

PROGRAMA DE PÓS-GRADUAÇÃO EM INFORMÁTICA

Nome: Frederico Guth

CPF: 273.723.818-86

Proposta de Projeto de Pesquisa

Título: Transferência de Aprendizado em Visão Computacional

Linha de Pesquisa: Sistemas de Computação

Área de Pesquisa: Visão Computacional

Transferência de Aprendizado em Visão Computacional

Frederico Guth

1 Introdução

Recentes avanços na área de Visão Computacional tornam possíveis aplicações que vêm merecendo atenção da mídia e público: são capazes de reconhecer de pessoas, lugares e objetos com acurácia super-humana[5], diagnosticar câncer de pele tão bem quanto dermatologistas[7], segmentar semanticamente cenas em tempo real para carros autônomos, localizar tumores em imagens de ressonância magnética, ver através de paredes usando sinais de rádio, entre tantas outras. Tal avanço apresenta um contraste extremo com como a comunidade se via há apenas 10 ou 20 anos:

”Apesar de como campo de pesquisa, [Visão Computacional] apresentar problemas interessantes e desafiadores, em termos de aplicações práticas bem sucedidas é decepcionante”

dizia T.S. Huang em 1996 [3].

O momento crucial para tal metóricoprogresso foi o resultado de Alex Krizhevsky et al.[4] no desafio *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC) de 2012 [1]. Em 8 anos de ILSVRC, o erro no reconhecimento de objetos diminuiu uma ordem de magnitude[5] e, em 2017, chegou a apenas 2,3%. Três desenvolvimentos simultâneos possibilitaram tal feito[2]: (a) redes convolucionais profundas, em que características visuais (*features*) são aprendidas dos dados ao invés de manualmente elaboradas; (b) barateamento do custo computacional para treinamento de algoritmos; (c) construção de bancos de imagens de larga escala com milhões de imagens e milhares de classes bem anotadas.

Neste contexto, é compreensível se iludir com o sensacionalismo e até pensar que *Deep Learning* para Visão Computacional seja uma área ”resolvida”—longe disso. Os melhores casos de sucesso foram desenvolvidos exclusivamente para uma tarefa, em um único domínio, ou seja, sobre a premissa básica que na prática, os dados com os quais o modelo será testado são do mesmo espaço de características (*feature space*) e possuem a mesma distribuição que os dados de treinamento[6]. Em outras palavras, temos algoritmos cada vez melhores que aprendem a encontrar padrões a partir de uma grande quantidade de dados rotulados, mas que ainda são deficientes na capacidade de generalizar para condições diferentes daquelas encontradas no treinamento.

Ao mesmo tempo que bancos de imagens de larga escala são um dos componentes chave para progresso, cada vez mais a necessidade de dados são um fator limitante para aplicação prática de inteligência artificial. Uma quantidade importante de comportamentos da natureza e atividades humanas obedecem a uma distribuição de cauda longa, o que torna bastante difícil a coleta representativa de dados. Há também o alto custo de rotulação: em diversas áreas de aplicação, e.g. na medicina, dados bem rotulados são extremamente difíceis de serem obtidos e pode-se levar anos para obter poucas dezenas de amostras. O mundo real é variado, os modelos encontrarão, na prática, diversos cenários para os quais não foram treinados.

Transferência de aprendizado nos permite lidar com a insuficiência de dados para treinamento, usando modelos pré-treinados para um outro domínio ou tarefa. Essa capacidade é absolutamente necessária para o uso de inteligência artificial em larga escala que vai além das tarefas e domínios para os quais a disponibilidade de dados rotulados é abundante. Em outras palavras, apesar de todo o sucesso, os modelos atuais atacam apenas os casos fáceis em termos de disponibilidade de dados. Para lidar com a cauda longa, precisamos transferir aprendizado.

2 Justificativa

- supervisionado inconteste sucesso - futuro depende de TL - quote Andrew Ng - x% dos n papers mais citados usam TL, mas não há um corpo didática bem consolidado. diferenças de nomenclatura - survey do Pan defasada

3 Objetivos

The authors were successful in advertising a promising idea in a very relevant problem. Due to weaknesses of the research, it is intriguing that it has already being cited 10 times and accepted to CVPR, whilst only in the Workshop, anyway. The fact it was sponsored by NVIDIA may explain some of this.

The main take a way is that tackling an important problem is an attention grabber.

This research could be better if it presented "apple to apples" comparisons. The insight is still relevant, though: domain adaptation by increasing variability of the input, what decreases the importance of the bias in the target domain of irrelevant features.

I believe it is worth investigating this problem and maybe trying to use the insight

4 Revisão da Literatura

- recente: cvpr 2018

- ζ transfer learning. survey ultrapassada.
- taskonomy

The authors were successful in advertising a promising idea in a very relevant problem. Due to weaknesses of the research, it is intriguing that it has already being cited 10 times and accepted to CVPR, whilst only in the Workshop, anyway. The fact it was sponsored by NVIDIA may explain some of this.

The main take a way is that tackling an important problem is an attention grabber.

This research could be better if it presented "apple to apples" comparisons. The insight is still relevant, though: domain adaptation by increasing variability of the input, what decreases the importance of the bias in the target domain of irrelevant features.

I believe it is worth investigating this problem and maybe trying to use the insight

5 Metodologia

The authors were successful in advertising a promising idea in a very relevant problem. Due to weaknesses of the research, it is intriguing that it has already being cited 10 times and accepted to CVPR, whilst only in the Workshop, anyway. The fact it was sponsored by NVIDIA may explain some of this.

The main take a way is that tackling an important problem is an attention grabber.

This research could be better if it presented "apple to apples" comparisons. The insight is still relevant, though: domain adaptation by increasing variability of the input, what decreases the importance of the bias in the target domain of irrelevant features.

I believe it is worth investigating this problem and maybe trying to use the insight

6 Plano de Trabalho

The authors were successful in advertising a promising idea in a very relevant problem. Due to weaknesses of the research, it is intriguing that it has already being cited 10 times and accepted to CVPR, whilst only in the Workshop, anyway. The fact it was sponsored by NVIDIA may explain some of this.

The main take a way is that tackling an important problem is an attention grabber.

This research could be better if it presented "apple to apples" comparisons. The insight is still relevant, though: domain adaptation by increasing variability of the input, what decreases the importance of the bias in the target domain of irrelevant features.

I believe it is worth investigating this problem and maybe trying to use the insight

7 Cronograma

The authors were successful in advertising a promising idea in a very relevant problem. Due to weaknesses of the research, it is intriguing that it has already being cited 10 times and accepted to CVPR, whilst only in the Workshop, anyway. The fact it was sponsored by NVIDIA may explain some of this.

The main take a way is that tackling an important problem is an attention grabber.

This research could be better if it presented "apple to apples" comparisons. The insight is still relevant, though: domain adaptation by increasing variability of the input, what decreases the importance of the bias in the target domain of irrelevant features.

I believe it is worth investigating this problem and maybe trying to use the insight

Bibliografia

- [1] Ian J. Goodfellow, Yoshua Bengio, and Aaron C. Courville. *Deep Learning*. Adaptive computation and machine learning. MIT Press, 2016.
- [2] Grant Van Horn and Pietro Perona. The devil is in the tails: Fine-grained classification in the wild. *CoRR*, abs/1709.01450, 2017.
- [3] T Huang. *Computer vision: Evolution and promise*, 1996.
- [4] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.
- [5] J. Deng L. Fei-Fei. Where have we been? where are we going?, 2017. [Online; accessada 28 de Junho de 2018].
- [6] Sinno Jialin Pan and Qiang Yang. A survey on transfer learning. *IEEE Trans. on Knowl. and Data Eng.*, 22(10):1345–1359, October 2010.
- [7] Saining Xie, Ross B. Girshick, Piotr Dollár, Zhuowen Tu, and Kaiming He. Aggregated residual transformations for deep neural networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, July 21-26, 2017*, pages 5987–5995, 2017.