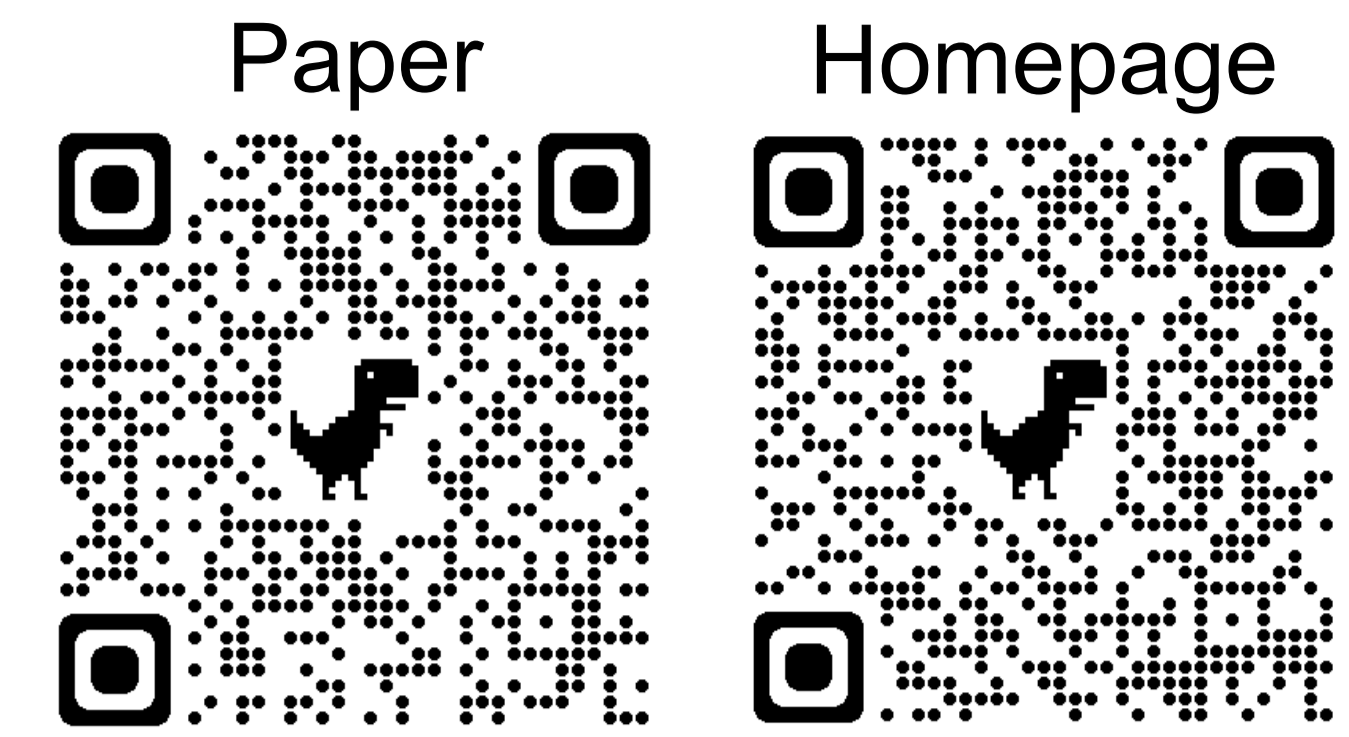# SSL-Net: A Synergistic Spectral and Learning-Based Network for Efficient Bird Sound Classification

Yiyuan Yang, Kaichen Zhou, Niki Trigoni, Andrew Markham
Department of Computer Science, University of Oxford
yiyuan.yang@cs.ox.ac.uk

Paper   Homepage

Motivation

- We present an efficient bird sound classifier by spectral features to mitigate distribution bias in the pretrained features and enhances performance with minimal labels.

- We put forth three fusion strategies to mitigate the distribution bias within spectral and learned features, thereby yielding enhanced representations for the classification task.

- Through comprehensive evaluations and ablation tests with real-world bird audio, our framework aids engineers in designing efficient and accurate classifiers with low effort.
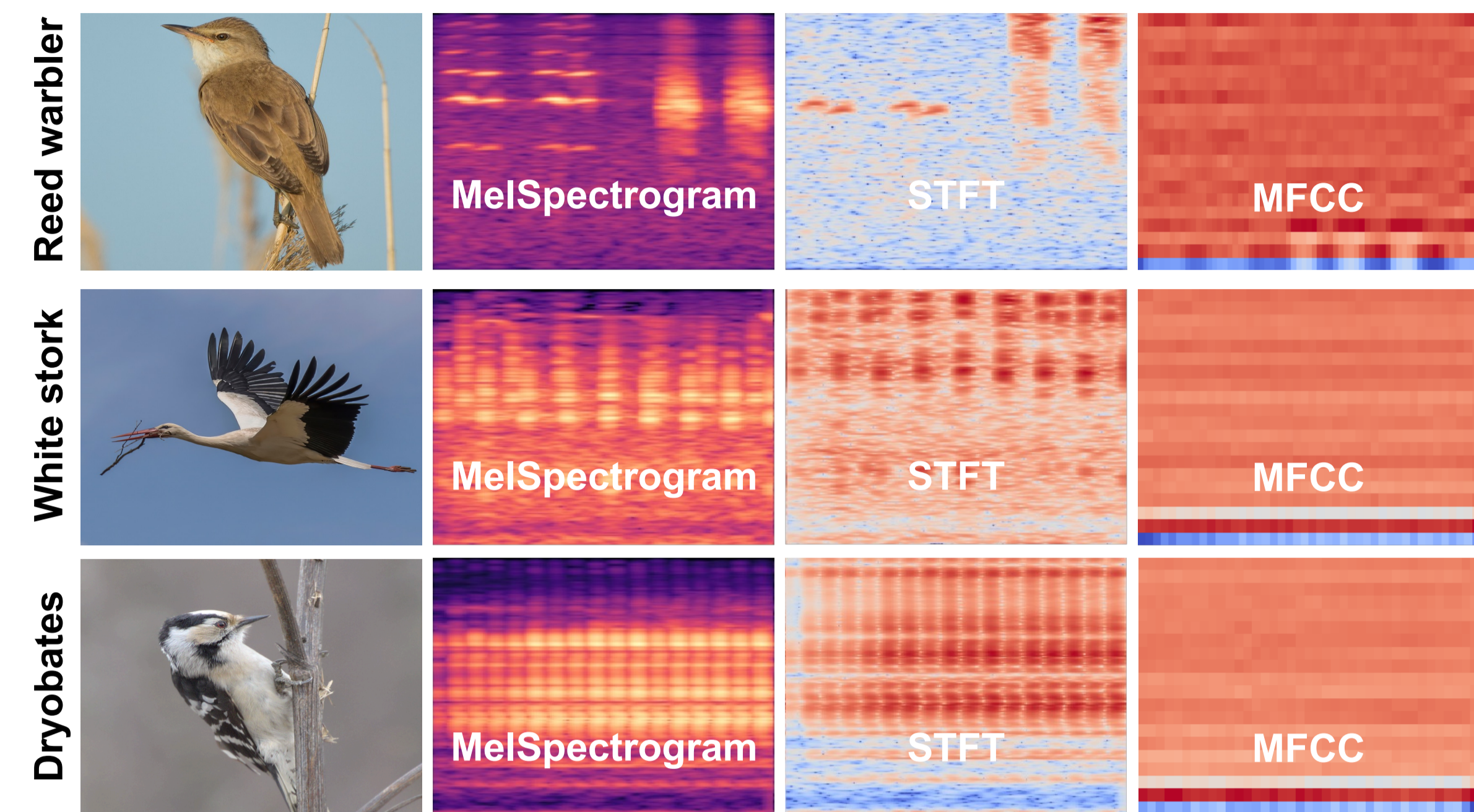
Methodology

- Synergize multiple features to yield efficient and accurate classification based on different fusion strategies.
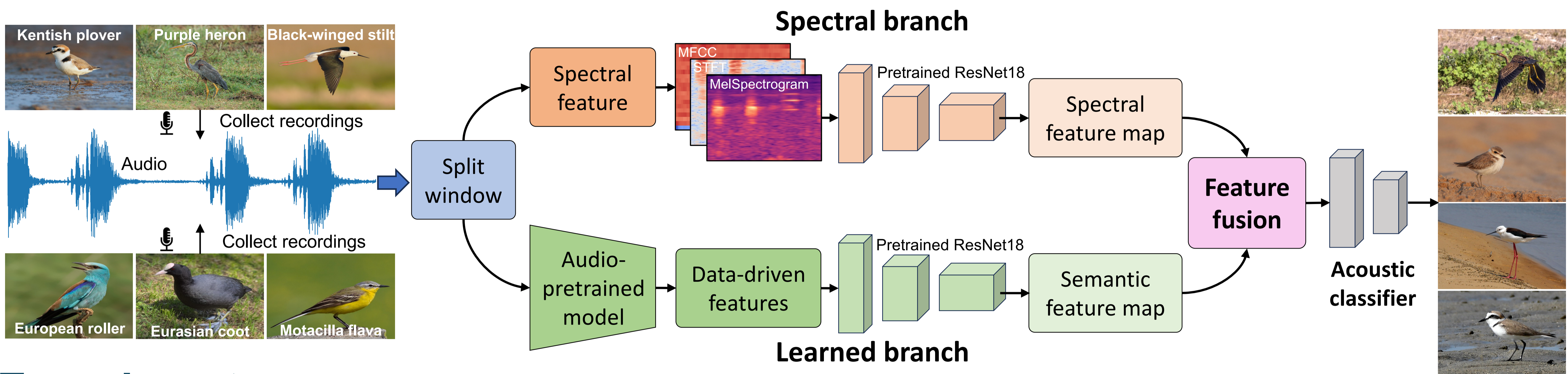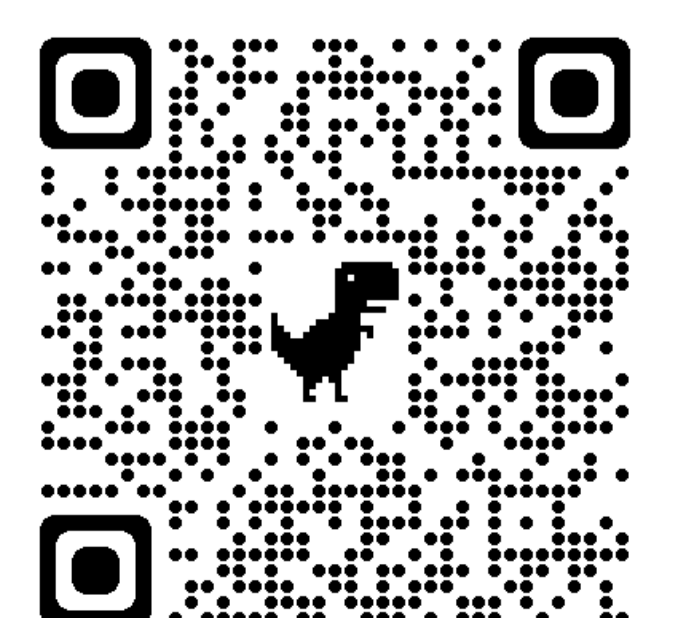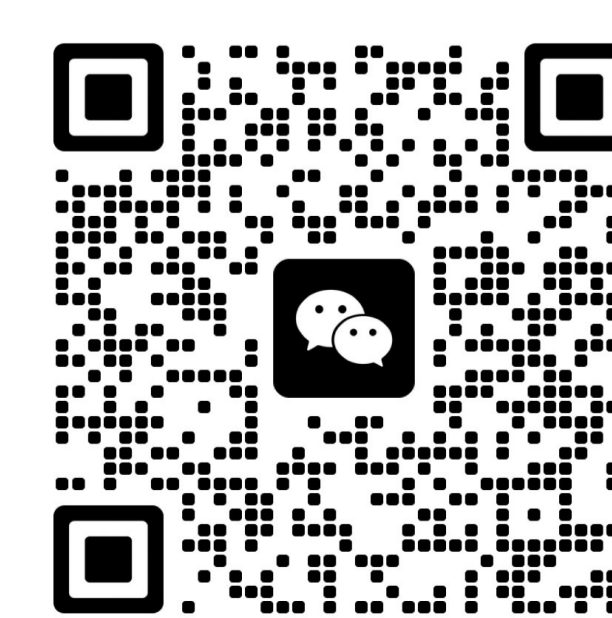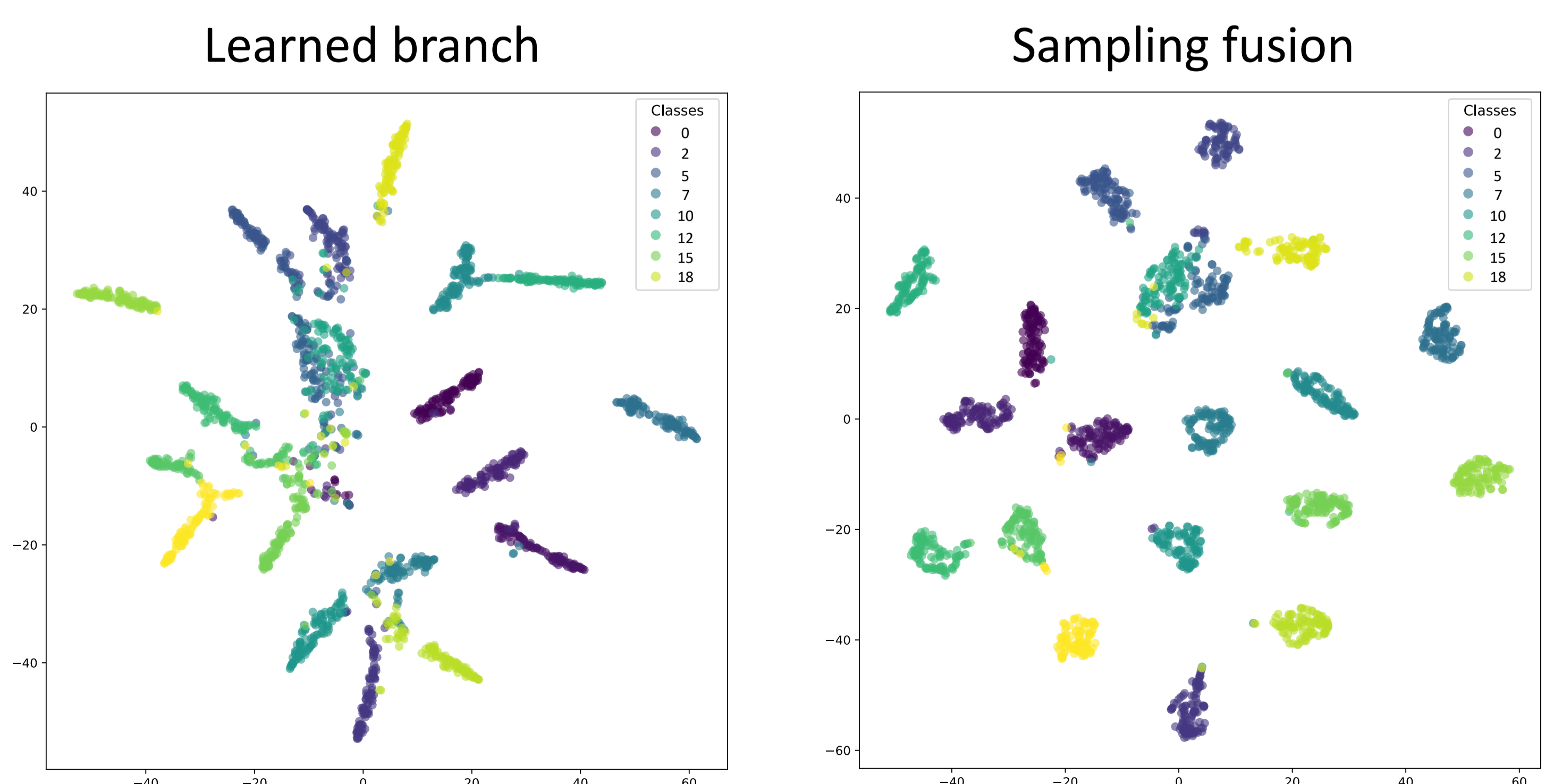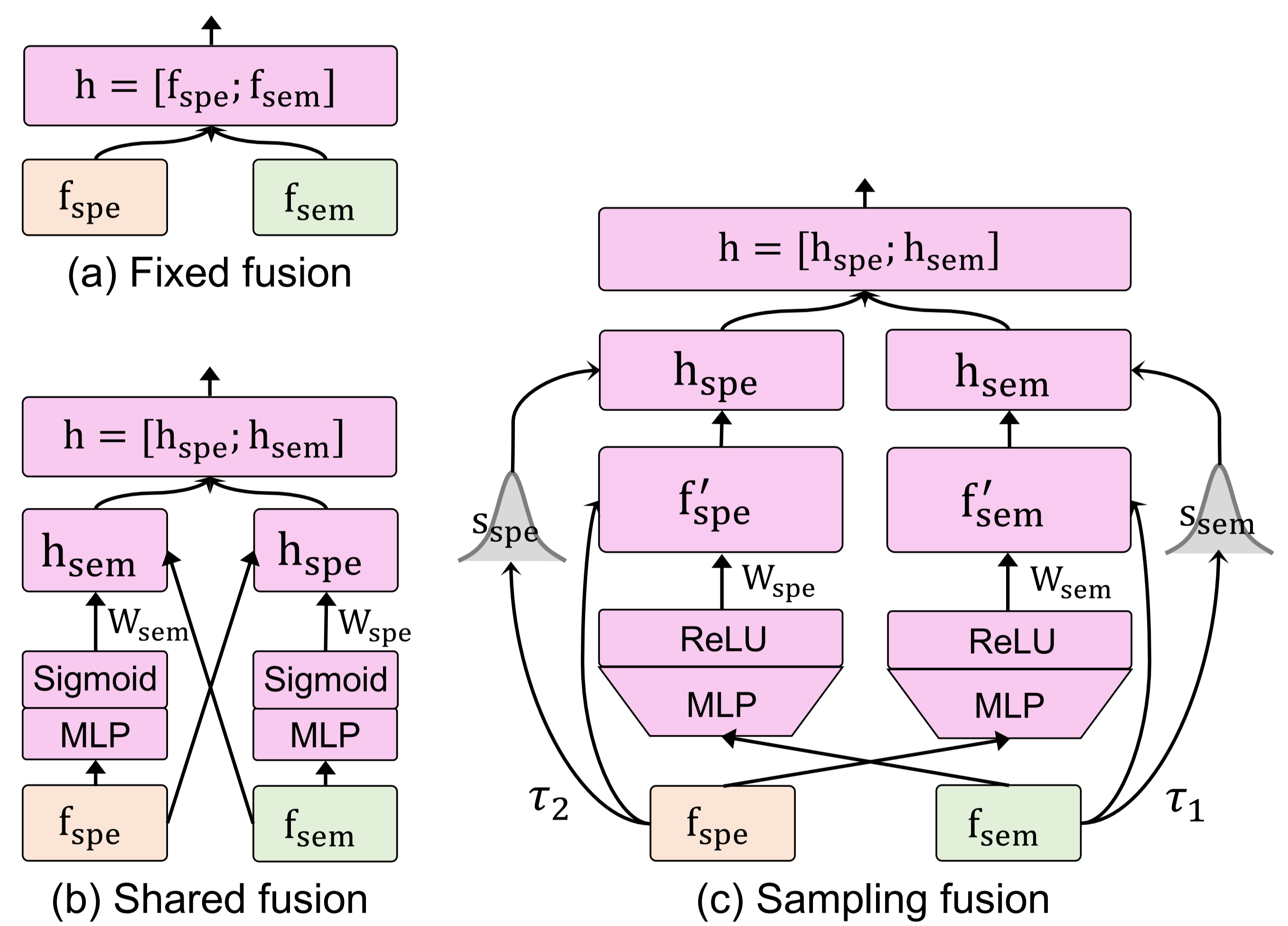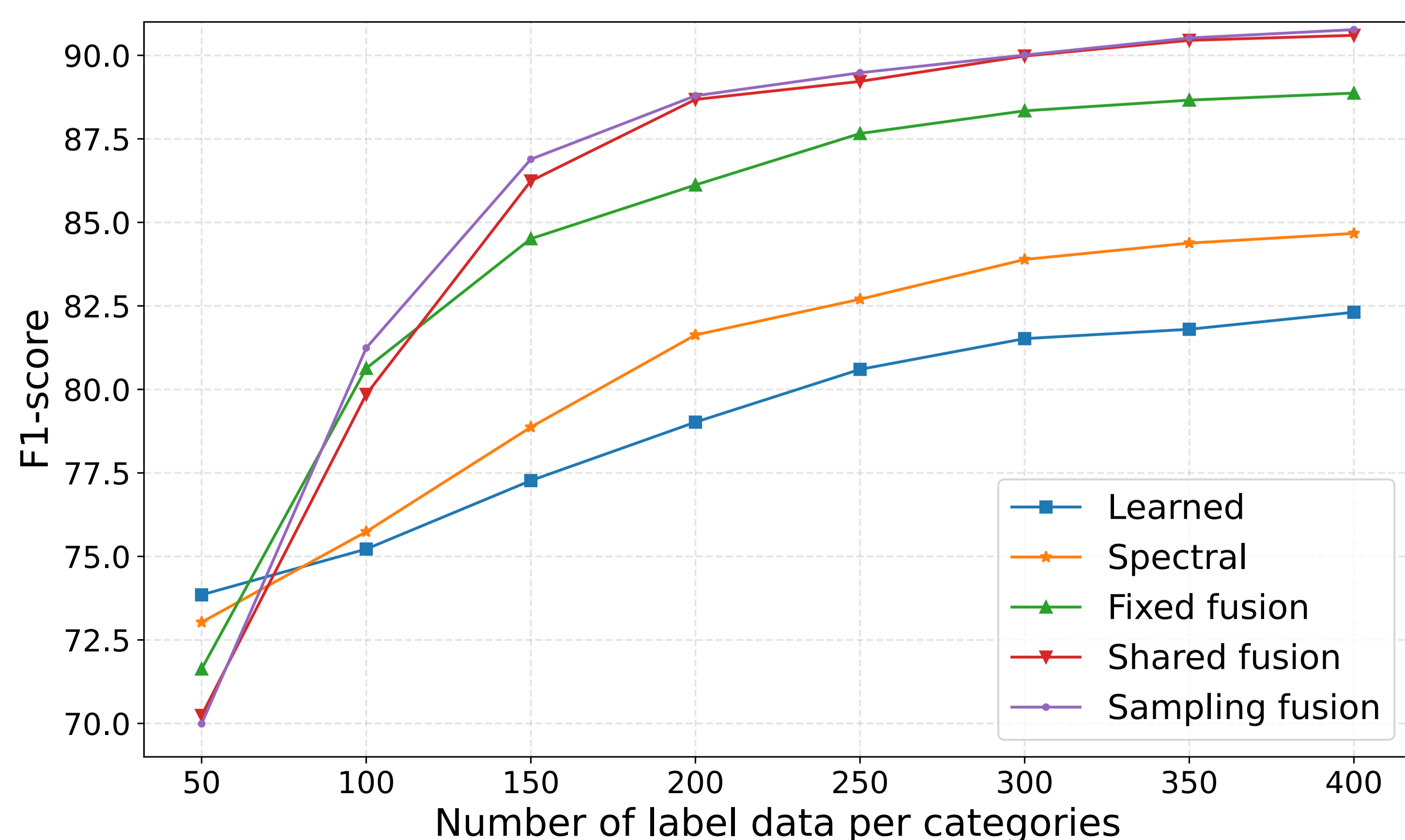

Experiments

- Totally 20 classes and 200 labeled samples per class in real dataset.

Branch	Model	Feature	Metric				Param (#)
			Acc↑ (%)	Prc↑ (%)	Rec↑ (%)	F1↑	
Baseline [21]	ResNet50	MEL	77.70	74.68	72.33	73.60	542,292
	ResNet18	MEL*	76.10	72.23	71.98	72.42	
Learned	ResNet18	LEAF	83.60	**86.16**	77.14	78.50	
		BEATs	83.23	80.56	78.61	79.02	
Spectral	ResNet18	MEL*	83.68	81.52	74.43	76.78	68,500
		STFT	83.45	83.25	73.98	74.99	
		MFCC	83.31	81.17	72.36	74.26	
		ALL	**84.02**	83.11	**82.70**	**81.63**	

Fusion	Branch	Metric				Param (#)
		Acc↑ (%)	Prc↑ (%)	Rec↑ (%)	F1↑	
Fixed	LEAF+BEATs	83.48	85.34	79.06	79.88	134,036
	Spectral+LEAF	83.33	87.78	84.24	83.71	
	Spectral+BEATs	83.65	86.92	86.28	86.12	
Shared	LEAF+BEATs	84.62	85.59	80.89	81.90	2,260,964
	Spectral+LEAF	**87.29**	90.38	85.81	86.53	
	Spectral+BEATs	86.54	89.60	88.29	88.68	
Sampling	LEAF+BEATs	83.37	84.79	80.03	81.59	4,262,708
	Spectral+LEAF	86.06	**91.98**	86.14	87.48	
	Spectral+BEATs	85.70	89.69	**88.77**	**88.79**	

(a) Fixed fusion

(b) Shared fusion

(c) Sampling fusion

Learned branch   Sampling fusion

WhatsApp   WeChat   Homepage