

# Metodologia de Pesquisa - Revisão sistemática

Uma revisão sistemática sobre classificação de mosquitos utilizando aprendizado de máquina

1<sup>st</sup> João Carlos Pandolfi Santana  
*Mestrando em Ciência da computação (of UFABC.)*  
Universidade Federal do ABC  
Santo André, SP  
joao.pandolfi@ufabc.edu.br

**Abstract—background aims methods results conclusion**

**Index Terms—**Transference learning, machine learning, computational vision, neural network, convolutional network, image classification cross validation ,classification, image segmentation, mosquito, mosquitos, insect, visão computacional, aprendizado de máquina, redes neurais, classificação de imagens, segmentação de imagem, insetos

## I. INTRODUCTION

A identificação eficiente de mosquitos é importante para evitar vetorização de doenças. Alguns exemplos de doenças que são vetorizadas por mosquitos são: Dengue, Chikungunya e Febre amarela. Uma forma de identificar o mosquito, é através de visão computacional, onde com uma imagem ou conjunto de imagens de entrada, seja possível identificar a espécie do inseto associando assim as doenças capazes de serem vetorizadas por ele. Uma técnica de visão computacional largamente adotada, são as redes neurais, capazes de identificar características de baixo e alto nível em imagens. Neste trabalho, utilizaremos um modelo de rede mista com o conceito de *transference learning*, no qual aproveita-se um conjunto de dados maior para treinamento de algum problema parecido e após um resultado suficientemente bom, adapta-se este aprendizado no problema chave, onde na maioria das vezes não se tem o volume de dados tão grande quanto o anterior.

Munidos destas técnicas, desenvolveremos um produto para identificação e classificação de mosquitos e seus possíveis vetores de doenças. Temos como objetivo migrar para uma plataforma mobile para facilitar o uso e acesso a ferramenta em regiões com baixo acesso a tecnologia de ponta. Também temos como objetivo contribuir com a instituição de pesquisa que nos cedeu os espécimes para coleta de imagens.

Para isso, esta revisão sistemática foi aplicada, visando encontrar trabalhos relacionados e materiais de apoio para o seguimento da pesquisa.

## II. PROPOSTA

### A. Problema de pesquisa

O problema abordado é a identificação por imagem de espécies de mosquito, de forma a agilizar e aumentar a eficiência de agentes de saúde e pesquisadores de campo em relação a epidemiologia e controle dos vetores de doença.

### B. Objetivos

O objetivo principal é fazer uma revisão da literatura, onde possa reunir os trabalhos relacionados e partir de um problema em aberto ou otimizar alguma etapa ineficiente. Outro objetivo é mapear as técnicas utilizadas e propor uma nova e mais eficiente para classificação de mosquitos. Tenho como objetivo final, desenvolver uma ferramenta mobile, no qual facilite a identificação do espécime capturado. Por fim, também pretendo revisar conceitos de aprendizado de máquina, que possam ser aplicados na pesquisa, como por exemplo, *transference learning* (transferência de aprendizado) onde possa utilizar sistemas especialistas em outras áreas, mas que possuam um certo grau de proximidade computacional do problema, de modo que sejam aplicados para otimização e melhoria no aprendizado do modelo.

### C. Hipótese

A hipótese inicial é que não haja muita pesquisa relacionada ao tema proposto, devido ao fato dos estudos serem voltados, na sua maior parte, a doenças e migração dos vetores não na identificação computacional dos espécimes, também levo em conta que pelos insetos serem muito pequenos e dependerem de captura para análise, os pesquisadores que fazem a coleta já são capacitados para identificar o espécime, não abrindo tanta abertura para o problema de classificação destes. No entanto, visto desenvolver uma ferramenta para facilitar e reduzir o viés desta classificação, facilitando assim a análise em regiões que não possuam muitos especialistas e também permitir um pessoal menos treinado a fazer a primeira etapa da pesquisa, a identificação do espécime.

### D. Métodos

Como a hipótese se baseia em não ter muita pesquisa relacionada ao tema proposto, procurarei conteúdo relacionado a classificação de imagem utilizando ferramenta computacional que pretendo usar na pesquisa, assim como procurar material relacionado a estrutura corporal e classificação morfológica dos insetos a serem estudados.

### E. Resultados esperados

Espero encontrar uma grande quantidade de conteúdo sobre classificação de imagens e conteúdo relacionado a morfologia de insetos. Assim como modelos e métodos previamente

validados de *transference learning* e visão computacional para basear minha pesquisa. Pretendo encontrar também, caso haja, pesquisas relacionadas ao tema proposto, de forma que possa seguir por caminhos consolidados e explorar conceitos ainda não estudados, assim, possa contribuir não só com o produto final produzido, mas também com técnicas computacionais novas ou aplicadas em áreas inexploradas.

### III. MÉTODOS

#### A. Critério de seleção das fontes de busca

O primeiro critério para seleção da fonte, foi a base ser internacional, assim eu tenho acesso a uma quantidade de artigos maior e consequentemente das maiores universidades e centros de pesquisa.

O segundo critério foram bases de dados relacionadas a publicações na área de ciência da computação, ou seja, não foram incluídas bases de dados que não tem a ver com o tema de pesquisa, esse critério foi adicionado para evitar ruído na pesquisa.

O terceiro e último critério, foi o uso do *google scholar* para refinamento final da pesquisa, onde o objetivo foi pesquisar com um espectro mais amplo, vai de encontro com o segundo critério onde o objetivo é exatamente não ficar preso no espectro definido. O objetivo é somente verificar a relevância e qualidade da *string* de busca, para garantir a assertividade, ou seja, se ao buscar numa base mais abrangente o resultado for próximo do encontrado na base mais específica, a *string* está boa.

#### B. Estratégia de busca

A primeira etapa, foi a definição das strings de busca. O critério de qualidade de cada *string*, foi definida pela quantidade de resultados em relação a correspondência do material apresentado com o tema da pesquisa. A sequência de *strings* utilizadas se encontram na seção III-C.

Após identificar as *strings* mais promissoras, utilizei a ferramenta de busca avançada *parsif.al*, onde adicionei os resultados em uma planilha no formato *.csv*. Filtrei os resultados obtidos pelo título com os critérios de *inclusão* III-D e *exclusão* III-E, onde os resultados podem ser observados na Tabela I.

Por final, utilizei a ferramenta *Google Scholar* para pesquisar artigos relacionados e verificar a qualidade das *strings* de busca, ao ponto que dependendo da correlação dos resultados obtidos com o tema de pesquisa, pude verificar se as buscas anteriores foram assertivas.

#### C. Strings de busca

Strings de busca utilizadas durante a revisão sistemática

- Parsifall
  - 1) "transference learning" AND ("neural network" OR "neural networks") AND (mosquito OR mosquitos OR insect OR insects)
  - 2) ("transference learning" OR "machine learning" ) AND ("mosquito" OR "mosquitos")

- 3) ("classification" ) AND ("mosquito" OR "mosquitos")
- 4) ("neural network" ) AND ("classification") AND ("mosquito" OR "mosquitos")
- 5) "transference" AND "learning" AND "machine"

#### • Google Scholar

- 1) transference learning machine
- 2) transductive learning on insects classification
- 3) image mosquito classification
- 4) transductive learning

#### D. Critérios de inclusão

- 1) Conter as palavras chave
- 2) Ser relacionado com:
  - Machine Learning
  - Transference Learning
  - Transductive Learning
  - Image Processing
  - Image Classification
  - Mosquito Classification
  - Insect Classification

#### E. Critérios de exclusão

- 1) Fugir dos temas
- 2) Não abranger sistemas de informação
- 3) Não utilizar machine learning
- 4) Não ter acesso ao pdf

#### F. Estratégia para seleção

Após o filtro pelo título dos trabalhos, fui para a leitura do *abstract*, onde utilizei novamente os critérios de *inclusão* e *exclusão* para selecionar os artigos que fizessem sentido. Por fim, após as três etapas de filtragem citadas, parti para leitura dos arquivos selecionados. Pesquisei no *google scholar* para não ficar na bolha das bases selecionadas e usar o algoritmo de sugestão do *google*.

Adicionei uma *label* indicando a importância

#### G. Avaliação da qualidade

- Quantidade de citações
- Correlação com o conteúdo pesquisado
- Profundidade da pesquisa
- Ferramentas utilizadas parecidas com as que pretendo usar
- Bons resultados
- Relevância da publicação
- Tipo de publicação
- Citação por artigo relevante, ou seja, um artigo relevante faz citação a este

TABLE I: Trabalhos selecionados

<i><b>Título</b></i>	<i><b>Autores</b></i>	<i><b>Publicação</b></i>	<i><b>Citações</b></i>	<i><b>Comentário</b></i>
Training of convolutional neural network using transfer learning for Aedes Aegypti larvae	Fuad M.	1 August 2018	5	Importância - alta
A vision-based counting and recognition system for flying insects in intelligent agriculture	Zhong Y.	9 May 2018	1	Importância - média
Classification MFCC feature from culex and aedes aegypti mosquitoes noise using support vector machine	Lukman A.	16 January 2018	1	Material de apoio
Butterfly species recognition using artificial neural network	Alhady S.	2018	0	Importância - alta
Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images	Rajaraman S.	2018	1	Importância - média
Automatic Identification of Malaria Using Image Processing and Artificial Neural Network	Kanojia M.	2018	0	Importância - média
Data driven prediction of dengue incidence in Thailand	Sumanasinghe N.	2018	1	Material de apoio
Artificial neural network as a classifier for the identification of hepatocellular carcinoma through prognostic gene signatures	Jujjavarapu S.	2018	0	Importância - média
Geometric morphometrics in mosquitoes: What has been measured?	Lorenz C.	October 2017	2	Importância - Alta
Elman neural network trained by using artificial bee colony for the classification of learning style based on students preferences	Shuib N.	1 September 2017	0	Material de apoio
Learning to count mosquitoes for the sterile insect technique	Ovadia Y.	13 August 2017	0	Material de apoio
Automatic insect recognition using optical flight dynamics modeled by kernel adaptive ARMA network	Li K.	16 June 2017	2	Importância - Média
Sequential minimal optimization algorithm with support vector machine for mosquito larvae identification	Yusoff M.	May 2017	0	Importância - Alta
A survey on image-based insect classification	Martineau M.	1 May 2017	9	Importância - Alta
Mosquito larva classification method based on convolutional neural networks	Sanchez-Ortiz A.	3 April 2017	0	Importância - Alta
Automated plasmodia recognition in microscopic images for diagnosis of malaria using convolutional neural networks	Krappe S.	2017	0	Importância - Média
Vision-based perception and classification of mosquitoes using support vector machine	Fuchida M.	2017	5	Importância - Muito Alta

Identifying Aedes aegypti mosquitoes by sensors and one-class classifiers	Souza V.	2017	0	Importância - Alta
Artificial Neural Network applied as a methodology of mosquito species identification	Lorenz C.	December 01, 2015	12	Importância - Muito Alta
Exploring Low Cost Laser Sensors to Identify Flying Insect Species: Evaluation of Machine Learning and Signal Processing Methods	Silva D.	1 December 2015	14	Importância - Alta
Flying Insect Classification with Inexpensive Sensors	Chen Y.	September 2014	29	Importância - Alta
A hierarchical artificial neural system for genera classification and species identification in mosquitoes	Venkateswarlu C.	2012	6	Importância - Muito Alta
Classification and identification of mosquito species using artificial neural networks	Banerjee A.	December 2008	21	Importância - Muito Alta
An aquatic insect imaging system to automate insect classification	Sarpola M.	2008	16	Importância - Alta
Classification MFCC feature from culex and aedes aegypti mosquitoes noise using support vector machine	Lukman A.	16 January 2018	1	Importância - Alta
Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images	Rajaraman S.	2018	1	Importância - Média
A survey on image-based insect classification	Martineau M.	1 May 2017	9	Importância - Muito Alta
Vision-based perception and classification of mosquitoes using support vector machine	Fuchida M.	2017	5	Importância - Muito Alta
The image processing handbook: Seventh edition	Russ J.	5 January 2016	5	Material de apoio
Identification of medical plants using genetic algorithm and recurrent neural network	Malarkhodi S.	2016	0	Material de apoio
Microscopic image segmentation using hybrid technique for dengue prediction	Ghosh P.	1 January 2016	0	Material de apoio
Artificial Neural Network applied as a methodology of mosquito species identification	Lorenz C.	December 01, 2015	12	Importância - Muito Alta
Mosquito vector monitoring system based on optical wingbeat classification	Ouyang T.	October 01, 2015	3	Importância - Muito Alta
A tool for developing an automatic insect identification system based on wing outlines	Yang H.	7 August 2015	11	Importância - Alta
Flying Insect Classification with Inexpensive Sensors	Chen Y.	September 2014	29	Material de apoio / Importância - Alta

On the role of pre and post-processing in environmental data mining	Gibert K.	2008	16	Material de apoio
Handbook of urban insects and arachnids	Robinson W.	1 January 2005	36	Material de apoio
Transductive learning for statistical machine translation	Nicola Ueffing National Research Council Canada Gatineau, QC, Canada nicola.ueffing@nrc.gc.ca Gholamreza Haffari and Anoop Sarkar Simon Fraser University Burnaby, BC, Canada			Material de apoio

<sup>a</sup>Sample of a Table footnote.

#### IV. DISCUSSÃO

Ao contrário do que previ na hipótese inicial encontrei bastante material relacionado a pesquisa, em contrapartida a maioria dos resultados das buscas foram sobre as doenças vetorizadas por mosquitos e estudos sobre migração e comportamento dos insetos. No quesito de aprendizado computacional, obtive bastante conteúdo de qualidade, o que irá ajudar bastante no desenvolvimento do projeto.

Tive bastante dificuldade ao selecionar conteúdos relacionados a computação, em vista que, muito material tem sido produzido na última década devido ao aumento do poder de processamento dos computadores e popularização das técnicas de aprendizado computacional. Para chegar aos artigos mais relevantes, dei ênfase aos mais citados e me vali de consultoria a professores da área para indicação e direcionamento da pesquisa.

Quanto aos trabalhos que têm como tema os insetos e sua morfologia, apesar de encontrar bastante artigo sobre o assunto, a maioria destes apontavam para livros, exceto os que tratavam de algum assunto específico estudado, como o [?] que trata da geometria morfométrica de mosquitos (no qual, acredito eu, ser uma chave para identificação de *features* da pesquisa da qual conduzirei). Os livros foram um pouco mais difíceis de se obter, em vista que são comprados, porém consegui acesso aos conteúdos e consegui adicionar a minha base de pesquisa.

#### V. CONCLUSÃO

A pesquisa foi mais árdua do que imaginei que seria, contudo, obtive resultados que se mostram promissores, devido ao fato de alguns deles serem diretamente relacionados com a minha proposta de pesquisa e apresentarem resultados positivos. Assim como consegui materiais de suporte para desenvolvimento direcionado das etapas que precisarei transpor para concluir a pesquisa.

Em relação a classificação de mosquitos, encontrei estudos que desenvolveram o que propus estudar, porém não englobam todo o escopo que havia definido, os que tornam uma ótima base de pesquisa e pontapé inicial, assim como prova de conceito que de viabilidade.

##### A. Agradecimentos

Agradeço aos professores da disciplina por me apoiarem e disponibilizarem material de apoio e tempo para que fosse possível o desenvolvimento deste trabalho.

#### ACKNOWLEDGMENT

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

#### REFERENCES

Please number citations consecutively within brackets [1]. The sentence punctuation follows the bracket [2]. Refer simply to the reference number, as in [3]—do not use “Ref. [3]” or “reference [3]” except at the beginning of a sentence: “Reference [3] was the first ...”

Number footnotes separately in superscripts. Place the actual footnote at the bottom of the column in which it was cited. Do not put footnotes in the abstract or reference list. Use letters for table footnotes.

Unless there are six authors or more give all authors’ names; do not use “et al.”. Papers that have not been published, even if they have been submitted for publication, should be cited as “unpublished” [4]. Papers that have been accepted for publication should be cited as “in press” [5]. Capitalize only the first word in a paper title, except for proper nouns and element symbols.

For papers published in translation journals, please give the English citation first, followed by the original foreign-language citation [6].

#### REFERENCES

- [1] G. Eason, B. Noble, and I. N. Sneddon, “On certain integrals of Lipschitz-Hankel type involving products of Bessel functions,” *Phil. Trans. Roy. Soc. London*, vol. A247, pp. 529–551, April 1955.
- [2] J. Clerk Maxwell, *A Treatise on Electricity and Magnetism*, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
- [3] I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in *Magnetism*, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.
- [4] K. Elissa, “Title of paper if known,” unpublished.
- [5] R. Nicole, “Title of paper with only first word capitalized,” *J. Name Stand. Abbrev.*, in press.
- [6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” *IEEE Transl. J. Magn. Japan*, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetism Japan, p. 301, 1982].
- [7] M. Young, *The Technical Writer’s Handbook*. Mill Valley, CA: University Science, 1989.

IEEE conference templates contain guidance text for composing and formatting conference papers. Please ensure that all template text is removed from your conference paper prior to submission to the conference. Failure to remove the template text from your paper may result in your paper not being published.