



Overcoming Data Sparsity in Acoustic Modeling of Low-Resource Language by Borrowing Data and Model Parameters from High-Resource Languages

Basil Abraham, S. Umesh, Neethu Mariam Joy

Indian Institute of Technology-Madras, India

{ee11d032, umeshs, ee11d009}@ee.iitm.ac.in

Abstract

In this paper, we propose two techniques to improve the acoustic model of a low-resource language by: (i) Pooling data from closely related languages using a phoneme mapping algorithm to build acoustic models like subspace Gaussian mixture model (SGMM), phone cluster adaptive training (Phone-CAT), deep neural network (DNN) and convolutional neural network (CNN). Using the low-resource language data, we then adapt the afore mentioned models towards that language. (ii) Using models built from high-resource languages, we first borrow subspace model parameters from SGMM/Phone-CAT; or hidden layers from DNN/CNN. The language specific parameters are then estimated using the low-resource language data. The experiments were performed on four Indian languages namely Assamese, Bengali, Hindi and Tamil. Relative improvements of 10 to 30% were obtained over corresponding monolingual models in each case.

Index Terms: speech recognition, low-resource, cross-lingual, data pooling, CNN, DNN

1. Introduction

With recent advancements in deep neural networks (DNN), there has been a significant improvement in the performance of speech recognizers. However such robust systems are mainly limited to popular languages like English, French etc. To build a robust speech recognizer in any language, large amount of transcribed speech data is required. In many of the under resourced languages like African and Indian languages, data sparsity is a critical problem in building good speech recognizers. In this paper, we try to overcome the afore mentioned problem by borrowing resources from other languages which has adequate transcribed speech data. Throughout this paper, we use the term *high-resource* language to refer to a language having abundant resources in terms of transcribed training data and *low-resource* language to the ones with limited training data.

Two broad approaches to handle data sparsity are cross-lingual acoustic modeling techniques [1, 2, 3, 4] and multilingual acoustic modeling techniques [5, 6]. In cross-lingual acoustic modeling, the model parameters from the high-resource acoustic models are borrowed and the language specific parameters are re-trained using the low-resource language data. In multilingual acoustic modeling, a common acoustic model is built using both low-resource and high-resource data with or without sharing the language specific parameters. However, in this case, acoustic model sharing among different languages becomes difficult due to difference in phone set across languages. Using a global phone set [7], or a knowledge based / data driven phone mapping [8, 9, 10] were

proposed to solve this issue. The recently proposed multilingual DNN [5, 11, 6] has shared hidden layers and individual softmax layers for each language. This was inspired from multilingual subspace Gaussian mixture model (SGMM) [12]. Yet another approach is to use cross-lingual tandem features [13, 14, 15, 16] extracted from the bottle-neck layer of a multi layer perceptron (MLP) trained on one or more high-resource languages.

In this paper, we propose two techniques to overcome the data sparsity problem. First, the problem of data insufficiency is addressed by pooling data from closely related languages. The pooled data along with the data from low-resource language is used for building acoustic models like continuous density hidden Markov model (CDHMM), SGMM [17], phone cluster adaptive training (Phone-CAT) [18], DNN [19] and convolutional neural network (CNN) [20]. To achieve further improvements, the data pooled models are then adapted towards the low-resource language. In the second approach, acoustic model parameters are borrowed from models (SGMM, Phone-CAT, DNN and CNN) built with high-resource language and are then further refined using low-resource data.

The experiments are performed with four Indian languages namely, Assamese, Bengali, Hindi and Tamil from the MANDI database. Both data pooling and cross-lingual model borrowing gave significantly improved performance over the baseline mono-lingual acoustic models built just with the low-resource language. The paper is organized as follows. The proposed methods are described in section 2. Details of the various experiments performed and their results are given in section 3. Conclusions are drawn in section 4.

2. Cross-lingual Acoustic modeling

In this section, we discuss the two cross-lingual techniques used in this paper, namely data pooling and borrowing model parameter.

2.1. Borrowing Data or Pooling Data

In many of the previous works [21, 22], the use of untranscribed data from the same language to improve the performance of the acoustic model in a low-resource language were studied. However, the use of transcribed data from closely related languages were not studied in detail. In this section we pool the data from a high resource language along with low resource language to increase the amount of training data. The acoustic model is built by using the phoneme-mapping obtained in [23]. The phoneme mapping is performed from low resource language to the high-resource language phones. In the preliminary work reported in [23], complete high-resource data were pooled with the low-resource language which resulted in poor performance after data pooling. In this section, we

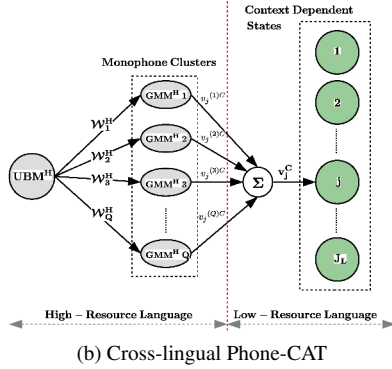
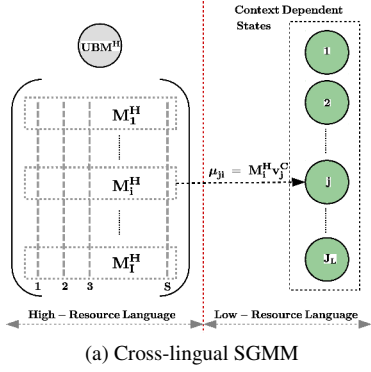


Figure 1: Cross-lingual Models

study the effect of the amount of data pooled from high-resource language on the performance. The performance of the data pooling will depend on the phonetic overlap between the languages as well as the similarity in environment noise.

2.1.1. Adaptation of Data Pooled Models

In the case of pooled model the phonetic tree is generated from the pooled data rather than from the low-resource language alone to have a richer tree. We then adapt this model to the low-resource data by retraining the language specific parameters using the low-resource data or training the output layer.

2.2. Borrowing Acoustic Models

The cross-lingual acoustic modeling techniques where acoustic model parameters trained with high-resource language are borrowed for improving the low-resource languages are described in [1, 2]. In this section the cross-lingual acoustic modeling using parsimonious models like SGMM, Phone-CAT and neural network models like DNN and CNN are discussed. The use of SGMM and Phone-CAT for cross-lingual modeling are described in [1, 23]. The block diagram representation of cross-lingual SGMM and cross-lingual-PhoneCAT are given in Figures 1a and 1b. In both SGMM and Phone-CAT, the acoustic parameters can be separated into global or language independent parameters and state specific or language dependent parameters. In SGMM, the universal background model (UBM), subspace projection matrices and the weight subspace forms the global parameters and the state specific vectors forms the language specific parameter. In Phone-CAT, the UBM, the maximum likelihood linear regression (MLLR) adaptation matrices and the phone Gaussian mixture model (GMM) clusters forms the global parameters and

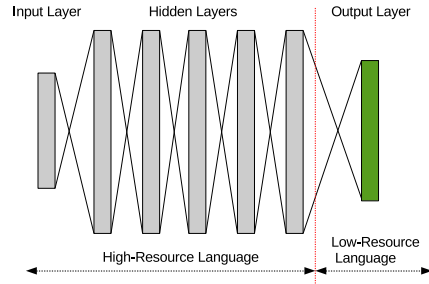
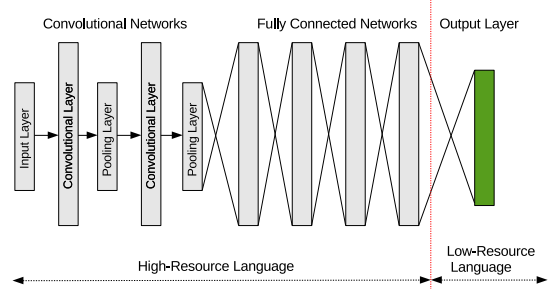
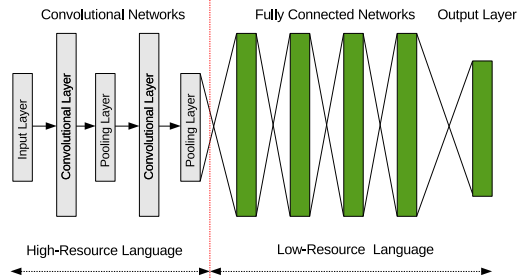


Figure 2: Cross-lingual DNN



(a) Setup 1



(b) Setup 2

Figure 3: Cross-lingual CNN

the state specific interpolation vectors and weight projection vectors forms the language specific parameters. In both these models the global parameters are borrowed and the language dependent interpolation vectors and weight projection vectors are estimated from the low-resource language data.

The use of DNN in cross-lingual acoustic modeling is discussed in [2]. First, a DNN is trained with high-resource language data with targets as the high-resource language tied states. The hidden layers of the high-resource language DNN are borrowed and the output layer with low-resource language tied states is then trained. The block schematic for cross-lingual DNN is shown in Figure 2.

In this paper, we also study the use of CNN for cross-lingual acoustic modeling. Two ways of using CNN for cross-lingual acoustic modeling are studied. The block schematic of the proposed methods are given in Figure 3a, 3b. The steps in training a CNN are described in [20]. The CNN model for the high-resource language was trained with two convolutional layers followed by four fully connected layers. The first cross-lingual CNN setup is similar to cross-lingual DNN where only the language specific output layer is trained keeping the hidden layers fixed. On the other hand, in the second setup, the high-resource language CNN model is used as a feature extractor for the fully connected layers which are then trained with the low-resource language data.

3. Experimental Setup

In this paper, experiments were performed with four Indian languages namely Assamese, Bengali, Hindi and Tamil from the MANDI database. For each language, data sets Train-low (≈ 2 hr) and Train-high (≈ 22 hr) were created to perform experiments in low-resource scenario. Experiments were performed with each pair of language to study the portability of data and model parameters for each of the language. The experiments were performed with Kaldi toolkit [24] to build the acoustic models.

3.1. Databases

MANDI database is a multilingual database consisting of six Indian languages. The database was collected for ‘‘Speech-based access to agricultural commodity prices’’, a Government of India project to build ASR systems in Indian languages to provide information about prices of agricultural commodities in different markets to farmers. The database contain Assamese, Bengali, Hindi, Marathi, Tamil and Telugu corpus. In each language corpus, speech was collected from the end-users in their native language. Each corpus contains mainly names of markets and commodities in the state were that language is commonly spoken. The data was mostly collected outdoor and varies from quiet to very noisy environments. In our work we used Assamese, Bengali, Hindi and Tamil corpus. Assamese and Bengali are Eastern Indo-Aryan origin, Hindi is Indo-Aryan Central origin and Tamil is of Dravidian origin. From the listening tests Tamil database was found to be most noisy followed by Hindi, Bengali and Assamese.

In every experiment in this paper, we consider one of the languages as a low-resource language and other languages as high-resource languages. When a language is considered as low-resource language the Train-low data set of that language is used for training and as high-resource language the Train-high data set is used. A common test set is used for each language irrespective of whether it is used as low-resource or high-resource language.

3.2. Baseline Experiments

The baseline results for the acoustic models built with Train-low data set from each language is given in Table 1, 2 and 3. The CDHMM model was built with 3 states for non-silence phones and 8 states for silence phones. For all languages the acoustic models used around 300 context-dependent states and 4 mixtures per state for the Train-low dataset. The SGMM and Phone-CAT models used around 1000 substates and 64 mixtures in UBM. The DNN models were trained with 6 layers and 2048 nodes for each hidden layers. The CNN model had 2 convolutional layers followed by 4 fully connected layers with 1024 nodes in each layer. The CDHMM, SGMM and Phone-CAT models were trained with MFCC features whereas, DNN and CNN models used 23 dimensional filter bank features.

3.3. Pooling Data

In this section, we discuss the experiments performed on a low-resource language by pooling data from other high-resource languages. For each low-resource language, varying amount of data ranging from 2 to 22 hour was pooled from any of the high-resource languages. For example, in the case of Tamil as low-resource language, the data was pooled from either Assamese, Bengali or Hindi corpus. To build the acoustic models in each case the speech transcription of low-resource language

Table 1: Results in (%WER) for the CDHMM model built for each low-resource language with varying amount of data pooled

Pooling Size (hours)	High-Resource Language	Assamese	Bengali	Hindi	Tamil
2	Baseline	37.26	18.46	14.59	33.81
	Assamese	-	17.09	13.43	33.61
	Bengali	31.39	-	13.21	31.49
	Hindi	26.10	16.21	-	31.84
	Tamil	40.24	18.19	15.13	-
3	Assamese	-	16.95	13.50	32.43
	Bengali	26.88	-	13.10	31.05
	Hindi	27.19	15.21	-	30.87
	Tamil	36.17	18.63	14.92	-
5	Assamese	-	17.68	14.01	33.91
	Bengali	24.29	-	13.19	29.91
	Hindi	25.55	15.73	-	29.62
	Tamil	34.56	18.27	15.27	-
22	Assamese	-	18.51	15.89	35.09
	Bengali	29.43	-	14.61	32.21
	Hindi	26.49	16.48	-	31.61
	Tamil	37.15	20.78	15.19	-

was converted to high-resource language phone set using the phone mapping algorithm described in [23]. While pooling data, low-resource language phones were mapped to high-resource language phone set so as to generate a richer phonetic context tree for building acoustic models. The acoustic models can be built once the speech data is pooled and transcriptions are mapped. In this experiment acoustic models like CDHMM, SGMM, Phone-CAT, DNN and CNN for each case with varying amount of pooled data were built. The recognition word error rate (WER) for CDHMM model for different amount of pooling data is given in Table 1.

The data pooled models gives improvements over the baseline acoustic models. However, better improvements can be achieved by adapting the pooled model towards the low-resource language. The low-resource language has ≈ 2 hrs of data, hence the model parameters can be re-trained. In the case of SGMM/Phone-CAT models the state specific interpolation vectors are re-estimated and in DNN/CNN the language specific output layer were retrained. The re-training gave improvements in most cases and did not have much effect for the cases where the phonetic overlap was poor.

3.4. Results of Pooling Data

The results for the different acoustic models built with pooled data are given in Table 2. In Table 1 the optimal amount of data that can be pooled for a low-resource language from different high-resource languages is given. From the results in Table 1, it is clear that pooling large amount of data from high-resource language deteriorate the recognition performance. The optimal amount of data varied between 2 to 5 hours. In majority of the data pooling experiments significant improvements were obtained over the baseline mono-lingual model. In most cases data pooling from closely related languages gave better performance than borrowing model parameters. In almost all cases using DNN or CNN gave better recognition performance compared to the parsimonious models. In the case of language combinations, Assamese and Tamil as low-resource language was benefited more by Hindi language. In the case of Bengali

Table 2: Performance in (%WER) of different acoustic models built with data pooling and adaptation

Acoustic Model Types	Low-Resource language			
	Assamese	Bengali	Hindi	Tamil
Mono-lingual SGMM	33.27	15.14	13.76	33.42
SGMM :w/Assamese	-	14.77	11.91	35.56
SGMM :w/Bengali	25.04	-	11.67	32.45
SGMM :w/Hindi	23.20	13.63	-	28.21
SGMM :w/Tamil	29.70	16.04	12.71	-
Mono-lingual Phone-CAT	33.74	15.16	13.50	33.59
Phone-CAT :w/Assamese	-	13.80	11.43	30.68
Phone-CAT :w/Bengali	26.65	-	11.79	30.21
Phone-CAT :w/Hindi	22.84	13.72	-	29.54
Phone-CAT :w/Tamil	30.37	14.99	12.65	-
Mono-lingual DNN	36.56	13.77	13.86	34.18
DNN :w/Assamese	-	11.38	11.01	36.57
DNN :w/Bengali	22.02	-	10.52	36.13
DNN :w/Hindi	22.96	9.26	-	35.91
DNN :w/Tamil	26.29	13.80	17.04	-
Mono-lingual CNN	41.26	13.99	13.51	36.08
CNN :w/Assamese	-	11.84	12.76	36.6
CNN :w/Bengali	21.90	-	11.59	29.05
CNN :w/Hindi	26.57	10.21	-	28.55
CNN :w/Tamil	32.33	13.24	12.70	-

*The first column has the format [Model-type] :w/[High-resource language].

as low-resource language Hindi was more useful and the reverse was also true. Pooling Tamil language data to other languages considered did not give much benefit as Tamil language is of Dravidian origin and Assamese, Bengali and Hindi are of Indo-Aryan origin. Another factor that affects the performance using data pooling or borrowing model parameters is the mismatch in the environmental conditions.

3.5. Borrowing Model Parameters

The cross-lingual experiments involving borrowing acoustic model parameters are described in this section. The experiments were performed with the parsimonious models like SGMM/Phone-CAT and neural network models like DNN/CNN. For each low-resource language the acoustic model parameters were borrowed from one of the high-resource languages. Experiments were performed for all the four languages with one of them as low-resource language and others as high-resource languages. The experiments were performed as described in Section 2.2.

3.6. Results for Borrowing Model Parameters

The recognition performance of cross-lingual SGMM, Phone-CAT, DNN and CNN are given in Table 3. In each case experiments were performed for each language as low-resource language. In Table 3 the model-type CNN1 and CNN2 refers to the cross-lingual CNN setup 1 and 2 respectively. The results showed consistent improvements in recognition performance of the corresponding low-resource language. The results for the hypothetical experiment where low-resource and high-resource language being the same are also given in Table 2 to show the upper bound for the improvements. DNN and CNN based cross-lingual models gave better performance compared to SGMM and Phone-CAT. The combination of low-resource and high-resource language corresponding to the best recognition

Table 3: Results in (%WER) for the various cross-lingual experiments with borrowing model parameters

Acoustic Model Types	Low-Resource Language			
	Assamese	Bengali	Hindi	Tamil
Mono-lingual SGMM	33.27	15.14	13.76	33.42
SGMM :w/Assamese	27.59	16.58	13.14	34.67
SGMM :w/Bengali	31.94	12.23	10.29	32.85
SGMM :w/Hindi	30.45	14.09	9.14	32.55
SGMM :w/Tamil	32.41	16.46	12.80	29.00
Mono-lingual Phone-CAT	33.74	15.16	13.50	33.59
Phone-CAT :w/Assamese	28.39	15.87	12.16	34.92
Phone-CAT :w/Bengali	30.25	12.06	11.06	31.52
Phone-CAT :w/Hindi	30.09	14.92	10.14	33.29
Phone-CAT :w/Tamil	32.25	15.26	12.50	27.67
Mono-lingual DNN	36.56	13.77	13.86	34.18
DNN :w/Assamese	48.00	12.55	13.29	32.01
DNN :w/Bengali	48.28	9.18	9.77	29.57
DNN :w/Hindi	50.27	10.89	8.97	28.85
DNN :w/Tamil	52.00	12.11	12.50	28.78
Mono-lingual CNN	41.26	13.99	13.51	36.08
CNN1 :w/Assamese	39.89	15.75	15.12	36.77
CNN1 :w/Bengali	36.91	12.53	11.93	34.35
CNN1 :w/Hindi	37.58	14.51	11.09	33.93
CNN1 :w/Tamil	39.03	15.07	16.19	32.60
CNN2 :w/Assamese	38.64	12.55	13.11	36.87
CNN2 :w/Bengali	37.15	11.79	12.75	35.74
CNN2 :w/Hindi	35.85	11.60	12.39	35.34
CNN2 :w/Tamil	38.32	12.16	12.79	35.88

*The first column has the format [Model-type] :w/[High-resource language].

performance is similar to data pooling experiments.

4. Conclusion

In this paper, two methods to overcome the data sparsity problem in acoustic modeling of low-resource language are proposed. The first method perform pooling of speech data from closely related high-resource languages and different acoustic models are built with the help of mapping low-resource language phones to high-resource language phones. In the second method the acoustic model parameters built with high-resource languages are borrowed and the language specific parameters of the corresponding acoustic model are estimated with the low-resource language. In both cases significant improvements were achieved over the corresponding acoustic model built with low-resource language alone. The use of CNN and DNN gave more recognition improvements compared to using parsimonious models like SGMM and Phone-CAT in the framework of data pooling and borrowing acoustic model parameters. Experiments were performed with four Indian languages and achieved a relative recognition improvements of 10 to 30% in all languages.

5. Acknowledgements

This work was supported in part by the consortium project titled "Speech-based access to commodity price in six Indian languages", funded by the TDIL program of DeITY of Govt. of India. The authors would like to thank consortium members involved in collecting Assamese, Hindi and Tamil corpus.

6. References

- [1] L. Lu, A. Ghoshal, and S. Renals, "Cross-lingual Subspace Gaussian Mixture Models for Low-resource Speech Recognition," *IEEE/ACM Transactions on Audio, Speech and Language Processing*, vol. 22, no. 1, pp. 17–27, Jan 2014.
- [2] J.-T. Huang, J. Li, D. Yu, L. Deng, and Y. Gong, "Cross-language knowledge transfer using multilingual deep neural network with shared hidden layers," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 7304–7308.
- [3] V.-B. Le and L. Besacier, "Automatic speech recognition for under-resourced languages: application to vietnamese language," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 17, no. 8, pp. 1471–1482, 2009.
- [4] L. Besacier, E. Barnard, A. Karpov, and T. Schultz, "Automatic speech recognition for under-resourced languages: A survey," *Speech Communication*, vol. 56, pp. 85–100, 2014.
- [5] N. T. Vu, D. Imseng, D. Povey, P. Motlicek, T. Schultz, and H. Bourlard, "Multilingual deep neural network based acoustic modeling for rapid language adaptation," in *Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 7639–7643.
- [6] J. Cui, B. Kingsbury, B. Ramabhadran, A. Sethy, K. Audhkhasi, Z. Tüske, P. Golik, R. Schlüter, H. Ney, M. J. Gales *et al.*, "Multilingual representations for low resource speech recognition and keyword search," *Context*, vol. 10, no. 10, p. 10.
- [7] T. Schultz and A. Waibel, "Experiments on cross-language acoustic modeling," in *INTERSPEECH*, 2001, pp. 2721–2724.
- [8] W. Byrne, P. Beyerlein, J. M. Huerta, S. Khudanpur, B. Marthi, J. Morgan, N. Peterek, J. Picone, D. Vergyri, and T. Wang, "Towards language independent acoustic modeling," in *Proc. ICASSP*, vol. 2, 2000, pp. 1029–1032.
- [9] V. B. Le and L. Besacier, "First steps in fast acoustic modeling for a new target language: application to vietnamese," in *Proc. ICASSP*, vol. 5, 2005, pp. 821–824.
- [10] K. C. Sim and H. Li, "Robust phone set mapping using decision tree clustering for cross-lingual phone recognition," in *Proc. ICASSP*, 2008, pp. 4309–4312.
- [11] Z. Tuske, J. Pinto, D. Willett, and R. Schluter, "Investigation on cross-and multilingual mlp features under matched and mismatched acoustical conditions," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 7349–7353.
- [12] L. Burget, P. Schwarz, M. Agarwal, P. Akyazi, K. Feng, A. Ghoshal, O. Glembek, N. Goel, M. Karafiát, D. Povey *et al.*, "Multilingual acoustic modeling for speech recognition based on subspace gaussian mixture models," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*. IEEE, 2010, pp. 4334–4337.
- [13] S. Thomas, S. Ganapathy, and H. Hermansky, "Cross-lingual and multi-stream posterior features for low resource lvcsr systems," in *INTERSPEECH*. Citeseer, 2010, pp. 877–880.
- [14] P. Lal and S. King, "Cross-lingual automatic speech recognition using tandem features," *Audio, Speech, and Language Processing, IEEE Transactions on*, vol. 21, no. 12, pp. 2506–2515, 2013.
- [15] N. T. Vu, F. Metze, and T. Schultz, "Multilingual bottle-neck features and its application for under-resourced languages," 2012.
- [16] K. Knill, M. J. Gales, S. P. Rath, P. C. Woodland, C. Zhang, and S.-X. Zhang, "Investigation of multilingual deep neural networks for spoken term detection," in *Automatic Speech Recognition and Understanding (ASRU), 2013 IEEE Workshop on*. IEEE, 2013, pp. 138–143.
- [17] D. Povey, L. Burget, M. Agarwal, P. Akyazi, K. Feng, A. Ghoshal, O. Glembek, N. K. Goel, M. Karafiát, A. Rastrow, R. C. Rose, P. Schwarz, and S. Thomas, "The Subspace Gaussian Mixture Model - A Structured Model for Speech Recognition," *Computer Speech & Language*, vol. 25, no. 2, pp. 404–439, 2011.
- [18] V. Manohar, B. S. Chinnari, and S. Umesh, "Acoustic Modeling using Transform-Based Phone-Cluster Adaptive Training," in *Proc. ASRU*, December 2013, pp. 49–54.
- [19] G. Hinton, L. Deng, D. Yu, G. E. Dahl, A. R. Mohamed, N. Jaitly, A. Senior, V. Vanhoucke, P. Nguyen, T. N. Sainath *et al.*, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Processing Magazine*, vol. 29, no. 6, pp. 82–97, 2012.
- [20] T. N. Sainath, A.-r. Mohamed, B. Kingsbury, and B. Ramabhadran, "Deep convolutional neural networks for lvcsr," in *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. IEEE, 2013, pp. 8614–8618.
- [21] T. Fraga-Silva, J.-L. Gauvain, L. Lamel, A. Laurent, V.-B. Le, A. Messaoudi, V. Vapnarsky, C. Barras, C. Becquey, D. Doukhan *et al.*, "Active learning based data selection for limited resource stt and kws," in *Annual Conference*, vol. 141, 2015, pp. 47–53.
- [22] Y. Qian, K. Yu, and J. Liu, "Combination of data borrowing strategies for low-resource lvcsr," in *Automatic Speech Recognition and Understanding (ASRU), 2013 IEEE Workshop on*. IEEE, 2013, pp. 404–409.
- [23] Basil Abraham, Neethu Mariam Joy, Navneeth K., and S. Umesh, "A Data-Driven Phoneme Mapping Technique Using Interpolation Vectors of Phone-Cluster Adaptive Training," in *Proc. SLT*, December 2014, pp. 36–41.
- [24] D. Povey, A. Ghoshal, G. Boulianne, L. Burget, O. Glembek, N. Goel, M. Hannemann, P. Motlicek, Y. Qian, P. Schwarz, J. Silovsky, G. Stemmer, and K. Vesely, "The Kaldi Speech Recognition Toolkit," in *Proc. ASRU*, December 2011.