

THE MLLP ASR SYSTEM FOR IWSLT 2015

M. A. del-Agua A. Martínez-Villaronga S. Piqueras A. Giménez A. Sanchis J. Civera A. Juan



MLLP | Machine Learning and Language Processing



UNIVERSITAT POLITÈCNICA DE VALÈNCIA

THE TRANSLECTURES-UPV TOOLKIT

Toolkit developed at UPV that allows the development of an end-to-end speech recognition system.

- DNN and HMM training and decoding systems.
- Support to Convolutional NNs (CNN).
- Support to Multilingual NNs.
- DNN speaker adaptation techniques.
- DNN sequence discriminative training.

ENGLISH ASR SYSTEM

Data Collection

- 245h of TED talks for acoustic training
- 3.7M ID words + 300B OOD words for language model

Acoustic Model

- ROVER combination of 5 subsystems

Name	Topology	Input feas.	Act. Func.	Train. Crit.
DNN-mmi	DNN	MFCCs	Sigmoid	MMI
DNN-sigm	DNN	MFCCs	Sigmoid	CE
DNN-relu	DNN	MFCCs	ReLU	CE
CNN-sigm	CNN	F. Bank	Sigmoid	CE
CNN-relu	CNN	F. Bank	ReLU	CE

- DNN speaker adaptation based on confidence measures

Language Model

- Linear interpolation of different n -gram models
- 200k words vocabulary

Results

- 13.3 WER on tst2015

GERMAN ASR SYSTEM

Data Collection

- 36 × 3 hours of GSC corpus for acoustic training
- 2.3B words for language model

Acoustic Model

- DNN hybrid acoustic model similar to the DNN-sigm English system
- DNN speaker adaptation based on confidence measures

Language Model

- Linear interpolation of different n -gram models
- 200k words vocabulary

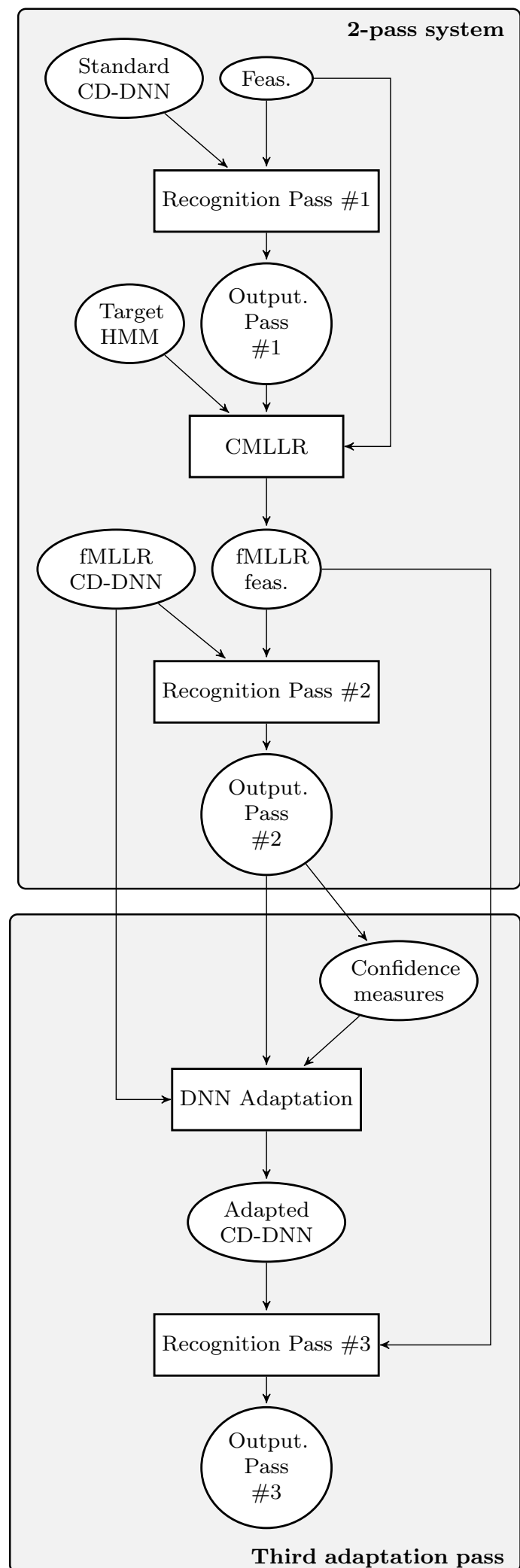
Results

- 43.3 WER on tst2015

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Union Seventh Framework Programme (FP7/2007-2013) under grant agreement no 287755 (transLectures) and ICT Policy Support Programme (ICT PSP/2007-2013) as part of the Competitiveness and Innovation Framework Programme (CIP) under grant agreement no 621030 (EMMA), the Spanish MINECO Active2Trans (TIN2012-31723) research project, the Spanish Government with the FPU scholarship FPU13/06241 and the Generalitat Valenciana with the VALi+d scholarship ACIF/2015/082.

SYSTEM ARCHITECTURE



CONCLUSIONS

- First competition of a TLK-based ASR system
- English system obtained quite competitive results
- Results in German were not as good as expected due to the lack of in-domain training data for AM and LM
- For future competitions we plan to use neural network based language models