

# Abordagens para aplicação de redes CNN com enfoque em bases reduzidas de imagens

Lucas Kovalczuk Paulino<sup>1</sup>

<sup>1</sup>Programa de Pós Graduação de Computação Aplicada

Universidade do Estado de Santa Catarina (UDESC)

lucaskpaulino@gmail.com

**Resumo.** *Este trabalho buscou classificar a pesquisa quanto a metodologia a ser aplicada, considerando aspectos como: objetivo da pesquisa, nível de maturidade, estilo de raciocínio lógico, quanto ao procedimento de coleta e, quanto à influência das variáveis. Este estudo busca contemplar todos os aspectos necessários para a elaboração de uma metodologia de estudo acerca do tema proposto.*

## 1. Resumo da pesquisa

O deep learning vem conquistando cada vez mais espaço na área de processamento de imagens. Dentro desta área de pesquisa destaca-se o uso das Redes Neurais Convolucionais (CNN). Dentre as desvantagens da utilização desta abordagem, destaca-se a necessidade de uma larga quantidade de imagens rotuladas, assim como um exacerbado poder computacional. Isto dificulta a utilização de arquiteturas de sucesso (como AlexNet, VGG e ResNet) em pequenas bases de imagens, sem que haja um *overfitting* das características. Afim de sanar estes empecilhos, propomos um método que utilize ferramentas do estado da arte, como transfer learning e data augmentation, afim de criar alternativas para resolução de problemas práticos relacionados ao uso de CNNs em pequenas bases de imagens.

## 2. Caracterização da pesquisa

### 2.1. Objetivo

Pode-se afirmar que, segunda a terminologia proposta por Tauchen (2013), a pesquisa quanto ao seu objetivo, define-se como uma **pesquisa descritiva**, pois tem como principal função identificar os fatores que colaboram ou são peças chave para a ocorrência dos fenômenos, e explicar “o que” teremos como resultado. Esta apresenta uma **hipótese** (M&L, 2005), a qual será de utilizar ferramentas computacionais do estado da arte, afim de obter os modelos de treinamento a serem testados e comparados a outros trabalhos, de maneira sistemática.

### 2.2. Nível de maturidade

Segundo Wazlawick (2010), o grau de maturidade de uma pesquisa pode ser determinado por diversos fatores. São definidas cinco categorias de maturidade, sendo que para este trabalho, é atribuído o **estilo 3** “Apresentação de algo Presumidamente Melhor”, proposto por Wazlawick (2008). Esta categoria exige que o trabalho seja comparado de forma quantitativa a outros trabalhos semelhantes da literatura, e aponta a importância do uso de métricas que permitam tais comparações. Deve-se manter o foco no cenário proposto,

e deixar claras as condições de reprodução e quais as melhorias obtidas (WAZLAWICK, 2008).

### 2.3. Raciocínio lógico

Primeiramente podemos classificar esta pesquisa, quanto ao seu raciocínio lógico, como define M&L (2005), sendo de cunho **dedutivo**, uma vez que busca a partir de um amplo conhecimento já desenvolvido na área (redes neurais convolucionais altamente treinadas, e com alta taxa de acerto), desenvolver a aplicação destas rede à bases de imagens reduzidas (situação nova/caso específico). A pesquisa dedutiva é aquela que busca sacrificar a ampliação do conteúdo para alcançar a certeza (M&L, 2005).

### 2.4. Procedimento de coleta

Já o método de execução da pesquisa bibliográfica, pode ser definida como uma **pesquisa de laboratório** (M&L, 2005), uma vez que as variáveis serão analisadas em ambiente controlado, dispondo dos materiais específicos. Fato que restringe a interferência de fatores externos com relação ao resultado a ser alcançado.

Quanto a amplitude de coleta dos dados, segundo PP&P (2007), esta se encaixa na categoria **descritiva**, pois busca trabalhos em profundidade e amplitude equivalentes, afim de encontrar resultados extremamente específicos relacionados a área de estudo.

### 2.5. Influência das variáveis

Como relatado por Cervo et al. (2007), a influência das variáveis divide-se em três categorias: variáveis independentes, dependentes e intervenientes.

Podemos citar as **variáveis independentes** como sendo:

- ♦ Taxa de aprendizado;
- ♦ Tamanho de lote;
- ♦ Número de épocas;
- ♦ Poder de processamento;
- ♦ Tipo de arquitetura.

A partir destas variáveis, serão então atribuídas as **variáveis dependentes**:

- ♦ Taxa de acerto;
- ♦ *Overfitting*;
- ♦ Tempo de processamento;
- ♦ Matriz de confusão.

Também estão presentes as **variáveis intervenientes**:

- ♦ Temperatura do ambiente;
- ♦ Banco de imagem tendencioso.

#### 2.5.1. Natureza das variáveis

Segundo afirmam PP&P (2007), existem dois tipos de variáveis quanto a sua natureza: a) qualitativa e; b) quantitativa. Neste trabalho, serão utilizadas principalmente as do tipo **quantitativas**, pois são valores exatos referentes as métricas (taxa de aprendizado, *overfitting*, tempo de processamento) e valores de “entrada” de dados (taxa de aprendizado, número de épocas). Assim como as variáveis: a) tipo de arquitetura e; b)

banco de imagens tendencioso; podem ser consideradas **qualitativas**, pois são uma escolha do autor baseada em um mapeamento sistemático da literatura.

### 2.5.2 Forma das variáveis

As variáveis **qualitativas** dividem-se entre cardinais e intervalares, assim como as variáveis **quantitativas**, são classificadas em duas classes: ordinais e categóricas (BARBETTA, 2006). No presente estudo, classificamos como **cardinais** e **discretas** sendo: tamanho do lote; número de épocas; poder de processamento; taxa de acerto, overfitting, tempo de processamento e; matriz de confusão. Sendo a taxa de aprendizado um valor **intervalar**. As presentes variáveis qualitativas podem ser consideradas **categóricas**, pois não apresentam relação de ordem.

## 3. Referências

BARBETTA, P. A. **Estatística Aplicada às Ciências Sociais**. 6a ed. Editora da UFSC. 2006.

CERVO, BERVIAN e SILVA, **Metodologia Científica**, 6a ed, Pearson Prentice Hall, 2007

MARCONI; LAKATOS. 5ª Ed. 2005

PACHECO Jr., W.; PEREIRA, V L. do V.; PEREIRA FILHO, H. do V. “**Pesquisa Científica sem Tropeços: Abordagem Sistêmica**”. São Paulo: Atlas. 2007. 978-85-224-4843-2.

TAUCHEN, J.. **Metodologia da Pesquisa: Como classificar as pesquisas com base em seus objetivos**. Disponível em <http://pt.scribd.com/doc/100638680/classificacao-das-pesquisas>. Acessado em Jun/2013

WAZLAWICK, Raul Sidnei. (2008). **Metodologia de Pesquisa para Ciência da Computação**. Rio de Janeiro, RJ, Editora: Elsevier, 2008. ISBN 978-85-352-3522-7. 159 pgs

WAZLAWICK, Raul Sidnei. **Uma Reflexão sobre a Pesquisa em Ciência da Computação à Luz da Classificação das Ciências e do Método Científico**. Revista de Sistemas de Informação da FSMA n. 6 (2010) pp. 3-10