

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: reconhecimento de pessoas, lugares e objetos com acurácia super-humana, segmentação semântica de cenas em tempo real possibilitando carros autônomos, diagnóstico e segmentação de tumores em imagens de ressonância magnética, sistemas capazes de colocar o rosto de uma pessoa em personagens de videos, visão 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"[3], dizia T.S. Huang em 1996.

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]. Três desenvolvimentos simultâneos possibilitaram tal feito[2]:

1. redes convolucionais profundas, em que características visuais (*features*) são aprendidas dos dados ao invés de manualmente elaboradas;
2. o barateamento do custo computacional de treinamento com GPUs;
3. a disponibilidade de bancos de imagens de larga escala com milhões de imagens e milhares de classes bem anotadas.

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%.

O momento crucial na melhoria meteórica dos resultados em reconhecimento de objetos se deu em 2012, no desafio *ImageNet Large Scale Visual Recognition Challenge* (ILSVRC)[1]. O time liderado por Alex Krizhevsky foi o primeiro a usar redes neurais convolucionais profundas (RCPs) na competição e ganhou por larga margem[4]. Desde então, técnicas baseadas em RCPs tem sido as mais bem sucedidas para este problema. É importante salientar, entretanto, que tal feito não seria possível se não houvesse uma banco de imagens de larga escala, manualmente anotada e com centenas de imagens por classe, como a ImageNet.

During the past five years we have witnessed dramatic improvement in the performance of visual recognition algorithms [42]. Human performance has been approached or achieved in many instances. Three concurrent developments have enabled such progress: (a) the invention of ‘deep network’ algorithms where visual computation is learned from the data rather than hand-crafted by experts [14, 27, 29], (b) the design and construction of large and well annotated datasets [8, 11, 13, 30] supplying researchers with a sufficient amount of data to train learning-based algorithms, and (c) the availability of inexpensive and ever more powerful computers, such as GPUs [33], for algorithm training.

- eyes tell the brain
- avanço dos últimos anos
- citar paper cancer e pele
- perspectiva de história recente com péssimos resultados. survey.
- hype torna difícil colocar o pé no chão e ver que há muito ainda a ser entendido
- mas datasets continuam caros, difíceis e sujeitos a viés.
- recente: cvpr 2018
- ¿ transfer learning. survey ultrapassada.
- aplicação carros autônomos, segmentação 3D de MRIs
- taskonomy

1.1 Bancos de Imagens para Reconhecimento de Objetos

Bancos de imagens são vitais para a pesquisa em reconhecimento de objetos. Pode-se dizer que foram um componente chave para o meteórico progresso obtido nos últimos anos, não apenas como fonte de dados para treinamento, mas também como meio de comparação de resultados de pesquisa. Diante deste dilema, uma maneira de se avaliar o viés de bancos de imagens é checar a generalização entre bancos de imagens: por exemplo, treinar com imagens Pascal VOC e testar com imagens da ImageNet.

Entretanto, a questão da generalização de domínio é tratada como um caso especial do reconhecimento de objetos chamado adaptação de domínio e é menosprezada na grande maioria dos desenvolvimentos de bases de imagens. Praticamente não se encontra comparações de resultados em-domínio (*in-domain*) e entre-domínios (*cross domain*) para os algoritmos baseados em redes neurais convolucionais profundas.

2 Justificativa

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 insights of Style Transfer

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

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; accessed 28 de Junho de 2018].