680	3082	9505	18171
Val						
1	331	277	245	486	1466	2805
2	314	157	182	349	1097	2099
3	287	133	154	349	973	1896
4	262	125	149	368	1081	1985
5	363	154	230	378	1241	2366
6	241	176	160	361	1287	2225
7	284	163	192	414	1139	2192
8	362	168	224	390	1200	2344
9	402	141	176	338	938	1995
10	396	265	228	488	1478	2855
Test						
1	274	109	132	328	1003	1846
2	276	118	163	372	966	1895
3	347	156	188	444	1145	2280
4	348	127	201	322	1045	2043
5	410	186	194	502	1346	2638
6	300	137	227	515	1326	2505
7	197	170	193	316	1106	1982
8	433	163	213	520	1305	2634
9	475	215	226	478	1390	2784
10	334	157	195	425	1436	2547