

Transfer Learning para Otimizar a Detecção de Cancros ou Lesões em Imagens Mamográficas

Trabalho de Conclusão de Curso

30 de Novembro de 2018

Orientando: Daniel Ambrósio Ferreira Júnior

Orientador: Prof. Dr. Marcelo A. C. Vieira

Departamento de Engenharia Elétrica e de Computação (EESC - USP)

1 Resumo

A Mamografia é um exame radiológico que avalia as condições da mama através da aplicação controlada de raios-x utilizando um aparelho chamado mamógrafo. Através da imagem gerada pelo aparelho, o médico é capaz de verificar a existência de lesões ou cânceres que se apresentam, geralmente, como nódulos ou calcificações. Para que o laudo definitivo seja dado, há a necessidade de que a imagem seja reavaliada por outro médico para que haja a confirmação do que foi primeiramente aferido. Entretanto, nos últimos anos surgiu uma tecnologia de detecção auxiliada por computador (*computer-aided detection* - CAD) que tem substituído a segunda opinião médica necessária para o laudo final. Tais sistemas computadorizados se valem de artifícios de Inteligência Artificial (IA), tais como as Redes Neurais Artificiais, para avaliarem as imagens médicas.

Redes Neurais Artificiais são métodos computacionais cujo modelo matemático é inspirado no sistema nervoso de organismos complexos, que adquirem conhecimento através de suas experiências. Uma de suas aplicações ocorre no reconhecimento de padrões de imagens para detecção de características de interesse. Realiza-se uma série de processos iterativos e convolucionais para que se defina um vetor de características a partir do qual há a detecção de determinada particularidade em uma imagem.

A grande dificuldade enfrentada por estudos que utilizam redes neurais para a identificação de padrões está no banco de dados necessário para que a rede seja implementada. A quantidade de imagens mamográficas que são disponibilizadas para os centros de pesquisa desenvolverem um bom sistema é muito pequena, uma vez que existem questões éticas envolvidas. Além disso, vale ressaltar que dentre as poucas imagens que estão disponíveis, uma quantidade menor ainda será realmente útil, uma vez que apenas uma parcela pequena do conjunto de imagens apresenta casos positivos de lesões ou cânceres.

Para que tal problema fosse resolvido, desenvolveu-se uma técnica chamada *Data Augmentation* que consiste basicamente em modificar as imagens existentes, como por exemplo através da atribuição de ruídos, rotações ou translações, para gerar mais imagens para o treinamento da rede. Contudo, recentemente tem surgido uma nova técnica

chamada *Transfer Learning* que tem uma proposta diferente para sanar o problema anterior.

O conceito de *transfer learning* é o de armazenar os conhecimentos que uma rede neural adquiriu ao resolver um problema e aplicar no auxílio da resolução de outros problemas relacionados. Portanto, a aplicação de *Transfer Learning* na detecção de patologias mamárias visa utilizar outras redes neurais pré treinadas a fim de melhorar os resultados na identificação de lesões ou cânceros, tornando-os mais eficientes do que os obtidos através da utilização de *Data Augmentation*.

2 Objetivos

O presente trabalho de conclusão de curso tem como objetivo o estudo e aplicação dos conceitos de *transfer learning* na melhora do treinamento de redes neurais capazes de identificar nódulos ou calcificações mamárias, propondo, portanto, uma alternativa para a utilização de *Data Augmentation*.

3 Plano de Trabalho e Cronograma de Execução

1. Investigação sobre redes neurais convolucionais e sua utilização através da linguagem de programação *python*.
2. Implementação dos códigos da rede neural artificial com o *Transfer Learning* para verificar a possibilidade de aplicação em imagens mamográficas.
3. Teste e análise do algoritmo a fim de verificar seu desempenho e possíveis erros de implementação.
4. Estudo sobre *Data Augmentation* e seus resultados.
5. Aquisição de material para avaliação e comparação entre as técnicas *Transfer Learning* e *Data Augmentation*.
6. Avaliação dos resultados a fim de verificar a possibilidade de utilizar o método proposto como alternativa para o *Data Augmentation*. Além disso, será verificado, se possível, a viabilidade de se utilizar esse método junto com o *Data Augmentation*.
7. Escrita da Monografia Final.

Será utilizada a biblioteca disponibilizada pela Google chamada *TensorFlow*.

Tarefa/Bimestre	Jan/Fev	Mar/Abr	Mai/Jun	Jul/Ago	Set/Out	Nov/Dez
1	x					
2		x	x			
3		x	x			
4			x	x		
5				x		
6				x	x	
7					x	x

4 Carga Semanal de Trabalho

As atividades anteriormente descritas foram planejadas para serem cumpridas com uma carga semanal de 12h de trabalho por semana.

5 Recursos Utilizados

- Será utilizada apenas uma IDE para compilar *python*. Caso haja necessidade, serão utilizadas outras linguagens de programação cujos softwares necessários para sua execução estão disponíveis nos computadores do Laboratório de Visão Computacional (LAVI) do Departamento de Engenharia Elétrica e de Computação da Escola de Engenharia de São Carlos (EESC/USP).
- As imagens mamográficas que serão utilizadas estão disponíveis no Laboratório de Visão Computacional (LAVI) da EESC/USP e foram cedidas pelo Instituto de Radiologia do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (InRAD/HCFMUSP).

Orientador: _____
Prof. Dr. Marcelo A. C. Vieira

Orientando: _____
Daniel Ambrósio Ferreira Júnior