

## IV Workshop de Dissertações de Mestrado do PPgSI (2017)

### FICHA DA PESQUISA<sup>i</sup>

DADOS GERAIS				
<b>Título do projeto de pesquisa</b>	Modelo de anotação automática de imagens baseado em agrupamento probabilístico			
<b>Orientando</b>	Raphael Becker Carnassale			
<b>Orientador(es)</b>	Patrícia Rufino Oliveira			
<b>Semestre no curso, na data do workshop</b>	<input type="checkbox"/> 2º semestre	<input type="checkbox"/> 3º semestre	<input checked="" type="checkbox"/> 4º semestre	<input type="checkbox"/> 5º semestre
<b>Qualificação</b>	<input checked="" type="checkbox"/> Qualificação já realizada em: 04/07/2017 <input type="checkbox"/> Realização da qualificação planejada para: dd/mm/aaaa			
<b>Defesa</b>	Prazo máximo para depósito: 29/08/2018		Realização da defesa planejada para: dd/mm/aaaa	
<b>Linha e Área de pesquisa</b>	Gestão e desenvolvimento de Sistemas: <input type="checkbox"/> BD <input type="checkbox"/> Gestão de SI <input type="checkbox"/> Eng. de Software		Inteligência de Sistemas: <input checked="" type="checkbox"/> IA <input type="checkbox"/> Rec. de Padrões <input type="checkbox"/> Proc. Gráfico	
<b>Área de aplicação</b>	<input type="checkbox"/> Ambientes Corporativos <input type="checkbox"/> Bioinformática <input type="checkbox"/> Biometria <input type="checkbox"/> Economia	<input type="checkbox"/> Educação <input type="checkbox"/> Educação a Distância <input type="checkbox"/> Internet <input type="checkbox"/> Jogos	<input type="checkbox"/> Linguagem Natural <input type="checkbox"/> Linguística <input type="checkbox"/> Processos de Negócio <input type="checkbox"/> Química	<input type="checkbox"/> Redes Sociais <input type="checkbox"/> Robótica <input type="checkbox"/> Saúde <input checked="" type="checkbox"/> [Anotação Automática de Imagens]
<b>Publicações associadas ao projeto de mestrado</b>				

DESCRIÇÃO DO PROJETO DE PESQUISA	
<b>Contextualização / motivação</b>	Utiliza-se um grande número de dados (imagens, características extraídas das mesmas) para anotar imagens hoje em dia. As maiores preocupações deste projeto de pesquisa são desempenho da anotação do modelo e tempo de execução.
<b>Problema de pesquisa</b>	O projeto de pesquisa busca remover algumas etapas de um modelo proposto para a Anotação Automática de Imagens, simplificando-o embora mantendo (ou melhorando) desempenho de anotação, incorporando ao mesmo um algoritmo de agrupamento probabilístico, e alterando seu vetor solução em sua fase de otimização.
<b>Objetivo geral da pesquisa</b>	Substituir o algoritmo de agrupamento com base em distância (Algoritmo de Agrupamento Gravitacional) pelo algoritmo de agrupamento probabilístico EICAMM. Utilizar um vetor solução com apenas valores reais no lugar de um vetor híbrido (com valores binários e reais) para a seleção de conjuntos de características.
<b>Trabalhos relacionados</b>	<p>“A multi-expert based framework for automatic image annotation” é modelo MEBF original, que este projeto de pesquisa possui intenção de adaptar.</p> <p>“Improvements on ICA mixture models for image pre-processing and segmentation” é o algoritmo de agrupamento probabilístico (EICAMM) a ser aplicado no modelo original.</p> <p>“Automatic segmentation of breast masses using enhanced ICA mixture model” é uma aplicação prática do EICAMM.</p>
<b>Justificativa e relevância</b>	O modelo MEBF realiza o agrupamento dos dados utilizando um algoritmo de agrupamento baseado em distância, para então estimar probabilidade das anotações das imagens pelas distâncias calculadas, o que pode não ser a forma mais eficiente de resolver o problema. Dada a grande quantidade de imagens e características extraídas das mesmas, que existe na área, um modelo que consiga realizar a anotação de forma mais eficiente tende obter um melhor desempenho na anotação (pois pode fazer mais combinações dos dados).
<b>Proposta para Solução</b>	O modelo MEBF realiza o agrupamento dos dados utilizando um algoritmo de agrupamento baseado em distância, para então estimar probabilidade das anotações das imagens pelas distâncias calculadas. Uma das hipóteses deste projeto de pesquisa é de que utilizar um algoritmo de agrupamento probabilístico irá reduzir etapas do modelo (diminuindo o tempo de execução do mesmo), mantendo seu desempenho de anotação.
<b>Dados</b>	Os dados utilizados serão características extraídas de dois conjuntos de dados (ESP-game, IAPR-TC12) e

	disponibilizados pelo instituto INRIA. Os mesmos dados foram utilizados para avaliar o modelo MEBF original.
<b>Validação</b>	Comparar desempenho do modelo original e adaptado, utilizando métricas como precisão, revocação, f-medida e tempo. Serão utilizados conjuntos de validação e teste.
<b>Limitações</b>	Não foram incluídas no escopo deste projeto áreas como geração de expressões de referência e linguagem natural, que são áreas de interesse que estavam previstas no início da pesquisa.
<b>Resultados esperados</b>	Contribuições científicas: Ao final deste projeto de pesquisa espera-se que sistemas de inteligência artificial estejam um passo mais próximos de realizar a anotação de imagens de forma mais similar à como os humanos o fazem. Contribuições tecnológicas:

<b>MÉTODO DE PESQUISA</b> [para cada item, marque todas as opções adequadas para seu projeto] (basear-se em: <a href="http://www.ufrgs.br/cursopgdr/downloadsSerie/derad005.pdf">http://www.ufrgs.br/cursopgdr/downloadsSerie/derad005.pdf</a> )	
<b>Gênero</b>	<input type="checkbox"/> Pesquisa teórica <input checked="" type="checkbox"/> Pesquisa prática <input type="checkbox"/> Pesquisa empírica <input type="checkbox"/> Pesquisa metodológica
<b>Natureza</b>	<input type="checkbox"/> Pesquisa básica/pura <input checked="" type="checkbox"/> Pesquisa aplicada
<b>Objetivo</b>	<input type="checkbox"/> Pesquisa exploratória <input type="checkbox"/> Pesquisa descritiva <input checked="" type="checkbox"/> Pesquisa explicativa <input type="checkbox"/> [outro – escrever]
<b>Abordagem</b>	<input checked="" type="checkbox"/> Pesquisa quantitativa <input type="checkbox"/> Pesquisa qualitativa <input type="checkbox"/> Pesquisa mista (quali-quant)
<b>Procedimento(s) técnico(s)</b>	<input checked="" type="checkbox"/> Pesquisa experimental <input type="checkbox"/> Pesquisa com <i>survey</i> <input type="checkbox"/> Pesquisa netnográfica <input type="checkbox"/> Pesquisa bibliográfica <input type="checkbox"/> Estudo de caso <input type="checkbox"/> Teoria fundamentada em dados ( <i>grounded theory</i> ) <input type="checkbox"/> Pesquisa documental <input type="checkbox"/> Pesquisa participante <input type="checkbox"/> Ciência do projeto ( <i>Design science research</i> ) <input type="checkbox"/> Pesquisa <i>ex-post-facto</i> <input type="checkbox"/> Pesquisa-ação <input type="checkbox"/> [outro – escrever] <input type="checkbox"/> Pesquisa de levantamento <input type="checkbox"/> Pesquisa etnográfica
<b>Fonte(s) de dados</b>	<input checked="" type="checkbox"/> pesquisa de laboratório <input type="checkbox"/> pesquisa de campo <input type="checkbox"/> pesquisa bibliográfica <input type="checkbox"/> [outro – escrever]
<b>Técnica(s) / Instrumento(s) de coleta de dados</b>	<input checked="" type="checkbox"/> medição <input type="checkbox"/> observação (direta / participante) <input type="checkbox"/> questionário <input type="checkbox"/> formulário <input type="checkbox"/> diário de campo / notas de campo <input type="checkbox"/> entrevista <input type="checkbox"/> <i>benchmark</i> <input type="checkbox"/> análise documental (ou de artefatos) <input type="checkbox"/> grupos focais <input type="checkbox"/> [outro – escrever]
<b>Técnica(s) de análise de dados</b>	Análise quantitativa: <input checked="" type="checkbox"/> Estatística descritiva <input type="checkbox"/> Análise de conteúdo <input checked="" type="checkbox"/> Estatística inferencial <input type="checkbox"/> Análise do discurso <input type="checkbox"/> [outro – escrever] <input type="checkbox"/> [outro – escrever]

	2017												2018											
	1	2	3	4	5	6	7	8	9	10	11	12	1	2	3	4	5	6	7	8	9	10	11	12
Revisão bibliográfica						x	x	x	x	x	x	x	x	x										
Implementação de algoritmos de agrupamento						x	x																	
Implementação das técnicas de busca por solução via otimização							x	x																
Implementação do modelo MEBF original								x	x															
Avaliação e sumarização dos resultados obtidos pelo modelo original									x	x														
Incorporação do algoritmo EICAMM											x	x												
Adaptação do vetor solução do modelo MEBF											x	x												
Avaliação e sumarização dos resultados obtidos pela proposta												x	x											
Realização de um estudo comparativo entre o modelo MEBF original e o modelo MEBF adaptado													x	x										
Escrita de artigos									x	x	x	x	x	x	x	x								
Escrita da dissertação									x	X	x	x	x	x	x									
Depósito da dissertação																	x							

Cronograma apresentado no exame de qualificação, aprovado em 04/07/2017.

