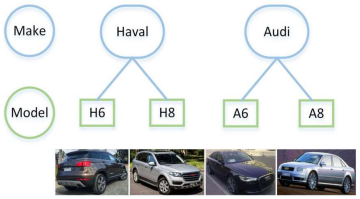
Semantic Bilinear Pooling for Fine-Grained Recognition

Xinjie Li, Chun Yang, Song-Lu Chen, Chao Zhu, Xu-Cheng Yin*
University of Science and Technology Beijing

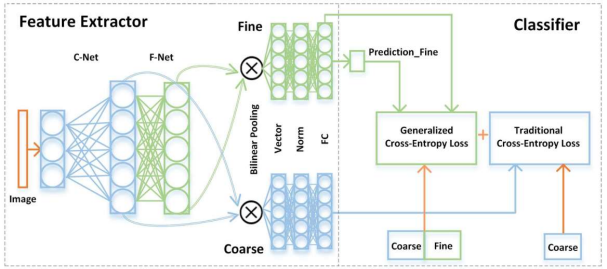
Introduction

- **Coarse-to-fine classification**
- **Hierarchical label tree**



Methods

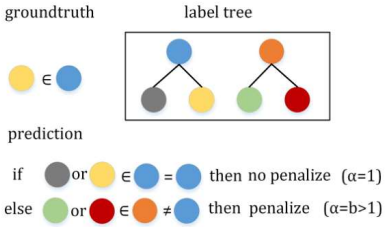
- **Two-Branch Network**



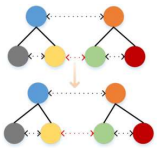
- **Generalized Cross-Entropy Loss Function**

$$Loss = -\sum_{i=1}^{n}\sum_{c=1}^{C} \alpha_i y_{ic} \log(a_{ic})$$

Illustration



Implicit meaning

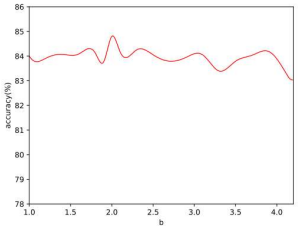


Experiments

- **Datasets**

Datasets	#Coarse	#Fine	#Train	#Val
CompCars [5]	75	431	16016	14939
StanfordCars [4]	49	196	8144	8041
CUBbirds [25]	70	200	5994	5794
Aircrafts [26]	70	100	6667	3333

- **Effect of α**



- **Effect of loss weight**

r	1 : 1	3 : 2	7 : 3	4 : 1	9 : 1
Accuracy(%)	83.3	84.6	84.8	84.0	82.9

- **Ablation Analysis**

Backbone	Method	TB	GCE	CompCars	StanfordCars	Birds	Aircrafts
VGG16	CBP [19]	–	–	∗94.0	∗90.8	84.0	∗87.4
	Ours w/o GCE	✓		94.3	91.3	84.3	88.2
	Ours w/o TB		✓	94.7	91.6	84.5	88.9
	Ours(CBP)	✓	✓	**95.2**	**91.9**	**84.8**	**89.3**
	iSQRT-COV [24]	–	–	∗96.3	92.5	87.2	90.0
	Ours w/o GCE	✓		96.7	92.9	87.4	90.6
	Ours w/o TB		✓	96.8	92.9	87.5	90.8
	Ours(iSQRT-COV)	✓	✓	**97.0**	**93.2**	**87.8**	**91.1**
ResNet50	iSQRT-COV [24]	–	–	∗96.9	92.8	88.1	90.0
	Ours w/o GCE	✓		97.3	93.5	88.5	90.5
	Ours w/o TB		✓	97.4	93.7	88.3	91.2
	Ours(iSQRT-COV)	✓	✓	**97.8**	**94.3**	**88.9**	**91.7**

- **Comparison with Semantic Methods**

Backbone	Method	S-Cars	Birds
VGG16	BGL [7]	86.0	75.9
	Ours(CBP)	91.9	84.8
	Ours(iSQRT-COV)	**93.2**	**87.8**
ResNet50	CLC [6]	–	79.3
	HSE [8]	–	88.1
	Ours(iSQRT-COV)	**94.3**	**88.9**

- **Comparison with State-of-the-Art Methods**

Backbone	Method	CompCars	StanfordCars	Birds	Aircrafts
VGG16	FCAN [2]	–	89.1	82.0	–
	LRBP [18]	–	90.9	84.2	87.3
	KP [20]	–	92.4	86.2	86.9
	iBCNN [17]	–	92.0	85.8	88.5
	G^2DeNet [22]	–	92.5	87.1	89.0
	HIHCA [21]	–	91.7	85.3	88.3
	MoNet [23]	–	90.8	85.7	88.1
	SWP [11]	95.3	90.7	–	–
	BCNN [3]	∗93.0	90.6	84.0	86.9
	CBP [19]	∗94.0	∗90.8	84.0	∗87.4
	iSQRT-COV [24]	∗96.3	92.5	87.2	90.0
	Ours(CBP)	95.2	91.9	84.8	89.3
	Ours(iSQRT-COV)	**97.0**	**93.2**	**87.8**	**91.1**
VGG19	RACNN [12]	–	92.5	85.3	88.2
	MACNN [31]	–	92.8	86.5	89.9
ResNet50	SWP [11]	97.5	92.3	–	–
	NTS [10]	–	93.9	87.5	91.4
	MAMC [9]	–	93.0	86.5	–
	DFL [32]	–	93.1	87.4	91.7
	KP [20]	–	91.9	84.7	85.7
	iSQRT-COV [24]	∗96.9	92.8	88.1	90.0
	Ours(iSQRT-COV)	**97.8**	**94.3**	**88.9**	**91.7**