Transfer Learning to Aid Dysarthria Severity Classification for Patients with Amyotrophic Lateral Sclerosis

T. Bhattacharjee[1], A. Jayakumar[1], Y. Belur[2], A. Nalini[2], R. Yadav[2], P.K. Ghosh[1]

[1]SPIRE Lab, EE Department, Indian Institute of Science, Bengaluru, India

[2]National Institute of Mental Health and Neurosciences, Bengaluru, India

Dysarthria in ALS

- **Dysarthria** due to **Amyotrophic Lateral Sclerosis (ALS)** critically impairs speech production.
- **Regular monitoring** of the disease condition is essential for effective disease management.
- Speech-Language Pathologists (SLPs) assess **dysarthria severity of an ALS patient** following the **speech measure of ALSFRS-R scale**.

Condition	Score
Normal speech processes	4
Detectable speech disturbance	3
Intelligible with repeating	2
Speech needs to be combined with nonvocal communication	1
Loss of useful speech	0

Drawbacks:

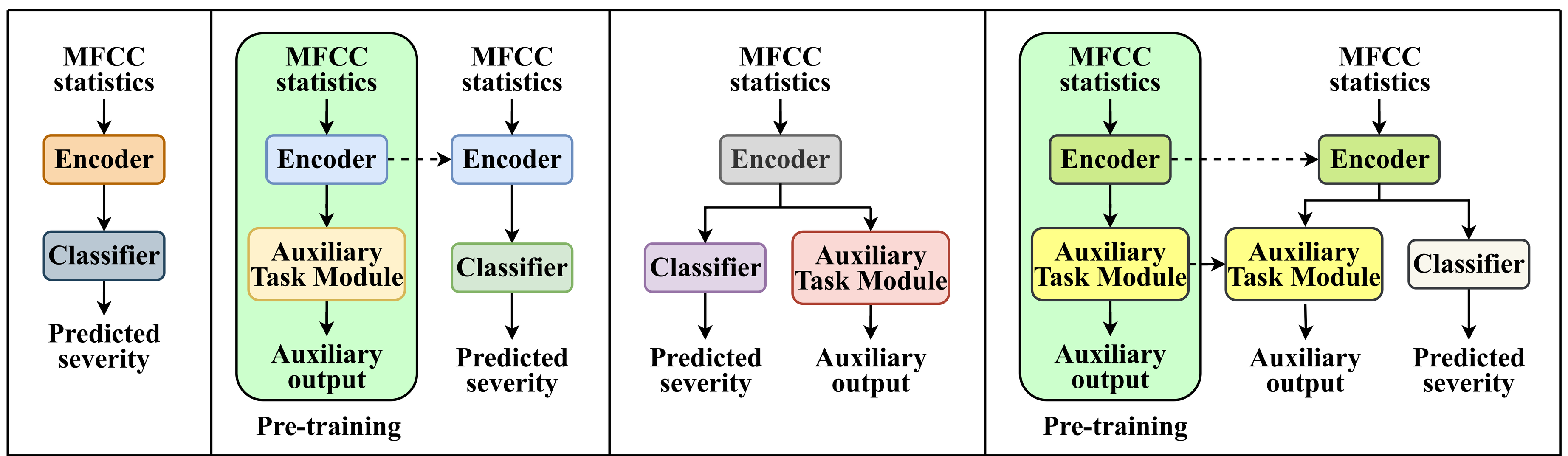
- Tedious and highly time-consuming
- Prone to subjective biases

Accurate and consistent automatic dysarthria severity prediction systems are the need of the hour.

State Of the Art

- Speech-based automatic methods are primarily restricted to the classification of ALS patients and Healthy Controls (HC).
- Only a few efforts have been reported in the domain of speech-based automatic dysarthria severity prediction for ALS.
- **Major Challenge** - Scarcity of data resources
  - Collecting speech data from patients with speech impairments is a delicate and laborious task.
  - Getting the collected data clinically annotated for dysarthria severity further adds to the difficulty.
- **Transfer learning** approaches have been explored for severity classification of dysarthria specific to Cerebral Palsy and Parkinson's Disease but **not ALS**.

Proposed Transfer Learning Approaches



(a) Single Task Direct Learning (STDL) | (b) Fine-Tuning (FT) | (c) Multi-Task Learning (MTL) | (d) Multi-Task Learning with Pre-training (MTLp)

	Task	Loss
Primary	**3-class (Normal vs. Mild vs. Severe) dysarthria severity classification** **Normal (N) :** ALSFRS-R 4 **Mild (M) :** ALSFRS-R 2-3 **Severe (S) :** ALSFRS-R 0-1	Cross-entropy
Auxiliary	MFCC Feature reconstruction (FR)	MSE
	Gender classification (GC)	Cross-entropy

- Transfer learning is performed with and without using auxiliary healthy datasets.
- For MTLp, two further sub-conditions are considered.

Condition	Pre-training	Network adaptation
MTLp1	ALS/auxiliary data	ALS data
MTLp2	auxiliary data	ALS + auxiliary data

- All Encoder, Classifier and Auxiliary Task Modules are implemented using Dense neural networks.

Acknowledgement: We thank Seena Vengalil and M S Keertipriya for their valuable assistance during data collection. We also thank the Department of Science and Technology (DST), Govt. of India for supporting this work.

Dataset

- All in-house data collections were performed at **NIMHANS, Bengaluru, India**.
- The mode of the dysarthria severity ratings given by three SLPs was considered as the final severity score.

Subject demography and recorded speech data duration

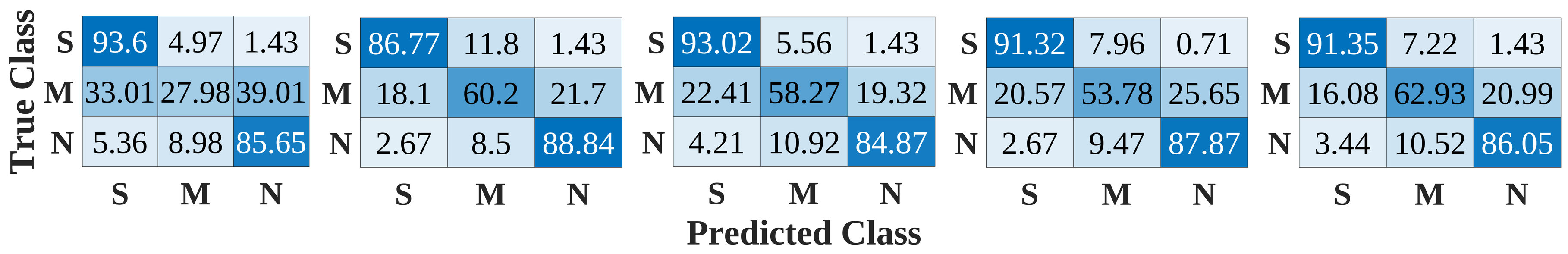
Dataset	ALS data					Auxiliary data		
						HC data	Indic TIMIT	TIMIT
Severity class	**Severe (S)**		**Mild (M)**		**Normal (N)**	**Healthy**	**Healthy**	**Healthy**
ALSFRS-R	**0**	**1**	**2**	**3**	**4**	**-**	**-**	**-**
#M:#F	9:13	12:6	15:5	11:9	27:13	67:21	39:41	438:192
Mean (SD) of age (years)	58.55 (1.14)	56.63 (1.20)	51.10 (1.08)	54.45 (1.04)	52.28 (0.76)	43.02 (9.13)	25.42 (6.05)	29.78 (8.09)
Speech duration (hours)	0.53	0.61	0.70	0.66	1.43	2.90	234.47	5.38
Speech task	Spontaneous					Spontaneous	Read	Read
Language	Bengali, Hindi, Tamil, Telugu, Kannada					Bengali, Hindi, Tamil, Telugu, Kannada	Indian English	American English

Results

Mean balanced classification accuracies in % (SD in bracket) obtained over 10-folds of random validation using different network training schemes; here, * indicates the approaches which outperform STDL at 1% significance level and # indicates that FR outperforms GC as the auxiliary task at 1% significance level (Wilcoxon signed-rank test is performed for all comparisons)

Auxiliary data	Auxiliary task	STDL			
-	-	69.08 (3.66)			
		FT	**MTL**	**MTLp1**	**MTLp2**
-	FR	77.14 (6.53)*	75.50 (3.91)*	**77.66 (3.47)***	-
	GC	74.30 (5.82)	75.44 (5.46)	73.17 (6.32)	-
HC data	FR	76.82 (4.98)*	74.88 (5.61)	76.28 (4.47)*	76.56 (6.37)
	GC	74.58 (4.39)*	73.70 (3.69)*	74.23 (7.16)	74.41 (4.78)*
Indic TIMIT	FR	**78.60 (6.52)***	75.88 (4.68)*	77.38 (3.75)*	75.41 (3.96)#*
	GC	71.22 (6.59)	75.45 (4.58)	71.56 (4.38)	71.02 (5.79)
TIMIT	FR	75.75 (5.34)*	**78.72 (6.89)***	75.75 (6.79)*	**80.11 (3.80)***
	GC	77.19 (4.26)*	77.52 (5.51)*	75.34 (3.66)*	76.60 (5.48)*

Confusion matrices (in %) averaged over 10-folds of random validation for STDL and the best performing configurations of the transfer learning approaches



(a) STDL (b) FT (c) MTL (d) MTLp1 (e) MTLp2

Conclusions

- All transfer learning schemes achieve **higher mean accuracies than STDL**.
- Transfer learning approaches significantly **improve** the performance on **classifying the mild class**.
- Average accuracies achieved using **feature reconstruction** tasks are **higher** than those obtained using **gender classification** tasks in almost all cases.
- Performances obtained **with or without** employing the **auxiliary datasets** are **statistically similar**.
- For any configuration of auxiliary task and dataset, the performances of all the **four transfer learning approaches** are found to be **statistically similar**.

Future Work

- To explore wider varieties of auxiliary tasks and network architectures
- To perform 5-class dysarthria severity classification

References

1. J. M. Cedarbaum et al., "The ALSFRS-R: a revised ALS functional rating scale that incorporates assessments of respiratory function," Journal of the neurological sciences, vol. 169, no. 1-2, pp. 13–21, 1999.
2. S. BN, J. Mallela, A. Illa, Y. BK, N. Atchayaram, R. Yadav, D. Gope, and P. K. Ghosh, "Speech task based automatic classification of ALS and Parkinson's disease and their severity using log mel spectrograms," in International Conference on Signal Processing and Communications (SPCOM), IEEE, 2020, pp. 1–5.
3. A. Wisler, K. Teplansky, J. R. Green, Y. Yunusova, T. Campbell, D. Heitzman, and J. Wang, "Speech-based estimation of bulbar regression in Amyotrophic Lateral Sclerosis," in Proceedings of the Eighth Workshop on Speech and Language Processing for Assistive Technologies, 2019, pp. 24–31.