



BranchConnect: Image Categorization with Learned Branch Connections

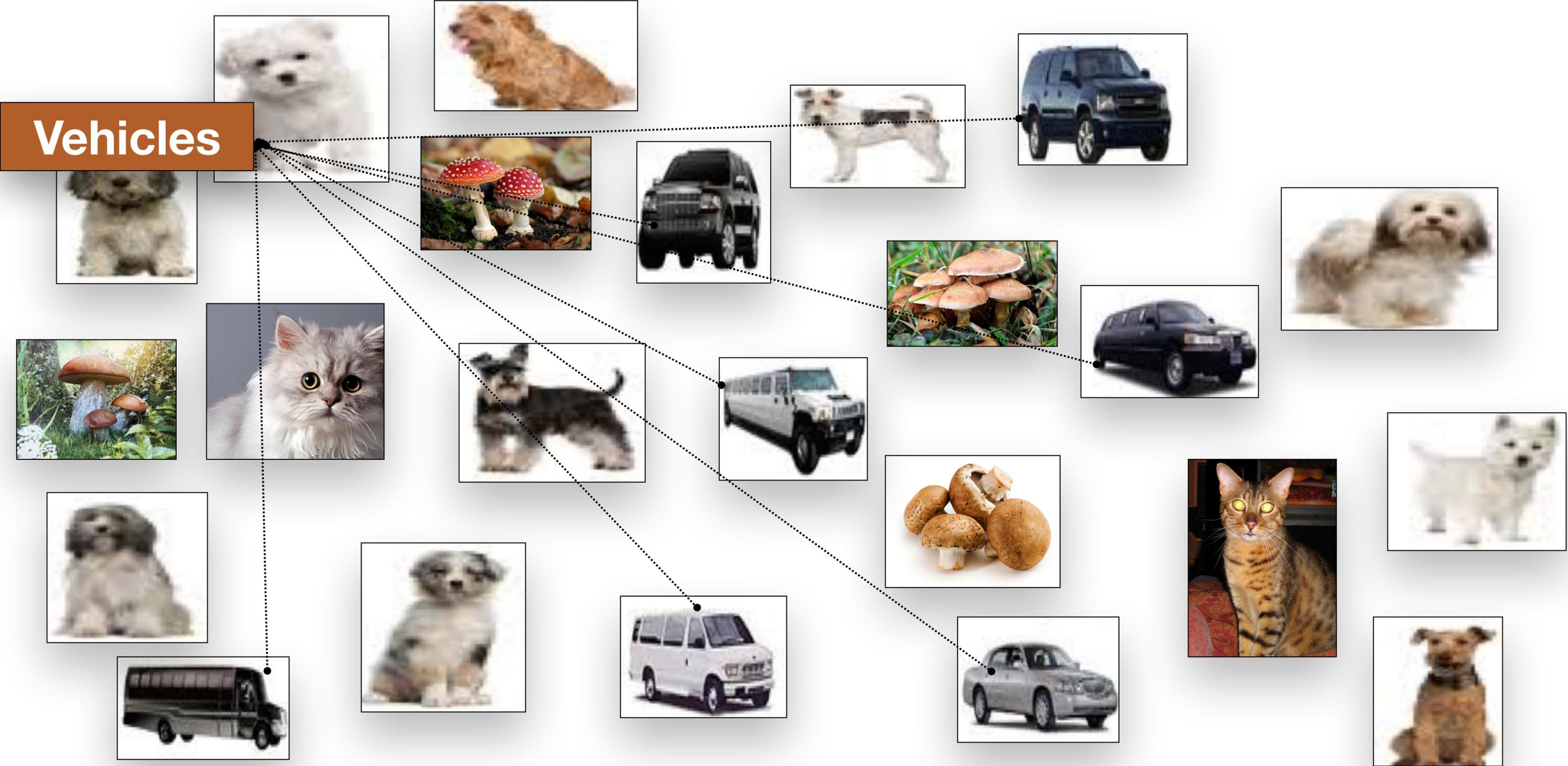
Karim Ahmed, Lorenzo Torresani



Classification in large scale Datasets



Classification in large scale Datasets

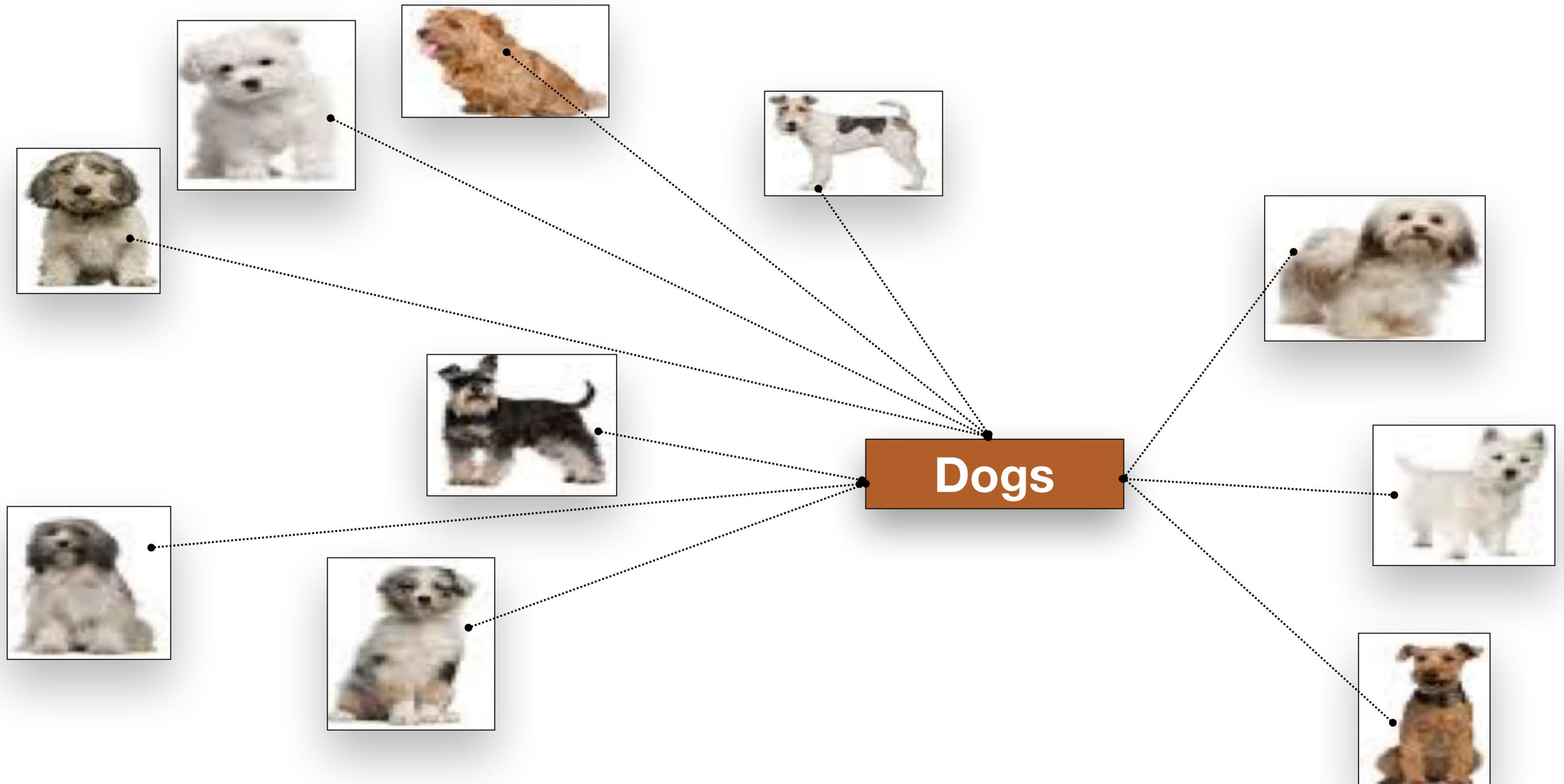


Classification in large scale Datasets

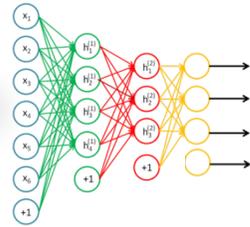
Cats



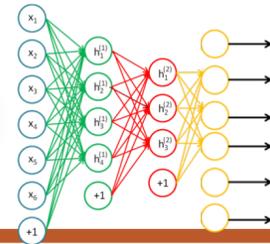
Classification in large scale Datasets



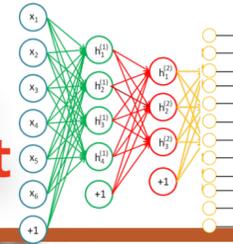
Mushrooms Expert



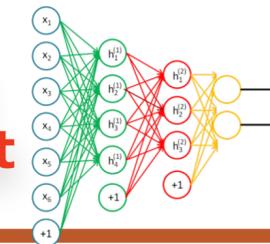
Vehicles Expert



Dogs Expert

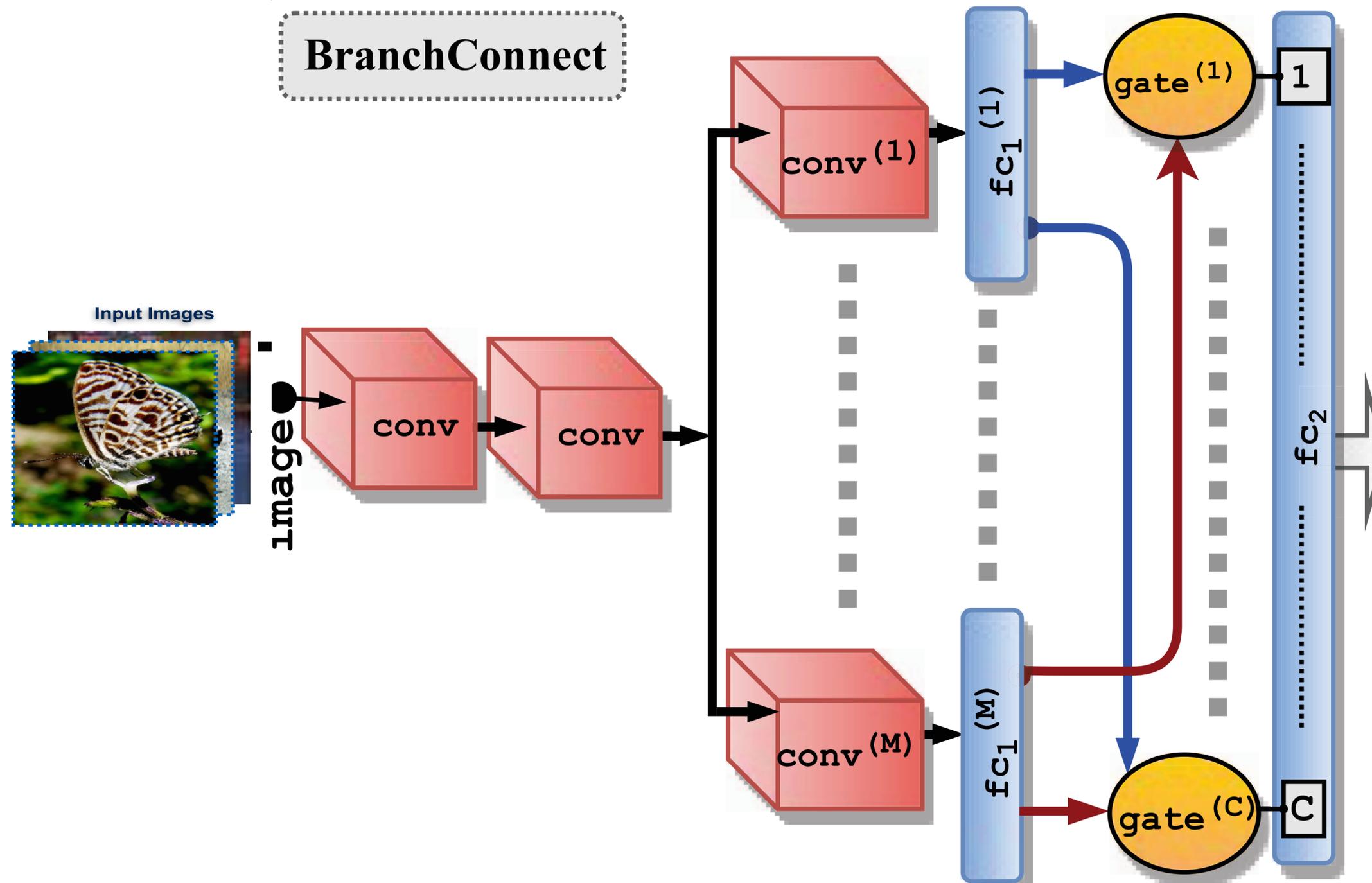


Cats Expert



Proposed Approach

End-to-end learning of separate **visual features (branches/experts)** for the different classes to distinguish.



- C classes
- M branches ($M < C$)
- C learned gates determine for each class the subset of features (branch/expert) to use.
- BranchConnect Can be built from any Convolutional architecture (AlexNet, NIN, ResNet, ..etc)

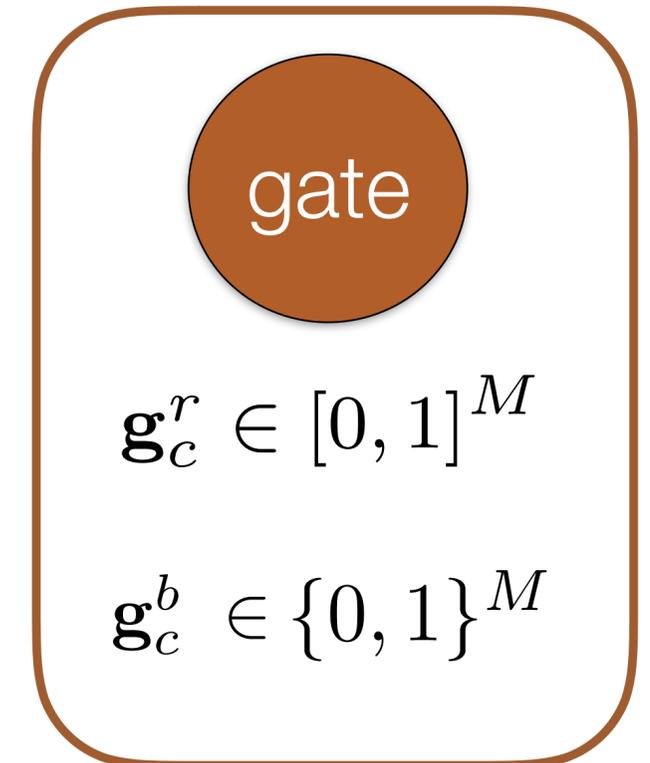
Technical Approach

Goal: learn branch connectivity to classes from data by optimizing training objective

□ Minimize $\ell(\theta, \mathbf{g})$ using backpropagation over \mathbf{g} and θ

□ During training, we update auxiliary real-valued gates $\mathbf{g}_c^r \in [0, 1]^M$

□ Constrain the number of active branch connections per class to be a constant, \mathbf{K} (a hyperparameter)



Technical Approach

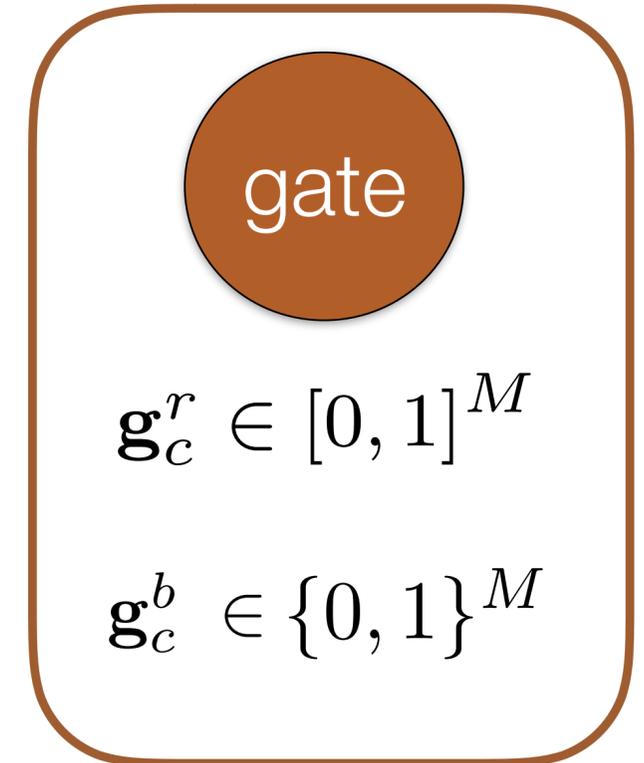
Forward propagation:

1. Stochastically binarize $\mathbf{g}_c^r \in [0, 1]^M$ into $\mathbf{g}_c^b \in \{0, 1\}^M$ s.t. $\sum_{m=1}^M g_{c,m}^b = K$
2. Perform forward pass using binary gates $\mathbf{g}_c^b \in \{0, 1\}^M$

$$F_c = \sum_{m=1}^M g_{c,m}^b \cdot E_m$$

c^{th} neuron in the last fully connected layer

m^{th} branch activations



Parameter update:

Compute $\frac{\partial \ell}{\partial g_{c,m}^b}$, and update real-valued masks $g_{c,m}^r$

$$g_{c,m}^r \leftarrow \text{clip}(g_{c,m}^r - \eta \cdot \frac{\partial \ell}{\partial g_{c,m}^b})$$

Datasets

CIFAR100

- 100 classes (**C = 100**)
- 32X32 images
- 50,000 training examples
- 10,000 testing examples.



CIFAR10

- 10 classes (**C = 10**)
- 32X32 images
- 50,000 training examples
- 10,000 testing examples.

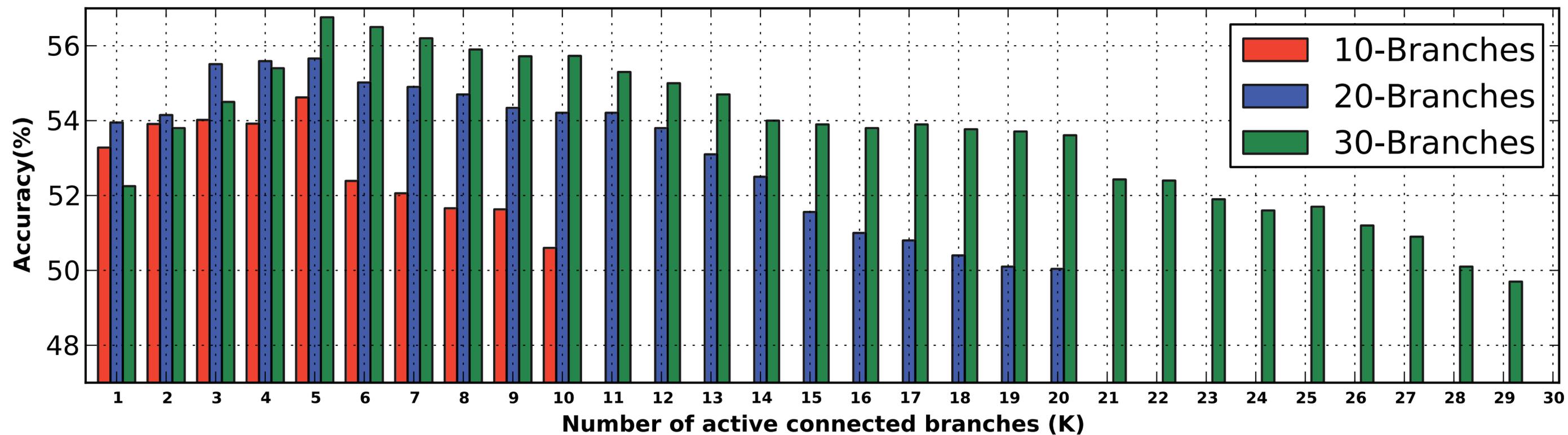
ImageNet

- 1000 classes (**C = 1000**)
- 256X256 images
- 1.28M training examples
- 50,000 validation examples.

IM  GENET



Experiments on CIFAR100



Experiments on CIFAR100

	Model	Accuracy %	
AlexNet	Baseline	44.3	
	BranchConnect: K=1, M=10	53.28	
	BranchConnect: K=5, M=10	54.62	+10.32%
Network IN Network	Baseline	64.73	
	BranchConnect: K=1, M=10	66.10	
	BranchConnect: K=5, M=10	66.45	+1.72%

Experiments on CIFAR100

		Model	Accuracy %	
ResNet56		Baseline	70.71	
		BranchConnect: K=1, M=10	71.24	
		BranchConnect: K=5, M=10	71.98	+1.72%
ResNet56-4X		Baseline	73.12	
		BranchConnect: K=1, M=10	75.55	
		BranchConnect: K=5, M=10	75.72	+2.6%

Experiments on CIFAR10

	Model	Accuracy %	
AlexNet	Baseline	76.86	
	BranchConnect: K=3, M=5	82.84	+5.98%
ResNet56	Baseline	92.04	
	BranchConnect: K=1, M=10	92.46	+0.42%

Experiments on ImageNet

	Model	Accuracy %	
AlexNet	Baseline	58.71	
	BranchConnect: K=5, M=10	63.49	+4.78%
ResNet50	Baseline	76.15	
	BranchConnect: K=5, M=10	77.39	
	BranchConnect: K=8, M=15	77.68	+1.53%
ResNet101	Baseline	77.37	
	BranchConnect: K=5, M=10	78.19	+0.82%

Thank you