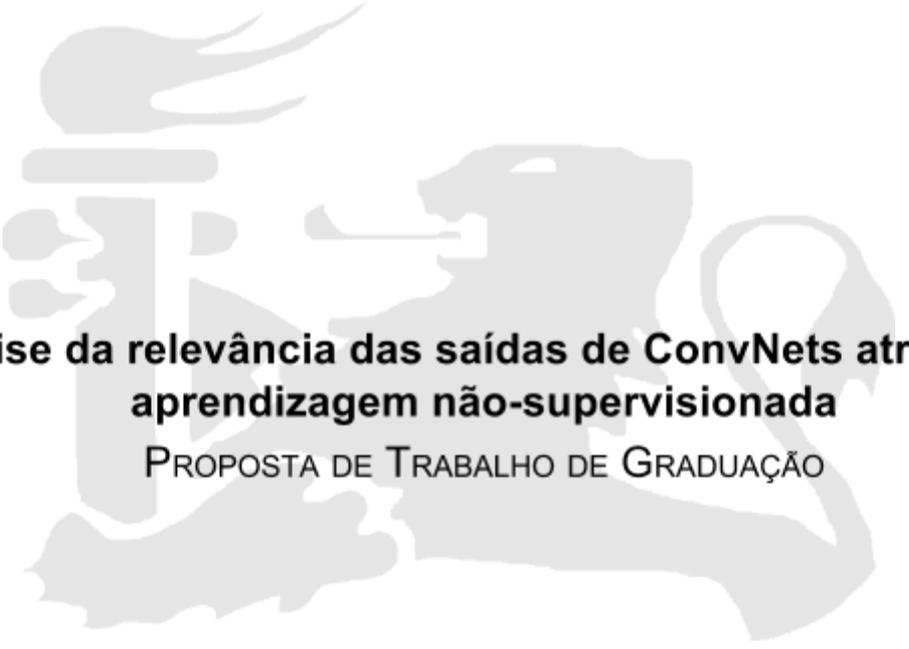




UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE INFORMÁTICA
GRADUAÇÃO EM ENGENHARIA DA COMPUTAÇÃO



**Análise da relevância das saídas de ConvNets através de
aprendizagem não-supervisionada**
PROPOSTA DE TRABALHO DE GRADUAÇÃO

Aluno: Heitor Rapela Medeiros (hrm@cin.ufpe.br)
Orientador: Hansenclever de França Bassani (hfb@cin.ufpe.br)
Área: Aprendizagem de Máquina

Resumo

A robótica nos últimos anos teve o crescimento impulsionado pelo avanço tecnológico e barateamento de sua produção, com isso, atividades repetitivas e complexas que eram executadas por pessoas estão sendo realizadas por robôs. Para fazer diversas atividades, os robôs necessitam detectar e reconhecer objetos de forma precisa, atualmente estão sendo estudadas várias técnicas de aprendizagem de máquina como redes neurais convolucionais com intuito de melhorar a visão do robô para com o mundo. O objetivo deste trabalho parte desta motivação para entender o funcionamento das camadas de autoencoders destas redes neurais, que vem dando ótimos resultados, como podemos ver na GoogLeNet, vencedora da competição da ImageNet [4]. Isso será feito através da análise das relevâncias através da técnica baseada em mapas auto-organizáveis LARFDSSOM [3], com o intuito de identificar quais as características mais importantes nos autoencoders para representar certas classes e com isto entender melhor o funcionamento do seu aprendizado.

Introdução

A robótica nos últimos anos teve o crescimento impulsionado pelo avanço tecnológico e barateamento de sua produção, com isso, atividades que eram executadas por pessoas estão sendo realizadas por robôs. As atividades desenvolvidas por robôs geralmente são repetitivas, demandam força e movimentos complexos. Exemplos destas atividades são vistas na competição Robocup@Home [1], que tem como objetivo desenvolver serviços e tecnologias robóticas assistivas para aplicações domésticas futuras, ou seja, produzir resultados relevantes para a sociedade, com o intuito de convencer o público à utilidade da robótica autônoma [1]. Dentre as sub-áreas contidas na competição, a Visão Computacional e o Reconhecimento de Objetos merecem atenção por facilitar a atuação dos robôs nos ambientes. O robô deve identificar e reconhecer objetos e pessoas. Os objetos geralmente são: mesas de jantar, cadeiras, sofás, armários, mesas com televisão, controle remoto, prateleiras com livros e geladeiras com latas e garrafas plásticas, que ficam na cozinha [2]. Vale ressaltar que a quantidade e o tipo dos objetos não são especificados previamente e que o local tem iluminação ambiente. Os objetos devem ser identificados e reconhecidos de maneira rápida e automática, no tempo de cada prova da competição. Visto que temos restrições de tempo de resposta, somos levados a utilizar técnicas mais robustas de aprendizagem de máquina, técnicas estas mais suscetíveis a variações no ambiente, se treinadas com tal objetivo, para realizar as atividades com mais adaptatividade, precisão e detalhe.

Para o problema proposto de detecção e reconhecimento de objetos, temos algumas abordagens supervisionada, as quais precisam de rótulos com a categoria dos objetos na etapa de treino, porém poucas no estado da arte que se baseiam em técnicas completamente não-supervisionada, ou seja, que não precisa destes rótulos. Atualmente, técnicas de detecção e reconhecimento de objetos mais robustas e com maiores taxas de acerto estão utilizando *Deep Learning* como método de aprendizagem. Como exemplo de *Deep Learning*, temos a GoogLeNet que é uma rede profunda, a qual foi treinada para a competição ImageNet. O fato da GoogLeNet utilizar várias convoluções paralelas em cada camada faz com que ela se diferencie das redes profundas tradicionais. Na GoogLeNet, as características são obtidas através dos autoencoders, que abstraem características básicas visuais, os quais são treinados, camada por camada, de maneira não-supervisionada. Após a obtenção das características pela rede GoogLeNet, no trabalho proposto, utilizaremos uma técnica de aprendizagem não-supervisionada baseada em Self-Organized Maps (SOM). A técnica escolhida se chama LARFDSSOM, proposta por BASSANI [3], o seu diferencial está em ser capaz de atribuir relevância às características e é utilizada para clusterização. Com isto, podemos descobrir quais as características tem mais relevâncias em determinadas classes, para podermos estudar como os autoencoders representam classes específicas.

Objetivos

O objetivo geral deste trabalho é estudar o funcionamento dos autoencoders nas redes neurais convolucionais. Para isto, utilizaremos o LARFDSSOM com o objetivo de dar relevância a certas características, e entender como os autoencoders escolhem as características mais importantes de maneira não-supervisionada para representar certas classes.

Objetivos específicos estão voltados em responder algumas perguntas como:

- Como funciona as camadas intermediárias dos autoencoders?
- Quais características são importantes no seu aprendizado para categorizar certas classes?

Metodologia

Como material de estudos e revisão bibliográfica utilizaremos artigos de área de aprendizagem de máquina, tendo como ponto de partida os artigos de [3], [5] e o livro [6].

Utilizaremos um banco de imagens Caltech-256 [7] e/ou em Imagenet [4], a princípio não vamos treinar a GoogLeNet, vamos utilizar a versão já treinada para extrair características de cada uma das imagens destas bases, em seguida utilizaremos o LARFDSSOM, para categorizar os objetos nas imagens e identificar as características relevantes para cada agrupamento. Faremos uma análise baseada em métricas comuns da área, com o objetivo de quantificar o quão bom está a categorização baseada neste método e o que afeta este resultado.

Referências

- [1] "@Home League - Robocup Wiki". wiki.robocup.org/@Home_League. N.p., 2017. Web. 3 Apr. 2017.
- [2] "Rules - Robocup@Home". robocupathome.org/rules. N.p., 2017. Web. 3 Apr. 2017.
- [3] Bassani, Hansenclever F., and Aluizio FR Araujo. "Dimension selective self-organizing maps with time-varying structure for subspace and projected clustering." *IEEE transactions on neural networks and learning systems* 26.3 (2015): 458-471.
- [4] "Imagenet". image-net.org. N.p., 2017. Web. 3 Apr. 2017.
- [5] Szegedy, Christian, et al. "Going deeper with convolutions." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2015.
- [6] "Deep Learning". deeplearningbook.org. N.p., 2017. Web. 3 Apr. 2017.
- [7] Griffin, G., Holub, A., and Perona, P. (2007). Caltech-256 object category dataset.

Possíveis Avaliadores

Prof. Um

Aluizio Fausto Ribeiro Araújo

Prof. Dois

Carlos Alexandre de Barros Mello

Assinaturas

Recife, ___ de _____ de _____

Heitor Rapela Medeiros

(Aluno)

Hansenclever de França Bassani

(Orientador)