

TraffSign: Multilingual Traffic Signboard Text Detection and Recognition for Urdu and English

Muhammad Atif Butt, Adnan Ul-Hasan, and Faisal Shafait

National Centre of Artificial Intelligence (NCAI), Pakistan.

School of Electrical Engineering and Computer Science (SEecs), National University of Sciences and Technology (NUST), Islamabad, Pakistan.

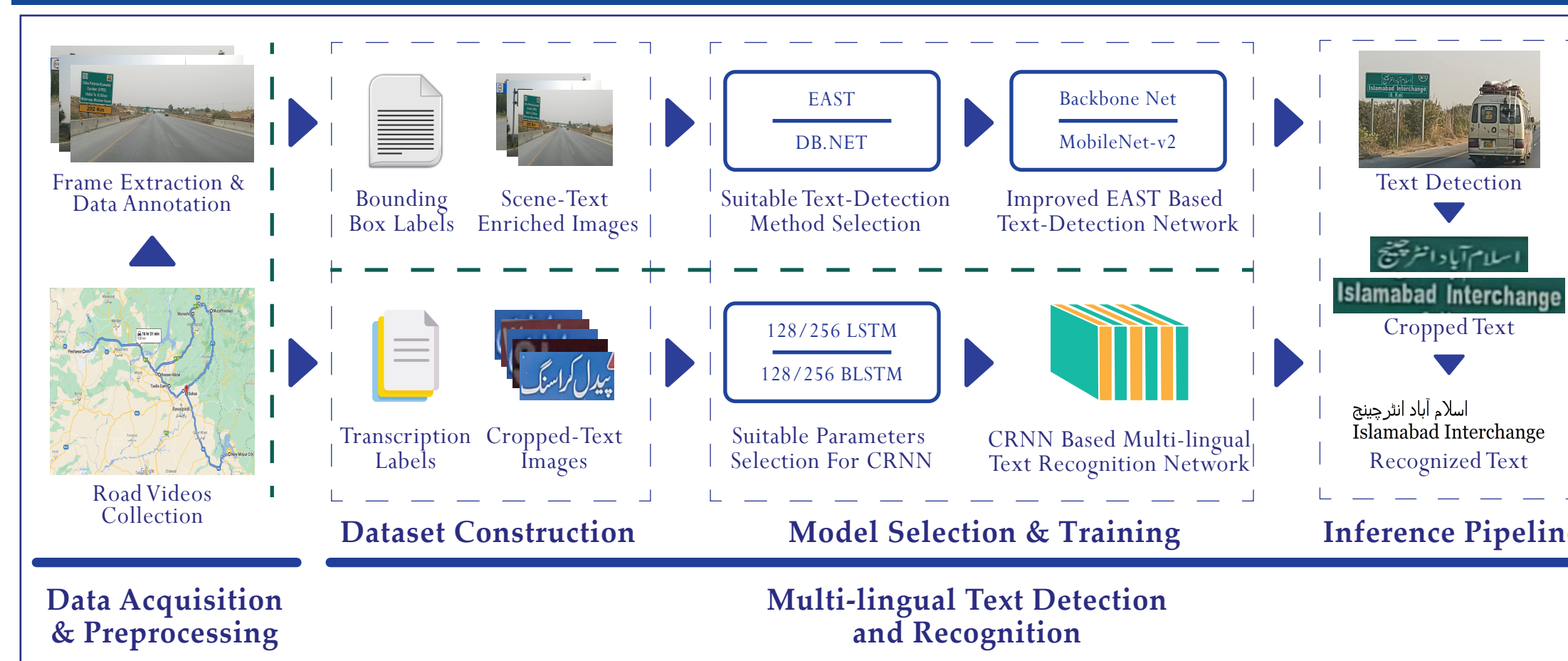


NATIONAL CENTRE OF
ARTIFICIAL INTELLIGENCE

Abstract

Scene-text detection and recognition methods have demonstrated remarkable performance on standard benchmark datasets. These methods can be utilized in human-driven/self-driving cars to perform navigation assistance through traffic signboard text detection and recognition. Existing datasets include scripts of numerous languages like English, Chinese, French, Arabic, German, etc. However, traffic navigation signboards in Pakistan and many states of India are written in Urdu along with the English translation to guide human drivers. To this end, we present Deep Learning Laboratory's Traffic Signboards Dataset (DLL-TraffSiD) to develop multi-lingual text detection and recognition methods for traffic signboards. In addition, we present a pipeline for multi-lingual text detection and recognition for an outdoor road environment. The results show that our presented system signified better applicability in text-detection and text recognition, and achieved 89% and 92.18% accuracy on the proposed dataset.

Method

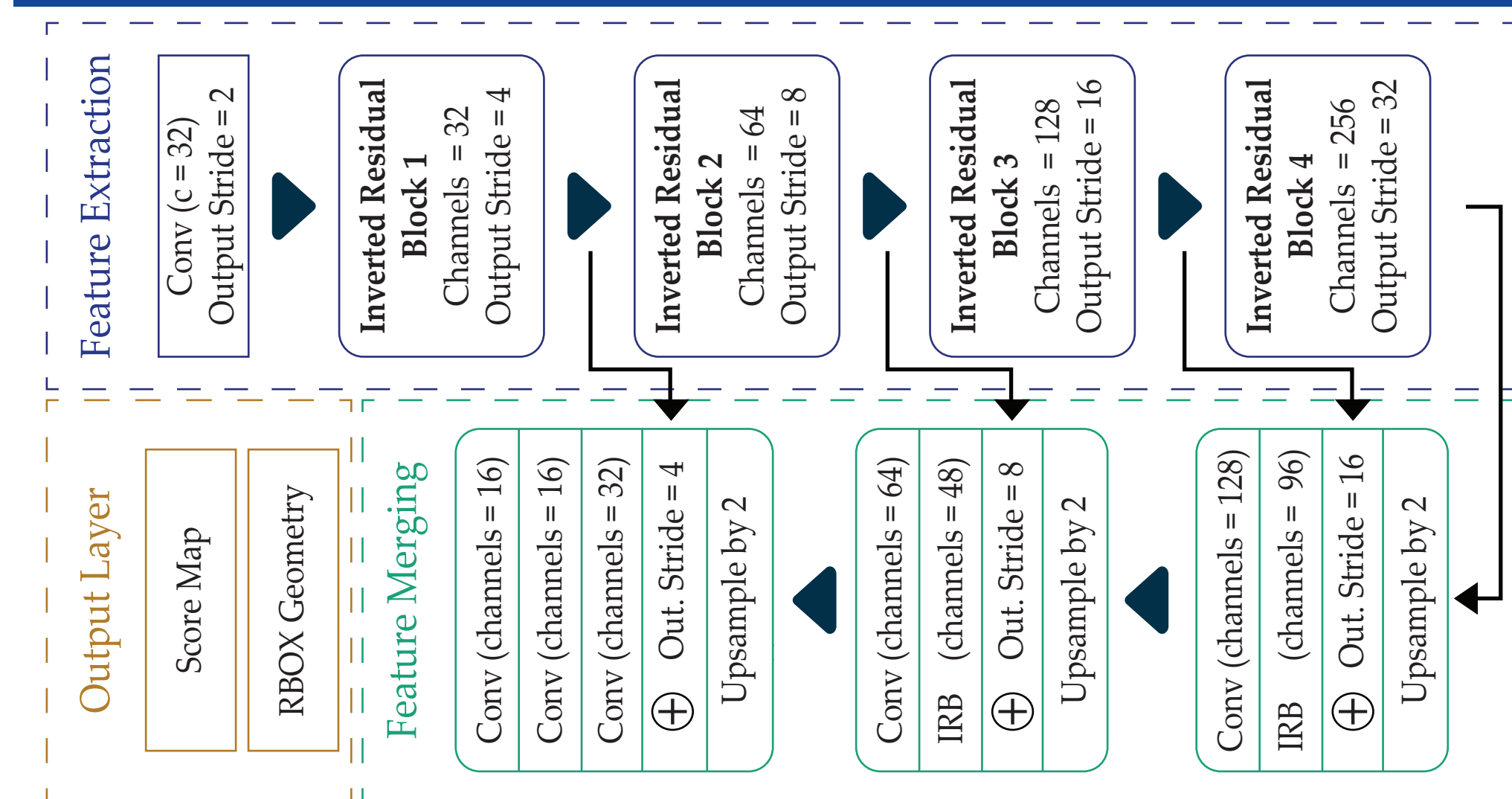


Results

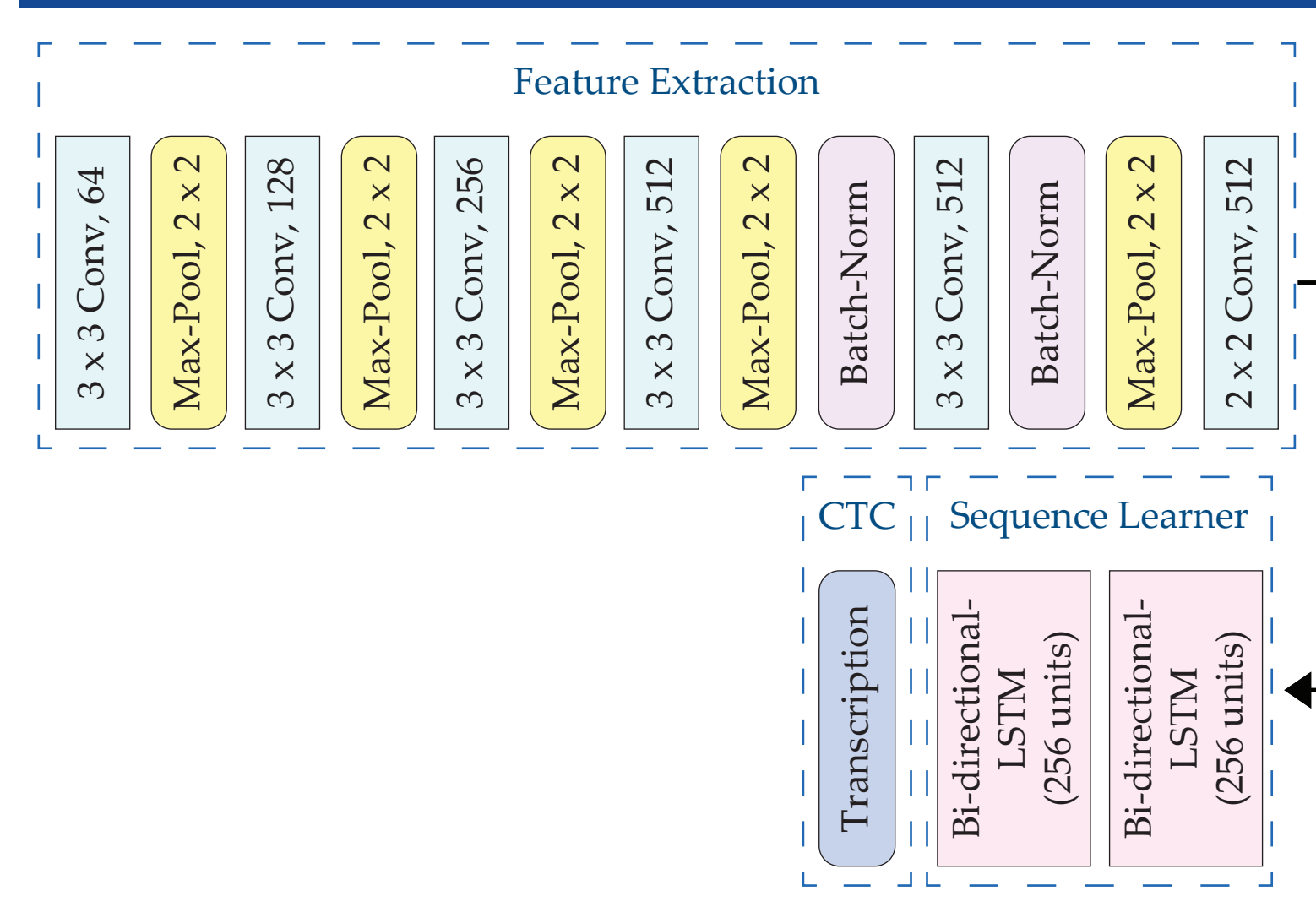
Model	Backbone	Learning Rate	Precision	Recall	F-Score
DB Net	FCN	Piece-wise	0.68	0.72	0.69
		0.0001	0.63	0.69	0.65
		0.001	0.61	0.66	0.62
EAST	PA-net	Piece-wise	0.51	0.39	0.44
		0.0001	0.54	0.42	0.46
	0.001	0.52	0.40	0.43	
	VGG-16	Piece-wise	0.43	0.27	0.33
		0.0001	0.41	0.25	0.31
	0.001	0.39	0.24	0.29	
ResNet-50	ResNet-50	Piece-wise	0.67	0.71	0.68
		0.0001	0.70	0.74	0.72
		0.001	0.64	0.68	0.69
Mobile-Net	Mobile-Net	Piece-wise	0.86	0.93	0.89
		0.0001	0.82	0.89	0.85
		0.001	0.77	0.84	0.86

Comparison of text detection networks on proposed multilingual dataset demonstrating performance metrics.

Text-Detection Architecture



Text Recognition Architecture



Architecture	Sequence Learner	Language	Hidden Units	WRR (%)
Baseline CRNN	LSTM	English	128	71.22
		Urdu		68.10
	BLSTM	English	128	74.87
		Urdu		72.29
	LSTM	English	256	75.41
		Urdu		72.66
BLSTM	English	256	78.34	
	Urdu		76.41	
Improved CRNN	LSTM	English	128	82.64
		Urdu		79.28
	BLSTM	English	128	86.71
		Urdu		85.56
	LSTM	English	256	89.07
		Urdu		86.43
BLSTM	English	256	92.18	
	Urdu		90.85	

Performance of proposed CRNN Architectures on proposed multi-lingual text recognition dataset.