Projeto de Conclusão de Curso

Aprendizado profundo com capacidade computacional reduzida: uma aplicação à quebra de captchas.

Diogo Felipe Félix de Melo

Área de Concentração: Aprendizado de Máquina.
Orientador(a): Pablo de Azevedo Sampaio

Recife, Maio/2018.

Documento de Projeto de Pesquisa

1 Identificacão

Aluno(a): Diogo Felipe Félix de Melo (diogoffmelo@gmail.com)
Orientador(a): Pablo de Azevedo Sampaio (prof.pablo.sampaio@gmail.com)
Título: Aprendizado profundo com capacidade computacional reduzida: uma aplicação à quebra de captchas.
Área de Concentração: Aprendizado de Máquina
Linha de Pesquisa: Redes Neurais de aprendizado profundo

2 Introdução

Modelos de aprendizado baseados em neurologia são conhecidos desde meados do século passado[1]. Das proposições iniciais até os dias de hoje, essa classe modelos tem evoluído em complexidade e técnicas de forma contínua, culminando em modelos com muitas camadas e níveis cada vez mais abstratos de representações (ver [2] para uma breve revisão histórica). Apesar dos avanços na área, foi apenas recentemente que modelos neurais começaram a redefinir o estado da arte, superando outras classes de algoritmos de aprendizado de máquina[3] e até mesmo alcançando performances sobre humanas[4]. Tais avanços foram possíveis devido a três fatores chaves: a viabilização de bases de treino cada vez maiores o aumento do poder computacional e o desenvolvimento de novas arquiteturas neurais, como redes convolucionais e redes recursivas.

A crescente melhoria de performance dos modelos de aprendizado profundo tem motivado estudos em áreas onde se é preciso distinguir computadores e humanos. CAPTCHAs [5] (do inglês Completely Automated Public Turing tests to tell Computers and Humans Apart) definem uma coleção de técnicas que tem como objetivo bloquear a ação de agentes autônomos na rede mundial de computadores. O subconjunto mais conhecido dessas técnicas talvez seja o de captchas baseados em texto[6]. Nesse tipo de desafio, uma imagem contendo uma sequência de caracteres é exibida. A validação é feita pela comparação entre o texto informado pelo usuário e a resposta correta. Em trabalhos recentes, foram relatadas acurácias próximos à humana em sequências formadas exclusivamente por números[7] ou por uma única fonte[8]. Para o problema geral de quebrar captchas baseados em texto, entretanto, modelos de aprendizado profundo ainda mostram desempenho inferior ao humano. Contudo, pesquisas recentes apontam para avanços claros nos próximos anos[9]. Em comum, esses modelos possuem a necessidade de muito poder computacional e/ou bases de dados extensivas. O treino dessas redes é tipicamente executado em clusters e/ou sistemas de computação sob demanda, com alto poder de paralelização e utilizando hardware de alto poder de processamento como GPUs e TPUs. Adicionalmente, As bases de treino comumente alcançam alguns

terrabytes e envolvem grandes operações de aquisição e/ou geração.

3 Problema de Pesquisa

Neste trabalho vamos estudar a viabilidade do treino e validação de redes de aprendizado profundo em computador com pode de processamento mais modesto do que os usualmente utilizados nos melhores resultados encontrados na literatura. Mas especificamente, investigaremos se é possível construir um modelo de aprendizado profundo para quebra de captchas de texto em um computador pessoal e ainda alcançar resultados próximos do estado da arte conhecido na literatura.

4 Justificativa

O estado da arte em redes de aprendizado profundo tem aberto portas para aplicações em áreas como processamento de texto [10], detecção de objetos [11] e jogos [4, 12]. Essas aplicações demandam por uma grande capacidade computacional, o que pode inviabilizar o acesso a essas novas tecnologias em realidades com orçamento mais baixo ou onde uma prototipação rápida e barata seja necessária. Com este trabalho esperamos demonstrar viabilidade da aplicação de modelos de aprendizado profundo para a quebra de captcha em realidades mais restritivas.

5 Objetivos

Objetivo Geral:

Testar a viabilidade do uso de modelos de aprendizado profundo em um computador pessoal.

Objetivos Específicos:

1. Investigar técnicas de aprendizado profundo aplicáveis ao problema.
2. Treinar e/ou validar modelos de aprendizado profundo em um computador pessoal para quebra de captcha.
3. Disponibilizar os resultados da experimentação de forma pública.

6 Etapas de Pesquisa

Durante a execução da pesquisa será realizada uma revisão da literatura as técnicas utilizadas em aprendizado profundo, com enfoque em problemas correlatos à extração

de texto de imagens (quebra de captcha). Experimentar modelos e arquiteturas progressivamente mais complexas, sendo o tempo de treino, uso de memória e acurácia na quebra de uma classe específica de captcha baseado em texto as variáveis de interesse.

Etapas:

1. Revisão literária.
2. Experimentação.
3. Confecção do TCC.
4. Apresentação do TCC.

7 Cronograma

Mês / Etapa	Maio	Junho	Julho
Revisão bibliográfica	X	X	X
Experimentos		X	
Confecção do TCC		X	X
Apresentação do TCC			X

Referências

[1] F. Rosenblatt. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, pages 65–386, 1958.

[2] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016.

[3] 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.

[4] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharshan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. Human-level control through deep reinforcement learning. *Nature*, 518(7540):529–533, February 2015.

[5] Luis von Ahn, Manuel Blum, Nicholas J. Hopper, and John Langford. Captcha: using hard ai problems for security. In *Advances in Cryptology, Eurocrypt*, volume 2656, pages 294–311, 05 2003.

[6] Y W Chow and W Susilo. Text-based captchas over the years. *IOP Conference Series: Materials Science and Engineering*, 273(1):012001, 2017.

[7] Ian J. Goodfellow, Yaroslav Bulatov, Julian Ibarz, Sacha Arnoud, and Vinay D. Shet. Multi-digit number recognition from street view imagery using deep convolutional neural networks. *CoRR*, abs/1312.6082, 2013.

[8] Dileep George, Wolfgang Lehrach, Ken Kansky, Miguel Lazaro-Gredilla, Christopher Laan, Bhaskara Marthi, Xinghua Lou, Zhaoshi Meng, Yi Liu, Huayan Wang, Alex Lavin, and D Scott Phoenix. A generative vision model that trains with high data efficiency and breaks text-based captchas. 358:eaag2612, 10 2017.

[9] Elie Bursztein, Jonathan Aigrain, Angelique Moscicki, and John C. Mitchell. The end is nigh: Generic solving of text-based captchas. In *WOOT*, 2014.

[10] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2*, NIPS'13, pages 3111–3119, USA, 2013. Curran Associates Inc.

[11] Joseph Redmon and Ali Farhadi. Yolo9000: Better, faster, stronger. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6517–6525, 2017.

[12] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, Yutian Chen, Timothy Lillicrap, Fan Hui, Laurent Sifre, George van den Driessche, Thore Graepel, and Demis Hassabis. Mastering the game of go without human knowledge. 550:354–359, 10 2017.