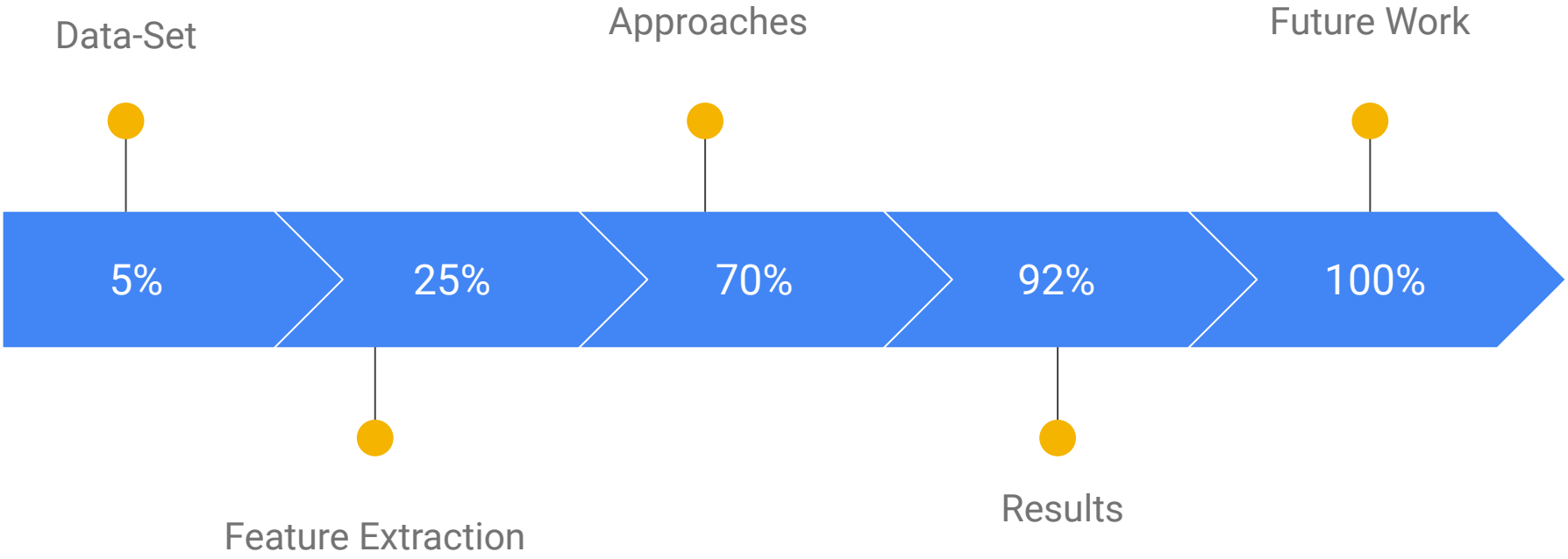
Low Resource Language ASR

Atul Sahay, Suraj Kumar and Nikhil Saini
(18305R003, 18305R008, 183059006)
CS 753 (Course Project)

OUTLINE


Data-Set
Approaches
Future Work
5%
25%
70%
92%
100%
Feature Extraction
Results

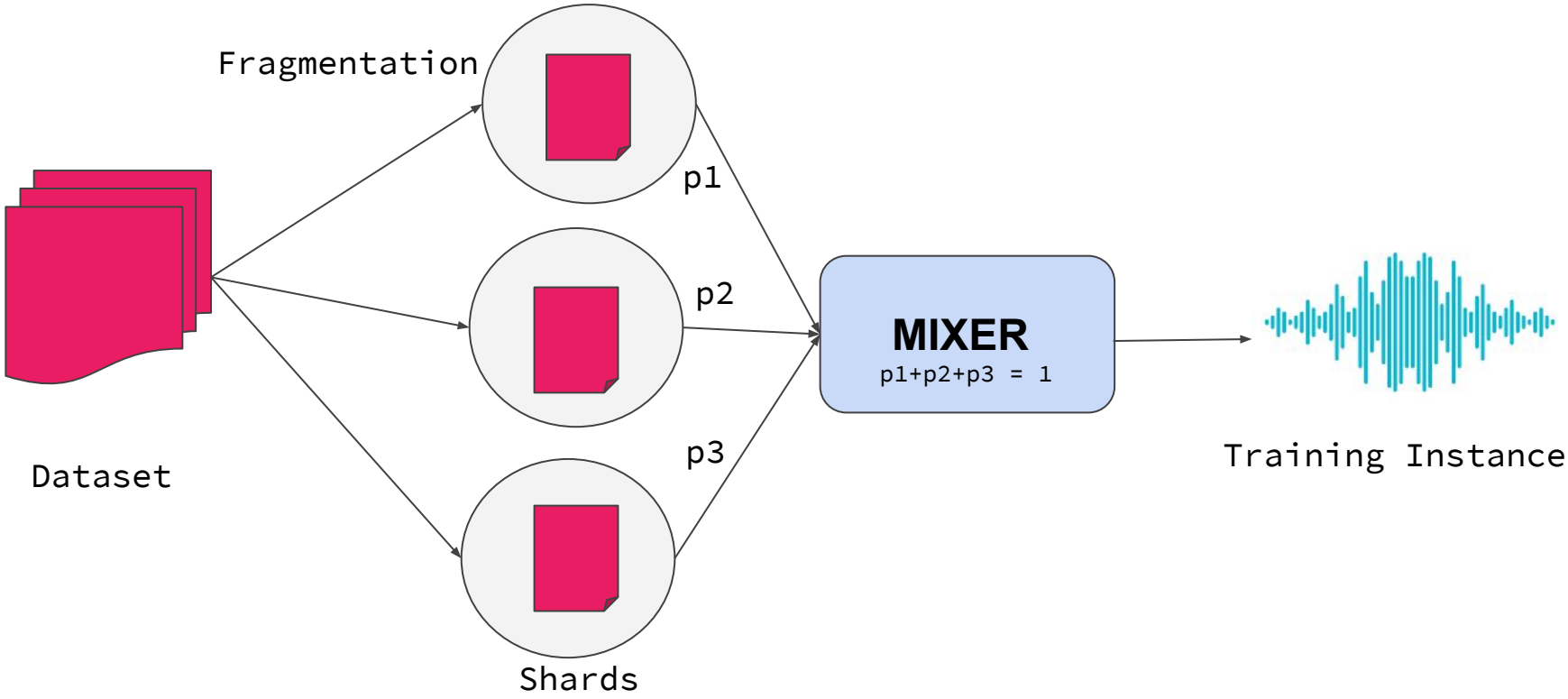
Datasets

OpenSLR Female (Gujarat)

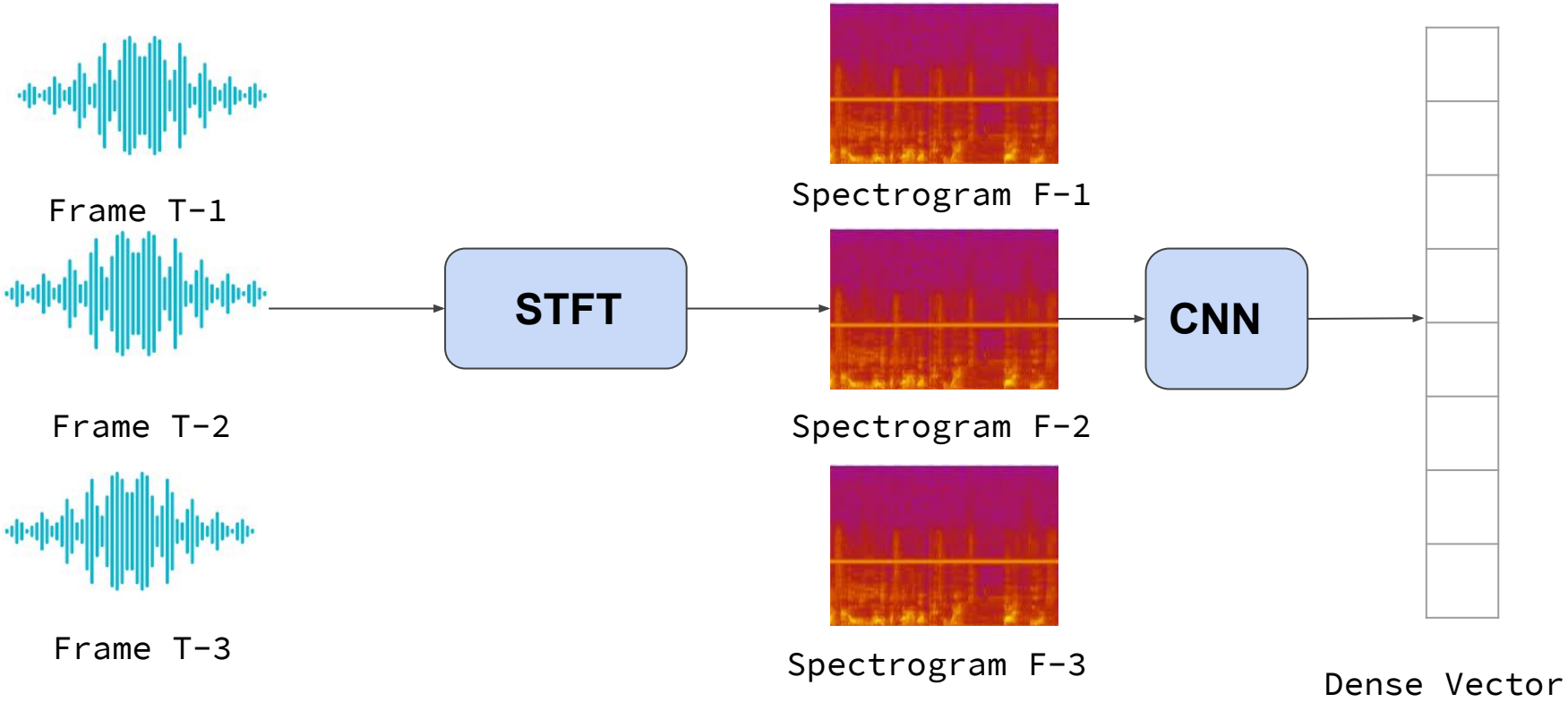
OpenSLR Male (Gujarati)

Microsoft Speech Corpus (Gujarati)

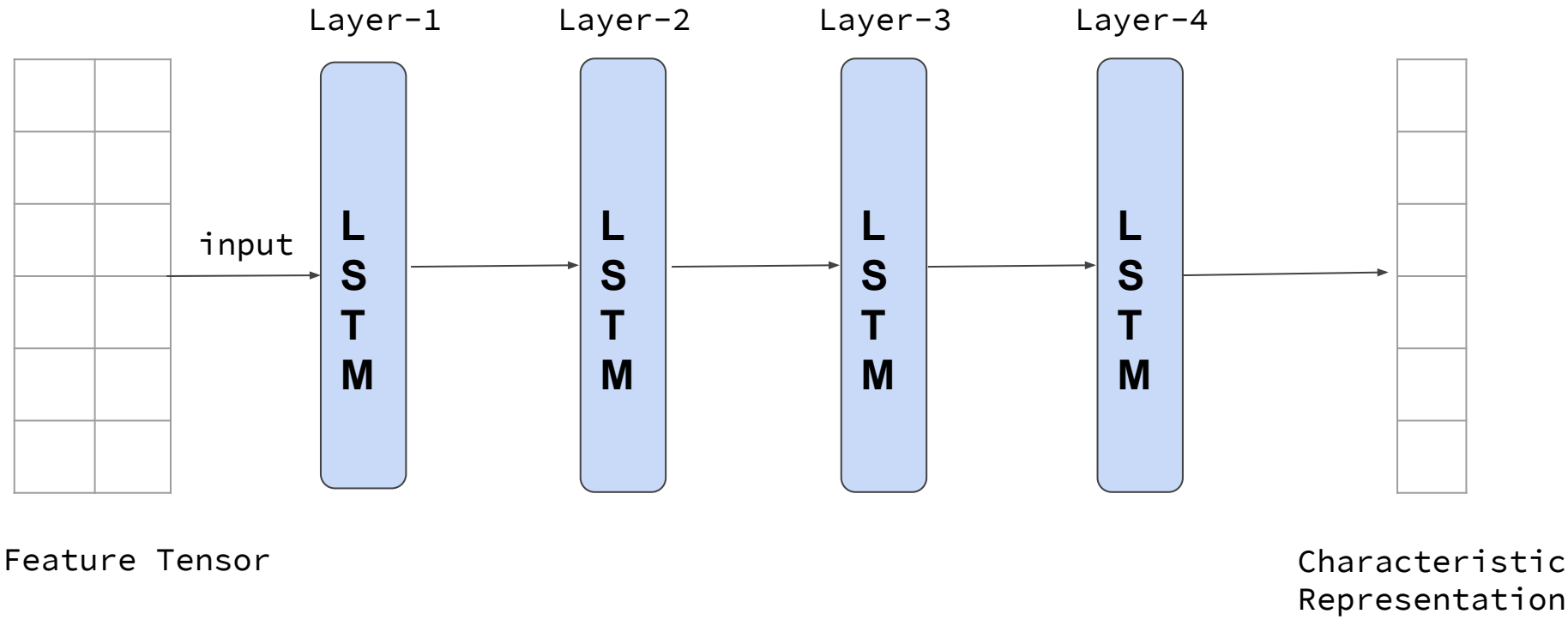
Segmentation



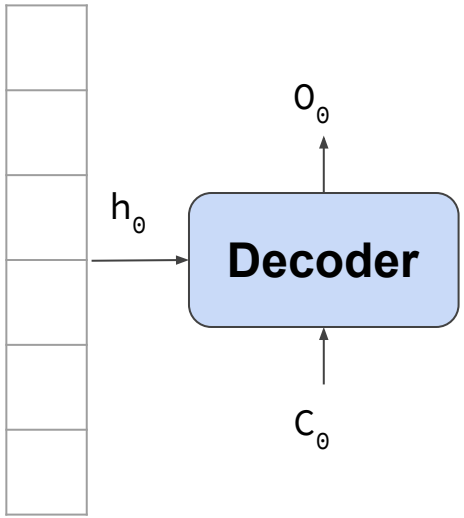
Feature Extraction



Encoding

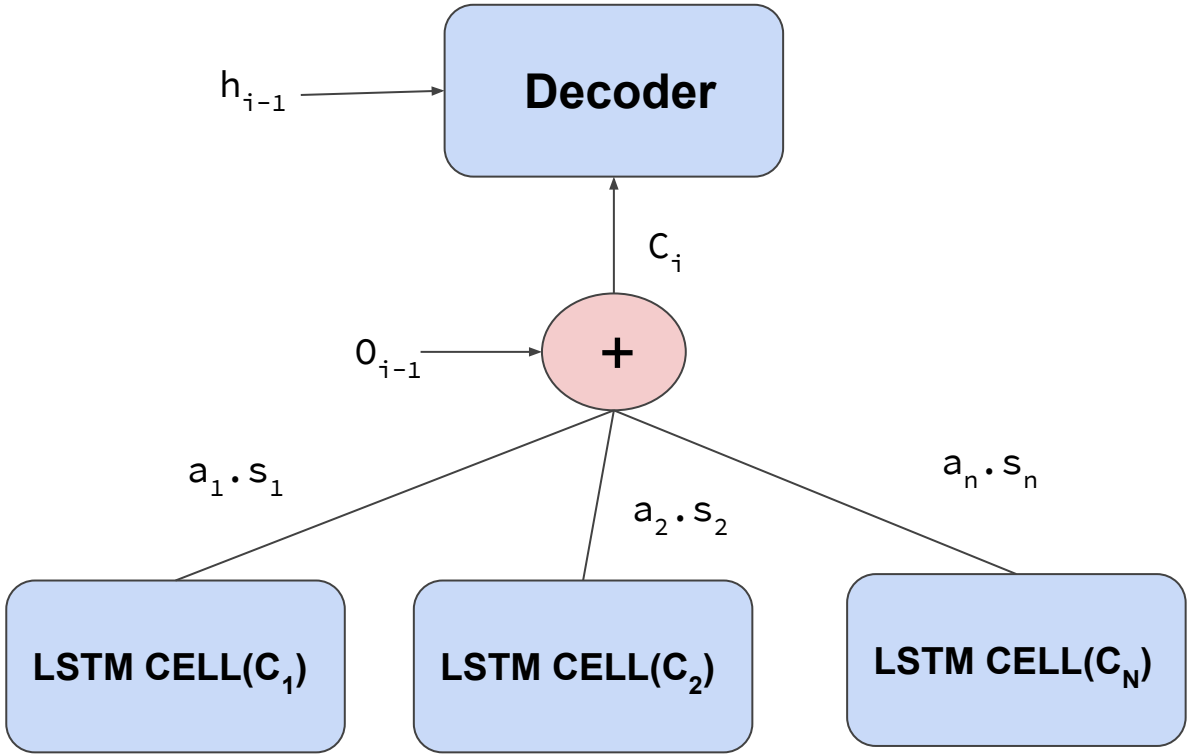

Layer-1
Layer-2
Layer-3
Layer-4
input
LSTM
LSTM
LSTM
LSTM
Feature Tensor
Characteristic Representation

Decoding

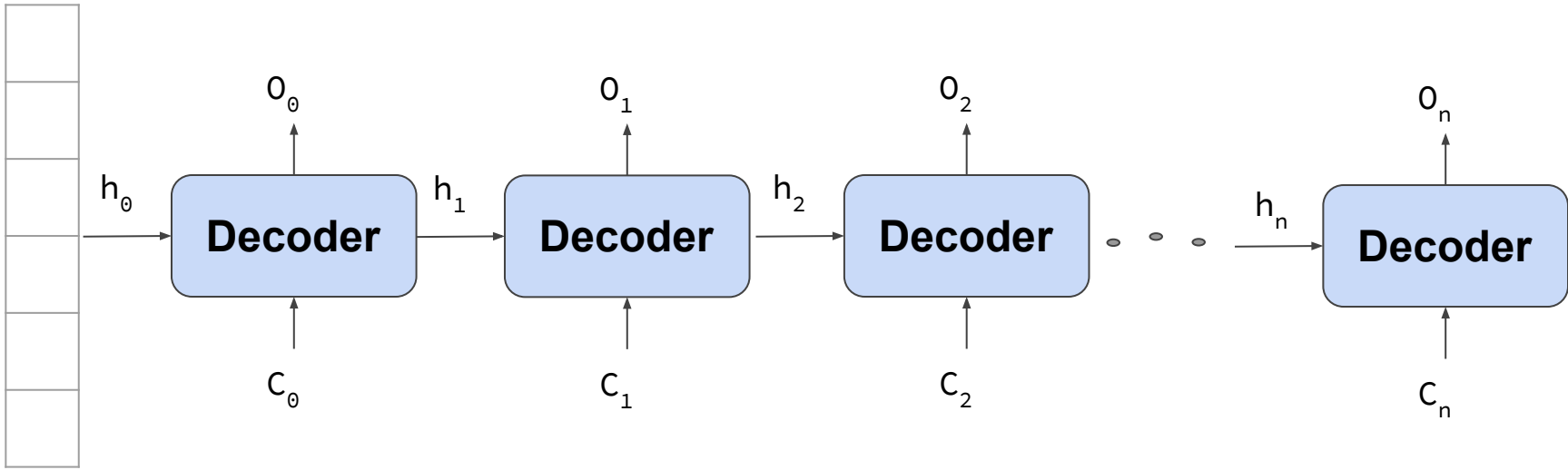


Characteristic
Representation

Decoding


h_{i-1}
Decoder
C_i
O_{i-1}
+
$a_1 . s_1$
$a_2 . s_2$
$a_n . s_n$
LSTM CELL(C_1)
LSTM CELL(C_2)
LSTM CELL(C_N)

Decoding



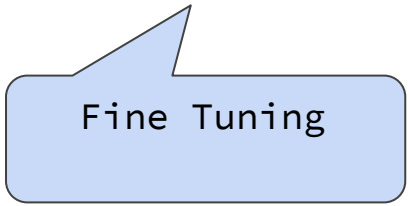
Characteristic
Representation

Results

Dataset	Time (hours)	BLEU
Microsoft Speech Corpus (Gujarati)	40	**7.28**
OpenSLR Male (Gujarati)	5	**0.17**
OpenSLR Female (Gujarati)	5	**0.14**

Results

Dataset	Time (hours)	BLEU
Microsoft Speech Corpus (Gujarati)	40	7.28
OpenSLR Male (Gujarati)	5	0.17
OpenSLR Female (Gujarati)	5	0.14
OpenSLR Male (Gujarati)	5	**4.30**
OpenSLR Female (Gujarati)	5	**2.77**

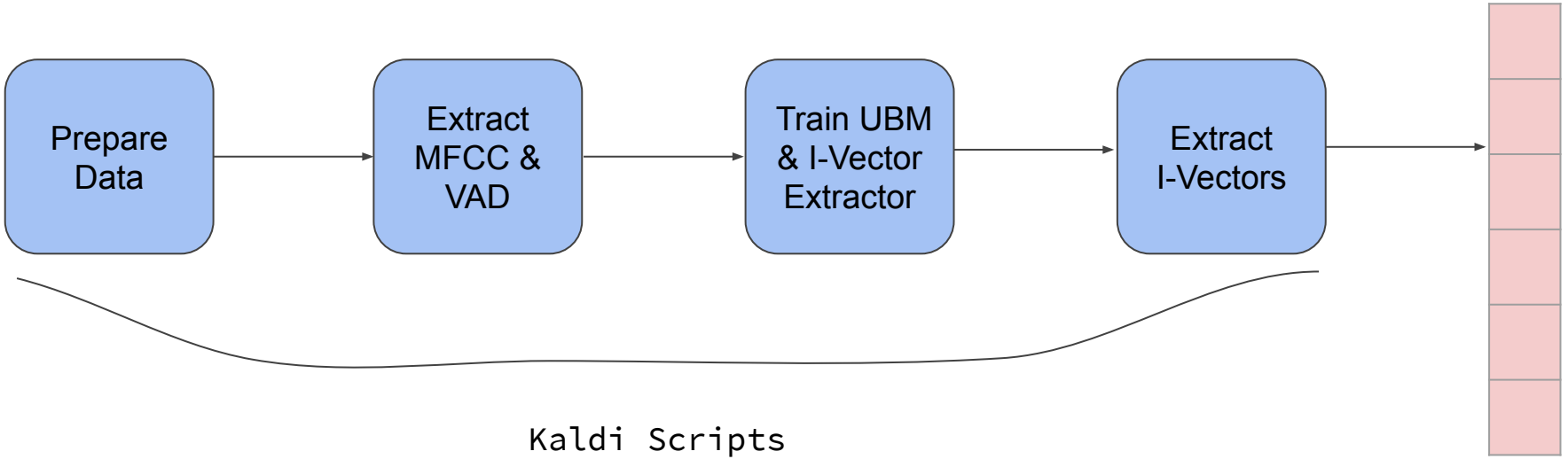


Results

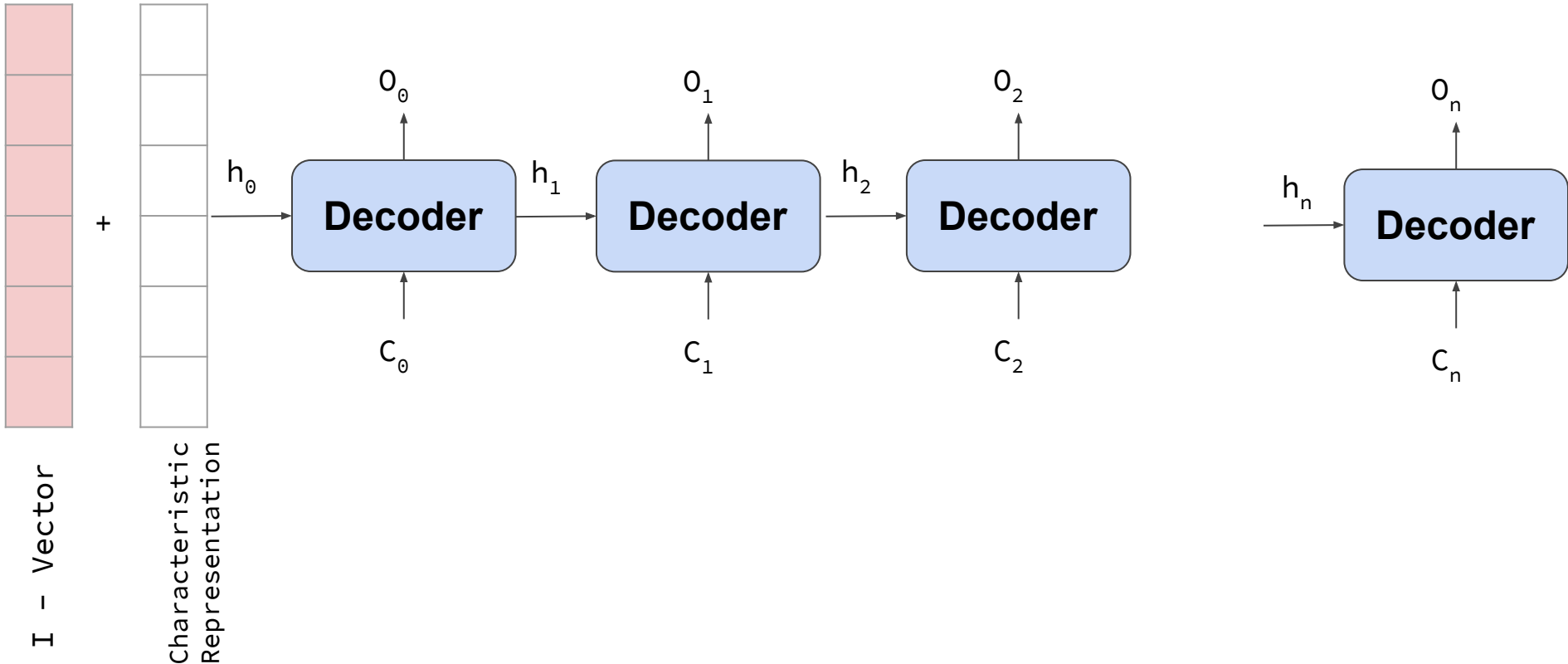
Dataset	Time (hours)	BLEU
Microsoft Speech Corpus (Gujarati)	40	7.28
OpenSLR Male (Gujarati)	5	0.17
OpenSLR Female (Gujarati)	5	0.14
OpenSLR Male (Gujarati)	5	4.30
OpenSLR Female (Gujarati)	5	2.77
OpenSLR Male (Gujarati)	5	**4.37**
OpenSLR Female (Gujarati)	5	**4.34**

Fine Tuning &
Weight Freezing

I-Vectors

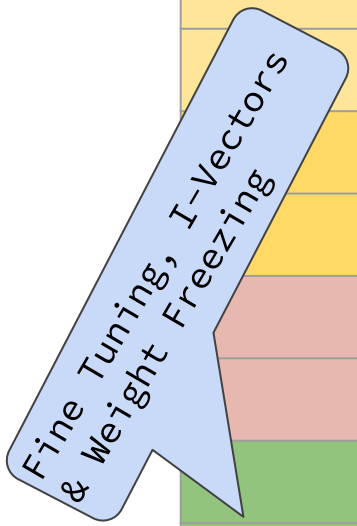

Prepare Data
Extract MFCC & VAD
Train UBM & I-Vector Extractor
Extract I-Vectors
Kaldi Scripts
i - Vector

Speaker Adaptation

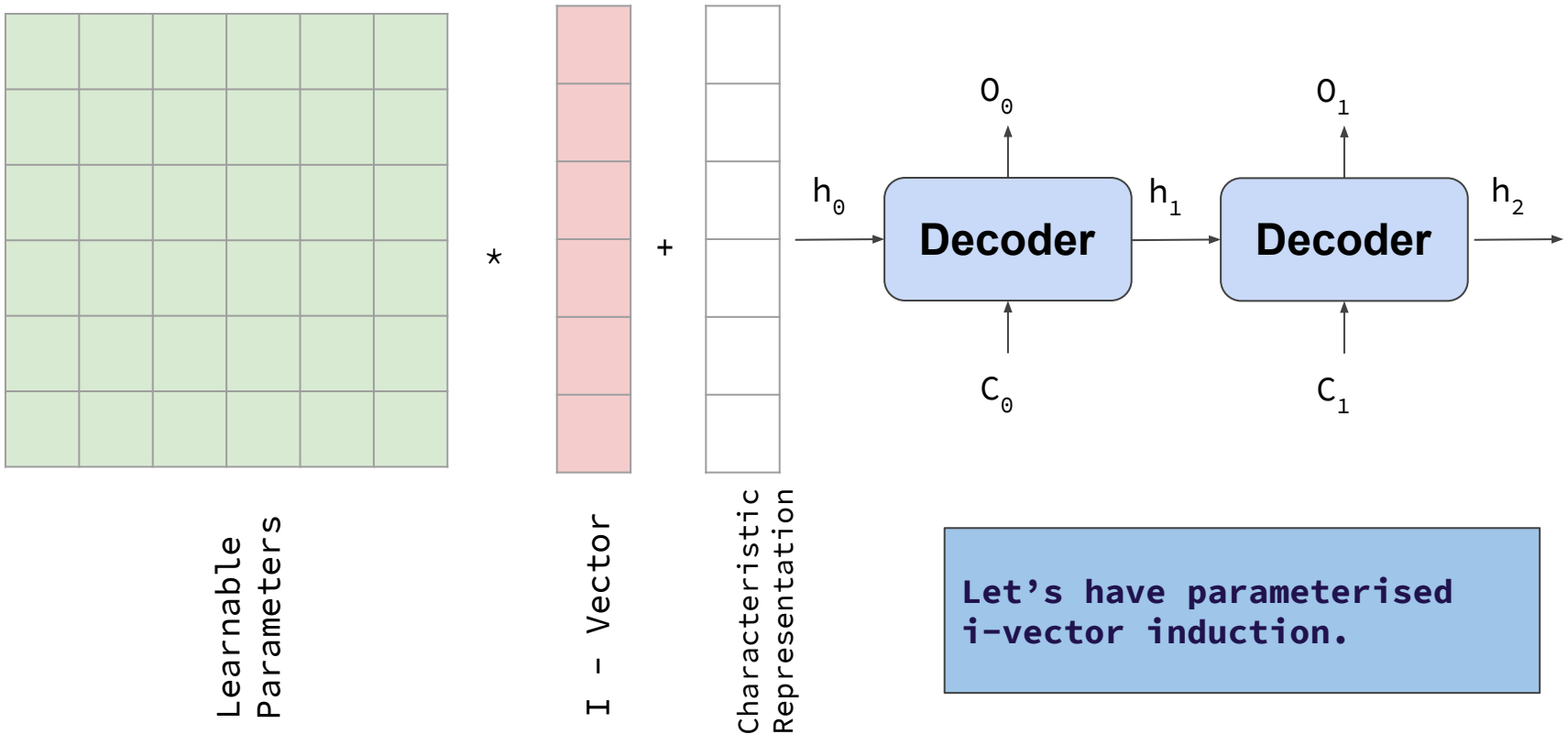


Results

Dataset	Time (hours)	BLEU
Microsoft Speech Corpus (Gujarati)	40	7.28
OpenSLR Male (Gujarati)	5	0.17
OpenSLR Female (Gujarati)	5	0.14
OpenSLR Male (Gujarati)	5	4.30
OpenSLR Female (Gujarati)	5	2.77
OpenSLR Male (Gujarati)	5	4.37
OpenSLR Female (Gujarati)	5	4.34
OpenSLR Male (Gujarati)	5	**4.37**
OpenSLR Female (Gujarati)	5	**4.34**



Future Work



Let's have parameterised i-vector induction.

Future Work

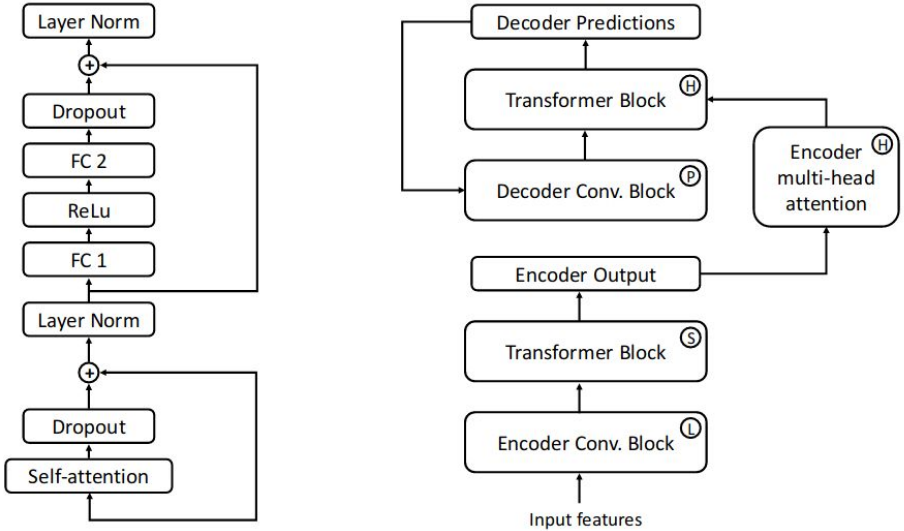


Figure 1: *Left: components of one transformer block. Right: Block diagram of the full end-to-end model*

*Transformers with convolutional context for ASR : https://arxiv.org/pdf/1904.11660.pdf

Other Tasks

1. Scripts for data structuring. (Link to github)
2. OpenNMT-py code change for:
    a. Freezing Weights of Encoder
    b. Transfer Learning
    c. I-Vector Feature Addition
3. Worked on fairseq code.
4. Kaldi I-Vector scripts

References

1. OpenNMT-py (https://github.com/OpenNMT)
2. I-vector Extraction (http://jrmeyer.github.io/asr/2017/09/29/challenge.html)
3. FairSeq (https://github.com/pytorch/fairseq)
4. Abdelrahman Mohamed, Dmytro Okhonko, Luke Zettlemoyer Transformers with convolutional context for ASR

Thank You