

PLANODE TRABALHO

Título do Projeto: Entomologia Forense: Identificação Genética de conteúdo Gastrointestinal de Moscas

Título do Plano de trabalho: Extração de Atributos em um Sistema de Visão Computacional para Análise de Larvas Necrofágicas

Orientador: HemersonPistori (pistori@ucdb.br)

Orientanda: Eudamara Barbosa da Silva Acosta **RA:** 148594

Curso: Engenharia da Computação **Semestre:** 7º

Resumo

Um grande desafio para identificar larvas ocorre quando se depara com uma grande variação de características, assim sendo este trabalho visa validar a partir de algoritmos de extração de atributos um sistema de visão computacional capaz de interpretar imagens microscópicas de larvas de forma automática ou semi-automática para o auxílio da entomologia forense, para que os peritos estimem origem do corpo e o intervalo pós morte.

1. Antecedentes e Justificativa

A entomologia forense é o estudo de insetos e outros artrópodes que se aplica a procedimentos legais como investigações médicos-criminais (HALL, 2001), que contribui em elucidações de crime a partir das informações relevantes do ciclo biológico e da diversidade dos grupos taxonômicos (SMITH, 1986). Para a entomologia forense existe uma carência de um banco de dados que possa ser referência para investigações criminais e estimativas de intervalo pós-morte. É um grande desafio o reconhecimento de larvas, pois estas são altamente articuladas, exibem um grau elevado de variações no tamanho, morfologia e cor, e algumas espécies são difíceis de distinguir visualmente (SMITH, 1986).

A ordem Diptera equivalente das moscas, mosquitos e afins, faz parte grupos de insetos mais diversos, tanto ecologicamente quanto em termos de

riquezas de espécies. Este grupo de insetos são os maiores responsáveis pela decomposição de um corpo, somente quando o corpo está muito ressecado que outros grupos de insetos são atraídos para continuar o processo (OLIVEIRA,2014).

A Visão Computacional é a área da Computação que tem como objetivo principal o estudo de imagem digital, a partir das imagens capturadas por câmeras, filmadoras, entre outros dispositivos, estas são transformadas por algoritmos para obter informações importantes, modelos, equações matemáticas através de um sistema computacional capaz de interpretar essas imagens de forma automática ou semi-automática (WINSTON, 1975).

Pode-se dizer também que a Visão Computacional é a área da Computação que se empenha em reproduzir em programas, a habilidade do ser humano em identificar ou resolver problemas através da visão. Para o ser humano é muito compreensível identificar objetos a partir de imagens, mas, para trazer essa realidade para o mundo computacional não é uma tarefa muito trivial (BALLARD, HARRY, 1982).

Na área de Visão Computacional existe o problema de extração e descrição de atributos de uma imagem, o qual ajuda muito na identificação de objetos de uma imagem. Existem algoritmos de extração de características como o SURF, que faz a extração e descrição de pontos relevantes na imagem, chamados pontos de interesse.

Neste plano de trabalho objetiva-se estudar técnicas para descrever pontos de interesse em imagens, baseados em diferentes atributos para identificação de gêneros de larvas necrofágicas. Atualmente este projeto contém um banco de imagens com dois gêneros. Experimentos serão realizados para coletar mais larvas e assim aumentar o banco de imagens.

Este trabalho é de grande importância para o projeto, pois é fundamental uma forma mais precisa de identificar as larvas necrofágicas, importante fator para a entomologia forense.

2. Objetivos

2.1 Geral

Desenvolver um sistema computacional capaz de interpretar imagens microscópicas de estruturas de larvas de forma automática ou semi-automática para o uso na entomologia forense.

2.2 Específicos

Para atingir o objetivo geral definido na seção 2.1, foram estabelecidos os seguintes objetivos específicos:

- Aprofundamento e atualização da revisão de literatura
- Criação de banco de imagens de larvas necrofágicas
- Desenvolvimento do módulo de Reconhecimento das estruturas morfológicas das larvas
- Validação do módulo de identificação
- Registro e divulgação de resultados

3. Revisão de literatura.

3.1 Entomologia Forense

Em termos de abundância de espécies, os insetos são o maior grupo taxonômico dentre os animais. Há mais de um milhão de espécies de insetos descritas, encontradas em considerável diversidade de habitats (SPEIGHT, 2008). Os corpos em decomposição apresentam-se ricos em matéria orgânica, essencial ao desenvolvimento dos insetos necrófagos (ANDERSON & CERVENKA, 2002). Portanto, os insetos podem ser vistos como participantes do processo de decomposição dos corpos (AMENDT, 2004).

A entomologia forense é uma sub-área da medicina legal que utiliza fundamentos zoológicos para auxiliar em investigações criminais. Na maior parte dos casos emprega-se a estimativa do intervalo pós-morte, através do conhecimento do ciclo de vida de insetos encontrados em cadáveres. (AMENDT, 2004).

Desde o século 13 aplica-se o estudo dos insetos para investigações no contexto da medicina legal na China (MCKNIGHT, 1981). Mesmo parecendo um longo tempo de estudos, e acúmulos de conhecimento, em tempos remotos havia baixa compreensão sobre o assunto entre os especialistas por não existirem formalmente conexões entre a entomologia e a medicina legal. Atualmente existe esta conexão e tem sido aceita em diversos países como ferramenta importante para a medicina legal (ANDERSON, 1996).

3.2 Técnicas de Extração de Atributos

Extração de atributos é uma das etapas de Visão Computacional em que se consegue analisar e extrair informações úteis ou de interesse da imagem processada. Esta etapa é a responsável pela mensuração dos atributos dos objetos pertencentes à imagem em estudo, que são utilizados para identificar uma classe de padrão. A seguir alguns algoritmos de extração de atributos.

3.2.1 SpeededUpRobustFeatures (SURF)

Os algoritmos SURF e SIFT são diferentes dos outros algoritmos de extração de atributos por conterem duas etapas, sendo elas a Detecção de Pontos de Interesse e descrição desses pontos, na segunda etapa são gerados os vetores de atributos. Uma das maneiras para se encontrar pontos interessantes em uma imagem é caracterizar a direção onde as mudanças estão acontecendo, por exemplo, se varia no eixo horizontal ou no eixo vertical. Em Detecção de Pontos de Interesse tem como objetivo a partir de uma imagem encontrar pontos dessa imagem que serão mais relevantes na hora de comparar essa imagem com outra para saber se essas imagens são iguais ou diferentes. O SURF é um algoritmo que visa encontrar pontos de interesse de uma maneira eficiente, utiliza-se como fundamento para fazer os cálculos das regiões a Matriz Hessiana (LOWE, 2004).

É importante observar que as imagens variam de escala, comparar um ponto de interesse de uma imagem com outra que sofreu mudança de escala, é um grande problema e uma das maneiras para resolver este problema é utilizando ferramentas computacionais para gerar versões da imagem em diferentes escalas, montando essas imagens em forma de pirâmide, colando como base a imagem original e acima dela suas versões com escalas

reduzidas. Na descrição dos pontos de interesse, para cada ponto de interesse gera-se um vetor de atributos, um vetor com valores que representa de alguma forma essa região. É importante que o descritor seja invariante a rotação, ou seja, mantenha a informação do ponto de interesse mesmo se a imagem sofrer uma rotação (BAY, 2006).

3.2.2 Histogramas de Palavras Visuais

Em Extração de Vetor utilizando Histogramas de Palavras Visuais (Bag of Visual Words - BOW) tem como objetivo transformar uma imagem ou palavra em um histograma, usando imagens. Procura-se identificar um cenário a partir de várias imagens pequenas, chamadas de pontos de interesse, pontos que serão relevantes na hora da classificação, é desejável que esses pontos sejam invariantes a algumas operações na imagem, como por exemplo, o ponto de interesse não pode transmitir uma informação diferente se sofrer uma rotação. Detectados os pontos de interesse, para cada um extrai-se um vetor de atributos e partir destes calcula-se a aproximação para cada ponto médio dos conjuntos detectados (BAHRI, ZOUAKI, 2013).

Pequenas imagens lisas variando apenas a cor não seria muito interessante para fazer uma comparação entre dois ambientes, para isso existem algoritmos que identifica pontos relevantes para classificação como o SURF e SIFT e outros que identificam cantos nas pequenas imagens, chamados de detectores ou extratores de pontos de interesse (CSURKA, 2004).

3.2.3 Matrizes de Coocorrência

O ser humano consegue identificar de maneira muito rápida um objeto, uma textura em uma imagem, diferente do campo computacional, a textura chega ao computador em forma de vetores numéricos. Em Matriz de Coocorrência a extração de atributos baseia-se na textura, sendo dispensável a cor.

A ideia da técnica é selecionar um ponto, um pixel e definir a partir deste ponto, uma distância e um ângulo. O pixel base, localizado no início do vetor

contem um tom de cinza e o outro pixel na ponta do vetor também tem tom de cinza, a matriz calcula a quantidade de vezes que determinado par de tons de cinza aparece para a distância e para o ângulo, ou seja, a coocorrência é em relação à tonalidade (STEIN, 2005).

Para cada distância e para cada ângulo se constrói uma matriz de coocorrência, a quantidade de distâncias e de ângulos depende do problema em questão. Após definidos estes parâmetros, percorre-se toda imagem com os pares de tons de cinza para cada ponto. Um ponto da imagem só pode ser desconsiderado, caso o par exceda o limite da margem. Espera-se que para cada textura diferente gere-se uma matriz de coocorrência com ângulos diferentes (NASCIMENTO, 2003).

Após obter uma matriz de coocorrência normalizada, ou seja, a soma das células é igual a um, aplica-se atributos de texturas, como por exemplo, Segundo Momento Angular, Entropia, Correlação, Homogeneidade, entre outros. Espera-se com os resultados obtidos pelos cálculos de textura, que para mesma textura não importando a luminosidade ou escala da imagem, chegar num número que represente aquela textura sendo este diferente do número para que represente outra textura. (NASCIMENTO, 2003)

3.3 Características Visuais das Larvas

A descrição das características das larvas, é uma das etapas antes da implementação do software, pois este será programado de acordo com as análises das imagens. Analisando as imagens segmentadas de *Lucilia* e *Chrysomya*, especificamente o crânio, pode-se observar algumas características que diferenciam um gênero do outro.

Encontram-se orifícios na parte posterior do crânio da *Chrysomya* (Figura 1A) e o mesmo não ocorre em *Lucilia* (Figura 1B).

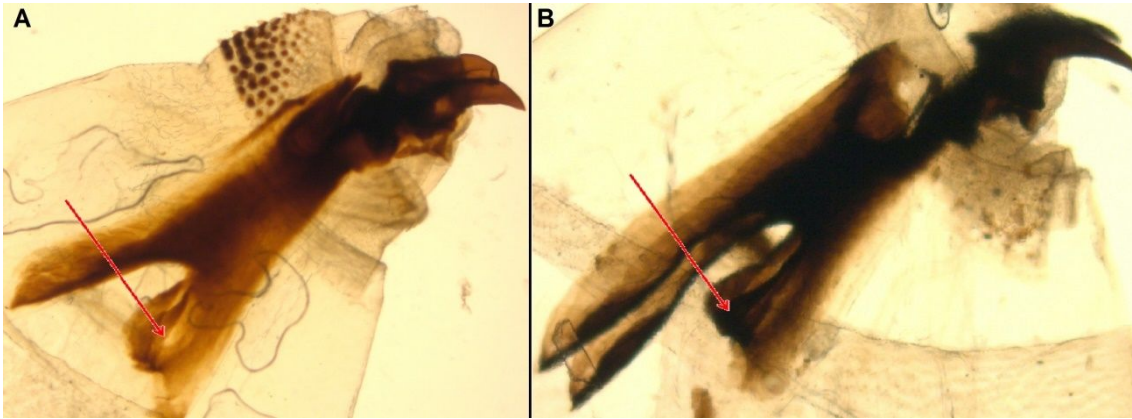


Figura 1: Imagens do crânio das larvas necrofágicas. A-*Chrysomya*; B -*Lucilia*.

Em *Chrysomya* há uma abertura menor na parte anterior no crânio (Figura 2A) em relação à *Lucilia* (Figura 2B).

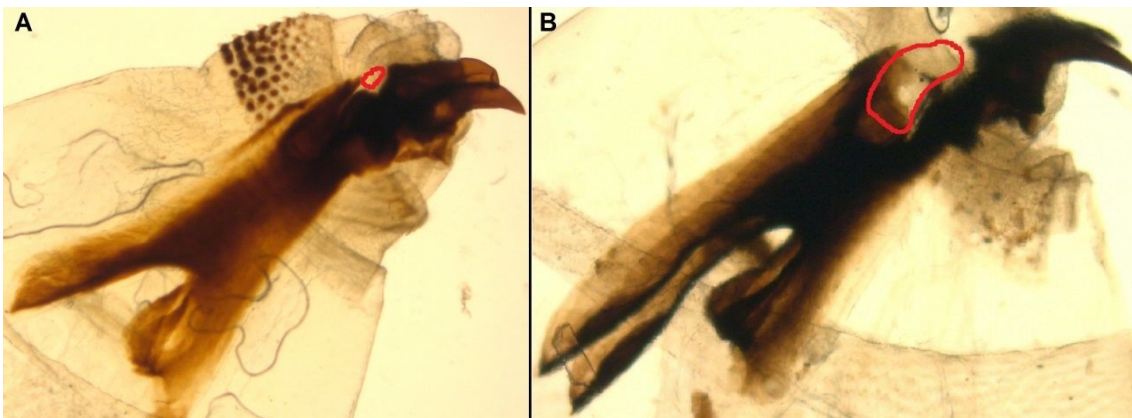


Figura 2: Imagens do crânio das larvas necrofágicas focando a abertura anterior no crânio. A-*Chrysomya*; B - *Lucilia*.

Em *Chrysomya* (Figura 3A) as pequenas manchas em forma de círculo na frente parece ser textura de escamas, mas em *Lucilia* (Figura 3-B) a forma não é exatamente um círculo, está mais próxima de oval, assumindo que a imagem não foi dimensionada, pois não há distorção em outras áreas localizadas da imagem.

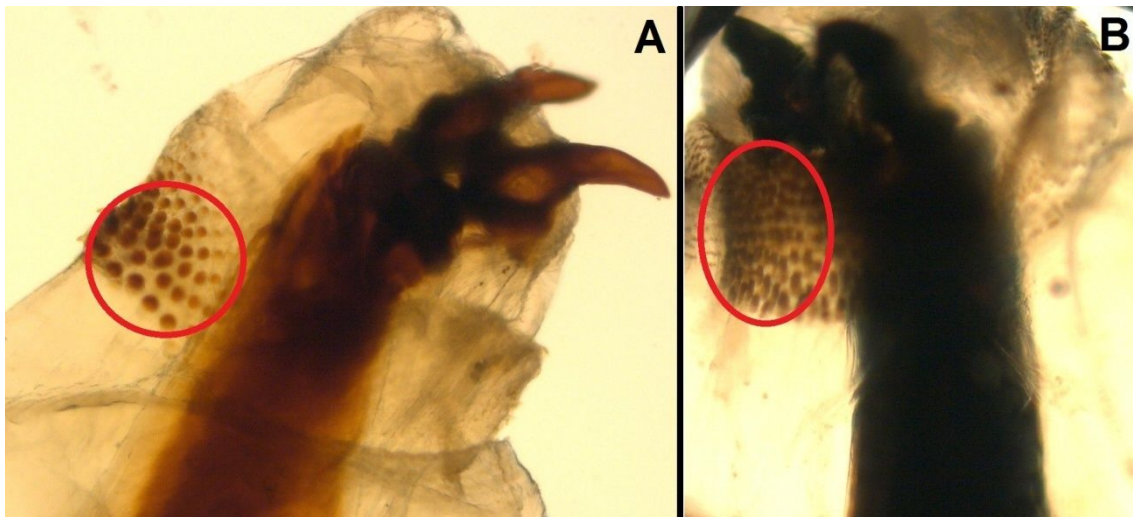


Figura 3: Imagens do crânio das larvas necrofágicas focando a abertura anterior no crânio. A- *Chrysomya*; B - *Lucilia*.

Olhando de outro ângulo e com uma escala de 10x, observa-se que os pontos circulares na região anterior de ambos os gêneros contêm certo relevo, mas em *Chrysomya* são mais pontiagudos (Figura 4A) do que em *Lucilia* (Figura 4B). Ocorre certa variação na cor desses pontos quando há mudança de luminosidade, esta mudança de luz pode ser resultado do flash da câmera.

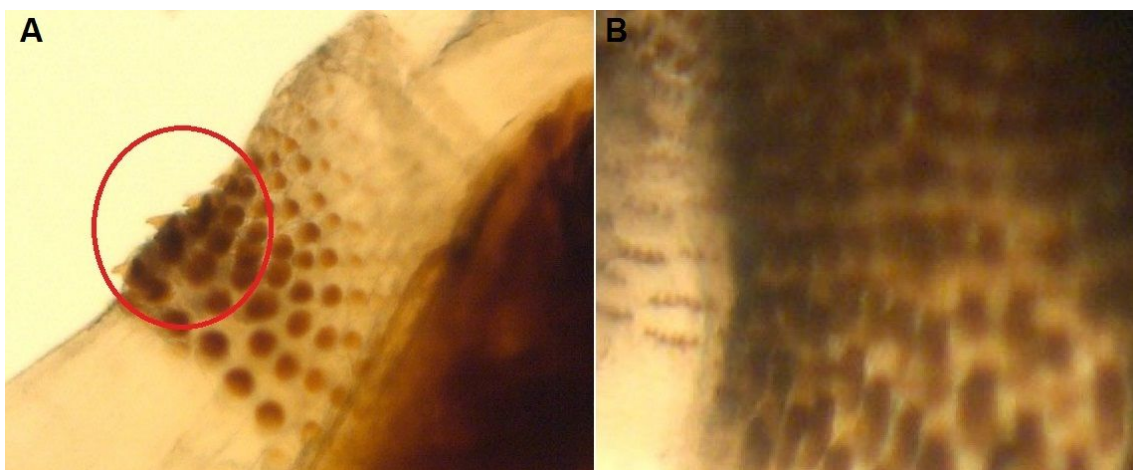


Figura 4: Imagens do crânio das larvas necrofágicas focando pequenas projeções pontiagudas. A- *Chrysomya*; B - *Lucilia*.

Em ambos os gêneros há uma mancha escura no centro da larva que também muda de tonalidade ao se observar várias imagens, em *Chrysomya* há

uma variação de largura (Figura 5-A), enquanto em *Lucilia* esta mancha apresenta uma forma mais reta do início ao fim (Figura 5-B).

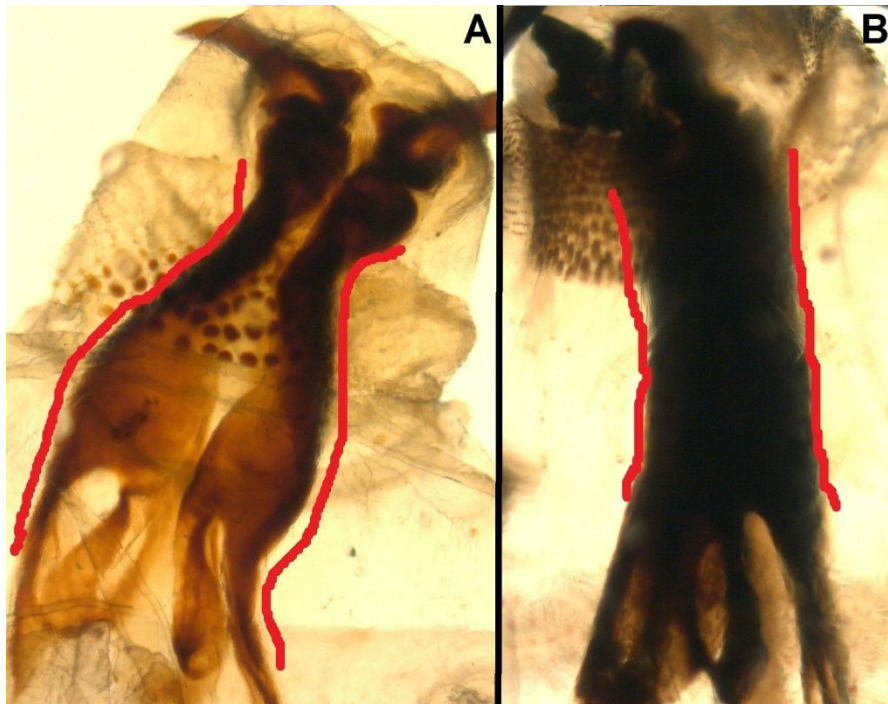


Figura 5: Imagens do crânio das larvas necrofágicas focando pequenas projeções pontiagudas. A- *Chrysomya*; B - *Lucilia*.

4. Metodologia

Para cada um dos objetivos específicos listados na Seção 2.2, serão apresentados a seguir os aspectos metodológicos que nortearão a execução desta proposta.

4.1. Aprofundamento e atualização da revisão de literatura

Através de consultas aos principais portais de periódicos mundiais, como IEEE Xplore, ACM DL, Science Direct e Scopus, serão identificados artigos com trabalhos correlatos nas áreas de Entomologia Forense, Larvas

Necrofágicas e Extração de Atributos. Estes artigos serão revisados para complementar o texto apresentado neste plano de trabalho.

4.2. Criação de banco de imagens de larvas necrofágicas

Para realizar os testes necessários para o módulo de Reconhecimento das larvas, será criado um novo banco de imagens com as estruturas das larvas, como espiráculos e espinhos em diferentes instares. Este banco conterá imagens de diferentes gêneros de larvas, capturadas com o microscópio digital LCD Micro Bresser com lentes de aumento de 5X. O banco de imagens será validado e aprimorado para constituir uma referência para identificação de larvas necrofágicas.

4.3. Desenvolvimento do módulo de Reconhecimento das estruturas morfológicas das larvas

O software será desenvolvido em Linguagem JAVA. Serão seguidas as regras definidas pelo grupo de pesquisa e desenvolvimento INOVISAO disponíveis no site do grupo¹. A metodologia de desenvolvimento de software do INOVISAO tem como base o SCRUM (SIMS; JOHNSON, 2011) com todo o material produzido sob controle de versões utilizando a ferramenta GitLab². O padrão de documentação de código é baseado no JavaDoc.

4.4. Validação do módulo de identificação

Os algoritmos escolhidos e implementados serão comparados entre si e também em relação ao desempenho humano utilizando o banco de imagens descrito na seção 4.2, serão convidados três peritos criminais para realizar o teste que consiste em avaliar o tempo e os acertos comparando a identificação dos peritos com a identificação realizada pelo software. Para cada algoritmo testado, serão calculados os desempenhos médios referentes às métricas de

¹O site do INOVISAO está em www.gpec.ucdb.br/inovisao e as instruções para desenvolvedores pode ser acessada através do link “trac”, neste mesmo site, ou diretamente em trac.gpec.ucdb.br.

²O software de controle de versões GitLab é apresentado em <http://git.inovisao.ucdb.br/>.

precisão, abrangência, medida-F e taxa de acerto ajustadas para problemas com mais de duas classes quando necessário. Para identificar se os algoritmos testados diferem estatisticamente em relação ao desempenho, considerando-se cada uma das métricas, serão utilizados o teste não-paramétrico proposto por Friedman (1940) e análise de variância (ANOVA), ambos disponíveis no software estatístico R³, versão 2.14.1, com cada bloco correspondendo a uma das classes do problema. Serão reportados os valores-p encontrados para cada métrica e o nível de significância necessário para descartar a hipótese nula. Para todas as métricas será realizado um pós-teste e os diagramas de caixa e valores-p dois a dois resultantes serão analisados. O pós-teste, também disponível no R, tem como base o teste de Wilcoxon com correção para FWER (Family-wise Error Rate) descrito por Hollander e Wolf (1999).

4.5. Registro e divulgação de resultados

Serão produzidos um resumo, um relatório final e no mínimo um artigo científico com resultados finais deste plano. O artigo será submetido para um evento ou revista da área de Visão Computacional. Será utilizada a ferramenta Latex⁴ para produção dos textos visando facilitar a adaptação dos mesmos para as regras utilizadas em periódicos e eventos da área da computação e que geralmente disponibilizam modelos em Latex.

5. Cronograma

	2015					2016							
	08	09	10	11	12	01	02	03	04	05	06	07	08
1.1. Identificação dos trabalhos correlatos.	X	X	X	X	X	X	X	X	X	X	X	X	X
1.2. Estudar técnicas de extração de atributos.	X	X	X	X	X	X	X	X	X	X	X	X	X
1.3. Revisão dos trabalhos correlatos	X	X	X	X	X	X	X	X	X	X	X	X	X

3O software R está disponível em <http://www.r-project.org/>. Para o teste de Friedman com análise post-hoc é necessário instalar o

4O editor de textos Latex é livre e gratuito e pode ser obtido em <http://www.latex-project.org/>

2.1. Validação do banco de imagens.	X	X												
2.2. Preparação do ambiente e captura de imagens	X	X	X											
2.3. Criação do novo banco de imagens.	X	X	X											
3.1. Preparação do banco de imagens		X	X	X										
3.2. Implementação do módulo Reconhecimento das estruturas morfológicas das larvas		X	X	X	X	X								
3.3. Documentação dos códigos gerados.				X	X	X								
4.1. Realização validação do módulo					X	X	X							
4.2. Ajustes dos parâmetros baseando no resultado da validação						X	X	X						
4.3. Realização dos testes								X	X	X	X			
5.1. Preparação do artigo, relatório final e resumo.										X	X	X	X	

6. Referências Bibliográficas

AMENDT J, KRETTEK R, ZEHNER R (2004) **Forensic entomology**. Naturwissenschaften91, 51-65.

ANERSON GS, VANLAERHOVEN SL (1996) **Initial studies on insect succession on carrion** in southwestern British Columbia. Journal of Forensic Science 41, 617–625.

BAHRI, A.; ZOUAKI, H. A **Surf-Color Moments For Images Retrieval Based On Bag-Offeatures**. European Journal of Computer Science and Information Technology. v. 1, p.11-22, 2013.

BALLARD, DANA HARRY: **Computer Vision**, Prentice Hall, 1982.

BAY, H.; TUYTELAARS, T.; GOOL, L.V. SURF: **Speeded Up Robust Features**. EuropeanConferenceon Computer Vision, v. 1, p. 404–417, 2006.

BRADSKI, G. The OpenCV Library.**Dr. Dobb's Journal of Software Tools**, 2000.

CERIGATTO, W. **Análise Faunística de Dípteros Necrófagos:Ecologia e Aplicação Forense**, 2009.

CSURKA, G.; DANCE, C.; FAN, L.; WILLAMOWSKI, J.; BRAY, C. Visual categorization with bags of keypoints. In: Workshop on Statistical Learning for Computer Vision, p. 59–74, 2004.

FRIEDMAN, M. A comparison of alternative tests of significance for the problem of m rankings. **The Annals of Mathematical Statistics**, v. 11, n. 1, p. 86–92, 1940.

HALL, M.; FRANK, E.; HOLMES, G.; PFAHRINGER, B.; REUTEMANN, P.; WITTEN, I. H. The WEKA Data Mining Software: An Update. **SIGKDD Explorations**, v. 11, n. 1., 2009.

HALL RD (2001) **Perceptions and status of forensic entomology** In: Byrd JH, Castner JL (eds) **Forensic entomology: the utility of arthropods** in legal investigations. CRC, Boca Raton, Fla., pp 1–15.

HOLLANDER, M.; WOLF, D. A. *Nonparametric Statistical Methods. 2nd Edition.* New York: John Wiley & Sons, 1999.

LIMA, L. **Aplicação de Visão Computacional para Detecção de estruturas de identificação de larvas de moscas de importância forense.** 2014.

LOWE, D. G. **Object recognition from local scale-invariant features.** Computer Vision, 1999. The Proceedings of the Seventh IEEE International Conference on.v.2, p. 1150-1157, 1999.

MCKNIGHT BE (1981) **The washing away of wrongs: forensic medicine** in thirteenth-century China. University of Michigan, Ann Arbor.

NASCIMENTO, J (2003): **Análise e Classificação de Imagens Baseadas em Características de Textura Utilizando Matrizes de Coocorrência.**

OLIVEIRA, W. **Identificação de DNA Vertebrado a Partir do Conteúdo Gastrointestinal de Larvas de Dípteras Necrofágicas.** Dissertação de Mestrado, Universidade Católica Dom Bosco. Campo Grande, 2014.

SIMS, C.; JOHNSON, H. L. **The Elements of Scrum.** Dymaxicon, 2011.

SMITH KGV (1986) **A Manual of Forensic Entomology**, Cornell University Press, Ithaca, NT, 205 pp.

SPEIGHT MR, HUNTER MD, WATT AD (2008) **Ecology of insects: concepts and applications.** Wiley-Blackwell, Oxford.

STEIN, T. **Avaliação de Descritores de Textura para Segmentação de Imagens.** Dissertação de Mestrado, Universidade Federal do Paraná. Curitiba, 2005.

WINSTON, P.H., **The Psychology of Computer Vision**, McGraw Hill, 1975