

Reconhecimento de alumínio e outros metais em meio ao lixo utilizando a base de dados de matérias MINC

Renato Nobre

15/0146698

Departamento de

Ciência da Computação
Universidade de Brasília

Khalil Carsten

15/0134495

Departamento de

Ciência da Computação
Universidade de Brasília

Resumo—

1. Introdução

O alumínio é o metal mais abundante na natureza e recentemente utilizado em diversas aplicações, como em embalagens, transmissões elétricas, construção civil e elementos estruturais de meios de transporte. O seu uso cada vez mais constante prova sua relevância para a vida moderna, e consequentemente cada vez mais haverá descartes intensos do material para o lixo. Portanto, facilitar sua reciclagem geraria maiores benefícios econômicos, sociais e políticos. De acordo com o artigo publicado na Folha de São Paulo o Brasil é o maior reciclagem de latas de alumínio do mundo e que esta atividade de venda e reciclagem dessas latas injetaram cerca de 845 milhões de reais na economia.

Este trabalho propõe no entanto uma maneira de facilitar a identificação do alumínio e metais dentro de outros tipos de lixo. Porém percebe-se que o reconhecimento de metais em imagens de lixo do mundo real é uma tarefa desafiadora. Os materiais que podem conter no lixo contém uma diversa gama de textura, geometria, luminosidade e agrupamento, que combinados geram o problema particularmente difícil. Para tentar superar essa dificuldade, foi utilizado um conjunto

de imagens em grande escala de materiais em diversos ambientes. Esse conjunto de imagens denominado, materiais em banco de dados de contexto, MINC (do inglês, *Materials in Context Database*), foi criado no departamento de ciência da computação da Universidade de Cornell e possui mais de três milhões de figuras de imagens. O MINC tem sua magnitude de materiais mais ampla que os bancos de imagens anteriores e é bem separado em 23 categorias diversas.

Utilizamos a CNN já treinada pelo MINC disponibilizada pelos autores de [2] no site OpenSurface. No entanto, notou-se que o MINC possuía em quase todas as imagens de treinamento um material bem separado do fundo e centralizado na imagem. Além disso, as imagens de treinamento para metais, o nosso foco nesse projeto, resumiam-se em eletro-domésticos, pias de metal, aço inox, aparelhos de cozinha no geral. Isso prejudicou muito a análise de imagens de entulho de lixo onde os materiais mais comuns estão entre: latas de alimentos enlatados, eletrônicos e latas de refrigerante. Para contornar esse problema tivemos que executar um refinamento na CNN com imagens que coletamos do repositório <https://github.com/garythung/trashnet> em que o autor treinou uma CNN somente para reconhecimento de lixos.

Em resumo, este trabalho fornece duas

contribuições:

- 1) Uma forma de separar metais do resto do lixo utilizando o banco de imagens MINC com uma rede neural co-evolucionária que possui uma arquitetura AlexNet
- 2) Métodos de analisar a imagem em segmentos menores na rede neural AlexNet

2. Modelo

Ao realizar trabalhos de reconhecimento de imagens, um grande obstáculo é a construção de uma base consistente para treinar um modelo de aprendizagem. Reconhecimento de materiais requer que uma grande quantidade de imagens de diversas classes sejam separadas e utilizadas, para superar essa dificuldade foi utilizado o banco de imagens MINC.

O MINC foi utilizado devido ao fato de consistir de mais de três milhões de imagens, com 23 categorias de materiais diferentes; seu tamanho é grande o suficiente para métodos de aprendizagem conseguir generalizar casos de teste; suas diversas categorias são representadas todas por uma diversa quantidade de conteúdo; há diversidade da disposição do conteúdo analisado na imagem.

A construção do MINC usou como base uma banco de imagens já existente, o FMD (do inglês, *Flickr Material Database*). No entanto, precisava de mais conteúdo para a criação de um banco de imagens consistente e para isso integrou fotos de um site contendo imagens de fotos profissionais de design de interior.

Após a criação do MINC, os autores de *Material Recognition in the Wild with the Materials in Context Database* utilizaram uma rede AlexNet [2] para testar a eficiência do seu banco de imagens na criação de um modelo de aprendizagem. As fotos utilizadas do MINC foram divididas em casos de teste e treino e foram inseridas no modelo de forma como mostrado na Figura 1, 2, 3, 4. Consequentemente, como grande parte do MINC foi feito utilizando fotos de interior de casas, ao analisar que contém diversos materiais em grande

quantidades, como imagens de entulho ou sucata de lixo mostrado na Figura 5, 6, 7, 8 a rede perdia muita precisão. O excesso de materiais e a diferença de uniformidade da textura e luminosidade geraram uma mudança significante na média de acerto da AlexNet.



Figura 1. Metal 1



Figura 2. Metal 2

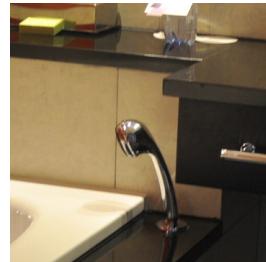


Figura 3. Metal 3

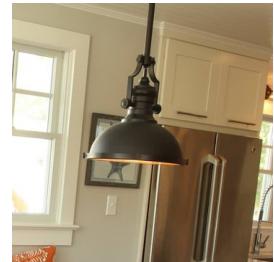


Figura 4. Metal 4

Torna-se evidente que ao analisarmos fotos de lixo com a AlexNet disponibilizada, a percentagem de acerto dos materiais metálicos reduz bastante. No entanto, como primeiro passo utilizamos de maneira simplista e direta a AlexNet pre-treinada, para uma classificação de metais usando como entrada *patches* de imagens com entulhos ou conjunto de materiais diversos. A Tabela 1 mostra o resultado desta classificação. Não obteve nenhum acerto e notas bem baixas, sendo 1 a maior nota possível e tal nota dividida entre todas as 23 classes. Portanto julgar uma imagem entre 23 possibilidade de classe se torna um tanto excessivo quando o objetivo é somente julgar se algo é ou não um metal.



Figura 5. Lixo 1



Figura 6. Lixo 2



Figura 7. Lixo 3



Figura 8. Lixo 4

Tabela 1. CLASSIFICAÇÃO ALEXNET MINC

Imagens	Resultado	Nota
Lixo 1	Vidro	0.0755
Lixo 2	Papel de Parede	0.0040
Lixo 3	Papel de Parede	0.0096
Lixo 4	Espelho	0.0316

3. Solução e Análise

4. Resultados

Para testar a solução proposta para o problema na seção anterior foram realizados diversos experimentos, com o objetivo de abranger todo o escopo do problema e analisar individualmente cada uma das implementações realizadas. Os experimentos foram realizados da forma que se descreve a seguir:

- Um experimento realizado com a AlexNet sem nenhuma modificação, do jeito fornecido pelo artigo [2]. Com esse experimento podemos usar de base para comparar o efeito das modificações feitas.

Tabela 2. TABELA DE ACERTO DA CLASSIFICAÇÃO

	Metal	Não Metal
Metal	0	6
Não Metal	0	6

- Experimento utilizando a técnica criada, Sliding Window, em cima da AlexNet não refinada, que verificará se a AlexNet realmente precisava de um refinamento ou se apenas uma abordagem mais local da imagem resolveria o problema.
- Um experimento realizado com a AlexNet refinada com a base de imagens TrashNet, para analisar se o efeito da grande quantidade de materiais pode ser tratado se a rede neural fosse treinada com arquivos com materiais diversos em grande quantidades.
- Experimento utilizando a técnica criada, Sliding Window, em cima da AlexNet refinada, para verificar o efeito na porcentagem de acerto com todas as mudanças implementadas.

Todos os experimentos foram realizados com o mesmo conjunto de teste, e seus resultados estão descritos e detalhados a seguir.

O primeiro experimento foi utilizado a AlexNet do jeito fornecido, pegando apenas uma janela da imagem e analisando em busca de metais. Como esperado, seu resultado foi extremamente insatisfatório nenhum metal foi encontrado em nenhuma das fotos fornecidas à rede neural (RN). Analisamos na tabela abaixo que em nenhum momento ele classificou algo como metal, dos 6 metais que continham ele classificou todos como não metais.

No segundo experimento foi realizado novamente o método Sliding Window, ainda com a AlexNet não refinada, para estudar se o refinamento era realmente necessário. Nota-se que não há nenhuma marcação, indicando que a RN não encontrou nada como metálico, um exemplo de uma imagem utilizada está apresentada abaixo.



Figura 9. Resultado da RN sem nenhum tratamento

Tabela 3. TABELA DE ACERTO DA CLASSIFICAÇÃO

	Metal	Não Metal
Metal	5	2
Não Metal	1	4

Portanto nota-se a necessidade de um refinamento na rede neural. Tal refinamento foi feito com o banco de imagens TrashNet. O terceiro experimento então é feito com a nova AlexNet refinada, mas no entanto, sem a técnica Sliding Window.

Os resultados apresentados por esse experimento já foram mais promissores, a rede neural classifica 5 metais de maneira correta, errando apenas em duas instâncias. E classificando também apenas um não metal de maneira errônea, acertando as outras quatro instâncias.

Para finalizar, realizou-se um ultimo experimento, usando a técnica Sliding Window desenvolvida, e a rede neural refinada. Os resultados estão dispostos nas imagens abaixo, organizadas pelo tamanho da janela e dos pulos utilizados. Percebe-se que os resultados apresentados realizam a identificação dos metais de maneira mais correta e eficiente.



Figura 10. Resultado da RN refinada com a técnica Sliding Window. Tamanho da janela de 64 pixelis com pulos de 32 pixelis



Figura 11. Resultado da RN refinada com a técnica Sliding Window. Tamanho da janela de 128 pixelis com pulos de 32 pixelis



Figura 12. Resultado da RN refinada com a técnica Sliding Window. Tamanho da janela 128 64 pixelis com pulos de 64 pixelis



Figura 13. Resultado da RN refinada com a técnica Sliding Window. Tamanho da janela de 256 pixels com pulos de 64 pixels



Figura 14. Resultado da RN refinada com a técnica Sliding Window. Tamanho da janela de 256 pixels com pulos de 64 pixels

5. Conclusão

O trabalho aborda um desafio intrigante na sociedade atual, a busca de matérias metálicas mediante a grande quantidade de lixo. A busca destes materiais para uma futura reciclagem traria benefícios diversos, tantos sociais, como econômicos e políticos. Para tratar esse desafio foi proposta uma análise de reconhecimento de imagens para detectar tais objetos.

O reconhecimento do metal nas imagens foi realizado utilizando uma rede neural com a arquitetura AlexNet. Essa rede neural foi treinada utilizando o banco de imagens MINC, portanto posteriormente refinada utilizando o banco TrashNet,

onde as imagens continham metais misturados a outros resíduos.

No entanto, o refinamento da rede neural não foi o suficiente para encontrar metais de maneira eficiente e consequentemente foi desenvolvida a técnica, Sliding Window, para análise de diversos trechos em uma mesma imagem.

Os resultados experimentais indicam que é possível sim realizar o reconhecimento de metais utilizando a AlexNet treinada com o MINC e refinada com a TrashNet. Mas apresentam resultados visivelmente melhores quando aplicados em conjunto com a técnica Sliding Window. Cabe a trabalho futuros realizar uma junção de mapa de probabilidades para os diversos tamanhos de janelas e pulos, pegando resultados cada vez mais precisos. Outros refinamentos na rede neural também está aberto a estudos.

Referências

- [1] H. Kopka and P. W. Daly, *A Guide to L^AT_EX*, 3rd ed. Harlow, England: Addison-Wesley, 1999.
- [2] S. Bell P. Upchurch N. Snavely K. Bala, Material Recognition in the Wild with the Materials in Context Database, Computer Vision and Pattern Recognition (CVPR), 2015.