ECCV 2018
European Conference on Computer Vision
8 – 14 September 2018 | Munich, Germany

NUS National University of Singapore

Dynamic Conditional Networks for Few-Shot Learning

Jian Zhao[1,2*], Fang Zhao[1*], Shuicheng Yan[1,3], Jiashi Feng[1]
* Indicates Equal Contribution
[1]National University of Singapore, [2]National University of Defense Technology, [3]Qihoo 360 AI Institute
zhaojian90@u.nus.edu, zhaofang0627@gmail.com, {eleyans, elefjia}@nus.edu.sg

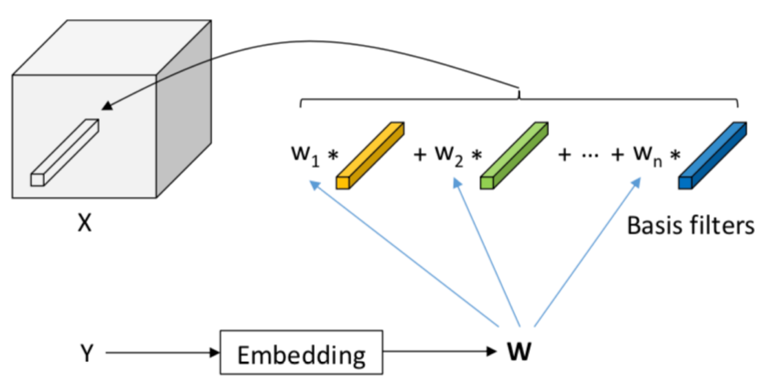
Abstract

This paper proposes a novel Dynamic Conditional Convolutional Network (DCCN) to handle conditional few-shot learning, *i.e*, only a few training samples are available for each condition. DCCN consists of dual subnets: DyConvNet contains a dynamic convolutional layer with a bank of basis filters; CondiNet predicts a set of adaptive weights from conditional inputs to linearly combine the basis filters. In this manner, a specific convolutional kernel can be dynamically obtained for each conditional input. The filter bank is shared between all conditions thus only a low-dimension weight vector needs to be learned. This significantly facilitates the parameter learning across different conditions when training data are limited. We evaluate DCCN on four tasks which can be formulated as conditional model learning, including specific object counting, multi-modal image classification, phrase grounding and identity based face generation. Extensive experiments demonstrate the superiority of the proposed model in the conditional few-shot learning setting.

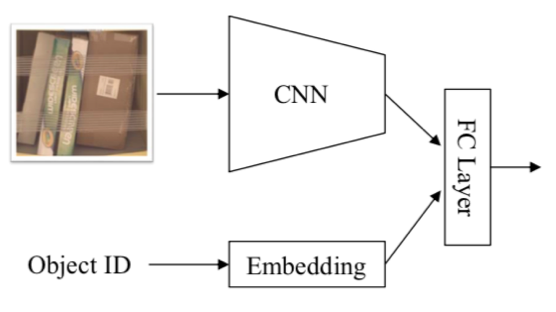
Contributions

- We present a novel and effective deep architecture, which contains a Dynamic Convolutional subNet (DyConvNet) and a Conditional subNet (CondiNet) that jointly per- form learning to learn in an end-to-end way. This deep architecture provides a unified framework for efficient conditional few-shot learning.
- The dynamic convolution is achieved through linearly combining the basis filters of the filter bank in the DyConvNet with a set of adaptive weights predicted by the CondiNet from conditional inputs, which is different from existing conditional learning approaches that combine the two inputs through direct concatenation.
- Our architecture is general and works well for multiple distinct conditional model learning tasks.

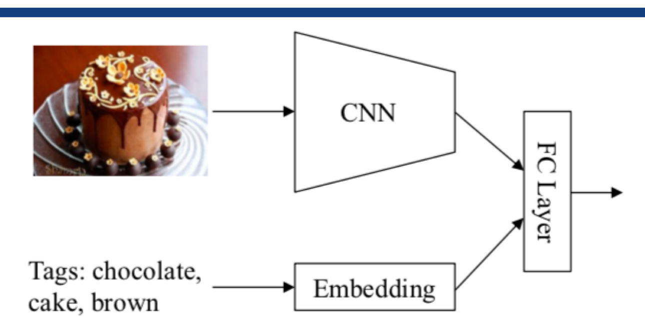
Method



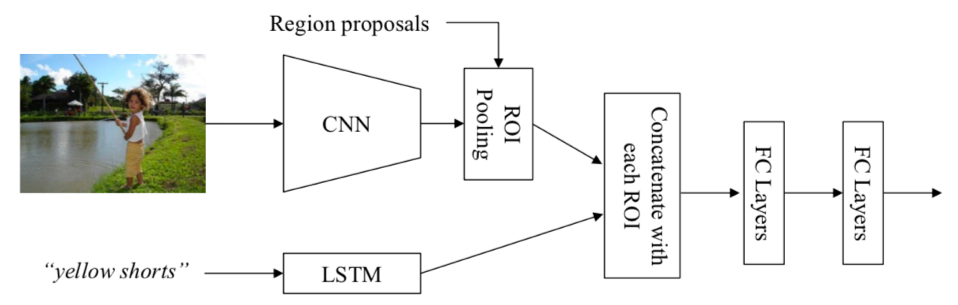
Dynamic convolutional layer of DyConvNet



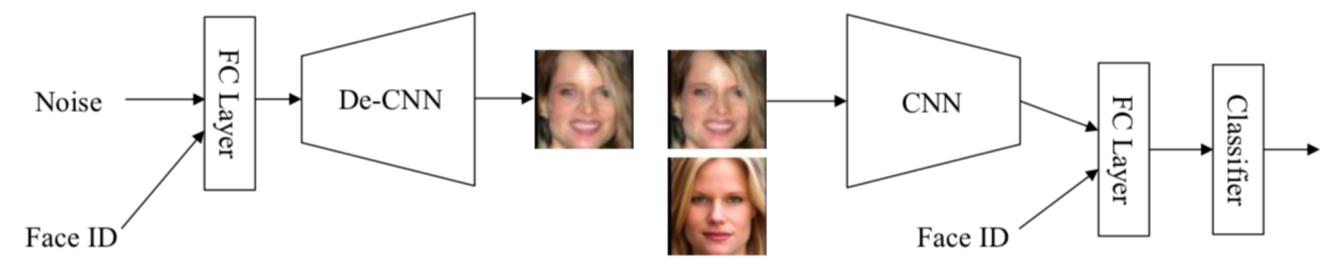
Plain Conditional Network



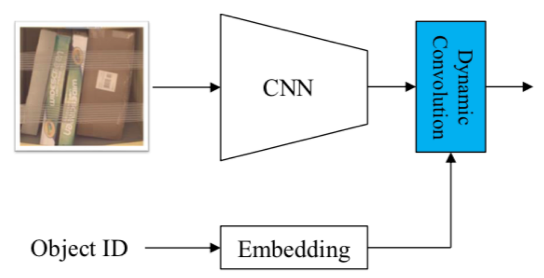
Plain Multi-modal Classification



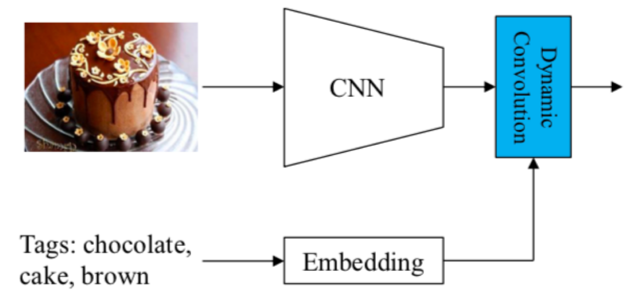
Dynamic Convolution Based Phrase Grounding



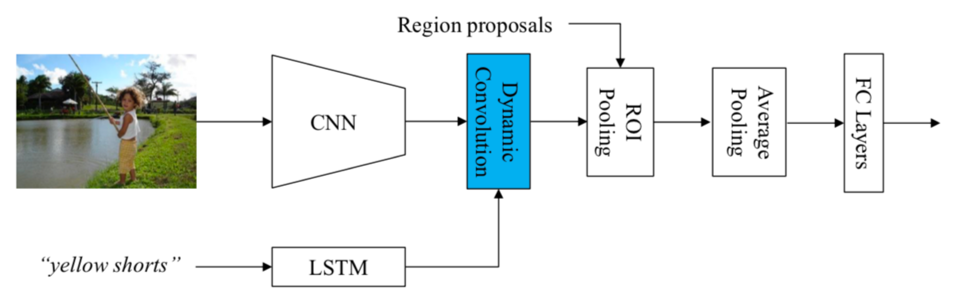
Conditional GAN



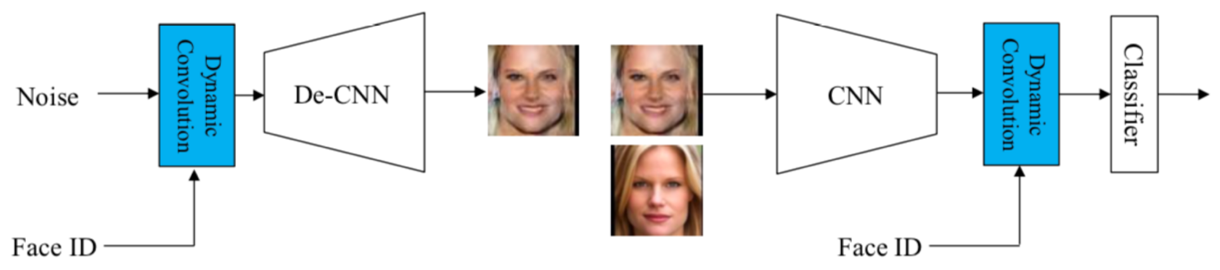
Dynamic Convolution Based Conditional Network



Dynamic Convolution Based Multi-modal Classification



Dynamic Convolution Based Phrase Grounding



Dynamic Convolution Based Conditional GAN

Results

Method	Dynamic			Plain
	4-D	8-D	16-D	
Identification	76.60%	**76.81%**	75.66%	74.48%
Verification	85.39%	**85.48%**	84.81%	84.87%

Specific Object Counting

Method	Acc		Precision@Coverage	
	Rank-1	Rank-5	P@C=0.99	P@C=0.95
Plain	45.70%	55.50%	5.00%	18.00%
Dynamic	**68.80%**	**74.80%**	**21.00%**	**71.00%**

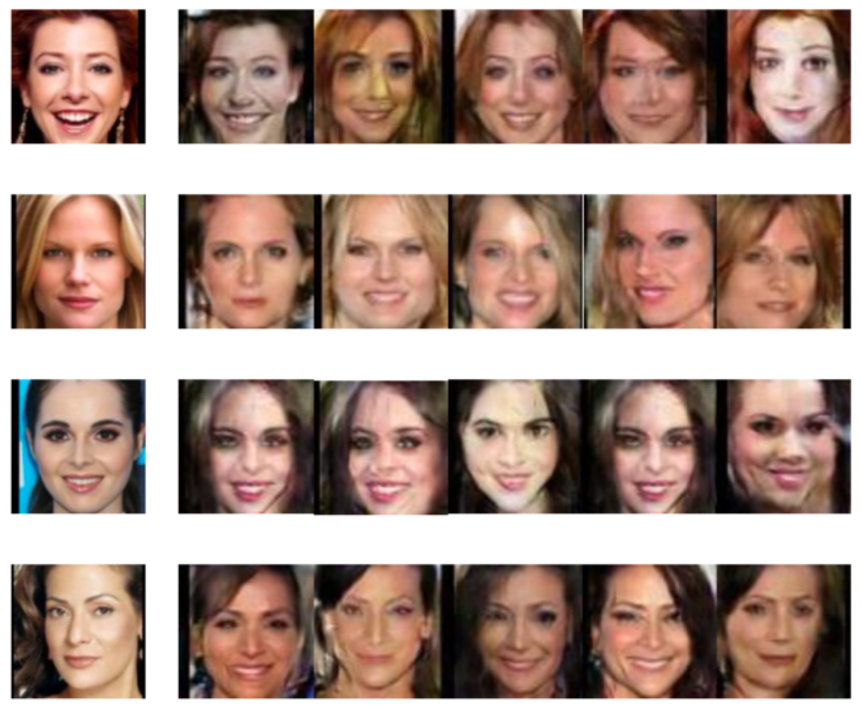
Identity Based Face Generation

Method		Dynamic			Plain
		32-D	64-D	128-D	
Tags	5k	65.17%	**65.53%**	65.20%	64.89%
	10k	65.13%	**66.06%**	65.60%	65.16%
	20k	65.49%	**65.77%**	65.43%	64.90%

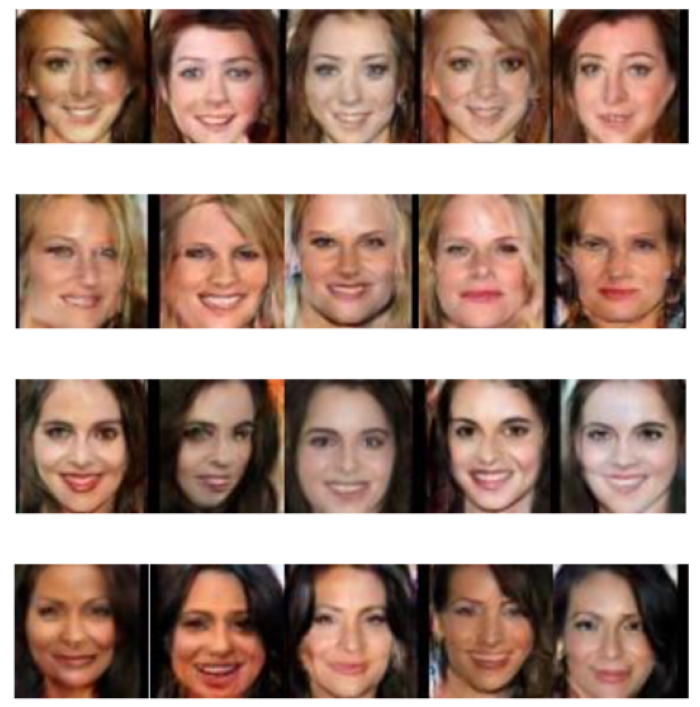
Multi-Modal Image Classification

Method	Dynamic			SMPL	Non-Linear SP	Ground R
	8-D	16-D	32-D			
Acc	50.18%	**50.65%**	50.52%	42.08%	43.89%	47.81%

Phrase Grounding



Conditional GAN



Dynamic Convolution Based Conditional GAN

Contact: Jian Zhao, Ph.D. candidate, Email: zhaojian90@u.nus.edu, Phone: (65) 9610 7176, Homepage: https://zhaoj9014.github.io/
Address: Vision and Machine Learning Lab, E4-#08-24, 4 Engineering Drive 3, National University of Singapore, Singapore 117583

Source Codes:
https://github.com/ZhaoJ9014/Dynamic-Conditional-Networks-for-Few-Shot-Learning.pytorch